AIR-FUEL RATIO CONTROL OF A SPARK IGNITION ENGINE

An Honors Thesis

Presented to

the Faculty of the Department of Mechanical Engineering

University of Houston

In Partial Fulfillment

of the Requirements for the Degree of

Bachelor of Science

in Mechanical Engineering

by

Jalal Yazji

May 2019

AIR-FUEL RATIO CONTROL OF A SPARK IGNITION ENGINE

Jalal Yazji

Approved:

Chair of the Committee
Karolos Grigoriadis, Professor,
Mechanical Engineering

Committee Members:

Jagannatha Rao, Associate Professor and Associate Chair,
Mechanical Engineering

Chad A.B. Wilson, Associate Instructional Professor,
Engineering Undergraduate Programs

Suresh K. Khator, Associate Dean,
Cullen College of Engineering

Pradeep Sharma, Professor and Chair,
Mechanical Engineering

AIR-FUEL RATIO CONTROL OF A SPARK IGNITION ENGINE

An Abstract

of an

Honors Thesis

Presented to

the Faculty of the Department of Mechanical Engineering

University of Houston

In Partial Fulfillment

of the Requirements for the Degree of

Bachelor of Science

in Mechanical Engineering

by

Jalal Yazji

May 2019

Abstract

Lean-burn operation in spark-ignition engines has shown promise in improving fuel economy and reducing harmful emissions in comparison with traditional stoichiometric operation. Close reference-tracking of the set air-fuel ratio profile is very crucial to healthy engine operation. What makes air-fuel ratio control challenging is the presence of a large variable time delay in the system's closed-loop, resulting mainly from the large distance traveled by the air-fuel mixture between the injection point and the exhaust. This thesis proposes modifications to an IMC-Smith predictor design employed to control the air-fuel ratio in a lean-burn engine. Matlab's Simulink provides a convenient platform to build dynamic models and simulate controllers, and for that reason, it is chosen to validate the proposed controller design and compare its performance to that of a PI controller and that of an IMC-Smith controller. Simulation results reveal the inadequacy of a basic PI controller in providing good reference tracking to a lean-burn profile. The proposed design shows very similar performance to a basic IMC-Smith controller in terms of overshoot and disturbance. However, its reduced settling time in comparison with the IMC-Smith controller (difference of up to 1.5s) renders it a more effective design at providing the desired level of reference-tracking.

Table of Contents

Table of Figures

Chapter 1: Introduction

Global concerns regarding the greenhouse effect, ozone layer depletion, and weather disasters continue to attract attention worldwide. Vehicle emissions are one of the largest contributors to environmental pollution. An ideal internal-combustion engine in a car would burn fuel to emit nothing more than water vapor and carbon dioxide, but practically, traditional engines are far from ideal and produce additional harmful substances through their exhaust gas. The rapid increase in car usage over the past century in the United States amplified pollution and raised public and governmental concerns calling for efforts to limit this pollution. Congress founded the Environmental Protection Agency (EPA) in 1970, with the aim of regulating polluting emissions in automobiles and limiting harm to the environment. In addition to pollution, high fuel consumption is another problem with internal combustion engines; in 2003, the United States consumed nearly 20 million barrels of oil per day. Gasoline consumed by cars and trucks constituted about 45% of total oil consumption [12].

The need to reduce emissions and improve fuel economy prompted car manufacturers to improve their engine design to try solving these two major challenges. The geometry of the engine internals and the functionality of the fuel injection system and exhaust system strongly dictate how efficiently the combustion process occurs, and their designs are constantly being modified. Moreover, the introduction of digital control and simulation to engine design has allowed for better visualization of engine dynamics and improved performance. Exploring novel and improved methods for controlling engine fuel consumption and emissions continues to be the subject of active research.

1.1 Outline of the Thesis

This thesis has been divided into seven chapters. An outline with a brief description of each chapter is presented in this section.

Chapter 2 is mainly concerned with introducing the reader to engine chemistry to establish the background of the air-fuel ratio control problem. It starts by explaining the defining features of engine categories and sub-categories in terms of functionality and application. The chapter narrows the focus on spark ignition engines and discusses their operating cycle. It continues to introduce the air-fuel ratio and discuss its correlation with a healthy engine operation. The chapter also lists the main engine pollutants and explores their formation conditions and negative effects on health and the environment. Applications for exhaust gas treatment are then discussed, with more focus on catalytic conversion. The chapter concludes with introducing lean-burn technology and its relevance to fuel economy and emission control.

Chapter 3 defines the air-fuel ratio control problem addressed by the thesis. The chapter starts with a qualitative description of the control objective followed by identifying the engine system to be controlled with all its components. The chapter then offers a dynamic modelling of the system with feedback.

Chapter 4 targets the problem of time delay in engines and investigates its adverse effects on the proper control of the air-fuel ratio. The chapter starts with an overview of time delay and continues to identify the causes of delay in the engine system. A root locus analysis is then performed using a first order Pade approximation to explain the destabilization effect that delay may induce on first order systems.

Chapter 5 describes in terms of dynamics and performance a few conventional control techniques, which have been previously utilized for the air-fuel ratio control problem. The chapter starts with an introduction of proportional-integral-derivative controls and their effects on closed loop dynamics. The internal model control principle is then described and explored as a method for tuning a proportional-integral-derivative controller. Time delay compensation is addressed in this chapter by introducing the Smith predictor technique. In addition, a link is established between this form of delay compensation and the internal model principle.

Chapter 6 proposes a new control design inspired by the techniques highlighted in the previous chapter. The approach for this design is discussed and the evaluation criteria are detailed. The chapter then presents the results in the form of three computer simulations to compare the performance of the control design to conventional control techniques previously used. The chapter concludes with a validation discussion of the proposed control design.

Chapter 7 extends discussion of the results and provides some remarks for future research on the air-fuel ratio control problem.

Chapter 2: Engine Chemistry and Emissions

2.1 Background

Our present-day transportation system composed of land vehicles, marine vessels, and aircraft, derives power mainly from converting chemical energy to mechanical energy through the burning of fossil fuels. This process takes place in what is commonly known as combustion engines. Though combustion engines may take different shapes and sizes and use different kinds of fuel, they can be divided into two main families: internal combustion and external combustion [11]. In external combustion engines, the fuel burns in a chamber outside the engine, and hence the name: external combustion. A working fluid acquires some heat from the combustion process, and transports it to a mechanism in the engine, where a portion of the heat is converted to usable mechanical energy [11]. On the other hand, in internal combustion engines (ICEs), both fuel ignition and the conversion of heat to mechanical work occur in the combustion chamber [11]. This duality eliminates the need for a heat transferring apparatus or a working fluid and helps reduce the size of the engine. Although there are many stationary applications for ICEs, mobile applications, such as automobiles and airplanes, are more dominant [11].

ICEs can be further divided into three categories: spark ignition engines (SI) and compression ignition engines (CI), which are both mainly used in land vehicles; and gas turbines, which can be found in aircraft and power generation applications [3]. Both SI and CI engines are considered "reciprocating engines," where the power conversion process is cyclic, and fuel combustion happens within specific cycles. In contrast, gas turbines are of the "continuous combustion" type [11].

Although SI and CI engines share many similarities, they have defining features related to operation and mechanism. The following list explains the main differences:

(1) Introduction of fuel: In SI engines, fuel and air are mixed together using a device called the carburetor to form a gaseous mixture, which is then injected into the combustion chamber through the throttle. CI engines do not have a carburetor, and the mixing of air with fuel happens in the combustion chamber rather than prior to injection. A fuel control valve regulates the amount of injected fuel [11].

(2) Ignition mechanism: SI engines require an electrical mechanism called the spark plug to initiate fuel combustion, and hence the name: spark ignition. Due to the high air compression ratios in CI engines, the combustion chamber reaches conditions of pressure and temperature large enough to combust fuel with no need for a special mechanism [11].

(3) Compression ratio: Compression ratios range from 5 to 10.5 for SI engines, and 12 to 20 for CI engines [11].

(4) Weight: Due to the high pressures generated within CI engines, their structure must be strong enough to withstand those pressures, which makes them considerably heavier than SI engines [11].

2.2 Operating Cycle of Spark Ignition Engines

SI engines may operate on a two-stroke power cycle or the more commonly-used four-stroke power cycle, which is illustrated in Figure 2.1. A typical four-stroke cycle starts with the piston descending inside the combustion chamber and drawing in the air-fuel mixture through the throttle (intake) valve. The intake valve is then closed, and the piston rises, compressing the mixture, following which, the spark plug is fired to ignite air and fuel together. The mixture expands upon combustion, which drives the piston back down, delivering power to the vehicle.

The cycle ends with the exhaust valve opening to eject the combustion products out of the cylinder, allowing for another quantity of air-fuel mixture to enter and repeat the cycle [3].

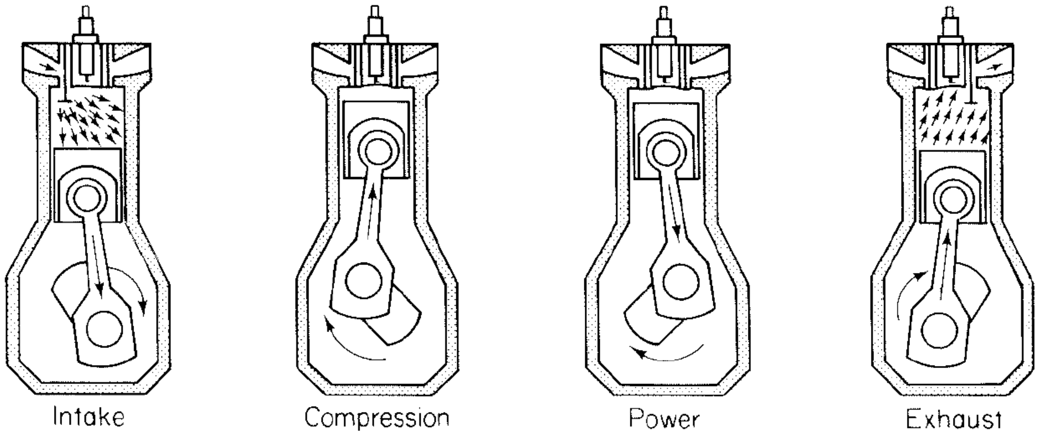


Figure 2.1: Four-stroke cycle of spark ignition engines [3].

2.3 The Air-Fuel Ratio

The composition of the air-fuel mixture in a SI engine can be described using a single parameter called the air-fuel ratio, defined as

$$AFR = \frac{A}{F}, \tag{2.1}$$

where A and F are the masses of air and fuel respectively. A more conventional description of the mixture composition compares the actual air fuel ratio to the stoichiometric ratio (at which, there is just enough oxygen to completely react with all the fuel present) and defines what is known as the relative AFR [5]:

$$\lambda = \frac{(A/F)_{actual}}{(A/F)_{stoich.}}. \tag{2.2}$$

Controlling the relative AFR is fundamental to a healthy engine operation, and many control strategies have been employed for that matter as will be discussed later in this thesis. A relative AFR value close to unity describes a near-stoichiometric engine operation. Smaller values describe a rich engine operation where there is excess fuel, while larger values describe a lean engine operation, where there is excess oxygen. Depending on the engine type and operating condition, the goal may be to set λ to either one of the three possible states.

2.4 Pollutants

The fossil fuels used in internal combustion engines are mixtures of different hydrocarbon compounds and are usually around 86 percent carbon and 14 percent hydrogen [5]. Hydrocarbons are known to combust with oxygen under the appropriate conditions. The combustion process is a fast exothermic reaction, and for a hydrocarbon fuel with an average molecular composition C_aH_b, the overall complete combustion equation with air [5] (assumed to be 1 part oxygen and 3.773 parts nitrogen) is

$$C_aH_b + \left(a + \frac{b}{4}\right)(O_2 + 3.773N_2) \longrightarrow aCO_2 + \frac{b}{2}H_2O + 3.773\left(a + \frac{b}{4}\right)N_2. \tag{2.3}$$

Equation (2.3) assumes stoichiometry and does not account for any incomplete reactions taking place. In practice, air and fuel inside the combustion chamber do not react fully as described by the ideal stoichiometric model, and other incomplete side reactions may take place producing unwanted substances, which are explored in this section.

2.4.1 Effect of Air-Fuel Ratio on Pollutant Formation

An important fact to note about Equation 3 is that there are no other reactants besides fuel and oxygen involved in the process, and thus, the air-fuel ratio places a large effect the overall efficiency of the process. Figure 2.2 illustrates how emission concentrations change for different compositions of the air-fuel mixture. Near stoichiometry, an upward trend can be observed in the concentrations of carbon monoxide and hydrocarbon emissions as the mixture moves from lean to rich in fuel. Nitric oxide concentration, however, decreases as the mixture becomes richer, near stoichiometry.

