Calibration reduction in internal combustion engine fueling control: modeling, estimation and stability robustness

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Jason Meyer, B.S., M.S.
Graduate Program in Mechanical Engineering The Ohio State University

2011

Dissertation Committee:
Prof. Stephen Yurkovich, Adviser
Prof. Marcello Canova
Prof. Yann Guezennec
Prof. Giorgio Rizzoni
Prof. Junmin Wang
ABSTRACT

Controlling the fuel injection system of an internal combustion engine is a challenging and multifaceted problem. Current control algorithms rely heavily on lengthy experimentally-based calibration techniques. Even with these extensive calibration processes, suboptimal performance is often achieved because the selected control gains depend on calibrator experience and intuition instead of objective metrics. The cost and manpower required to calibrate a fueling controller can be staggering. Recent advances in air path actuation technologies such as the variable geometry turbocharger for diesel engines and variable cam timing for gasoline engines have expanded the dimensionality of the fueling control problem, further increasing the calibration burden of conventional controllers.

This dissertation presents model based alternatives to the current calibration-heavy fueling controllers used in production gasoline and diesel engines. The novelty of these controllers is derived from both the underlying plant models and the application of estimation/control theory to these models. One of the most unique aspects of these control problems is the time varying delay which characterizes the plant (i.e. the engine air path system). From a deep understanding of the physical phenomena involved, new types of air path models that directly capture the oxygen transport and mixing dynamics are developed. An experimental validation demonstrates that
such a model can even account for cylinder-to-cylinder response variations caused by asymmetrical exhaust runner lengths.

Advanced model based estimation and control techniques are applied to these models to develop more effective methods of estimating the most important dynamic variables (i.e. in-cylinder oxygen concentration and cylinder imbalance) and controlling fuel delivery in diesel and gasoline engines. During the design processes, particular emphasis is placed on stability robustness, and comprehensive stability proofs are provided for the controllers and estimators developed. Because of the practical issues involved in dynamically measuring the in-cylinder oxygen concentration, the in-cylinder oxygen concentration estimator and the diesel fueling controller are validated in simulation using the engine modeling software GT-Power. Experimental validations demonstrate the capabilities of the cylinder imbalance estimator and comprehensively quantify the performance of the gasoline fueling controller through direct comparisons to the existing production controller. Above even their performance and robustness benefits, the most compelling reasons to adopt these estimators and controllers over their production counterparts are their modest calibration requirements.
ACKNOWLEDGMENTS

Firstly, I would like to thank my adviser Prof. Stephen Yurkovich for providing guidance and invaluable insight throughout my graduate school experience. You are truly committed to your graduate students and always make yourself available. I would like to thank Dr. Sai Rajagopalan, Dr. Yiran Hu, Mr. Chris Hoops and Mr. John Davis for their collaborative efforts on this work. A special thanks to the support provided by Dr. Kenneth Dudek and Dr. Vivek Sujan.

I really appreciate all of the opportunities that everyone at the Center for Automotive Research has provided me. I am especially grateful for the time that Prof. Giorgio Rizzoni spent mentoring my teaching. I want to thank all of the research and support staff at CAR for their great work and willingness to help. Dr. Midlam-Mohler in particular has been an integral part of my success. Lastly I would like to thank all of my committee members as a whole. Your suggestions have helped to guide the final stages of my research.
VITA

May 19, 1985 ...............................Born - Dayton, OH

2007 ........................................B.S. Mechanical Engineering, The Ohio State University

2008 ........................................M.S. Mechanical Engineering, The Ohio State University

PUBLICATIONS


FIELDS OF STUDY

Major Field: Mechanical Engineering
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>iv</td>
</tr>
<tr>
<td>Vita</td>
<td>v</td>
</tr>
<tr>
<td>List of Tables</td>
<td>ix</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xi</td>
</tr>
<tr>
<td>Nomenclature</td>
<td>xvi</td>
</tr>
<tr>
<td>Chapters:</td>
<td></td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Organization</td>
<td>4</td>
</tr>
<tr>
<td>2. Background and Motivation</td>
<td>9</td>
</tr>
<tr>
<td>2.1 Feed-forward Diesel Fueling Control</td>
<td>10</td>
</tr>
<tr>
<td>2.2 Air-to-Fuel Ratio Control in Gasoline Engines</td>
<td>15</td>
</tr>
<tr>
<td>2.3 Cylinder Imbalance Estimation and Control</td>
<td>20</td>
</tr>
<tr>
<td>2.3.1 Crankshaft Speed Fluctuation Based Imbalance Estimation and Control</td>
<td>22</td>
</tr>
<tr>
<td>2.3.2 Cylinder Pressure Based Imbalance Estimation and Control</td>
<td>24</td>
</tr>
<tr>
<td>2.3.3 Exhaust Oxygen Concentration Based Imbalance Estimation and Control</td>
<td>25</td>
</tr>
<tr>
<td>2.4 Engine Modeling</td>
<td>27</td>
</tr>
<tr>
<td>2.4.1 Conventional Air Path Modeling Approach</td>
<td>28</td>
</tr>
<tr>
<td>2.4.2 Oxygen Storage Based TWC Model</td>
<td>38</td>
</tr>
</tbody>
</table>
6. AFR Switching Period Control in a Gasoline Engine ................. 181
   6.1 Additional Nomenclature ........................................ 184
   6.2 Pre-Catalyst AFR Switching Period Controller Description ..... 190
   6.3 Post-Catalyst Setpoint Controller Schemes ...................... 197
      6.3.1 Switching Voltage Setpoint Controller ................... 198
      6.3.2 Rich/Lean Switching Period Ratio Controller ............. 199
   6.4 Stability Study ................................................ 201
      6.4.1 Global Stability and Convergence .......................... 202
      6.4.2 Local Asymptotic Stability ............................... 208
      6.4.3 Stability Robustness ..................................... 217
      6.4.4 Implications on Control Gain Selection .................... 222
   6.5 Calibration Procedures ........................................ 224
      6.5.1 Engine Model Components ................................. 224
      6.5.2 Optimization Procedure ................................. 227
   6.6 Experimental Setup ............................................ 229
   6.7 Post-Catalyst Setpoint Controller Comparison .................. 230
   6.8 Selecting the Periodic Waveform Signal ......................... 239
      6.8.1 Waveform Shape Selection ................................ 241
      6.8.2 Dither Period ......................................... 245
      6.8.3 Dither Amplitude ........................................ 248
      6.8.4 Signal Selection Summary ................................ 249
   6.9 Performance Comparison: Switching Period Controller Versus Production Controller ................. 251
   6.10 Summary ..................................................... 255

7. Conclusions ..................................................... 257

Bibliography ..................................................... 265
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Cylinder impulse response $R^2$ comparison: HFEOD model versus measured</td>
</tr>
<tr>
<td>3.2</td>
<td>Cylinder-1 imbalance $R^2$ comparison: Models (HFEOD and ideal) versus measured</td>
</tr>
<tr>
<td>3.3</td>
<td>Cylinder-1 imbalance $R^2$ comparison: Models (HFEOD and ideal) versus measured</td>
</tr>
<tr>
<td>4.1</td>
<td>Assumptions used to derive the LPV air path model</td>
</tr>
<tr>
<td>4.2</td>
<td>Characteristics of the nominal input and output disturbance trajectories</td>
</tr>
<tr>
<td>4.3</td>
<td>RMS in-cylinder oxygen concentration estimation errors in percent oxygen (input disturbances only)</td>
</tr>
<tr>
<td>4.4</td>
<td>RMS in-cylinder oxygen concentration estimation errors in percent oxygen (input and output disturbances)</td>
</tr>
<tr>
<td>4.5</td>
<td>Instantaneous in-cylinder oxygen concentration estimation error percentile statistics (input and output disturbances)</td>
</tr>
<tr>
<td>5.1</td>
<td>Fueling table $R^2$ fit comparison (only optimal in-cylinder conditions)</td>
</tr>
<tr>
<td>5.2</td>
<td>Fueling table $R^2$ fit comparison (all in-cylinder conditions)</td>
</tr>
<tr>
<td>5.3</td>
<td>Drive cycle averaged combustion performance results: Ideal air path response</td>
</tr>
<tr>
<td>5.4</td>
<td>Drive cycle averaged combustion performance results: Shifted air path actuator commands</td>
</tr>
</tbody>
</table>
5.5 Drive cycle averaged combustion performance results: Dynamic air path actuator errors ........................................... 175