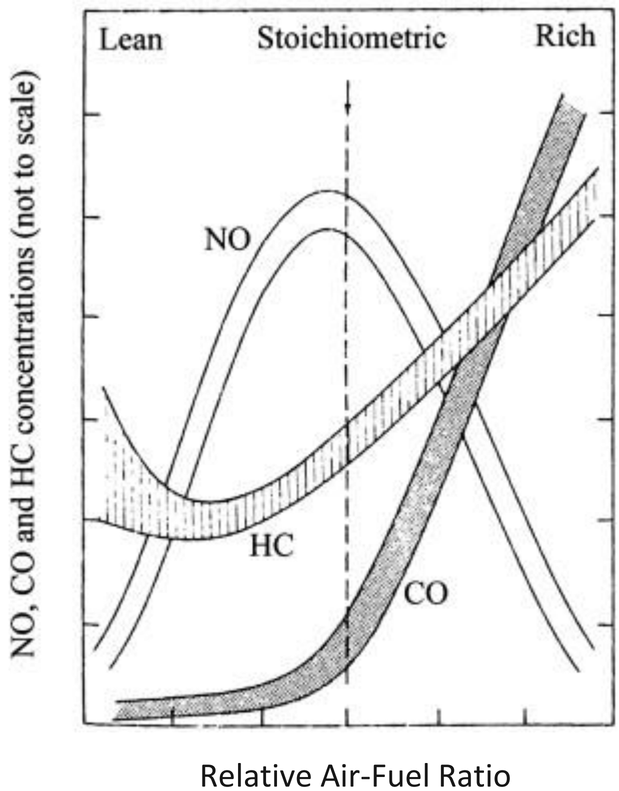


Figure 2.2: Variation in emission concentrations with the relative air-fuel ratio [5].

2.4.2 Nitrogen Oxides

Nitrogen oxides (NO_x) are contributors to the formation of ground-level ozone and fine particles, which are associated with many negative health effects. Additionally, high concentrations of these oxides can be harmful to vegetation.

The high pressure and temperature conditions in SI engines during combustion trigger the oxidation of nitrogen to form both nitric oxide (NO) and nitrogen dioxide (NO_2). NO is more predominant and constitutes about 98% of nitrogen oxide emissions. In addition to the air-fuel ratio, spark timing and the burned gas fraction of the unburned mixture both affect the formation rate of nitrogen oxides [5].

2.4.3 Carbon Monoxide

Hemoglobin, which is the principal oxygen-carrying substance in blood, has a much stronger chemical affinity for carbon monoxide (CO) than for oxygen (O_2). So when CO is inhaled, it bonds with hemoglobin, inhibiting the transport of O_2 through blood, and consequently leading to what is known as carbon monoxide poisoning, which can be fatal. CO is also a contributor to the greenhouse effect and climate change.

Like NO, CO also forms during the high temperature and pressure conditions of combustion in SI engines. Rich air-fuel mixtures yield higher concentrations of CO as there is not a sufficient amount of oxygen present to fully burn all the fuel and produce the far less harmful carbon dioxide (CO_2) [5].

2.4.4 Unburned Hydrocarbons

Hydrocarbons (HC) contribute to the formation of ground-level ozone and are linked with many health problems. The presence of unburned HC in the exhaust gas of an SI engine is linked to an incomplete combustion of fuel. The air-fuel ratio largely affects the concentrations of HC emissions. However, other factors related to combustion and fuel injection dynamics have been identified: (1) the escape of fuel into crevice volumes in the cylinder which helps shield some of this fuel from the propagating flame during combustion; (2) the absorption of fuel vapors by oil layers in the cylinder; (3) misfire, which is defined as the failure of the spark plug to ignite the air-fuel mixture during the combustion stroke [5]. Proper control of the air-fuel ratio and the optimization of spark timing and cylinder geometry are crucial for countering the problem of incomplete fuel combustion.

2.5 Exhaust Gas Treatment

Harmful exhaust emissions can be removed through the oxidation of CO and HC and the reduction of NO_x. Adjustment of the general operation of an engine to favor emission reduction is insufficient and can adversely affect performance and efficiency. Thus, an external mechanism with emission-removal capabilities must be utilized. Two main applications have been employed on SI engines for exhaust gas treatment: thermal reactors and catalytic converters [3].

2.5.1 Thermal Reactors

The oxidation of CO and HC continues to occur as they are passing through the exhaust path but at a much slower rate due to rapid cooling. Oxidation times can be elongated through the installation of a thermal reactor, which is an enlarged exhaust manifold that attaches directly to the cylinder head. Thermal reactors attempt to maintain the exhaust gases at high temperatures for an extended time window, and by that, rapid oxidation can extend to the exhaust stage and more CO and HC can be converted [5]. The challenge with thermal reactors is their inability to effectively reduce NO_x emissions. The introduction of ammonia as a reducing agent has been proposed. However, non-catalytic reduction by ammonia only works for a narrow temperature window, and the proper control of ammonia flow remains challenging, which places many limitations on the application of thermal reactors for exhaust gas treatment [3].

2.5.2 Catalytic Converters

The use of catalysts allows for a successful conversion of emissions at lower temperatures than those in the combustion chamber environment. The catalytic convertor is an exhaust-treatment device that consists of an active catalytic material (usually a noble metal or a metal oxide) housed in a special casing that is specifically designed to direct the exhaust gas through the catalyst bed for treatment [5]. The design of a pellet-type converter is illustrated in Figure 2.3. Catalytic performance is highly dependent upon the concentrations of oxygen and the different emissions in the exhaust gas, or more broadly, the air-fuel ratio set to the engine. Oxidation of CO and HC species is favored in lean operation when the exhaust gas is rich in oxidizing agents, while the reduction of NO_x is favored in rich operation when the mixture is rich in reducing agents (such as CO) [3].

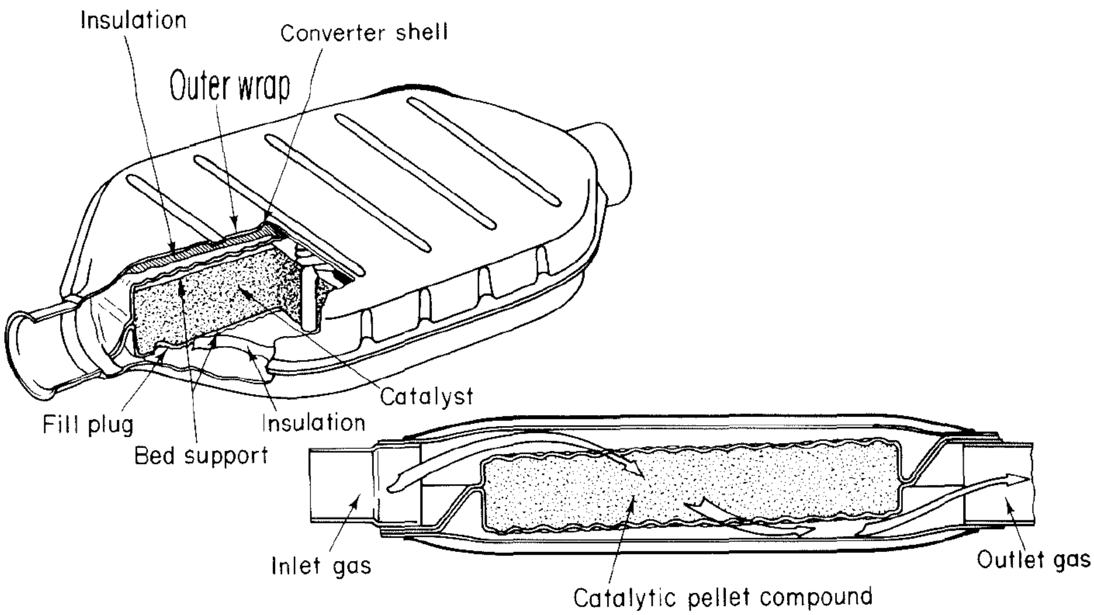


Figure 2.3: Pellet-type catalytic converter [3].

Catalytic converters can take one of two schemes: the dual-bed converter and the three-way catalyst. The dual-bed scheme requires rich engine operation to successfully reduce NO_x, while external air is added to oxidize CO and HC. Although this application can convert all three pollutant species, the requirement of running rich operation reduces engine efficiency appreciably. Three-way catalyst technology eliminates this problem by running the engine at stoichiometry, which was proven effective at removing all three pollutants simultaneously, and hence the name: three-way catalyst or TWC. Figure 2.4 demonstrates TWC's heightened effectiveness at stoichiometric point.

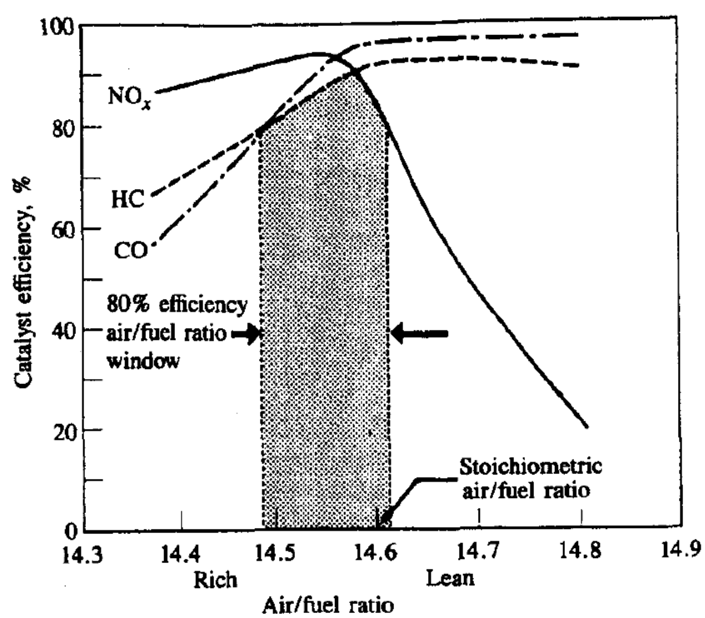


Figure 2.4: Emission conversion efficiency of TWC [5].

2.6 Lean-Burn Technology

SI engines made compatible with lean-burn operation offer improved fuel economy and emit far less CO and HC, as compared with conventional SI engines, which operate at stoichiometric point. Lean-burn engines may operate at air-fuel ratios as large as 65:1. The challenge that comes with lean-burn engines is their increased NO_x emissions as the TWC loses its NO_x-reduction efficiency with leaner air-fuel mixtures. To counter this problem, a device called the lean NO_x trap or LNT, has been developed and installed downstream of the conventional TWC. LNT absorbs and sotres NO_x while the engine is operating lean until the concentration threshold is met, at which point, the engine switches briefly to rich operation, and LNT is purged to regenerate its capacity. The released NO_x is then reduced by the rich air-fuel mixture. Controlling the storage and purge cycles of LNT is crucial for a healthy engine operation and requires precise feedback control [6].

Chapter 3: Control Problem Identification

3.1 Control Objective

The goal of the ongoing research on air-fuel ratio control for lean-burn technology is to develop a controller that can closely track an air-fuel ratio profile for a spark ignition lean-burn engine. The reference air-fuel ratio is set to achieve optimum three-way catalyst (TWC) and lean-NOx trap (LNT) emission-reduction performance.

3.2 System Description

The system to be controlled is characterized by the dynamics of the air-fuel mixture path. As presented in Figure 3.1, the air-fuel ratio of the mixture is set by the fuel injector. After the combustion process, the air-fuel exhaust travels a long distance to reach the exhaust gas treatment segment composed of both the TWC and LNT modules. A universal exhaust gas oxygen (UEGO) sensor reports to the controller a measurement of the relative air-fuel ratio λ downstream of the LNT module and directly before the tailpipe. The controller in turn regulates the amount of fuel fed by the injector, and thus the air-fuel ratio of the mixture entering the engine [12]. The path of the mixture between the injector point and the UEGO sensor point downstream of the engine constitutes open-loop dynamics, while the reading signal communicated by the UEGO sensor back to the controller constitutes the feedback path, and both paths collectively form the closed-loop system.

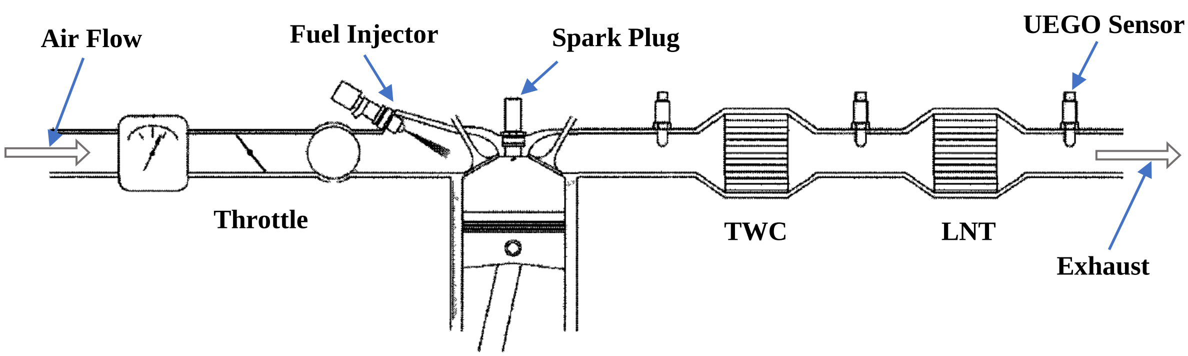


Figure 3.1: Air-fuel path (distance between engine components not to scale) [13].

3.3 Modeling Engine Dynamics

The general structure of the engine feedback control system is shown in Figure 3.2. The controller input is the error signal, or the difference between the air-fuel ratio reading and the desired value. The output of the controller serves as input to the engine system, which may also be subject to disturbance inputs such as fuel injector uncertainty and canister purges.

The UEGO sensor has a highly nonlinear overall profile and loses sensitivity in regions away from the stoichiometric point. Near the stoichiometric region, however, the sensor readings follow a near-linear profile, so a linear model is assumed in this thesis for simplicity [4].

The engine system undergoes complex chemical reactions which may magnify the complexity of the dynamics. However, it is proven experimentally that simple first order lag of the form

$$G(s) = \frac{1}{\tau_s s + 1}, \tag{3.1}$$

is valid to estimate the complex engine dynamics [8]. In (3.1) τ_s is the engine time constant, which depends on the engine speed. The system is not limited to this first order transfer function and is usually accompanied by a variable time delay, which yields the delayed system transfer function

$$G(s) = \frac{1}{\tau_s s+1} e^{-\theta s}, \tag{3.2}$$

where θ is the varying delay value. The modelling and effects of time delay on the engine system will be explored thoroughly in the following chapter.

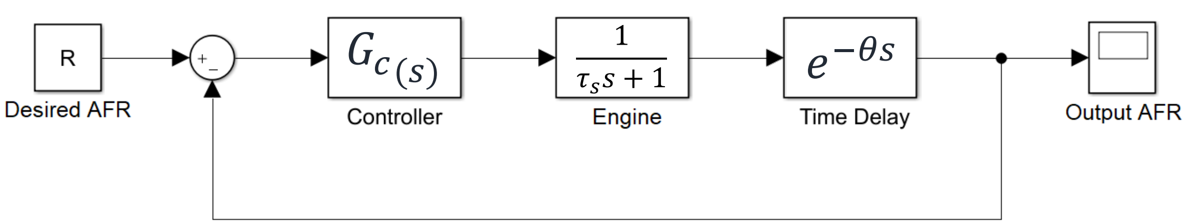


Figure 3.2: General diagram of the engine feedback control system.

Chapter 4: Time Delay Analysis

4.1 Background

Time delay in control systems is defined as the time spent to acquire information needed for decision-making, to process this information, and then to make and execute decisions based on the acquired information [9]. Time delay can have adverse effects on dynamic systems as it affects the damping characteristics and may make an otherwise stable system unstable [8].

All physical sensors, actuators, and controllers exhibit delay to a certain extent: Sensors take time to collect data and transmit signals; controllers take time to process sensor data and send control signals; and actuators take time to perform their functions to shape the system dynamics. This time lag results in different components acting on signals or commands corresponding to a past time, and not to the present state of the system.

In systems that combine a complex arrangement of such components, the effect of delay becomes more apparent as the overall delay combines the different delays produced by the individual components in the system [9]. Power transmission systems, chemical processes, networks, and teleoperation systems are all examples of systems with a considerable level of delay [8].

4.2 Time Delay Modelling in Spark Ignition Engines

The overall time delay in the SI engine system is composed of two main parts: gas transport delay and cycle delay [2]. Gas transport delay is defined as the time required for the exhaust gas to reach the tailpipe UEGO sensor after having passed through the engine and any other equipment installed downstream of the engine (TWC or LNT, for example) [2]. This delay can be approximated by

$$\tau_g = \frac{\alpha}{\dot{m}}, \tag{4.1}$$

where $\dot{m}$ is the air mass flow rate, and α is a constant that is determined experimentally [2].

The engine itself contributes to what is known as cycle delay, which is estimated by one engine cycle due to the four strokes in operation [2]. This delay can be approximated by

$$\tau_c = \frac{720}{\left(\frac{360}{60}\right)N} = 120/N, \tag{4.2}$$

where N is the engine speed in rpm [2]. The overall delay is the sum of the two delays

$$\tau = \tau_g + \tau_c. \tag{4.3}$$

Because this delay depends on the engine's operating condition, it varies with time. Engine speeds typically vary between 600 to 6000 rpm and thus the delay value may vary by a factor as large as 10 over an operating cycle in the engine [4].

4.3 Stability of Delay-Free Systems

Because the SI engine system can be modeled as a simple first order system, it can be shown that a feedback loop with no time delay can never have instability regardless of the controller gain. The overall transfer function of the system with respect to the reference input is

$$\frac{Y(s)}{R(s)} = \frac{k * \frac{1}{\tau_s + 1}}{1 + k * \frac{1}{\tau_s + 1}}, \tag{4.4}$$

where k is the proportional controller gain and τ_s is the engine time constant. The corresponding characteristic equation is

$$\tau_s + 1 + k = 0, \tag{4.5}$$

which has a single pole that is of the form

$$p = -\frac{1 + k}{\tau_s}. \tag{4.6}$$

This pole is always negative (given that the controller gain is positive), and thus, the stability condition is met regardless of the controller gain. Figure 4.1 provides a graphical illustration of this condition-independent stability by plotting the root locus of the delay-free engine and showing that all the poles exist in the left half plane.

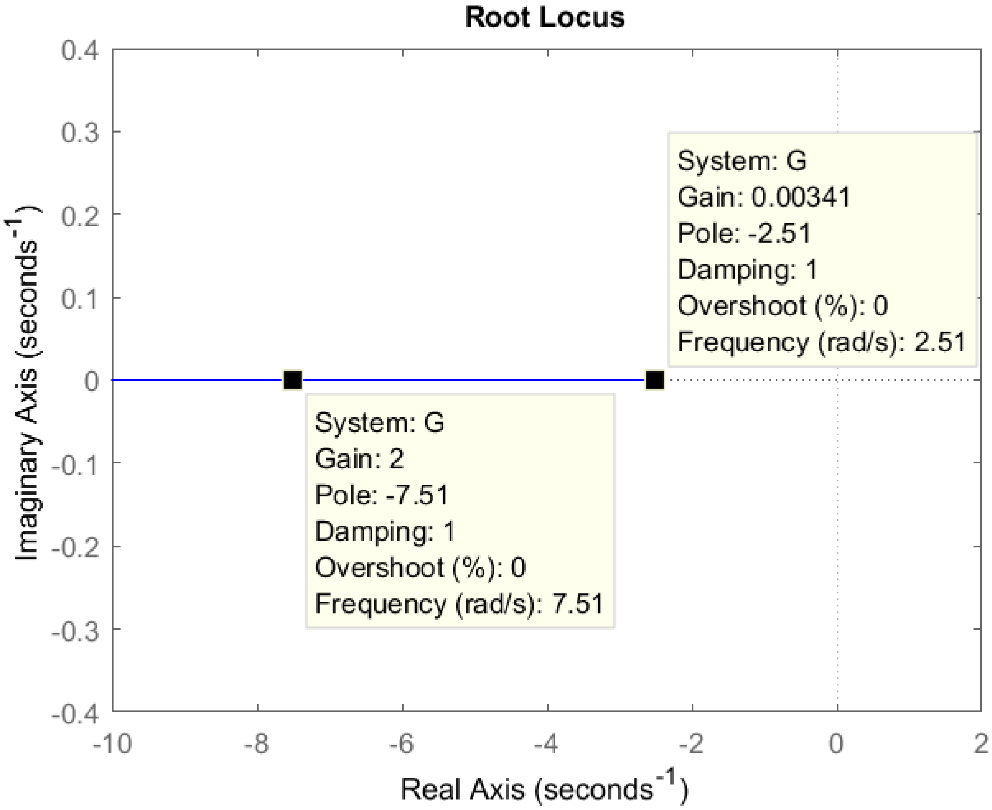


Figure 4.1: Root locus plot of the delay-free engine system.

4.4 Analytical View of Delay-Induced Instability

Although time delay in the engine system is variable, a simple analysis using pure unvarying time delay can be performed to demonstrate the effect of a non-varying pure delay on system poles. The analysis requires linearizing the delay term in the engine system. For this purpose, a first order Pade approximation [7] is used to estimate delay, which is presented as

$$e^{-\theta s} \cong \frac{-\frac{\theta}{2}s+1}{\frac{\theta}{2}s+1}. \tag{4.7}$$

Substituting this expression into the overall feedback transfer function yields

$$\frac{Y(s)}{R(s)} = \frac{k * \frac{1}{\tau_s + 1} * \frac{-\frac{\theta}{2}s + 1}{\frac{\theta}{2}s + 1}}{1 + k * \frac{1}{\tau_s + 1} * \frac{-\frac{\theta}{2}s + 1}{\frac{\theta}{2}s + 1}}, \tag{4.8}$$

which can be further simplified to the form

$$\frac{Y(s)}{R(s)} = \frac{k * (-\frac{\theta}{2}s + 1)}{\frac{\tau_s\theta}{2}s^2 + \left(\tau_s + \frac{\theta}{2} - \frac{\theta}{2}k\right)s + 1 + k}. \tag{4.9}$$

From the denominator of the transfer function, the characteristic equation of this transfer function will be

$$\frac{\tau\theta}{2}s^2 + \left(\tau_s + \frac{\theta}{2} - \frac{\theta}{2}k\right)s + 1 + k = 0\,, \tag{4.10}$$

which is a second order equation, for which the poles are of the form

$$s = \frac{-\theta \pm \sqrt{(\theta(k-1) - 2\tau_s)^2 - 8\theta(k+1)\tau_s} - \mathrm{k}\theta + 2\tau_s}{2\tau_s\theta}. \tag{4.11}$$

A simple substitution of $\theta = 2\ seconds$ and k = 1.5 yields two poles with imaginary parts and positive real parts

$$p_{1,2} = 0.1250\ \pm\ 2.4969i\,. \tag{4.12}$$

The presence of positive real parts corresponds to instability in the output and shows time delay's ability to destabilize the system.

4.5 Computational View of Delay-Induced Instability

To further demonstrate the instability generated by time delay in the engine system, a root locus plot was produced using Matlab as shown in Figure 4.2, for delay values ranging from 0.3 to 2.7 seconds, with 0.1-second time intervals. A first order Pade approximation is used to approximate delay. As can be seen from Figure 4.2, the marginal-stability controller gain k, decreases with increasing time delay, which means that higher delay values place more stability restrictions on a given controller.

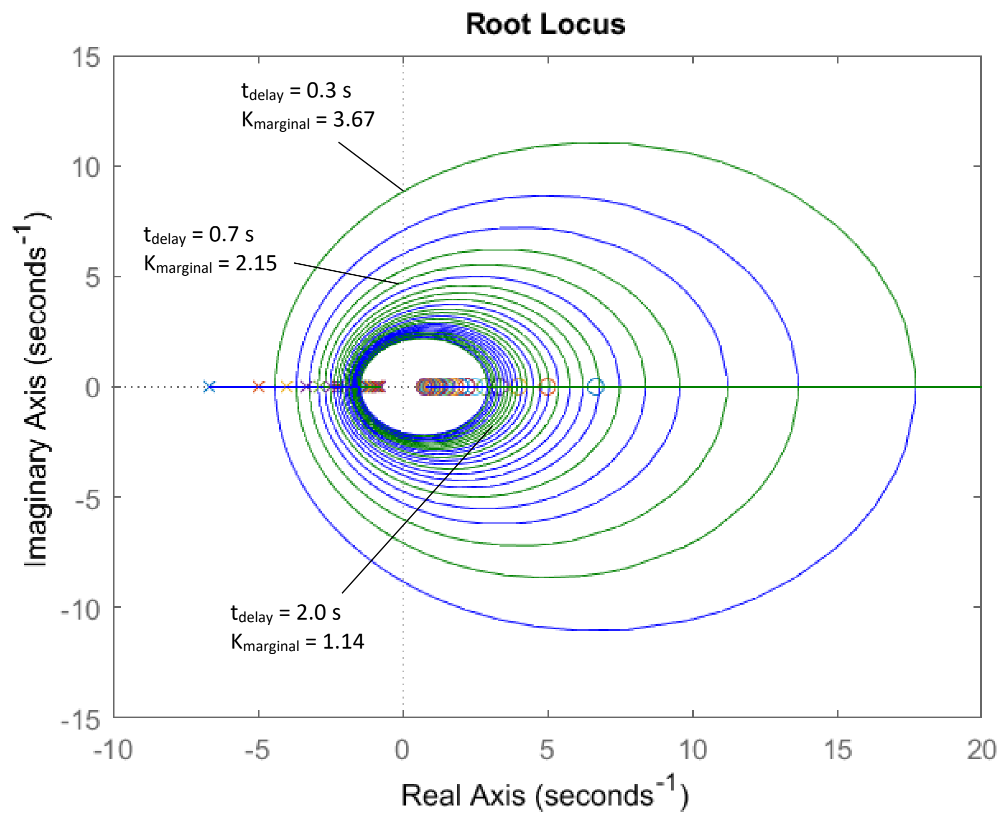


Figure 4.2: Root locus plot of the engine system with delay values ranging from 0.3 to 2.7 seconds.