6.1 EQR error statistics for a simulated FTP cycle ..................... 231

6.2 EQR and emissions statistics for a FTP-72 cycle (all values relate to stock and are in percentages) ......................... 252
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Current air path and fueling control approach</td>
<td>12</td>
</tr>
<tr>
<td>2.2</td>
<td>Model of a flexible crankshaft in a six cylinder engine from [1]</td>
<td>23</td>
</tr>
<tr>
<td>2.3</td>
<td>Diagram of zero-dimensional air path model of a turbocharged engine with EGR</td>
<td>35</td>
</tr>
<tr>
<td>2.4</td>
<td>Diagram of one-dimensional air path model of a turbocharged engine with EGR</td>
<td>38</td>
</tr>
<tr>
<td>3.1</td>
<td>Effect of diffusion rate on the step response of one-dimensional gaseous system</td>
<td>46</td>
</tr>
<tr>
<td>3.2</td>
<td>Diagram of an air path model which models the filling and emptying dynamics with a zero-dimensional model and models the oxygen transport and mixing dynamics with a lumped plug flow oxygen dynamics model</td>
<td>53</td>
</tr>
<tr>
<td>3.3</td>
<td>Cell splitting diagram</td>
<td>59</td>
</tr>
<tr>
<td>3.4</td>
<td>Steady-state BSFC comparison: Experimental versus GT-Power</td>
<td>64</td>
</tr>
<tr>
<td>3.5</td>
<td>Steady-state NO\textsubscript{x} comparison: Experimental versus GT-Power</td>
<td>65</td>
</tr>
<tr>
<td>3.6</td>
<td>Transient fresh air flow rate comparison: Experimental versus GT-Power</td>
<td>66</td>
</tr>
<tr>
<td>3.7</td>
<td>Transient EGR flow rate comparison: Experimental versus GT-Power</td>
<td>67</td>
</tr>
<tr>
<td>3.8</td>
<td>Transient ([O_2]_{exh}) comparison: Experimental versus GT-Power</td>
<td>67</td>
</tr>
</tbody>
</table>
3.9 Transient NO\textsubscript{x} comparison: Experimental versus GT-Power ........ 68
3.10 Trapped residual fraction comparison: lookup function versus GT-Power 68
3.11 Oxygen concentration prediction comparison: GT-Power versus lumped oxygen dynamics model ........................................... 70
3.12 Oxygen concentration prediction error comparison: Lumped oxygen dynamics model versus traditional oxygen concentration model .... 71
3.13 Example 4-2-1 exhaust system separated into 100 cells ............... 73
3.14 Measured versus model predicted impulse response of cylinder-1 at idle 80
3.15 Measured versus model predicted impulse response of cylinder-3 at idle 81
3.16 Measured versus model predicted impulse response of cylinder-1 under moderate load ......................................................... 82
3.17 Measured versus model predicted impulse response of cylinder-3 under moderate load ......................................................... 83
3.18 Measured versus model predicted EGO response of a cylinder-1 imbalance error at idle .......................................................... 85
3.19 Measured versus model predicted EGO response of cylinder-3 imbalance error at idle .......................................................... 85
3.20 Evolution of the input dependency vector .................................. 92
3.21 Output dependencies: Event prior to commanding fuel to cylinder-1 at idle .......................................................... 96
3.22 Output dependencies: Event when fuel is commanded to cylinder-1 at idle .......................................................... 97
3.23 Influence on the output of a single fueling input from cylinder-1 .... 98
3.24 Influence on the output of a single fueling input from cylinder-3 .... 99
3.25 Summation of the input and output contributions during step changes in mass flow rate with a fixed engine speed ........................................... 100

3.26 Summation of the input and output contributions during step changes in engine speed with a fixed mass flow rate ................................. 101

4.1 Control diagram of the hybrid feed-forward/feedback estimator ........... 127

4.2 Illustration of the shifting and scaling applied to the nominal EGR disturbance trajectory ................................................................. 129

4.3 In-cylinder oxygen concentration RMS estimation error probability density function (input disturbances only) ..................................... 132

4.4 In-cylinder oxygen concentration RMS estimation error cumulative probability function (input disturbances only) ............................... 132

4.5 In-cylinder oxygen concentration RMS estimation error probability density function (input and output disturbances) ......................... 135

4.6 In-cylinder oxygen concentration RMS estimation error cumulative probability function (input and output disturbances) ....................... 135

4.7 Estimated in-cylinder oxygen concentration comparison ..................... 136

4.8 In-cylinder oxygen concentration estimation error comparison ............. 137

4.9 Estimated in-cylinder oxygen concentration comparison zoomed to show the effects of modeling the transport delays ............................... 138

4.10 Identification of a cylinder-1 imbalance error applied at event 300 during idle (baseline) ................................................................. 146

4.11 Identification of a cylinder-1 imbalance error applied at event 300 during idle, with first-order filter (tuned) .................................................. 147

4.12 Condition number surface ............................................................... 148

5.1 Conventional air path and fueling control approach ............................ 154

5.2 Proposed air path and fueling control approach .................................... 156

xiii
5.3 Fueling parameter calibration example .......................... 159
5.4 ISFC Comparison: Ideal air path response ......................... 164
5.5 BSNOx Comparison: Ideal air path response ....................... 165
5.6 In-cylinder oxygen concentration variations: Cases A - E ........ 167
5.7 In-cylinder oxygen concentration variations: Cases E - I .......... 168
5.8 Trade-off between cycle averaged BSNO\textsubscript{x} and cycle average ISFC .......................... 170
5.9 Trade-off between cycle averaged BSNO\textsubscript{x} and cycle averaged ISFC .......................... 171
5.10 Instantaneous ISFC difference (Conventional - [O\textsubscript{2}]\textsubscript{cyl} based) probability density functions ........................................ 173
5.11 Instantaneous BSNO\textsubscript{x} difference (Conventional - [O\textsubscript{2}]\textsubscript{cyl} based) probability density functions ........................................ 174
5.12 Distribution of (ISFC, BSNO\textsubscript{x}) pairs for the oxygen concentration based fueling controller ........................................ 176
5.13 Distribution of (ISFC, BSNO\textsubscript{x}) pairs for the conventional fueling controller ........................................ 177
5.14 Trade-off between cycle averaged BSNO\textsubscript{x} and cycle averaged ISFC: Dynamic air path errors ........................................ 178
6.1 Pre-catalyst switching period and post-catalyst setpoint controllers. 192
6.2 Effect of the rich/lean switching period ratio setpoint on the dither signal 200
6.3 Response of the late switching compensator when dither period is properly selected ($T_{\text{dither}} = 20, d_{\text{exh}} = 5, \delta = 0, A = 0.01$) .................................. 210
6.4 Response of the late switching compensator when dither period is too short ($T_{\text{dither}} = 20, d_{\text{exh}} = 10, \delta = 0, A = 0.01$) .................................. 210
6.5 Response of the late switching controller with proper and improper proportional gain selection ($T_{dither} = 20, d_{exh} = 3, \delta = 0, A = 0.015$). 211

6.6 EQR tracking performance: 15/10 rich/lean switching period ratio (reference - dashed, measured - solid) 233

6.7 EQR tracking performance: 10/15 rich/lean switching period ratio (reference - dashed, measured - solid) 233

6.8 EQR tracking performance: 700 mV switching voltage (reference - dashed, measured - solid) 234

6.9 EQR tracking performance: 300 mV switching voltage (reference - dashed, measured - solid) 234

6.10 Effect of the pre-catalyst switching voltage on the post-catalyst EGO voltage 236

6.11 Evolution of the post-catalyst voltage without post-catalyst control 238

6.12 Post-catalyst voltage regulation performance: Switching voltage setpoint of 680 mV 238

6.13 Pre-catalyst switching voltage setpoint commanded by the post-catalyst controller 239

6.14 Effect of the dither waveform on the EQR distribution over a drive cycle 245

6.15 Effect of the dither waveform on idle emissions 246

6.16 Effect of the dither period on idle emissions 247

6.17 Effect of the dither period on the EQR distribution over a drive cycle 248

6.18 Effect of the dither amplitude on idle emissions 250

6.19 Effect of the dither amplitude on the EQR distribution over a drive cycle 250

6.20 EQR response comparison: FTP cycle 253

6.21 EQR response comparison: 10% EQR disturbance at event 200 and 800 255
### NOMENCLATURE

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFR</td>
<td>Air-to-fuel ratio.</td>
</tr>
<tr>
<td>DoE</td>
<td>Design of experiments.</td>
</tr>
<tr>
<td>DPF</td>
<td>Diesel particulate filter.</td>
</tr>
<tr>
<td>ECU</td>
<td>Engine control unit.</td>
</tr>
<tr>
<td>EGO</td>
<td>Exhaust gas oxygen sensor (switching type).</td>
</tr>
<tr>
<td>EGR</td>
<td>Exhaust gas recirculation.</td>
</tr>
<tr>
<td>EQR</td>
<td>Fuel/air equivalence ratio.</td>
</tr>
<tr>
<td>FIR</td>
<td>Finite impulse response.</td>
</tr>
<tr>
<td>FTP</td>
<td>Federal test procedure.</td>
</tr>
<tr>
<td>HCCI</td>
<td>Homogeneous charge compression ignition.</td>
</tr>
<tr>
<td>HFEOD</td>
<td>High fidelity exhaust oxygen dynamics.</td>
</tr>
<tr>
<td>ISS</td>
<td>Input to state stability.</td>
</tr>
<tr>
<td>LPV</td>
<td>Linear parameter varying.</td>
</tr>
<tr>
<td>MAF</td>
<td>Mass air flow rate.</td>
</tr>
<tr>
<td>RLS</td>
<td>Recursive least squares.</td>
</tr>
<tr>
<td>RMS</td>
<td>Root mean squared.</td>
</tr>
<tr>
<td>RPM</td>
<td>Revolutions per minute.</td>
</tr>
<tr>
<td>SCR</td>
<td>Selective catalytic reduction.</td>
</tr>
<tr>
<td>SOI</td>
<td>Start of injection.</td>
</tr>
</tbody>
</table>
TWC Three-way catalyst.
UEGO Universal exhaust gas oxygen sensor (linear sensor).
VGT Variable geometry turbocharger.
\( \Gamma_{KF} \) Process noise matrix of the nonlinear four state oxygen dynamics model.
\( \Phi_{err} \) Vector containing the measured EQR deviations.
\( \Phi_{imb} \) Vector containing the cylinder imbalance errors.
\( \Phi_{KF} \) State matrix of the nonlinear four state oxygen dynamics model.
\( \Psi_{KF} \) Input matrix of the nonlinear four state oxygen dynamics model.
\( \Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4 \) Summation of the coefficients of the FIR exhaust oxygen dynamics model.
\( \Theta_s \) Stored oxygen fraction within a TWC.
\( \alpha_0 \) Parameter related to the ratio of the current EGR flow rate to the maximum EGR flow rate.
\( \alpha_1 \) Parameter defined as the ratio of the cylinder mass before fuel injection to after fuel injection.
\( \alpha_2 \) Parameter defined as the ratio of the EGR flow rate to the charge flow rate.
\( \beta \) General weighting coefficient for a plug flow model.
\( \beta_{egr} \) Weighting coefficient for a plug flow model of the intake system.
\( \beta_{exh} \) Weighting coefficient for a plug flow model of the exhaust system.
\( \beta_{int} \) Weighting coefficient for a plug flow model of the intake system.
\( \beta_{ISS} \) Class KL function used in ISS stability proof.
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>δ&lt;sub&gt;egr&lt;/sub&gt;</td>
<td>Total number of cells that the exhaust gas entering the EGR system will occupy.</td>
</tr>
<tr>
<td>δ&lt;sub&gt;exh&lt;/sub&gt;</td>
<td>Total number of cells that the combustion gases entering the exhaust system will occupy.</td>
</tr>
<tr>
<td>δ&lt;sub&gt;int&lt;/sub&gt;</td>
<td>Total number of cells that the charge gas entering the intake system will occupy.</td>
</tr>
<tr>
<td>η&lt;sub&gt;c&lt;/sub&gt;</td>
<td>Compressor isentropic efficiency.</td>
</tr>
<tr>
<td>η&lt;sub&gt;t&lt;/sub&gt;</td>
<td>Turbocharger isentropic efficiency.</td>
</tr>
<tr>
<td>η&lt;sub&gt;e&lt;/sub&gt;</td>
<td>Volumetric efficiency.</td>
</tr>
<tr>
<td>ε</td>
<td>Maximum EGR weighting factor.</td>
</tr>
<tr>
<td>γ&lt;sub&gt;1&lt;/sub&gt;, γ&lt;sub&gt;2&lt;/sub&gt;, γ&lt;sub&gt;3&lt;/sub&gt;, γ&lt;sub&gt;4&lt;/sub&gt;</td>
<td>Parameters used in the ISS stability proof.</td>
</tr>
<tr>
<td>γ&lt;sub&gt;ISS&lt;/sub&gt;</td>
<td>Class K function used in ISS stability proof.</td>
</tr>
<tr>
<td>γ&lt;sub&gt;res&lt;/sub&gt;</td>
<td>Trapped residual fraction.</td>
</tr>
<tr>
<td>γ&lt;sub&gt;hr&lt;/sub&gt;</td>
<td>Specific heat ratio.</td>
</tr>
<tr>
<td>κ</td>
<td>Hybrid oxygen concentration estimator correction factor weighting.</td>
</tr>
<tr>
<td>λ&lt;sub&gt;f&lt;/sub&gt;</td>
<td>RLS forgetting factor.</td>
</tr>
<tr>
<td>ν</td>
<td>Filter constant for the delay based oxygen dynamics model.</td>
</tr>
<tr>
<td>ω</td>
<td>Coefficient used to design LPV state estimator.</td>
</tr>
<tr>
<td>φ</td>
<td>EQR.</td>
</tr>
<tr>
<td>φ&lt;sub&gt;cat&lt;/sub&gt;</td>
<td>EQR entering the catalyst.</td>
</tr>
<tr>
<td>φ&lt;sub&gt;err&lt;/sub&gt;</td>
<td>Measured EQR deviation.</td>
</tr>
<tr>
<td>φ&lt;sub&gt;in&lt;/sub&gt;</td>
<td>EQR of the pre-combustion mixture.</td>
</tr>
<tr>
<td>φ&lt;sub&gt;O₂&lt;/sub&gt;</td>
<td>Predicted EQR at the oxygen sensor.</td>
</tr>
</tbody>
</table>
\( \rho_{cv} \) Density within a control volume.

\( \rho \) Density (one dimensional control volume).

\( \tau_{\text{evap}} \) Fuel evaporation time constant.

\( \tau_{O_2} \) Time constant of the oxygen sensor in the time domain.

\( \tau_{tc} \) Time constant of the turbocharger system.

\( \theta \) Sum of the parameters which define the LPV oxygen dynamics model.

\( \chi_{\text{fuel}} \) Fraction of injection fuel which adds to the fuel film.

\( \zeta \) EGR mixing filter constant.

\( A \) Sum of the closed loop state matrices scaled by their respective parameters.

\( A_0, A_1, A_2 \) State matrices of the LPV oxygen dynamics model.

\( \bar{A}_0, \bar{A}_1, \bar{A}_2 \) Close loop state estimation matrices in the LPV state estimator.

\( A_o \) Orifice cross sectional area.

\( A_t \) Turbocharger effective area.

\( A \) Cross sectional area of one dimensional control volume.

\( \text{AFR}_s \) Stoichiometric air-to-fuel ratio.

\( B_0, B_1, B_2 \) Input matrices of the LPV oxygen dynamics model.

\( C_0, C_1, C_2 \) Output matrices of the LPV oxygen dynamics model.

\( C_d \) Discharge coefficient.

\( C_{imb} \) Steady-state cylinder imbalance matrix.

\( C_{in-cyl} \) Output matrix corresponding to the in-cylinder oxygen concentration.

\( C_t \) Discharge coefficient of the turbocharger.
\( \text{CTWC} \) Oxygen storage capacity of the TWC.

\( \text{CA}_{50} \) Crank angle at which 50\% of the total heat has been released.

\( D \) Diffusion constant for a plug flow model.

\( D \) Diffusivity constant.

\( D \) Diffusion coefficient.

\( D \) Diffusion coefficient.

\( \text{EGR}^{q_{egr}} \) Vector containing the oxygen concentrations of the cells within the EGR control volume after the \( q_{egr} \)th cell has entered the system.

\( \text{EGR}_{i}^{q_{egr}} \) Oxygen concentration within the \( i \)th cell of the EGR control volume after the \( q_{egr} \)th cell has entered the system.

\( \text{EXH}^{q_{exh}} \) Vector containing the oxygen concentrations of the cells within the exhaust control volume after the \( q_{exh} \)th cell has been exhausted.

\( \text{EXH}_{i}^{q_{exh}} \) Oxygen concentration within the \( i \)th cell of the exhaust control volume after the \( q_{exh} \)th cell has been exhausted.

\( F_{egr}^{j} \) Contribution from the first partition of the \( j \)th group of exhaust cells to the \( j \)th cell entering the EGR system.

\( F_{int}^{j} \) Contribution from the first partition of the \( j \)th group of EGR cells to the \( j \)th cell entering the intake system.

\( G_{0}, G_{1}, G_{2} \) Square matrices used to design LPV state estimator.

\( G_{int}^{q_{int}} \) Contribution from the EGR gases on the \( q_{int} \)th cell entering the intake system.
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_{KF}$</td>
<td>Output matrix of the nonlinear four state oxygen dynamics model.</td>
</tr>
<tr>
<td>$INT^{q_{int}}$</td>
<td>Vector containing the oxygen concentrations of the cells within the intake control volume after the $q_{int}^{th}$ cell has entered the system.</td>
</tr>
<tr>
<td>$INT_i^{q_{int}}$</td>
<td>Oxygen concentration within the $i^{th}$ cell of the intake control volume after the $q_{int}^{th}$ cell has entered the system.</td>
</tr>
<tr>
<td>$IO$</td>
<td>Vector containing the coefficients of the FIR exhaust oxygen dynamics model.</td>
</tr>
<tr>
<td>$K^*$</td>
<td>Mass flow rate entering a control volume in units of control volumes per event.</td>
</tr>
<tr>
<td>$K_{KF}$</td>
<td>Kalman filter gain.</td>
</tr>
<tr>
<td>$K_{rls}$</td>
<td>RLS gain matrix.</td>
</tr>
<tr>
<td>$L_0$, $L_1$, $L_2$</td>
<td>Observer gain matrices of the LPV state estimator.</td>
</tr>
<tr>
<td>$L^j_{egr}$</td>
<td>Contribution from the last partition of the $j^{th}$ group of exhaust cells to the $j^{th}$ cell entering the EGR system.</td>
</tr>
<tr>
<td>$L^j_{int}$</td>
<td>Contribution from the last partition of the $j^{th}$ group of EGR cells to the $j^{th}$ cell entering the intake system.</td>
</tr>
<tr>
<td>$M_{ab}$</td>
<td>Mixing coefficient for the confluence junction of pipes $a$ and $b$.</td>
</tr>
<tr>
<td>$M_1$</td>
<td>Matrix used in the LPV state estimator stability proof.</td>
</tr>
<tr>
<td>$M^j_{egr}$</td>
<td>Contribution from the middle partition of the $j^{th}$ group of exhaust cells to the $j^{th}$ cell entering the EGR system.</td>
</tr>
<tr>
<td>$M^j_{int}$</td>
<td>Contribution from the middle partition of the $j^{th}$ group of EGR cells to the $j^{th}$ cell entering the intake system.</td>
</tr>
</tbody>
</table>
$N$  
Engine speed.

$[O_2]_{air}$  
Oxygen concentration of ideal air on a mass basis.

$[O_2]_{amb}$  
Oxygen concentration of ambient air on a mass basis.

$[O_2]_{comb}$  
Oxygen concentration after combustion on a mass basis.

$[O_2]_{cv}$  
Oxygen concentration of a generic control volume on a mass basis.

$[O_2]_{cyl}$  
In-cylinder oxygen concentration on a mass basis.

$[O_2]_{egr}$  
Oxygen concentration at the confluence junction where the EGR system meets the intake system.

$[O_2]_{in}$  
Oxygen concentration of the gas entering a control volume (only one inlet).

$[O_2]_{in,i}$  
Oxygen concentration of the $i^{th}$ stream entering a control volume.

$[O_2]_{int}$  
Oxygen concentration at the intake ports.

$[O_2]_{out,j}$  
Oxygen concentration of the $j^{th}$ stream exiting a control volume.