Chapter 5: Overview of Control Techniques

5.1 Proportional-Integral-Derivative Control (PID)

Proportional-integral-derivative (PID) control is arguably the most common type of control in the process and robotics industries. Though it is considered relatively simple, PID control has proven effective when used with low order and non-varying systems in terms of both cost reduction and control performance. Another advantage of PID control is that it does not require a full understanding of the controlled plant, as the performance is dictated by how well PID gains are tuned [1]. The development of more advanced computing software, such as Simulink, made PID simulating and tuning far easier than before, which further popularized the use of this control technique.

5.1.1 Transform Equation

The basic structure of PID control is [4]

$$u(t) = k_P e_{(t)} + k_I \int_0^t e(\tau) d\tau + k_D \frac{de(t)}{dt}, \tag{5.1}$$

where e is the error as a function of time, k_P is the proportional gain, K_I is the integral gain, and k_D is the derivative gain [4]. In the Laplace domain, PID can be written as the transfer function

$$D(s) = k_P + \frac{k_I}{s} + k_D s\,. \tag{5.2}$$

An equivalent but more practical representation of the PID transfer function is of the form

$$D(s) = k_P \left[1 + \frac{1}{\tau_I s} + \tau_D s\right], \tag{5.3}$$

where τ_I is the integral term time constant, while τ_D is the derivative term time constant. This expression is more representative of the physical structure of PID control, which will be more fully explained in a later section of this chapter.

Figure 5.1 illustrates the structure of PID control as a block diagram. Upon entering the PID controller, the error signal $e_{(t)}$ diverges into the three branches of the controller to be processed, and following which, the signals are united again into one signal forming the controller output, which then feeds into the plant as input $u_{(t)}$.

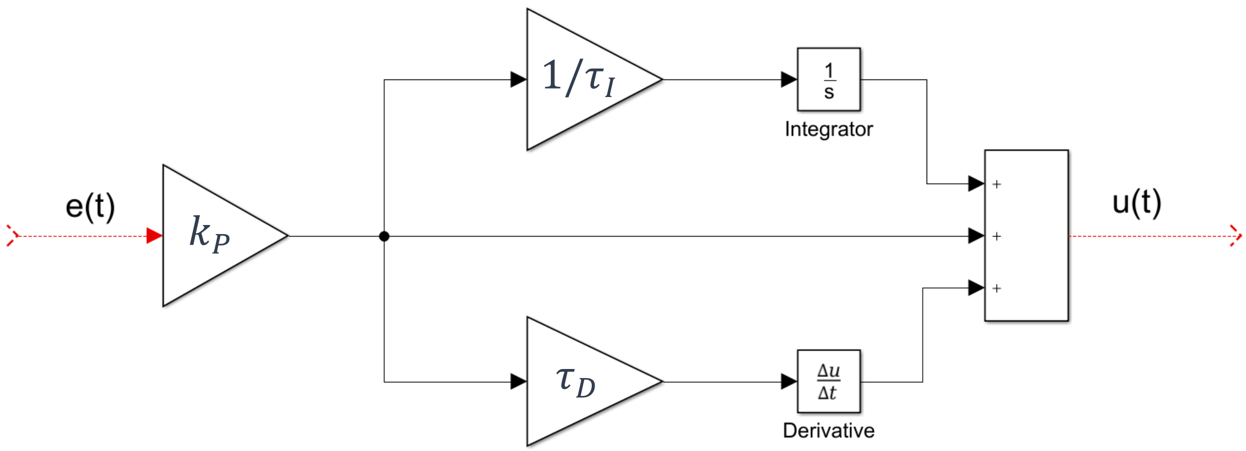


Figure 5.1: PID control diagram.

5.1.2 Dynamics

PID may not always appear with all three active branches. Partial combinations like the proportional-integral (PI) or the proportional-derivative (PD) may suffice depending on the application. To help illustrate the functionality of each of the three branches of PID, the effect of some combinations on a second order system of the form

$$G(s) = \frac{A}{s^2 + a_1 s + a_2} \tag{5.4}$$

is analyzed here. This analysis is followed by an overview of their effects when combined.

Proportional control (P), as the name suggests, multiples the error signal by a constant, and the result is fed back into the plant [4]. Proportional feedback with the second order transfer function yields the following characteristic equation

$$s^2 + a_1 s + a_2 + k_P = 0 \,, \tag{5.5}$$

which shows that the proportional gain manipulates the constant term, allowing the designer to control the natural frequency but not the damping of the system [4]. A higher proportional gain helps to reduce steady-state error but may produce unwanted transient oscillations [4]. This deficiency renders proportional control alone inadequate for many applications.

Integral control (I) works to integrate the error signal over the time period that begins upon process initiation. For this reason, integral control is said to keep record of the history of the plant. P and I together form the commonly-used proportional-integral control (PI). Applying PI feedback control to the transfer function (5.4) gives the following third order characteristic equation [4]

$$s^3 + a_1 s^2 + a_2 s + Ak_P s + Ak_I = 0 \,. \tag{5.6}$$

The P and I control gains make possible the manipulation of the coefficient of s and the constant term, but not the coefficient of s^2, which requires the addition of the derivative term [4].

In contrast with integral control, which records the past of the error signal, derivative control (D) differentiates the error signal and by that, predicts the future behavior of the error. The characteristic equation that gathers all three PID components is of the form [4]

$$s^3 + (a_1 + Ak_D)s^2 + (a_2 + Ak_P)s + Ak_I = 0 \,, \tag{5.7}$$

which grants the designer freedom to manipulate the coefficients of s and s^2 in addition to the constant term.

To further illustrate the dynamics of the PID controller, the studied second-order transfer function is placed in a feedback schematic with a PID control and a reference step input. The plant is subjected to a disturbance input in the form of a 0.35-magnitude step that activates 10

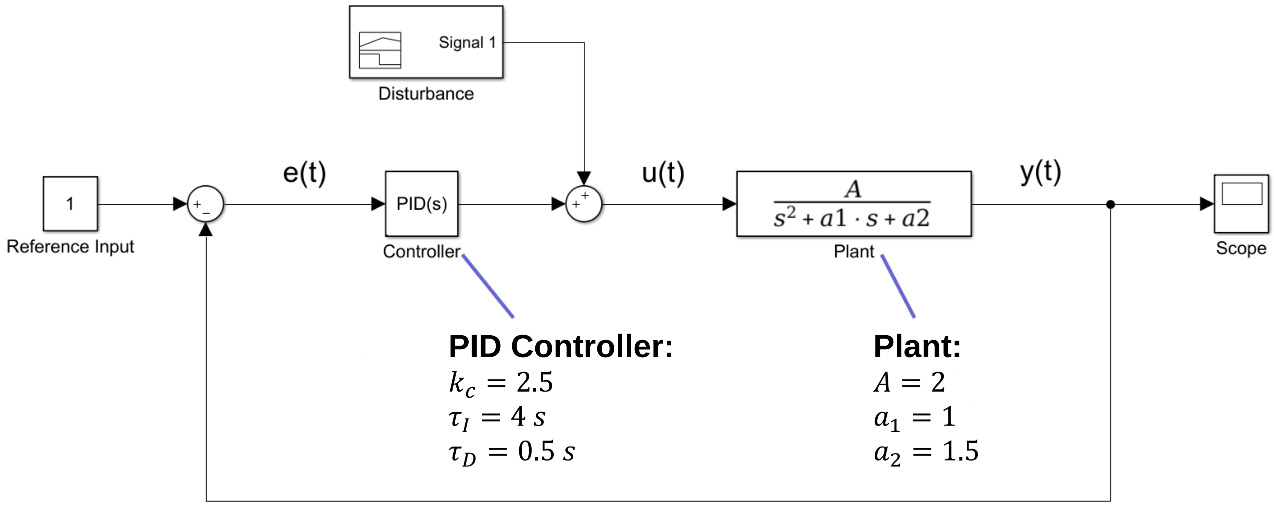


Figure 5.2: Second order plant with PID feedback control.

seconds after process initiation. The block diagram and the selected plant parameters and tuned control gains are presented in Figure 5.2. The output response in Figure 5.3 displays the PID control's ability to achieve good reference tracking and disturbance rejection. Methods for tuning PID gains to achieve optimum control will be discussed later in this chapter.

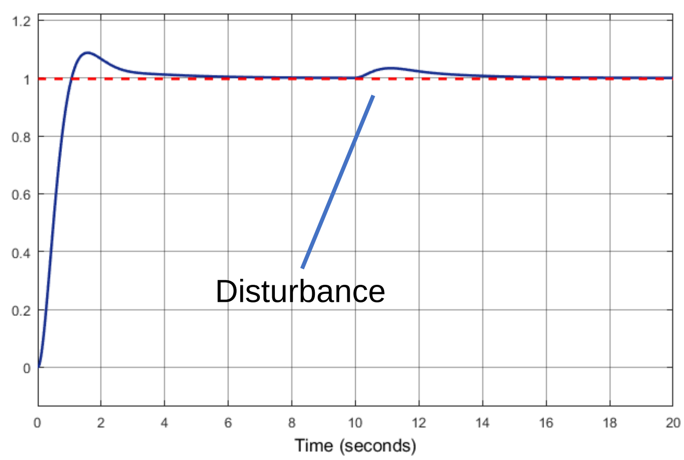


Figure 5.3: Sample output response with tuned PID control.

The signals exiting each of the three PID branches behave differently. Figure 5.4 illustrates the behavior of each control signal with time as the error signal decays to zero. In steady state, it is important to note that the control outputs from both P and D reach zero, as does the error signal. However, this is not the case with the I signal, which settles at a non-zero constant value that results from integrating the error over time. This phenomenon explains the integral control's ability to eliminate the effect of disturbances in steady state, as it can manipulate its control output to absorb disturbances and bring the plant output back to reference.

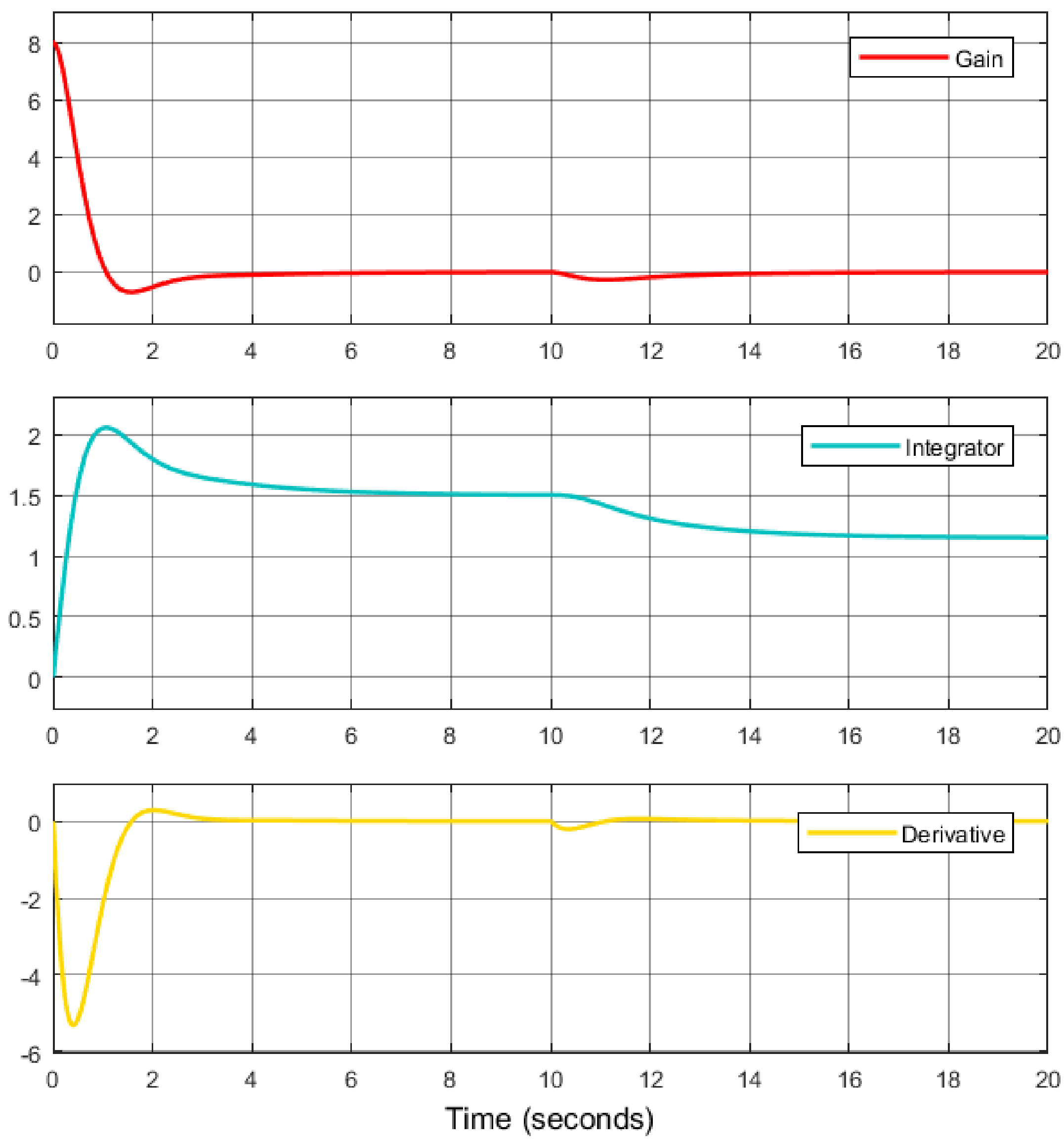


Figure 5.4: Signal outputs of tuned PID controller.

5.2 Internal Model Control (IMC)

The concept behind internal model control (IMC) is based on the internal model principle, which states that "control can be achieved only if the control system encapsulates, either implicitly or explicitly, some representation of the process to be controlled" [10]. If, for example, the designer managed to acquire full knowledge of the process being controlled so as to build a controller that exactly reverses the effect of the process, as shown in Figure 5.5, then the output will always match the reference input at any point in time without the need to collect any

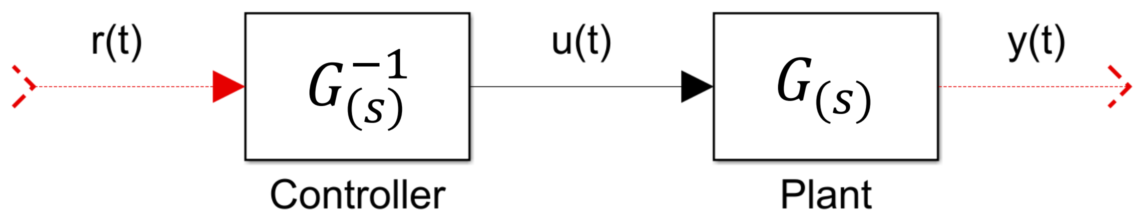


Figure 5.5: Ideal internal model open loop control.

feedback from the output [10]. However, this ideal scenario is strictly hypothetical as it is difficult to predict a flawless representation of any particular process. And even if the exact representation of that process was attainable, shaping a controller around a reversed process may create realizability issues (due to producing a transfer function whose order of the numerator is higher than that of the denominator). Consequently, realistic models are flawed, and feedback is required to correct for deviations [10].

5.2.1 Dynamic Model

IMC control aims to estimate the process $G_{(s)}$ with an internal model $\tilde{G}_{(s)}$ and then compare their outputs and use the difference as a feedback error signal. The controller $Q_{(s)}$ is derived based on the internal model and will be discussed in this section. A block diagram representation of IMC control is presented in Figure 5.6.

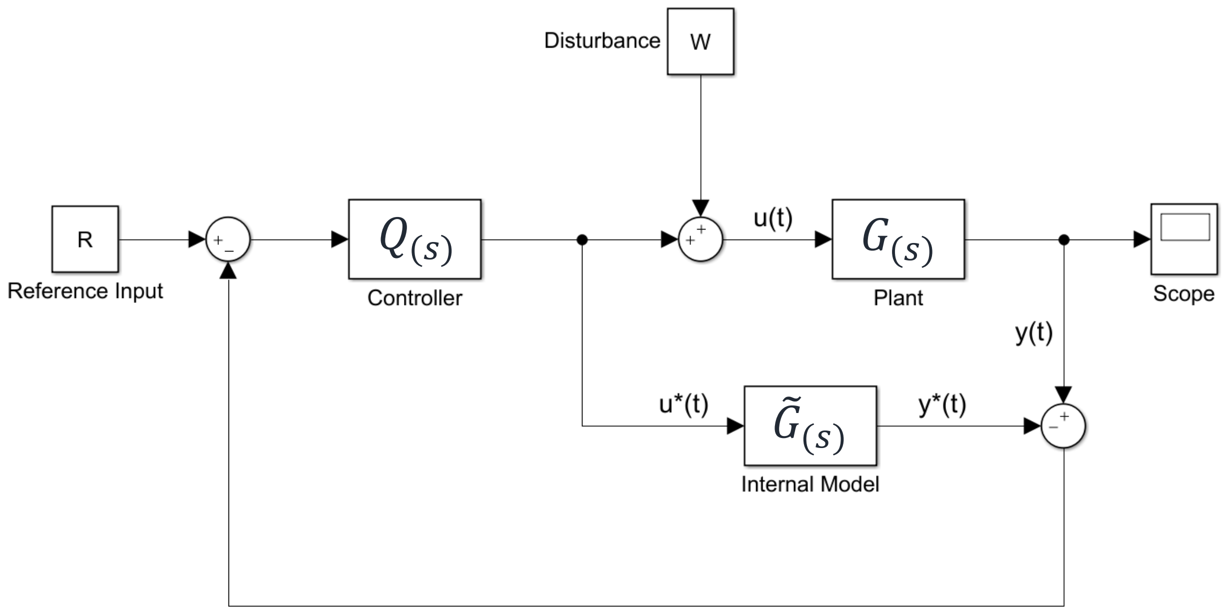


Figure 5.6: IMC control block diagram.

To better understand what constitutes a good choice for the control, the output can be derived in terms of the system transfer functions, the reference input, and the disturbance. The physical plant has the transfer function

$$G(s) = \frac{Y(s)}{U(s)}, \tag{5.8}$$

while the model is of the form

$$\tilde{G}(s) = \frac{Y^*(s)}{U^*(s)}. \tag{5.9}$$

The plant input is the sum of the model input and the disturbance, so the plant transfer function can be rewritten as

$$G(s) = \frac{Y}{U^* + W}. \tag{5.10}$$

The input to the controller is composed of the plant and model outputs, and the reference input, and thus the controller's transfer function takes the following form

$$Q(s) = \frac{U^*}{R - Y^* - Y}\,. \tag{5.11}$$

With some algebraic manipulations and simplifications, the plant output can be derived as

$$Y(s) = \frac{R + W\left(\frac{1}{Q} - \tilde{G}\right)}{\frac{1}{G}\left(\frac{1}{Q} - \tilde{G}\right) + 1}. \tag{5.12}$$

In the presence of a reference input, the desire from designing any control is to achieve good reference tracking in the output, so to acquire an expression of how $y(t)$ will behave in steady state, the final value theorem can be applied to (5.12) as

$$y_{ss} = \lim_{t\to\infty} y(t) = \lim_{s\to 0} \mathrm{sY(s)}\,. \tag{5.13}$$

By analyzing the limit, it can be inferred that the steady state value of the output can only go to the reference value, R, if the following equality is true

$$\frac{1}{Q(s)} - \tilde{G}(\mathrm{s}) = 0, for\ s = 0\,. \tag{5.14}$$

Based on this condition, an intuitive approach is to set $Q(s)$ equal to $\tilde{G}^{-1}(\mathrm{s})$ but inverting the fraction turns the poles zeros and the zeros poles, and in the case of positive zeros in $\tilde{G}(\mathrm{s})$, the resultant transfer function for Q(s) becomes unstable. For this reason, any unstable terms are replaced by stable ones but of the same structure to guarantee that (5.14) is met.

5.2.2 IMC as a Method for Tuning PID

The IMC control principle can be regarded as a way to tune PID controllers. Considering, for example, a first order transfer function (resembling the engine transfer function) with a time delay term that is of the form

$$G(s) = \frac{1}{\tau_s + 1} * e^{-\theta s}\,, \tag{5.15}$$

which yields the following internal model function when a first order Pade approximation is applied to linearize the delay term

$$\tilde{G}(s) = \frac{1}{\tau_s + 1} * \frac{-\frac{\theta}{2}s + 1}{\frac{\theta}{2}s + 1}\,. \tag{5.16}$$

To find the appropriate IMC controller, (5.16) is inverted and the positive-zero delay numerator $(-\frac{\theta}{2}s + 1)$ is replaced with a first order stable filter term, producing the following controller

$$Q(s) = (\tau_s + 1) * \frac{\frac{\theta}{2}s + 1}{\tau_F s + 1}\,, \tag{5.17}$$

where τ_F is the time constant of the filter [7]. Before any PID tuning parameters can be derived, the block diagram of the system must be manipulated to look like a conventional feedback diagram with only one controller and one plant transfer function forming the closed loop. An alternative representation of IMC control is shown in Figure 5.7, where the new controller $G_c(s)$ is composed of both $Q(s)$ and $\tilde{G}(\mathrm{s})$ [7], and has the following simplified transfer function

$$G_c(s) = \frac{Q(s)}{1 - \breve{G}(s)Q(s)}\,. \tag{5.18}$$

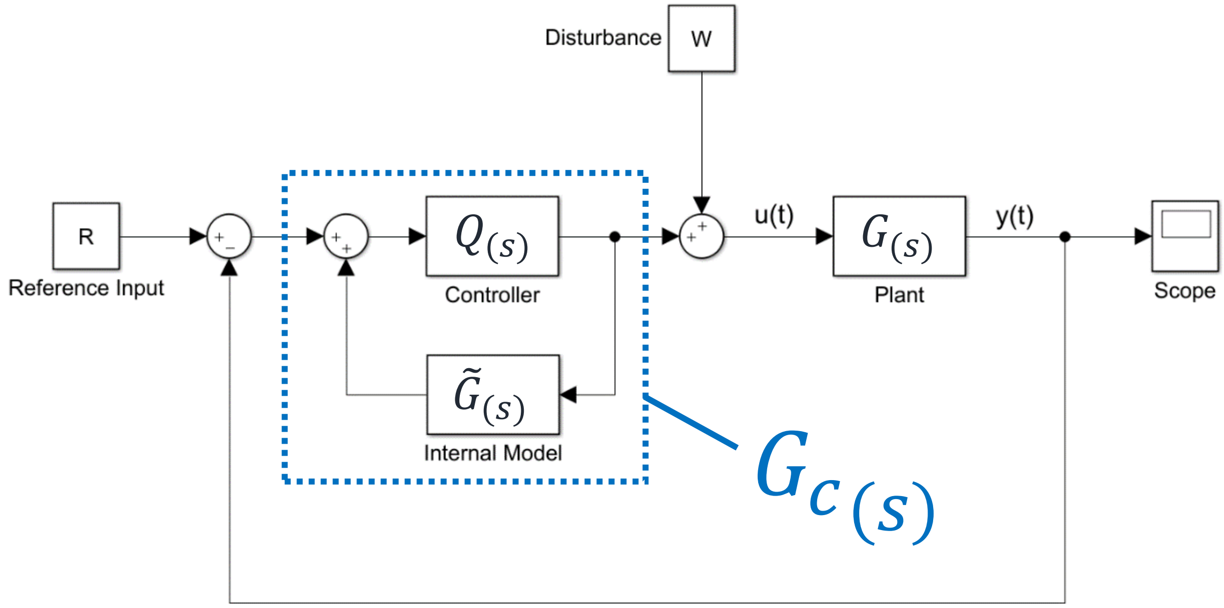


Figure 5.7: Alternative IMC control block diagram.

After substituting the appropriate transfer functions and with some algebraic simplifications, the resultant expression for $G_c(s)$ matches the form of a PID controller

$$G_c(s) = k_P \left[1 + \frac{1}{\tau_I s} + \tau_D s\right], \tag{5.19}$$

where the controller parameters have the following values [7]

$$k_P = \frac{2\tau_s + \theta}{2(\tau_F + \theta)}, \tag{5.20a}$$

$$\tau_I = \tau_s + \frac{\theta}{2}, \text{ and} \tag{5.20b}$$

$$\tau_D = \frac{\tau_s \theta}{2\tau_s + \theta}. \tag{5.20c}$$

5.3 The Smith Predictor

The presence of time delay in closed loop systems often introduces stability and performance issues, and this becomes more severe when the delay value is comparable to the system's time constant. O. J. M. Smith pioneered a new method for delay compensation called the Smith predictor, which is characterized by taking the time delay term outside the closed loop. This modification allows for basing the control design on the plant alone apart from the time delay term [4]. In an ideal setting, where the controlled process and the time delay value are known with great precision, the effect of delay on performance becomes irrelevant.