$\mathcal{P}$  
Positive definite ISS Lyapunov matrix.

$\mathcal{P}_1$  
Positive definite ISS Lyapunov matrix $\mathcal{P}$ with an engine cycle index $\overline{k}$.

$\mathcal{P}_2$  
Positive definite ISS Lyapunov matrix $\mathcal{P}$ with an engine cycle index $\overline{k} + 1$.

$P_{KF}$  
Kalman filter prediction error covariance matrix.

$P_c$  
Compressor power.

$P_{rls}$  
RLS recursion matrix.

$P_t$  
Turbocharger power.
\( \dot{Q}_{\text{gen}} \)  Heat generate rate.

\( \dot{Q}_{\text{ht}} \)  Heat transfer rate into the control volume.

\( Q_{KF} \)  Kalman filter process noise covariance matrix.

\( R \)  Universal gas constant.

\( R^2 \)  Correlation metric.

\( R_{KF} \)  Kalman filter measurement variance.

\( \dot{R}_{O_2} \)  Rate at with oxygen is produced by chemical reactions.

\( S_0, S_1, S_2 \)  Symmetric positive definite matrices used to design LPV state estimator.

\( Sec_{q_{exh}}^{\xi,i} \)  Oxygen concentration of the \( i^{\text{th}} \) cell in pipe \( \xi \) of the exhaust system after the \( q_{exh}^{\text{th}} \) gas packet has been exhausted.

\( Sec_{\xi,i} \)  Event averaged oxygen concentration of the \( i^{\text{th}} \) cell in pipe \( \xi \) of the exhaust system.

\( Sec_{n_{pipes},n_{pipes}} \)  Event averaged oxygen concentration at the location of the exhaust oxygen sensor.

\( SEC_{\xi}^{q_{exh}} \)  Matrix containing the input dependency vectors of the cells in pipe \( \xi \) of the exhaust system after the \( q_{exh}^{\text{th}} \) gas packet has been exhausted.

\( SEC_{\xi}^{q_{exh}}(i,:) \)  Input dependency vector of the \( i^{\text{th}} \) cell in pipe \( \xi \) of the exhaust system after the \( q_{exh}^{\text{th}} \) gas packet has been exhausted.

\( SEC_{\xi}^{q_{exh}}(i,j) \)  The \( j^{\text{th}} \) element of the input dependency vector of the \( i^{\text{th}} \) cell in pipe \( \xi \) of the exhaust system after the \( q_{exh}^{\text{th}} \) gas packet has been exhausted.

xxiii
\( SEC_{n_{\text{pipe}}} \) Event averaged input dependency vector at the location of the exhaust oxygen sensor.

\( T_a \) Ambient temperature.

\( T_{cv} \) Control volume temperature.

\( T_{egr,cv} \) Lumped EGR control volume temperature.

\( T_{exh,cv} \) Lumped exhaust control volume temperature.

\( T_{int,cv} \) Lumped intake control volume temperature.

\( T_{in,i} \) Gas temperature of inlet stream \( i \).

\( T_{out,j} \) Gas temperature of outlet stream \( j \).

\( T_r \) Gas temperature at flow restriction.

\( U \) Extended vector of commanded EQR values.

\( V \) ISS Lyapunov function.

\( V \) Generic square matrix.

\( V_{cv} \) Volume within a generic control volume.

\( V_d \) Displacement volume.

\( V_{egr,cv} \) Total volume of the lumped EGR control volume.

\( V_{exh,cv} \) Total volume of the lumped exhaust control volume.

\( V_{int,cv} \) Total volume of the lumped intake control volume.

\( W \) Generic full rank matrix.

\( Y \) Vector of measured AFR values.

\( Z_{pd} \) Generic positive definite matrix.

\( Z_{psd} \) Generic positive semidefinite matrix.

\( c_{\text{imb}} \) Steady-state cylinder imbalance vector.

\( c_p \) Constant pressure specific heat.
\(c_{\text{p,air}}\) Constant pressure specific heat of air.

\(c_{\text{p,comb}}\) Constant pressure specific heat of the products of complete combustion.

\(c_v\) Constant volume specific heat.

\(c_{v,\text{air}}\) Constant volume specific heat of air.

\(c_{v,\text{comb}}\) Constant volume specific heat of the products of complete combustion.

\(d_{\text{egr}}\) Transport delay of the EGR system.

\(d_{\text{egr,max}}\) Imposed upper bound on the estimated transport delay of the EGR system.

\(d_{\text{exh}}\) Transport delay of the exhaust manifold system.

\(d_{\text{exh,uego}}\) Transport delay between the exhaust ports and the post-turbocharger UEGO sensor.

\(d_{\text{exh,max}}\) Imposed upper bound on the estimated transport delay of the exhaust manifold system.

\(d_{\text{fuel,exh}}\) Fixed delay between injecting fuel and exhausting combusted gases.

\(d_{\text{fuel,int}}\) Fixed delay between injecting fuel and inducting charge.

\(d_{\text{int,fuel}}\) Fixed delay between inducting charge and injecting fuel.

\(d_{\text{int,exh}}\) Fixed delay between induction and exhaust strokes.

\(d_{\text{int}}\) Transport delay of the intake manifold system.

\(d_{\text{int,max}}\) Imposed upper bound on the estimated transport delay of the intake manifold system.

\(f\) Orifice flow correction factor.
$f_{body}$  Body force (one dimensional control volume).

$f_{corr}$  Hybrid oxygen concentration estimator correction factor.

$h$  Specific enthalpy.

$i o_j$  $j^{th}$ coefficient of the FIR exhaust oxygen dynamics model.

$k_{rate}$  Oxygen storage/release rate of the TWC.

$m_{air,\text{trapped}}$  Trapped unburned air mass.

$m_{cv}$  Mass within a generic control volume.

$m_{film}$  Mass of fuel film.

$m_{fuel,\text{trapped}}$  Trapped fuel mass.

$m_{res}$  Mass of trapped residuals.

$m_{\text{trapped}}$  Total trapped cylinder mass.

$\dot{m}_c$  Mass flow rate through the compressor.

$\dot{m}_{cat}$  Mass flow rate entering the catalyst.

$\dot{m}_{charge}$  Charge mass flow rate.

$\dot{m}_{egr}$  Mass flow rate entering the EGR system.

$\dot{m}_{egr,max}$  Maximum mass flow rate of EGR gas.

$\dot{m}_{eng}$  Mass flow rate into the engine.

$\dot{m}_{exh}$  Mass flow rate entering the exhaust system.

$\dot{m}_{fuel}$  Mass flow rate of fuel injected directly into the cylinders.

$\dot{m}_{fuel,\text{in}}$  Mass flow rate of fuel into the fuel film.

$\dot{m}_{fuel,\text{inj}}$  Mass flow rate of fuel injected into the intake ports.

$\dot{m}_{fuel,\text{out}}$  Mass flow rate of fuel out of the fuel film.

$\dot{m}_{in}$  Mass flow rate entering a control volume (only one inlet).

$\dot{m}_{in,i}$  Mass flow rate entering a control volume from inlet $i$. 
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\dot{m}_{out,j}$</td>
<td>Mass flow rate exiting a control volume from outlet $j$.</td>
</tr>
<tr>
<td>$\dot{m}_r$</td>
<td>Mass flow rate through a restriction.</td>
</tr>
<tr>
<td>$\dot{m}_t$</td>
<td>Mass flow rate through the turbocharger.</td>
</tr>
<tr>
<td>$n_\xi$</td>
<td>Number of cells in pipe $\xi$.</td>
</tr>
<tr>
<td>$n_{cyl}$</td>
<td>Number of cylinders.</td>
</tr>
<tr>
<td>$n_{egr}$</td>
<td>Number of cells used to model the EGR control volume.</td>
</tr>
<tr>
<td>$n_{events}$</td>
<td>Number of events per engine cycle.</td>
</tr>
<tr>
<td>$n_{exh}$</td>
<td>Number of cells used to model the exhaust control volume.</td>
</tr>
<tr>
<td>$n_{in}$</td>
<td>Number of inlet streams entering a control volume.</td>
</tr>
<tr>
<td>$n_{int}$</td>
<td>Number of cells used to model the intake control volume.</td>
</tr>
<tr>
<td>$n_{out}$</td>
<td>Number of outlet streams exiting a control volume.</td>
</tr>
<tr>
<td>$n_{pipes}$</td>
<td>Number of pipes in the exhaust system.</td>
</tr>
<tr>
<td>$n_{total}$</td>
<td>Total number of cells in the exhaust model.</td>
</tr>
<tr>
<td>$n_{window}$</td>
<td>Size of the output window used by the FIR model.</td>
</tr>
<tr>
<td>$p_{cv}$</td>
<td>Pressure within a generic control volume.</td>
</tr>
<tr>
<td>$p_d$</td>
<td>Pressure downstream of a restriction.</td>
</tr>
<tr>
<td>$p_{egr,cv}$</td>
<td>Lumped EGR control volume pressure.</td>
</tr>
<tr>
<td>$p_{exh,cv}$</td>
<td>Lumped exhaust control volume pressure.</td>
</tr>
<tr>
<td>$p_{exh,cv}$</td>
<td>Lumped intake control volume pressure.</td>
</tr>
<tr>
<td>$p_u$</td>
<td>Pressure upstream of a restriction.</td>
</tr>
<tr>
<td>$\dot{q}_{gen}$</td>
<td>Heat generation rate.</td>
</tr>
<tr>
<td>$\dot{q}_{ht}$</td>
<td>Heat transfer rate into the control volume.</td>
</tr>
<tr>
<td>$q_{in}$</td>
<td>Volumetric flow rate entering a control volume.</td>
</tr>
<tr>
<td>$r_{egr}$</td>
<td>Ratio of the outgoing exhaust cells to the incoming EGR cells.</td>
</tr>
</tbody>
</table>
$r_{int}$ Ratio of the outgoing EGR cells to the incoming intake cells.