5.3.1 Dynamics

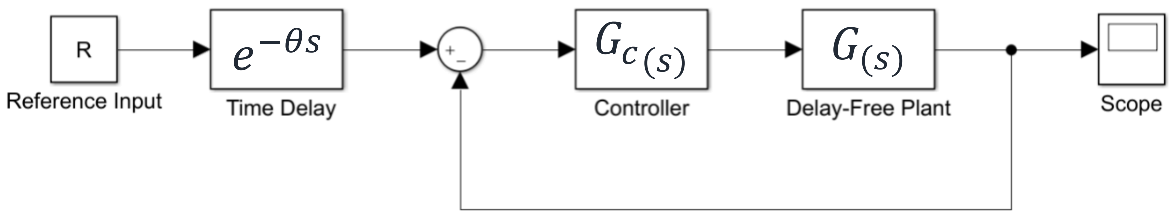
The Smith predictor builds a dummy model of the system (taking $G(s)$ to be the delay-free plant) with its controller, where the time delay term is transferred out of the closed loop and into the reference input [4]. It then transforms this model back to the actual state of the system and modifies the controller accordingly (now called Smith regulator), so that both systems are identical. This transformation is presented in Figure 5.8.

The goal behind this transformation is to equate the response of the dummy model (delay-free closed loop) with the response of the actual system [4] as

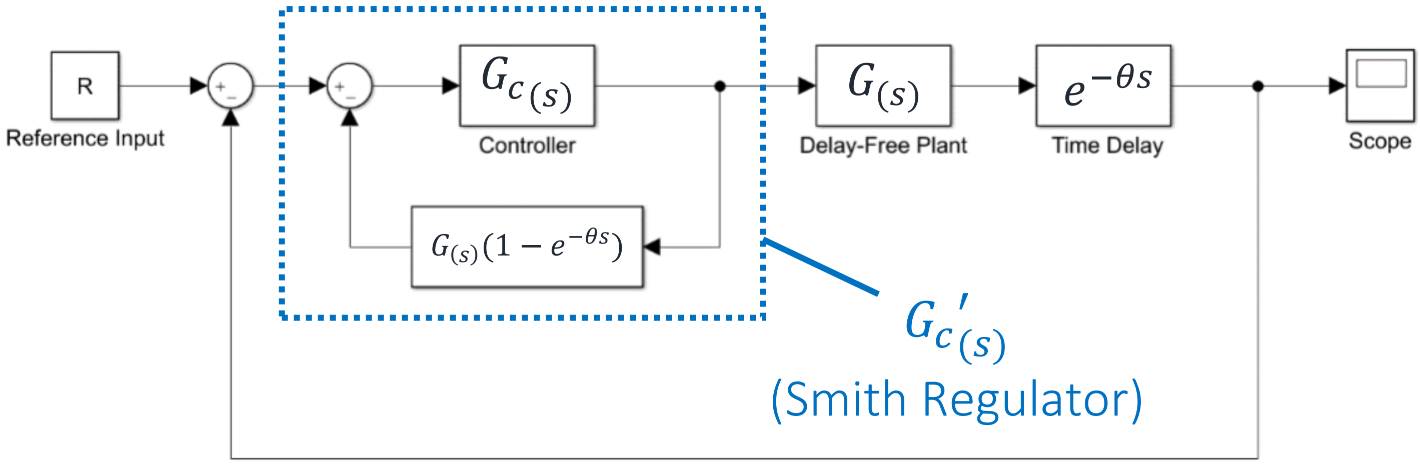
$$\underbrace{\frac{G_c(s)G(s)}{1+G_c(s)G(s)}e^{-\theta s}}_{\text{Dummy Transfer Function}} = \frac{Y(s)}{R(s)} = \underbrace{\frac{G_c{}'(s)G(s)e^{-\theta s}}{1+G_c{}'(s)G(s)e^{-\theta s}}}_{\text{System Transfer Function}}, \tag{5.21}$$

and accordingly, the Smith regulator that can satisfy this condition can be derived [4] as

$$G_c{}'(s) = \frac{G_c(s)}{1+G_c(s)[G(s)-G(s)e^{-\theta s}]}. \tag{5.22}$$



(a)



(b)

Figure 5.8: Smith transformation from the delay-free closed loop (a) to the actual system representation with compensated control (b).

5.3.2 Smith Compensation with the Internal Model

The internal model principle and the Smith predictor are closely tied together. Smith compensation requires collecting information from the output of the plant both with and without the delay. And because time delay is usually introduced by the internal dynamics of the process and/or by the sensor used to communicate the output, acquiring the non-delayed output is not feasible, and building a model estimator (internal model) becomes crucial [4].

A feedback structure with internal model estimation and Smith compensation is presented in Figure 5.9. The diagram may resemble an IMC control structure. However, it is important to note that the controller $G_{c_{(s)}}$ is in part composed of the system's internal model and has an identical from to the one derived in section 5.2.3.

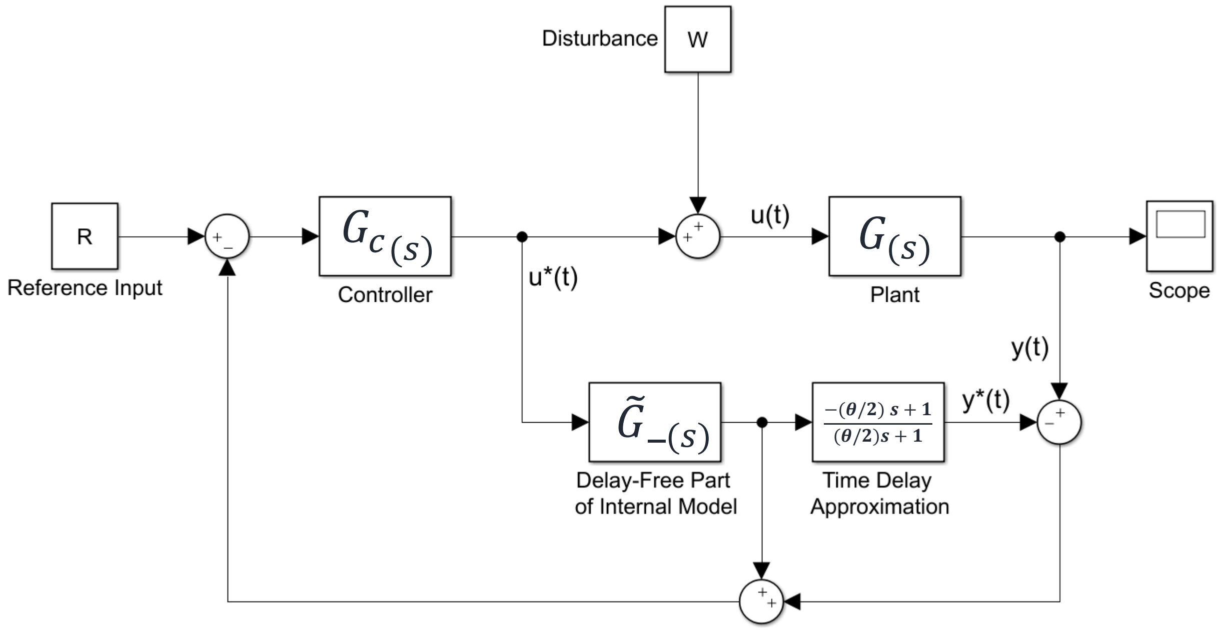


Figure 5.9: IMC with Smith predictor control block diagram.

Chapter 6: Controller Development and Evaluation

6.1 Approach

A conventional IMC controller for a first order system constructed in PID form utilizes two constant parameters: the system delay approximation and the filter time constant. Due to the large variability in the engine system, a satisfactory control performance dictates variability in the control design to compensate for that of the system. Therefore, improving IMC performance requires exploring its effect on the engine at different operating conditions.

The addition of a time delay compensation using a Smith predictor, in theory, helps eliminate the oscillatory transient effects introduced by delay on the AFR output. However, time delay estimation in the engine application comes with uncertainty, preventing the complete elimination of transients. The first order filter helps fill this deficiency while also slowing down the response. The correlation between the time delay value and the function of the filter is explored, and it is found that the larger the time delay value in the close loop, the larger the filter time constant needs to be to suppress transients. A controller design based on the combined IMC and Smith predictor techniques is inspired by this observation. In this design, a direct relationship is established between the time delay estimation and the filter time constant and also the overall control gain, as detailed in the following section.

6.2 Controller Design

The proposed controller follows a similar structure to that of an IMC-tuned PID controller

$$G_c(s) = k_P\left[1 + \frac{1}{\tau_I s} + \tau_D s\right], \tag{6.1}$$

where

$$k_P = k_\theta * \frac{2\tau_s + t_d}{2(\tau_F + t_d)}\ , \tag{6.2a}$$

$$\tau_I = \tau_s + \frac{t_d}{2}, \text{ and} \tag{6.2b}$$

$$\tau_D = \frac{\tau_s t_d}{2\tau_s + t_d}. \tag{6.2c}$$

The overall control gain k_P is scaled by another gain, k_θ, which is equivalent to the value of the estimated time delay saturated between 1.2 and 2. The filter time constant is also made a function of the time delay estimation θ and is expressed as

$$\tau_F = 0.8 * \theta, \text{ saturated between 0 and 1.3.} \tag{6.3}$$

In (6.3) τ_s is the engine's internal model time constant and is set to 0.4, while t_d is a constant that is set to 0.4. A smith predictor structure is also introduced, and the overall feedback system diagram can be seen in Figure 6.1.

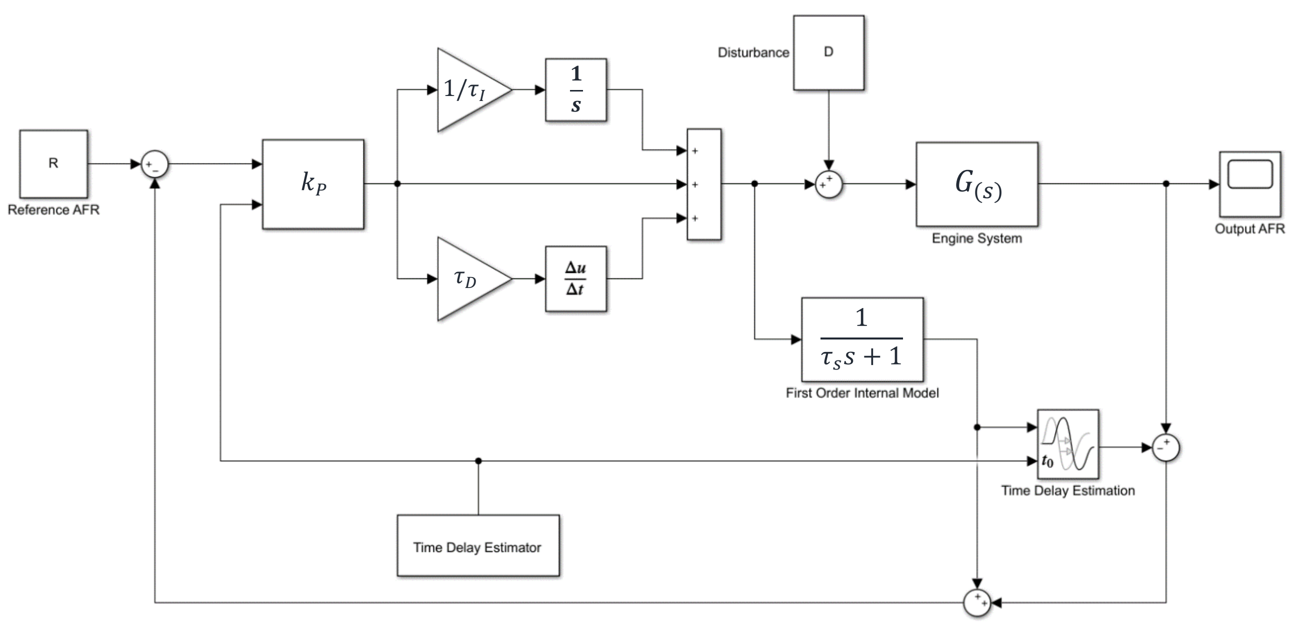


Figure 6.1: Proposed controller diagram.

6.3 Evaluation Parameters

Controller evaluation is performed in Simulink using a model that attempts to mimic the behavior of a lean-burn engine. Close tracking of a command AFR is very critical for engine operation, which makes reference-tracking the main criterion in judging controller performance. The controller must also exhibit robustness, that is, it must maintain satisfactory performance under the wide range of the engine operation. In the simulation, three operating conditions are being varied to test performance: the reference AFR, engine time constant, and time delay value. The reference AFR is modeled for lean-burn operation, where the air-fuel mixture is held lean for a period of time followed by a quick switch to rich conditions and then a return to lean conditions. The engine time constant is varied between three different simulations but is assumed constant over the period of each simulation. A time delay profile based on experimental data is used [8]. The delay estimator used for the proposed controller design is set to have an uncertainty of ±20%. The actual and estimated time delay profiles used in the simulation can be seen in

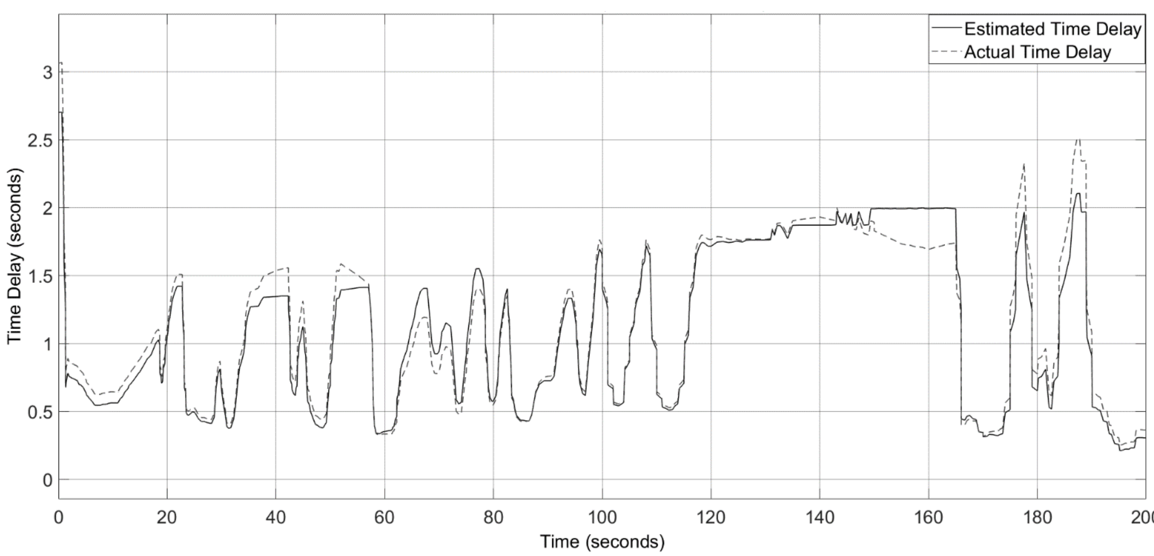


Figure 6.2: Time delay profile used for controller validation.

Figure 6.2. Additionally, the controller's ability to reject disturbances is tested with the disturbance profile shown in Figure 6.3.

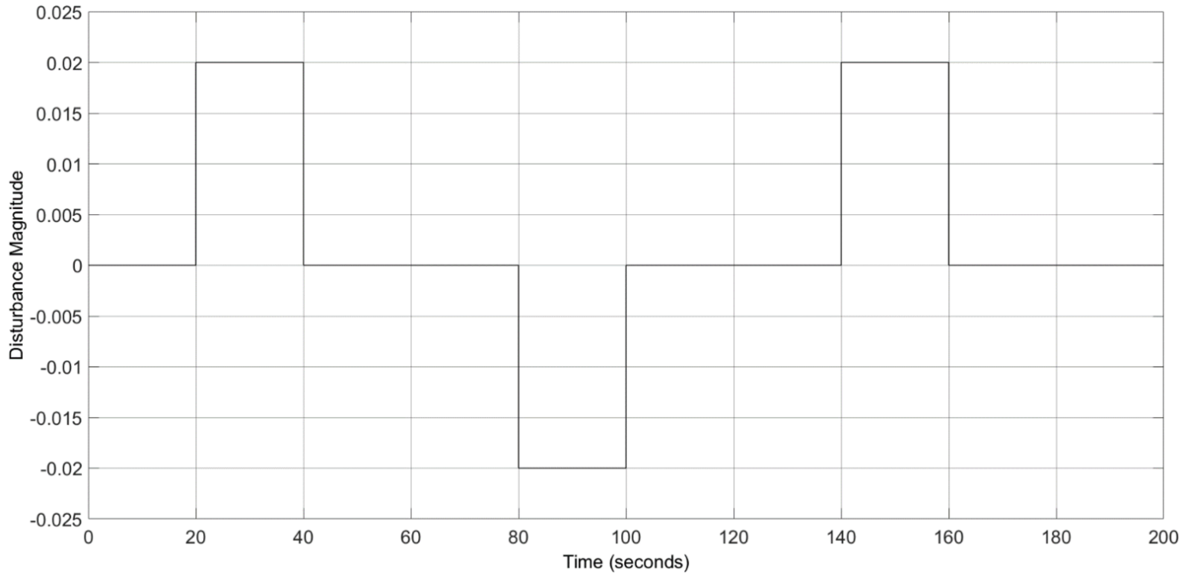


Figure 6.3: Disturbance profile used for controller validation.

In addition, comparison with conventional controllers may be a helpful addition to the evaluation process. For this reason, the performance of the proposed controller is compared with that of an IMC-tuned PID controller paired with a Smith predictor and that of a simple PI controller. Both controllers used for comparison are tuned for optimum performance.

6.4 Results and Discussion

Controller performances are compared in Figures 6.4, 6.5, and 6.6, where the engine time constant is set to 0.3, 0.4, and 0.5 seconds respectively.

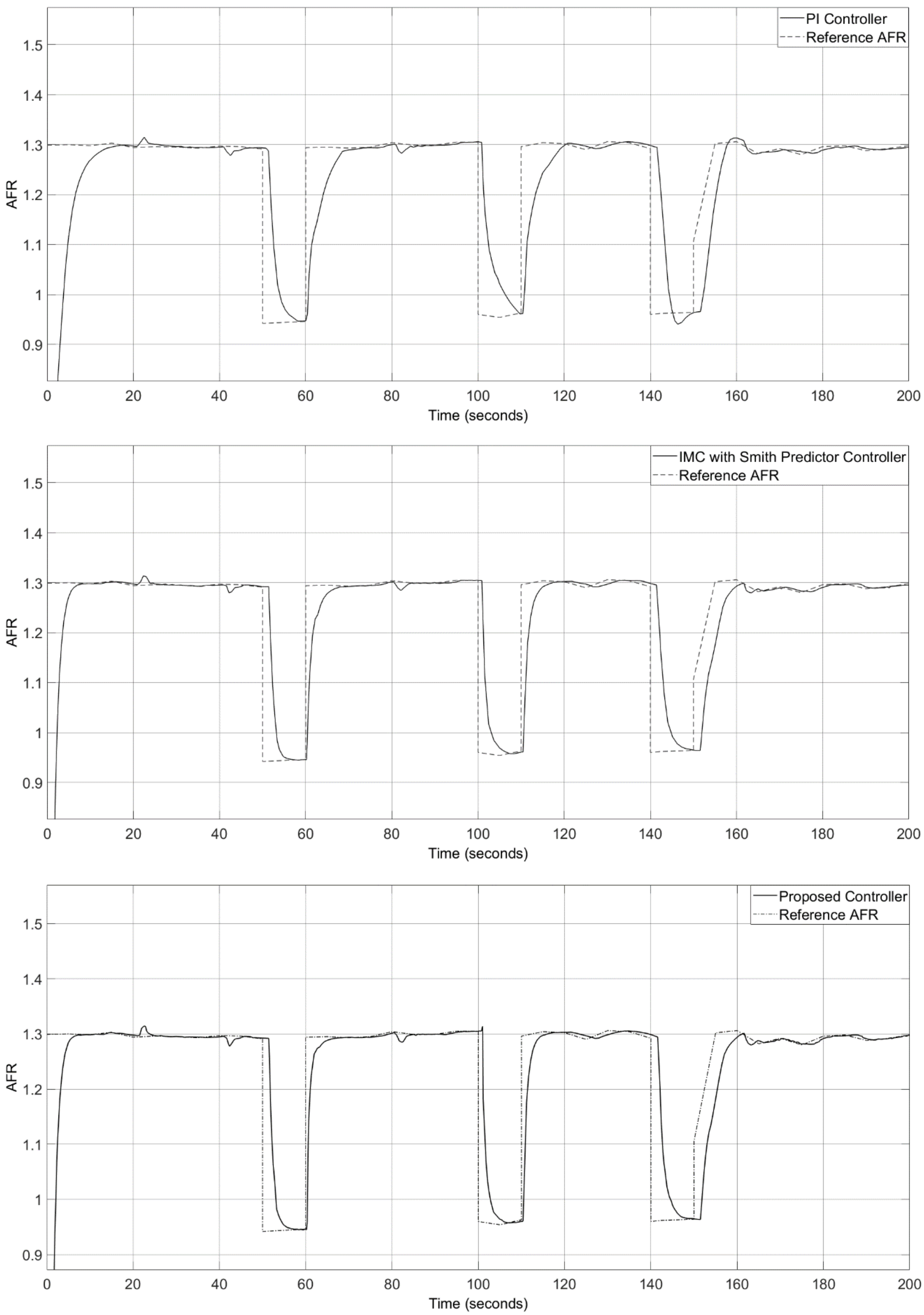


Figure 6.4: Performance comparison with an engine time constant of 0.3s.

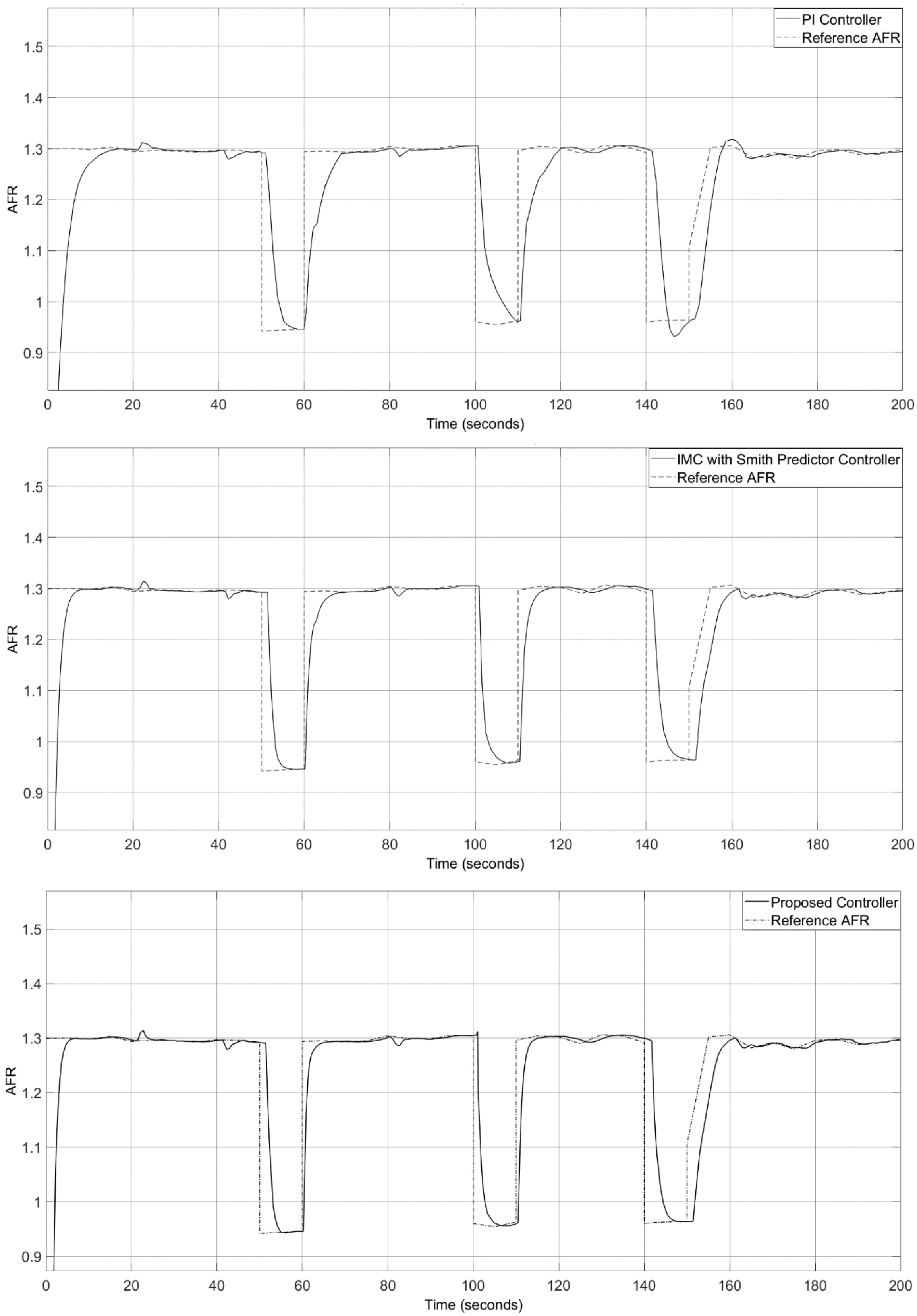


Figure 6.5: Performance comparison with an engine time constant of 0.4s.