$\dot{r}_{O_2}$ Rate at with oxygen is produced by chemical reactions.

$u$ Specific internal energy.

$u_{KP}$ Input vector of the Kalman filter oxygen dynamics model.

$u_{LVP}$ Input vector of the LPV oxygen dynamics model.

$u_{window}$ Vector of commanded EQR values.

$v$ Gas velocity.

$v_d$ ISS lumped disturbance term (inputs and outputs).

$w$ Output error disturbance variable.

$x$ State vector of the LPV oxygen dynamics model.

$y$ Output of the LPV oxygen dynamics model (exhaust oxygen concentration).

$y_{meas}$ Measured value.

$\bar{y}_{meas}$ Average measured value.

$y_{pred}$ Predicted value.

$z$ Distance variable.

**INDEX VARIABLES**

$\vartheta$ Generic event offset.

$e$ Continuous event domain variable.

$i$ Generic index.

$j$ Generic index.

$k$ Engine event index.

$k$ Engine cycle index.

$q_{egr}$ Flow index for the EGR system.
\( q_{\text{exh}} \) Flow index for the exhaust system.

\( q_{\text{int}} \) Flow index for the intake system.

\( s \) EGO switching event index.

\( t \) Continuous time domain variable.

\( \xi \) Exhaust pipe designator.

**SPECIAL MODIFIERS**

\( \Delta \) Change in a variable.

\( \dagger \) Nominal value of variable.

\( \ast \) Per event value of variable.

\( \hat{\ast} \) Estimated value of variable.

\( \sim \) Error in the variable.

\( \swarrow \) Value of a variable after rounding up.
CHAPTER 1

Introduction

To meet the increasing stringent federal emission standards and the customer’s desire for increased power and fuel efficiency, internal combustion engines have adopted many new technologies such as variable geometry turbochargers (VGT), fuel injectors capable of multiple injections and variable cam timing actuators. In the rush to be first to market and gain a competitive advantage, the existing control and calibration approaches used by the engine manufacturers are simply modified to accommodate the new technologies rather than designing entirely new approaches which take full advantage of the technologies. Lookup tables which were formerly functions of two variables have become functions of three and four variables. Increasing the dimensionally in this manner allows for most of the benefits of the new technologies to be realized but at the expense of a significantly more expensive and time consuming calibration process.

By coupling physically based models with a model based control architecture, it is possible to achieve equivalent or better performance than the current production approaches with significantly less calibration effort. Instead of relying exclusively on experimental data, an engine “truth” model with sufficient fidelity can supplement experimental methods by helping to characterize and predict the response of an
engine in an offline setting. On the other hand, a computationally efficient “control-oriented” engine model can predict the response of an engine in real time, by integrating available measurement and input information, to sufficient accuracy so that when combined with modern control techniques, very accurate estimators and stability guaranteed controllers can be designed using systematic approaches which require very little if any experimental data.

One of the biggest obstacles in the entire design process is the synthesis of the control-oriented model. Typically, high order truth models are simplified by making assumptions to reduce the model order and by linearizing all nonlinearities. The resulting control-oriented model must necessarily be simple to enable real time execution in a production engine control unit (ECU) and to facilitate control design. Current approaches, however, tend to oversimplify their control-oriented models and often make inappropriate assumptions. In an attempt to account for some of the unmodeled dynamics, control-oriented models commonly consist of parameters which vary as a function of the operating conditions (e.g. engine speed). This type of parameter scheduling can help improve the model accuracy, but many times the models are so oversimplified that they rather poorly approximate the physical systems they are meant to represent. The resulting controllers designed based on these models often use gain scheduling methods to match the parameter scheduling methods used in the model. Calibrating such a model can controller can require extensive effort and the performance may not be adequate when the controller is implemented on a physical engine.
This dissertation focuses on finding a unified, production-ready model based approach to fueling control in internal combustion engines that best satisfies the compromises between model accuracy, model order, control design feasibility, control performance, computational complexity and most importantly calibration effort. To demonstrate its robustness, this approach has been applied to both gasoline engines and diesel engines. Because of the distinct differences in the physical hardware and techniques used to produce combustion, the current production approaches to fueling control in diesel engines are vastly different than the current production approaches to fueling control in gasoline engines. In a diesel engine, fuel is injected directly into the combustion chamber in several distinct stages, whereas in a conventional gasoline engine, fuel is injected upstream of the combustion chamber at the intake ports. To produce low emissions, the timing and duration of the fueling pulses injected into a diesel engine are adjusted while keeping the total fuel mass constant. Conversely, gasoline engines produce low emission by adjusting the injected fuel mass to ensure that the combustion mixture is stoichiometric.

Even though diesel and gasoline engines have all of these differences, the biggest challenges in controlling the fueling processes for either type of engine is accurately estimating the oxygen concentration of the trapped gas charge. To this end, three different truth models of varying fidelity have been developed to approximate the relevant physical phenomena within the fluid paths of an internal combustion engine including most notably the transport delay. Using a systematic approach which preserves the dominant dynamics including the transport delay, a series of control-oriented models have also been constructed. Through the application of modern control theory, advanced control systems based on these models are separately developed for diesel
engines and for gasoline engines to achieve the performance, robustness and stability requirements. Furthermore, the high fidelity models are used to alleviate the overall calibration burden by using model aided calibration procedures instead of experimentally based calibration procedures. Both the models and the estimators/controllers developed from them represent significant contributions to the modeling, control and automotive communities.

1.1 Organization

The overall objective of an internal combustion engine is to produce torque by burning fueling. In general, the combustion process is dependent on three factors: air, fuel and the combustion trigger. Although all three of these factors are important, this dissertation focuses exclusively on controlling the fueling portion of this recipe. To better define the importance and scope of this control problem, Chapter 2 provides a thorough overview of the current practices in the fields of diesel fueling control and gasoline fueling control. Typically a primary fueling controller determines a nominal fueling mass and a cylinder imbalance controller adjusts the nominal fueling mass on a cylinder-by-cylinder basis to ensure uniform performance. Sections 2.1 and 2.2 describe the current “nominal” fueling control techniques which have been applied to diesel engines and those which have been applied to gasoline engines; Section 2.3 describes the current cylinder imbalance techniques. Many of these approaches rely on subsystem level models of an engine either to help design a model based controller or to estimate the control inputs. Section 2.4 provides an overview of these models with an emphasis on current air path models.
An air path model is often used to predict the trapped charge mass, the in-cylinder oxygen concentration and the oxygen concentration of the exhaust stream at the location of oxygen sensor. The modeling portion of this dissertation which is described in Chapter 3 focuses on the oxygen transport and mixing dynamics within the air path system for the specific applications of estimating the oxygen concentration inside the cylinders and at the oxygen sensor. To model these dynamics, a plug flow approximation is applied to a series of control volumes used to represent the air path system. The derivation of the equations which form the basis of a plug flow based oxygen dynamics model is presented in Section 3.1. Each one of the model based control applications (diesel fueling control, gasoline fueling control and cylinder imbalance control) has different modeling requirements necessitating the construction of three different models: a lumped oxygen dynamics model, a high fidelity exhaust oxygen dynamics model and a finite impulse response oxygen dynamics model. Sections 3.2-3.4 are devoted to describing these models.

As with most truth models, these plug flow oxygen dynamics models must be simplified before they can be used in a real time estimation application. Chapter 4 outlines the process by which a lumped oxygen dynamics model can be simplified to assist in real time in-cylinder oxygen concentration estimation and a finite impulse response oxygen dynamics model can be recast to enable online cylinder imbalance estimation. In Section 4.1, two reduced order oxygen dynamics models are derived allowing for a linear parameter varying state estimator and an extend Kalman filter to be created. Since the in-cylinder oxygen concentration cannot be easily measured, the performance of these estimators is compared in simulation to a traditional in-cylinder oxygen concentration estimator. In a similar manner, Section 4.2 introduces
a method of configuring a finite impulse response oxygen dynamics model into a time varying vector representation that is paired with a recursive least squares estimator. The estimation potential of this technique is demonstrated experimentally.

Without an estimate of the in-cylinder conditions, a conventional diesel fueling controller is unable to account for the dynamic behavior of the air path system. Chapter 5 addresses this issue and proposes an alternative fueling control architecture based on a dynamic estimate of the cylinder contents. Section 5.1 outlines this new diesel fueling control method and explains its benefits, in particular its robustness to errors in the air path control system. In the proposed fueling controller, three-dimensional tables rather than two-dimensional tables are used to predict each of the fueling parameters. Numerous variables such as the trapped oxygen mass and the combustion mixture AFR could be used in this fueling control structure. Through an empirical study, the results of which are presented in Section 5.2, it was determined that the variables that best capture the variation in the desired values of the fueling parameters across all conditions are the in-cylinder oxygen concentration, the engine speed and the total fuel mass. The proposed fueling controller with these table indexing variables has been implemented in simulation using the in-cylinder oxygen concentration estimator designed in Section 4.1 and is directly compared to a conventional fueling controller in Section 5.3. The benefits of an in-cylinder oxygen concentration based diesel fueling controller especially in terms of robustness are further highlighted by this study.

In contrast to the table based feed-forward fueling controllers which characterize diesel engines, the design focus of a gasoline fueling controller is the feedback loop(s). Chapter 6 proposes a new unified pre-catalyst and post-catalyst fueling
control system that transforms the typical regulation problem into an exhaust gas oxygen (EGO) sensor switching frequency control problem. Using some additional nomenclature conveniently organized in Section 6.1, the pre-catalyst portion of the new control formulation is defined in Section 6.2 and the post-catalyst portion is defined in Section 6.3. The robust stability characteristics of this control system are analyzed in Section 6.4. Aside from its stability properties, the two most compelling reasons for using this controller architecture are its significantly condensed experimental calibration requirements and its ability to maintain the catalyst efficiency at a high level. By combining a high fidelity exhaust oxygen dynamics model with other subsystem level models, the calibration procedures outlined in Section 6.5 explain how nearly all of the control gains can be calibrated in simulation. The primary motivation for the unique EGO switching frequency design of this controller is the link between periodic oxygen concentration oscillations and the conversion efficiency of a three-way catalyst. More specifically, previous studies have shown that when all other variable are held constant dithering the air-to-fuel ratio of the combustion mixture increases the efficiency range of a three-way catalyst (TWC). To verify that the benefits of dithering the air-to-fuel ratio (AFR) can be realized under real world driving conditions and to determine the best type of dither signal, an experimentally based investigation was performed. Section 6.8 provides all of the testing details as well as the results which confirm that oscillating the AFR is indeed beneficial. In the final validation described in Section 6.9, the AFR switching frequency controller is directly compared to the existing production controller. These results show that in terms of every relevant metric an AFR switching frequency controller achieves equivalent or better performance.
To close, Chapter 7 summarizes the modeling, estimation and control contributions of this dissertation. The improved model accuracy achieved by explicitly capturing the oxygen transport and mixing dynamics has in turn produced more effective estimators and controllers. Harmonizing these advanced models with model based estimation and model based control techniques also streamlines the calibration process that is particularly burdensome for current direct tuning approaches.
CHAPTER 2

Background and Motivation

To properly motivate the research presented in this dissertation, an overview of past fueling control research is now presented. For a conventional diesel engine, fuel is typically delivered based on a feed-forward mapping which does not actively account for the state of the air path system. The most commonly encountered disturbances alter the behavior of the air path system, so in addition to producing poor transient performance this type of control structure lacks robustness.

Fuel injection in gasoline engines incorporates both feed-forward and feedback control techniques. A wide range of closed loop control systems based on oxygen oxygen concentration feedback have been developed using different model based control techniques. Nearly all of these approaches suffer from the same two shortcomings. Firstly, the error signals which these controllers try to minimize do not directly correspond to the overall objective of minimizing the tailpipe emissions. Secondly, the models used to develop these control laws do not capture the oxygen mixing and transport dynamics within the exhaust system.

Most fueling controllers (both diesel and gasoline) rely on a separate controller to correct cylinder imbalances. Cylinder imbalance is often estimated by analyzing the speed fluctuations of the crankshaft, analyzing the cylinder pressure traces or
analyzing the exhaust oxygen concentration. The first two approaches are best suited for identifying torque imbalances whereas the latter approach is best suited for identifying AFR imbalances. Current oxygen sensor based approaches rely exclusively on empirical data.

All of these current deficiencies can potentially be eliminated with a model based control approach when an appropriate model is used. Many of the systems which comprise an engine are modeled sufficiently well; however, the oxygen transport and mixing dynamics of the air path system are consistently ignored or oversimplified. Developing a new modeling approach to properly capture these dynamics is the foundation of this dissertation.

2.1 Feed-forward Diesel Fueling Control

The diesel combustion process strongly depends on the contents of the cylinder (fresh air mass and residual gas mass), the homogeneity of the mixture, the injected fuel mass and the fuel injection timing. The majority of the injected fuel is delivered in a single pulse called the main injection. Injecting fuel prior to the main injection (pilot injection) adds energy to the combustion mixture thereby significantly reducing the ignition delay of the main injection, especially at low temperatures [2, 3]. As a result, the majority of the heat release occurs closer to top-dead-center which results in a higher thermodynamic efficiency. However, the improvement in fuel consumption can also be accompanied by an unwanted increase in NO\textsubscript{x} emissions [4]. Injecting fuel after the main injection (post injection) adds energy to the system after the primary combustion process has occurred. This energy increases the temperature of the combustion gases at the end of the combustion cycle without increasing the peak
combustion temperature which is directly linked to NO\textsubscript{x} production. A higher end-of-cycle temperature promotes higher soot oxidation rates and an overall decrease in the engine-out soot levels [4]. Higher end-of-cycle temperatures are also very important to the efficiency of aftertreatment components such as selective catalytic reduction (SCR) catalysts.

For a diesel engine to produce combustion which provides the optimal trade-off between performance, noise, fuel economy and emissions, both the cylinder contents and the injected fuel must be precisely controlled. To this end, an air path controller ensures that the cylinders contain the correct masses of fresh air and residual gases, while a fueling controller determines the optimal start of injection (SOI) timings and the optimal fuel injection quantities for the given cylinder contents. Consider the conventional diesel engine control diagram shown in Figure 2.1. Such a control system consists of a supervisory controller, a setpoint determination system including a feed-forward fueling controller and an air path controller [5]. At the highest level, the supervisory controller selects the system operating mode depending on the dynamics of the system (i.e. steady-state versus transient operation), the ambient air density (e.g. low density at altitude), the states of the aftertreatment systems (e.g. the soot level within a diesel particulate filter (DPF)) and other off-nominal considerations (e.g. turbocharger surge prevention).

Depending on the state of the control system, the relative importance of performance, noise, fuel economy and emissions changes. When the engine is operating in steady-state, the fuel consumption rate might be of high importance, whereas during a transient the torque response might be the most important consideration. To satisfy
each of these trade-off problems, both the air path setpoints and the fueling parameters are scheduled as functions of not only the driver torque request and the speed of the engine but also the system operating mode. This structure is typically realized by using different sets of two-dimensional fueling parameter and air path setpoint tables for each system operating mode. In contrast to the open loop approach used to control the fuel system, the air path controller uses sensor feedback measurements to achieve the desired air path setpoints as quickly as possible.

Many different approaches, including model based control [6, 7], Lyapunov based control [8], linear parameter varying control [9], gain-scheduled PI control [10] and adaptive control [11] have been used to simultaneously control the two degree-of-freedom air path actuation system (exhaust gas recirculation (EGR) valve position and VGT vain position) which has become an industry standard for diesel engines.
An air path controller tries to regulate a pair of air path variables to a pair of desired setpoints using measurement feedback. In one of the more common realizations, the VGT is adjusted to achieve a desired manifold pressure, whereas the EGR valve is adjusted to achieve a desired fresh air flow rate.

Several other setpoint variables including EGR flow rate, charge flow rate and EGR fraction have also been considered with the commonality between the variables being that each setpoint variable corresponds to a direct sensor measurement or can be readily derived from sensor measurements. One of the more interesting alternatives to the traditional setpoints is the exhaust oxygen concentration because it depends on both the air and fuel systems. Two studies that have considered this variable, [12] and [13], corrected setpoint errors exclusively using air path control action. Although these two air path controllers differ slightly from one another, both approaches use the same overall engine control structure which prevails throughout the literature. In all of the proposed approaches the fueling parameters and air path setpoints are both scheduled based on engine speed, the desired torque/fueling and the system operating mode (i.e. steady-state, transient, altitude etc.). Within this conventional control structure, the fueling parameters do not consider the instantaneous state of the air path system. As shown in [14], the choice of the controlled air path variable pair does not have a significant effect on engine performance when the fuel injection parameters are scheduled in this manner.

The response time of the air path system is much slower than the fuel delivery system; however, the conventional control structure does not take the lag of the air system into account when selecting the fueling parameters. Both the fueling parameters and the air path setpoints are scheduled on the same inputs: engine speed,
total fueling and the system operating mode. Because of this pairing, the fueling control system is not directly tied to the instantaneous cylinder conditions but rather indirectly tied to the desired air path setpoints. As a result, the fueling parameters selected for a given set of transient air path setpoints must produce desirable combustion throughout the transient air path process. Depending on the original conditions, the actual instantaneous value of each air path setpoints could be either higher or lower than their desired value. Using a single value for such a wide range of possible operating conditions is an oversimplification that can significantly compromise the maximum achievable performance levels. Also, because the fuel delivery system is autonomous with respect to the air path system, corrective feedback control is limited exclusively to the air path system. Under certain situations, a desired air path setpoint may not be realizable. Although the achieved fresh air and EGR flow rates might be known, this information is not passed to the feed-forward fueling controller thus reducing the combustion performance.

In addition to these performance limitations, a conventional diesel engine controller system is also very difficult to calibrate, requiring copious amounts of experimental data. To calibrate even a single set of air path setpoint and fueling parameter tables, a multidimensional optimization must be performed at numerous engine torque (fuel mass) and engine speed pairs. For a fuel delivery system which has pilot, main and post injections, five fueling parameters (main SOI, pilot SOI, post SOI, pilot quantity and post quantity) must be optimized. The amount of experimental data required to simultaneously optimize five such fueling parameters and the two air path setpoints is staggering; a full factorial design of experiments (DoE) which considers only five different values for each variable would constitute 78,125 data points. Not
only must this procedure be repeated for other torque and engine speed pairs but also for different system operating modes. Clearly, engine manufacturers have developed calibration procedures which require more reasonable amounts of experimental data; nevertheless, calibrating a conventional fueling controller is both expensive and time consuming.

An alternative to the current control structure is to determine each of the fueling parameters using a single lookup table rather than having one table for each different system operating modes. Because each different system operating modes uses a different set of air path setpoints, a higher dimensional lookup table would be required. It has been well established that the in-cylinder oxygen concentration has a strong effect on the combustion process. As the oxygen concentration decreases due to the addition of a dilutant such as recirculated exhaust gas, the heat release rate decreases. To offset this effect, the start of combustion must be advanced by advancing the start of injection timings. Explicitly capturing the effect of the different air path setpoints with a third variable such as the in-cylinder oxygen concentration has the potential to reduce the fueling system calibration effort. To achieve an equivalent level of performance, though, the in-cylinder oxygen concentration must be dynamically estimated with high accuracy. Although the automotive community has invested much effort in developing efficient methods for estimating the in-cylinder oxygen concentration, there are still significant opportunities for improvement.