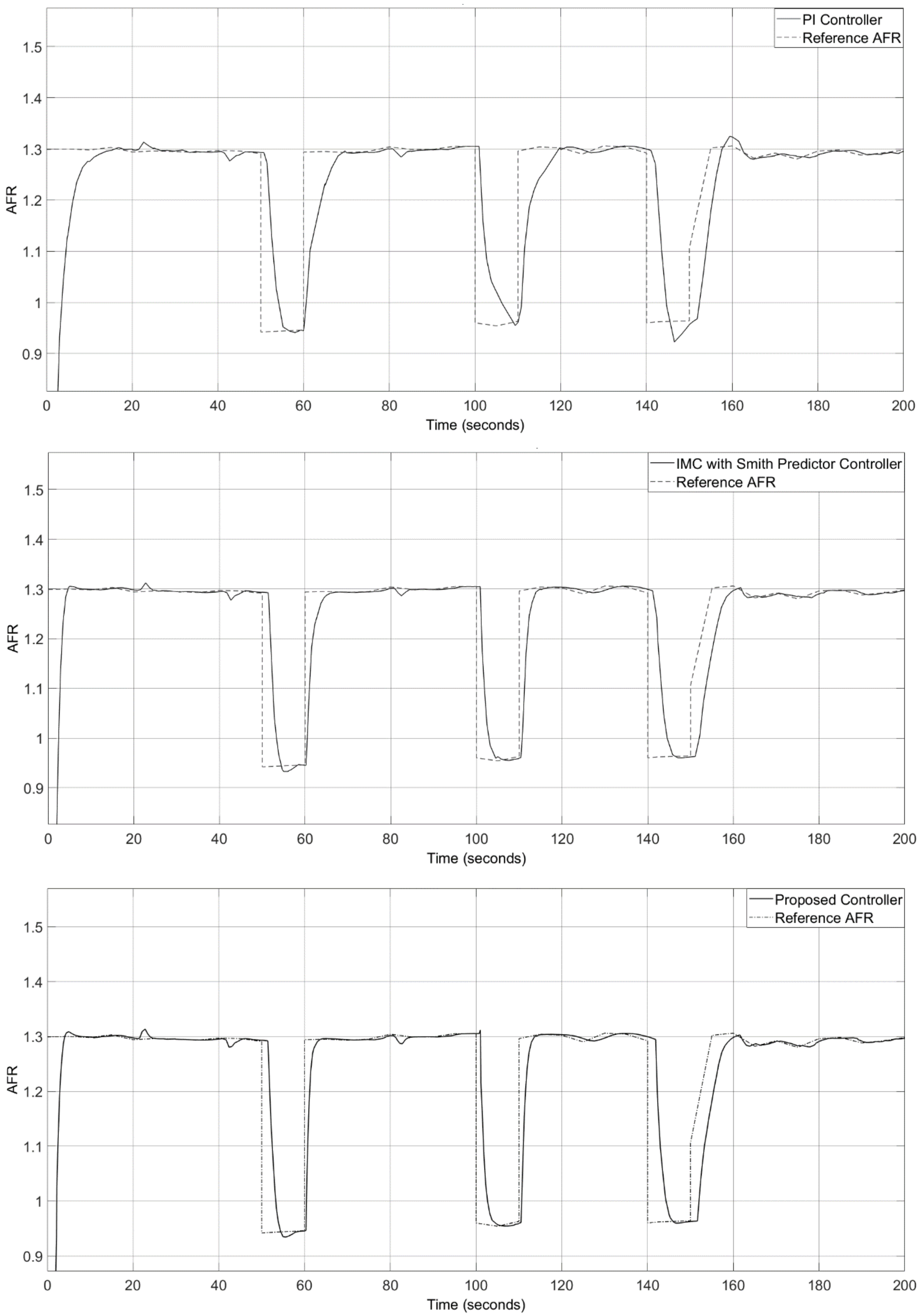


Figure 6.6: Performance comparison with an engine time constant of 0.5s.

Controller performance data has been further extracted from the simulation and tabulated in Table 6.1. The settling time and percent overshoot values presented are averages of the three simulated lean-rich-lean switches.

Table 6.1: Controller performance data.

Engine Time Constant = 0.3s

	Lean to Rich Switch		**Rich to Lean Switch**		**Disturbance**	
	Settling Time (s)	Percent Overshoot (%)	Settling Time (s)	Percent Overshoot (%)	Effect Duration (s)	Percent Overshoot (%)
PI	9.27	0.95	10.07	0.35	12.90	2.31
IMC + Smith	6.80	0.00	6.67	0.00	8.00	1.27
Proposed Design	6.77	0.00	6.57	0.00	7.50	2.31

Engine Time Constant = 0.4s

	Lean to Rich Switch		**Rich to Lean Switch**		**Disturbance**	
	Settling Time (s)	Percent Overshoot (%)	Settling Time (s)	Percent Overshoot (%)	Effect Duration (s)	Percent Overshoot (%)
PI	9.63	1.15	9.00	0.44	9.40	1.19
IMC + Smith	7.80	0.35	6.57	0.00	7.50	1.15
Proposed Design	5.23	0.26	5.93	0.00	7.50	1.23

Engine Time Constant = 0.5s

	Lean to Rich Switch		**Rich to Lean Switch**		**Disturbance**	
	Settling Time (s)	Percent Overshoot (%)	Settling Time (s)	Percent Overshoot (%)	Effect Duration (s)	Percent Overshoot (%)
PI	9.03	1.23	9.00	0.59	9.60	1.23
IMC + Smith	7.17	1.79	5.50	0.00	9.00	1.19
Proposed Design	7.43	2.32	4.53	0.00	8.60	1.23

As seen in the simulation results and the performance data, all three controllers seem to provide good disturbance-rejection and stable performance. The PI controller shows some overshoot over the entire range of the engine time constant, while the other two controllers only start to develop overshoot near the higher end of the range. The proposed controller design and the IMC with Smith predictor controller display similar behavior in providing close reference-tracking to the AFR command, while the PI controller fails to offer the same level of tracking. To better visualize the comparison between the two similar controllers, Figures 6.7 and 6.8 provide a zoomed window on a time interval in the engine operating range, containing a lean-rich-lean

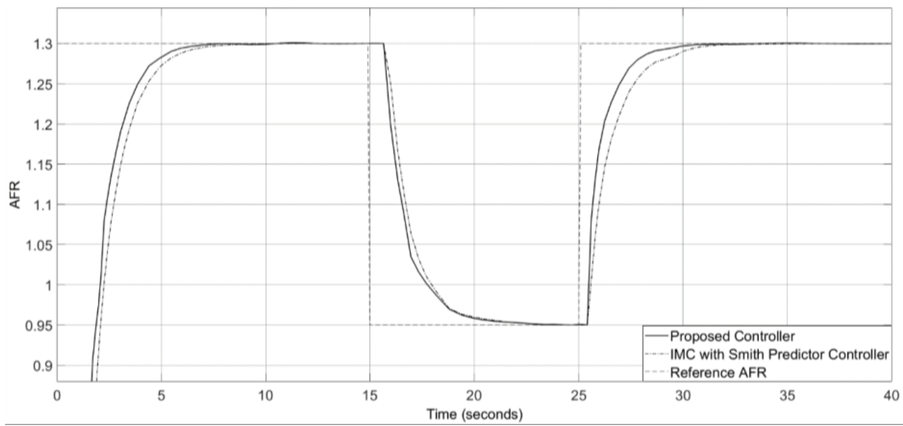


Figure 6.7: Comparison between IMC/Smith predictor and the proposed controller design for an engine time constant of 0.3s.

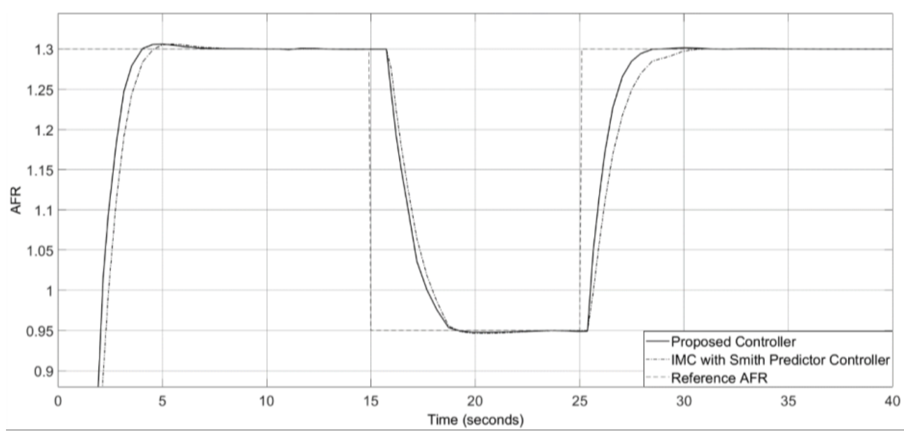


Figure 6.8: Comparison between IMC/Smith predictor and the proposed controller design for an engine time constant of 0.5s.

switch. As clearly seen in the figures, the proposed controller design seems to have a shorter settling time than that of the IMC- Smith predictor controller, without introducing any significant transient effects, which indicates the former’s better reference-tracking. This can also be observed in the performance data and especially for the engine time constant of 0.4s.

Chapter 7: Conclusion

This thesis proposes a control design, based on a combined internal model (IMC) and Smith predictor feedback controller, to provide reference-tracking to a lean-burn air-fuel ratio profile in a spark-ignition (SI) engine. The choice for this method stems from the ability of this controller to counter the effect of time delay on closed-loop dynamics.

The proposed design is validated through simulation and is compared to the performance of a fully-tuned PI controller and IMC-Smith controller. The results reveal the inadequacy of a PI controller in providing the desired level of control to the engine system due to the transients present in the system response. The considerably larger presence of overshoot in the PI controller is mainly attributed to the lack of time delay compensation, as opposed to the other two controllers which employ a Smith estimator. Overshoot, however is still present to an extent despite the Smith regulation, which is explained by that the Smith regulator requires an exact time delay measurement to eliminate transient effects, which is unattainable yet in the engine system. A more precise estimation of the time delay would provide a more proper control of the air-fuel ratio. In comparison with a Smith-IMC controller, the proposed design provides better reference tracking with introducing very negligible overshoot percentage. It is possible to further reduce the settling time of the proposed controller design by increasing the overall control gain or weakening the effect of the IMC filter, but doing so may induce unwanted overshoot within the rapid lean-rich-lean intervals.

Although the proposed design has been validated through simulation and using a dynamic setting that is derived experimentally, the validation process does not stop at this stage. The comparison performed for this thesis only considers one experimental variable time delay profile, so the results may appear different, should a researcher simulate the proposed controller with a different delay profile. In addition, to further investigate the advantages and disadvantages of this design, the controller must be constructed and tested in an experimental setting. Aside from control performance, additional parameters to evaluate include: ease of implementation, compatibility with an actual engine, and control effort (or associated power consumption).

References

[1] Bai, Y. (2013). "Studies on SI engine Simulation and Air/Fuel Ratio Control Systems Design," Ph.D. Thesis, School of Engineering and Design, Brunel University.

[2] Ebrahimi, B., Tafreshi, R., Mohammadpour, J., Franchek, M., Grigoriadis, K., & Masudi, H. (2014). "Second-order Sliding Mode Strategy for Air-Fuel Ratio Control of Lean-Burn SI Engines," IEEE Transactions on Control Systems Technology, 22(4), 1374-1384.

[3] Flagan, Richard C. & Seinfeld, John H. (1988). *Fundamentals of Air Pollution Engineering*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey. ISBN 0-13-332537-7. http://resolver.caltech.edu/CaltechBOOK:1988.001

[4] Franklin, G., Emami-Naeini, A., & Powell, D. (1994). Feedback control of dynamics systems. Estados Unidos: Addison-Wesley.

[5] Heywood, J. B. (1988). *Internal Combustion Engine Fundamentals*, Singapore: McGraw Hill International Editions.

[6] Nikolaou, M. & Grigoriadis, K. (n.d.). "Simple Robust Control of Lean-Burn Air-Fuel Ratio for a Spark Ignition Engine."

[7] Rivera, E. D. Internal Model Control: A Comprehensive View. Arizona State University, Arizona, 1999

[8] Salavati, S., Grigoriadis, K., & Franchek, M. (2018). "Robust Delay-Dependent Criteria for Uncertain Time-Delay Systems via IMC with Application to Lean-Burn Engine Control," International Journal of Control.

[9] Sipahi, R., Niculescu, S., Abdallah, C. T., Michiels, W., & Keqin Gu. (2011). "Stability and Stabilization of Systems with Time Delay," IEEE Control Systems, 31, 38-65.

[10] Tham, M., (2002). *Introduction to Robust Control*, Newcastle University.

[11] W., G. P., Gill, P. W., Smith, J. H., & Ziurys, E. J. (n.d.). *Fundamentals of Internal Combustion Engines*, Calcutta: Oxford and IBH Pub.

[12] Zhang, Y. (2005). “Advanced Linear Parameter Varying Control for Systems with Delays, Saturation and Implementation Constraints,” Ph.D. Thesis, Cullen College of Engineering, University of Houston.

[13] Zope, R. A. (2011). “Model-based Estimation and Control in Spark Ignition Engines using LPV Techniques,” Ph.D. Thesis, Cullen College of Engineering, University of Houston.