2.2 Air-to-Fuel Ratio Control in Gasoline Engines

In gasoline engine combustion control, maximizing the conversion efficiency of a TWC should be the most important control objective because minimizing the tailpipe
emission is the primary target of AFR control. It has been well documented that a TWC has the highest steady-state conversion efficiencies when the incoming exhaust gases are produced from stoichiometric combustion. When the AFR of the gases at the inlet to a catalyst fluctuates between lean and rich values, the catalyst switches between oxygen storage and oxygen depleting modes. The oxygen stored in a catalyst during lean operation is depleted during rich operation. The instantaneous storage capacity of a catalyst relative to its maximum storage capacity is very important to the instantaneous conversion efficiency of the catalyst. For lean operation, a catalyst stores oxygen until the maximum storage capacity is reached, during which time the effective conversion efficiency of the catalyst remains high. Similarly, the conversion efficiency during rich operation is high as long as the stored oxygen has not been completely depleted. These physical phenomena have been modeled in a variety of ways; however, most catalyst models only succeed at predicting the low frequency trends such as the cumulative emission mass over a drive cycle. When a catalyst is modeled using simple integrator models as in [15], the cumulative tailpipe emissions predictions over a FTP-72 (Federal Test Procedure) drive cycle can have errors as large as 38%. More complicated thermodynamic models which capture both spacial and temporal variations such as [16] can produce cumulative errors as large as 200%. Similar type models like [17] have shown better performance but only for step inputs. When faced with this level of uncertainty, experimental testing must be relied on heavily. Although TWC models work well for diagnostics [18], they are ill suited for model based control.
Two types of oxygen sensors, binary exhaust gas oxygen (EGO) sensors and universal exhaust gas oxygen (UEGO) sensors, can be used to provide a feedback measurement of the composition of the combustion mixture. A binary EGO sensor indicates whether the mixture is rich or lean, whereas a UEGO sensor actually indicates the AFR magnitude of the combustion mixture. Near stoichiometry, which is the desired operating region, an EGO sensor has a much higher sensitivity than a UEGO sensor. As long as an accurate AFR estimator can be designed, an EGO based control scheme can reject disturbances as well as a UEGO based control scheme. Given that the higher sensitivity near stoichiometry provided by an EGO sensor enables more precise control during steady-state operation, lower cost EGO based control schemes can perform as well as UEGO based schemes. For these reasons many of the gasoline powered automobiles in production today still use EGO sensors.

Virtually all AFR control structures used in practice and proposed by those in academia consist of two separate controllers, one based on a pre-catalyst oxygen sensor and one based on a post-catalyst oxygen sensor. The control objective for the pre-catalyst controller is almost always to regulate the pre-catalyst AFR to stoichiometry. Conversely, post-catalyst controllers try to regulate the oxygen stored in the TWC to some nominal value. Most of the control authority resides with the pre-catalyst AFR controller. If the pre-catalyst AFR could be regulated perfectly to stoichiometry, then the tail-pipe emissions would be minimized. However, even the best pre-catalyst AFR controllers still produce deviations from stoichiometry which are often significant. Moreover, while trying to regulate to stoichiometry, most pre-catalyst AFR controllers produce uncontrolled limit-cycle type behavior. Oscillating the AFR around stoichiometry extends the conversion efficiency range of a
TWC [19, 20]; however, uncontrolled oscillations due to limit cycle behavior should be avoided.

For sinusoidal AFR oscillations, the theoretical investigation presented in [21] found that both the amplitude and the frequency of the sinusoidal signal affected the average conversion efficiency of a TWC. The experimentally based work presented in [22] explored how the shape of a triangle wave affects the conversion efficiency of a catalyst. Excluding a few pulse-like changes in the fueling, the performance evaluations shown in [22] were all conducted during steady-state operation. Although these papers motivate the use of tracking based AFR controllers, they only studied the impact of different periodic waveforms under fixed artificial conditions. Under real world conditions where the composition of the exhaust, the exhaust flow rate and the oxygen stored in the catalyst are all changing, it is significantly more difficult to predict how periodically dithering the AFR of the combustion mixture in different manners affects the conversion efficiency of a TWC.

Although university based AFR control research has evolved over the years especially with advent of the TWC, the control objective of most AFR controllers still does not directly match the overall objective of emissions reduction. Most advanced control formulations such as $H_\infty$ controllers [23], Lyapunov based controllers [24], model-predictive control [25] and linear parameter varying controllers [26] require UEGO sensors and are developed using overly simplified plant models. The $H_\infty$ control law developed in [23] for example is based on a fixed exhaust mixing and delay model. Similarly, [24] uses a simple discrete state space model and [26] uses a first order lag model. None of these models account for the time varying nature of the
transport delay within the exhaust system. This type of unmodelled dynamics almost certainly leads to uncontrolled oscillations and can potentially cause instability.

Sliding mode controllers possess a degree of robustness, and several formulations such as [27], [28] and [29] only require a binary EGO sensor. However, all three of these formations rely heavily on the accuracy of empirical models. Even with a very accurate engine plant model, sliding mode controllers struggle to precisely regulate the AFR during steady-state. Chattering is an inherent problem common to all sliding mode controllers. For systems such as an internal combustion engine which are characterized by a time varying delay, this chattering behavior is even more difficult to overcome. As a result, sliding mode based AFR controllers are able to reject disturbances well but have a difficult time precisely regulating the EGO switching frequency once stoichiometry is reached.

Because these controllers are not able to precisely regulate the AFR trajectory to stoichiometry, the average AFR may be slightly rich or slightly lean even when the controller is meeting its control objective. These effects are compounded by sensor errors and slow sensor dynamics, particularly for UEGO sensors. The composition of the exhaust stream can cause oxygen sensor deception due to varying diffusion rates of the gas constituents, thus indicating an AFR or switching point different from the actual AFR incident on the sensor [30–32]. The resulting drift in the oxygen stored within the TWC may not be corrected by the post-catalyst controller until a breakthrough occurs. Even papers which present complete (both pre-catalyst and post-catalyst) AFR control architectures such as [33] and [34] are prone to the same problems because they still use pre-catalyst controllers which regulate to an AFR set-point.
In practice, manufacturers predominately use gain scheduled PI controllers to regulate the pre-catalyst and post-catalyst AFR. Although this type of control structure is very simple, it is extremely difficult to calibrate because the control must be robust to such a wide range of AFR disturbances. To tune the gain scheduled proportional and integral terms, many hundreds of drive cycles must be driven experimentally. Even with all of this data, the methods used to select the final controller gains are very subjective and vary from calibration engineer to calibration engineer. In lieu of an analytical proof, engine manufacturers try to verify the closed loop stability of an AFR controller with even more experimental testing.

2.3 Cylinder Imbalance Estimation and Control

Cylinder imbalance is a general term which is used to describe any condition in which the contents of the cylinders are not all uniform. Cylinder imbalance is undesirable and can cause additional noise, driveline oscillations, a reduction in driveability and/or an increase in the emissions production. Many different sources can contribute to cylinder imbalance including breathing differences and fuel delivery differences. Because the physical characteristics of the intake and exhaust runner systems are different for each cylinder, the resulting three-dimensional fluid dynamics and wave dynamics can potentially cause cylinder-to-cylinder differences in the inducted mass or the trapped residual mass. For engines with external EGR, these same effects can cause a maldistribution of the burned recirculated gases. Conversely, manufacturing differences between fuel injectors and/or aging effects can cause the delivered fuel mass to vary between cylinders.
Although many engines now have flexible valve actuation technologies such as variable cam timing and multiple cam profiles, these techniques uniformly affect all of the cylinders. Fully flexible valve actuation via electromechanical or hydraulic valve actuation could potentially be used to independently control the air induction and exhausting processes for each cylinder; however, these technologies are not likely to be implemented in a production engine in the near future. Because modern engines use separate fuel injectors for each cylinder, the fueling system can be controlled independently for each cylinder. Normally the same fuel commands are given to the fuel injectors for each of the cylinders.

When fuel is delivered via port injection (common in gasoline engines), the variable of primary importance is the AFR of the combustion mixture. For these types of gasoline engines, the primary goal of a cylinder imbalance controller is to ensure that the combustion mixture within each cylinder has the same AFR. Regardless of whether cylinder imbalance is caused by a charge imbalance or a fuel imbalance, the mass of fuel delivered by each injector can be adjusted to correct an AFR imbalance error.

For direct injection engines, both the injected fuel mass and the injection timing are important. In diesel engines, fuel can be delivered in several distinct injection pulses. The injection timing and quantity of each injection is important. These added degrees of freedom provide better control of the combustion process. Depending on the operating conditions, different injection strategies are used; therefore, the ability to distinguish between a fueling imbalance and a charge imbalance is important. The three most prominent approaches to estimating cylinder imbalance are based on three different feedback measurement variables: crankshaft speed, in-cylinder pressure and
exhaust oxygen concentration. The crankshaft speed and in-cylinder pressure based approaches focus on estimated torque (fuel) imbalances whereas the exhaust oxygen concentration based approaches focus on estimation AFR imbalances.

2.3.1 Crankshaft Speed Fluctuation Based Imbalance Estimation and Control

When the fuel inside the cylinder is combusted, the pressure within the cylinders increases leading to an increase in the downward force on the piston head. Through the connecting rod linkage, a torque is transferred to the crankshaft causing the crankshaft to accelerate. By analyzing the crankshaft speed fluctuations that result from the combustion in each cylinder, it is possible to identify the torque produced by each cylinder. To make this prediction, the crankshaft is commonly modeled as a series of lumped inertias connected to each other via spring and damping elements [1,35–38]. In this approach, each cylinder is connected to a separate lumped inertia making the total number of lumped inertias equal to the number of cylinders. Figure 2.2 taken from [1] shows an illustration of a such a model for a six cylinder engine. Similar lower order models use only a single lumped inertia to model the crankshaft [39,40]. Using a lumped inertia crankshaft model, the combustion torque can be estimated using model inversion techniques [37], unknown input observer techniques [1], Kalman filtering [39] and frequency analysis techniques [35,36,38].

Alternatively, the crankshaft speed measurement can be also be used to identify cylinder imbalance by directly analyzing the measurement signal. In [40], for example, the difference between the maximum and minimum speed of the engine strokes is used to identify cylinder imbalance for a four cylinder engine. By studying the elapsed time in a certain fixed region of the crank angle domain, torque imbalances were
identified in [41, 42]. Although crank speed fluctuation based diagnostics can be very effective, they are only able to detect torque imbalances. As long as the net AFR of the combustion mixture is lean, the torque produced during combustion is relatively insensitive to changes in the charge mass; therefore, these approaches are not particularly adept at estimating cylinder imbalance caused by charge imbalance. On the other hand, variations in the start of combustion which can be caused by start of injection errors for example can cause the torque production to change. This type of fuel injection timing error is difficult to distinguish from a fuel mass injection error. Crankshaft speed fluctuation based cylinder balancing controllers typically assume that all identified torque imbalances are caused by fuel mass injection errors and compensate them accordingly.
2.3.2 Cylinder Pressure Based Imbalance Estimation and Control

An in-cylinder pressure transducer can facilitate direct control of the combustion process on a cylinder by cylinder basis. Unlike crank speed fluctuation based techniques, cylinder pressure information can be used to identify both fuel mass injection errors as well as fuel injection timing errors. Most cylinder pressure based combustion controllers, however, focus only on regulating the injection timing. For engines which use alternative combustion modes such as homogeneous charge compression ignition (HCCI), cylinder pressure based techniques are very desirable because the combustion process must be regulated tightly to avoid unstable combustion. A cylinder pressure measurement can be used to directly calculate the apparent heat release rate or the cylinder pressure signal can be analyzed directly. Because of the computational requirements, this type of approach is typically reserved for engines which use alternative combustion modes [43–47] although it has been studied in conventional diesel engines as well [48]. Using the apparent heat release rate, the crank angle in which fifty percent of the total heat has been released, CA$_{50}$, can be calculated. Most proposed cylinder pressure based combustion controllers use CA$_{50}$ as the combustion control variable because the desired value of CA$_{50}$ can be mapped relatively easily and the CA$_{50}$ can be directly controlled by changing the injection timings [43–47].

Alternatively, the cylinder pressure signal can be used to directly identify similar type variables without the intermediate and computationally taxing step of calculating the apparent heat release rate. In such an approach, the measured cylinder pressure is compared to the motored pressure to estimate the pressure rise caused by combustion. Using this signal, the start of combustion can be approximated as the
crank angle at which the pressure difference exceeds some predefined pressure [49] or the center of combustion can be estimated by calculating the center of gravity of the signal [50]. Although these variables are not as closely tied to combustion performance, they are effective alternatives to their heat release counterpart variables. Even without the extra step of calculating the apparent heat release rate, cylinder pressure based cylinder imbalance controllers have significant computational and sampling requirements. When also considering the sensor costs and their susceptibility to noise, engine manufacturers have yet to adopt a cylinder pressure based approach to combustion control.

2.3.3 Exhaust Oxygen Concentration Based Imbalance Estimation and Control

By comparing the instantaneous fluctuations in the measured exhaust oxygen concentration to the cycle average value of this measurement, it is possible to identify cylinder imbalance provided the effect that each cylinder has on the output oxygen concentration measurement can be predicted. Although numerous model based control techniques have been applied to pre-catalyst and post-catalyst AFR control, model based control has not yet been widely applied to oxygen concentration based cylinder imbalance control. Most of the cylinder imbalance controllers presented in the open literature are based on simple input/output relationships. More specifically, the oxygen sensor measurement (output) is represented as a weighted summation of the oxygen concentrations of the combustion mixtures for each cylinder (input). Physically, each weighting factor represents the ratio of the mass of combustion gases expelled by a particular cylinder that pass by the oxygen sensor to the total mass of combustion gases that pass by the oxygen sensor over a designated window of time.
These weighting factors are dependent on the location within the engine cycle (which cylinder is firing) and the exhaust flow rate. In steady-state, however, the weighting factors repeat every engine cycle. Since each cylinder produces combustion once every engine cycle, the period in engine events is equal to the number of cylinders, \( n_{cyl} \). For this reason, most of the proposed cylinder imbalance models are represented using a square matrix of size \( n_{cyl} \times n_{cyl} \) which contains the weighting factors for steady-state operation [51–54].

For a model of this structure to represent the physical engine, the total mass and the mass of each species, where the mass of exhaust gas expelled by each cylinder is considered a separate species, must be conserved. As a direct result of the conservation of mass law, the row-wise summation of the weighting factors should equal unity. Similarly, the conservation of species law requires that the column-wise summation of the weighting factors must be unity. Even though these requirements are necessary to produce an accurate and unbiased cylinder imbalance estimate, they are often not considered. It is important to note that when calibrating a matrix of weighting factors using experimental data, these requirements are not automatically satisfied, primarily because of measurement noise.

Another major shortcoming of matrix based cylinder imbalance models is the extensive calibration burden. Because the coefficients within these matrices correspond to steady-state operation, a new matrix must be calibrated for each zone of the operating space. Other cylinder imbalance models take different forms but still require zone based calibrations [55–58]. Even with a very fine grid characterization of the zones, the accuracy of a smooth model which is a continuous function of the operating parameters cannot be matched. Nearly all of the proposed cylinder imbalance
controllers and estimators have been designed specifically for gasoline engines. In a diesel engine the exhaust oxygen sensor is generally placed after the turbocharger, making the transport delay even more significant. Nevertheless exhaust oxygen concentration based cylinder imbalance is possible even in turbocharged diesel engines. The investigatory testing conducted in [59] and [60] confirm this for a fixed operating condition.

2.4 Engine Modeling

The differences between the feed-forward diesel fueling control problem and the feedback AFR control problem require slightly different modeling approaches. The modeling objective of the diesel engine model used in this dissertation is to dynamically estimate the oxygen concentration inside the engine cylinders. To make this prediction, the air path of a diesel engine must be modeled. In a typical air path model, the fluid and thermal dynamics of the intake, exhaust and EGR fluid paths as well as the rotational dynamics of the turbocharger/compressor are modeled. Most commonly the intake, exhaust and EGR systems are approximated with zero dimensional control volumes whose states (temperature, pressure, and oxygen concentration) are governed by the first law of thermodynamics, the conservation of mass law, the conservation of species law and the ideal gas law. To calculate the mass flow rate entering and exiting these control volumes, a wide range of algebraic relationships such as the orifice flow equation are utilized. The mass flow rates through the turbocharger and compressor are sometimes predicted using empirically based efficiency maps scheduled on the turbine speed. To predict the turbine speed, Newton’s second
law is applied to a lumped rotational inertia model of the turbocharger/compressor system.

The objective of the gasoline engine model used in this dissertation is to predict the oxygen concentration of the gas charge trapped inside the engine cylinders by using feedback measurements of the oxygen concentration within the exhaust system. The same techniques used to model the air path of a diesel engine are also used to model the air path system of a gasoline engine; however, other systems must also be modeled to best incorporate the available measurement information. In a diesel engine, the exhaust oxygen concentration is often measured using a single oxygen sensor placed downstream of the turbocharger. In a gasoline engine, an oxygen sensor is placed both upstream and downstream of the TWC. To predict the response of the post-catalyst oxygen sensor, the TWC must be modeled. A port injected gasoline engine must also account for the puddling and evaporation dynamics of the fuel. A description of the most common approaches used to model each of these systems follows.

### 2.4.1 Conventional Air Path Modeling Approach

The dynamic behavior of the compressible gases within an engine air path system can be characterized using three-dimensional, one-dimensional or zero-dimensional models. Both three-dimensional and one-dimensional models are described by coupled sets of nonlinear partial differential equations. The complexity of these models makes them ill suited for control applications. Thus in the automotive control community, air path systems are typically modeled using a simple zero-dimensional approach. The next two subsections will provide an overview of zero- and one-dimensional air path modeling techniques.
Zero-Dimensional Air Path Models

The zero-dimensional modeling process begins by combining large sections of piping into lumped control volumes. To model an engine with external EGR for example, three control volumes which represent the intake manifold, the EGR system and exhaust manifold are often used. As the pressure inside one of these control volumes increases, the control volume fills with more gas (i.e. the mass increases). Conversely, as the pressure inside this control volume decreases, the gases empty out of the control volume (i.e. the mass decreases). These so called filling and emptying dynamics are modeling by applying the first law of thermodynamics, the conservation of mass law and the ideal gas law to each of the control volumes. This well established approach [9,11–14,61–73] gives rise to the following general state equations

\[
\frac{dp_{cv}}{dt} = \frac{\gamma hr}{V_{cv}} \left( \frac{\dot{Q}_{ht} + \dot{Q}_{gen}}{c_p} + \sum_{i=1}^{n_{in}} \dot{m}_{in,i} T_{in,i} - \sum_{j=1}^{n_{out}} \dot{m}_{out,j} T_{out,i} \right) \tag{2.1}
\]

and

\[
\frac{dT_{cv}}{dt} = \frac{RT_{cv}}{p_{cv}V_{cv}} \left( \frac{\dot{Q}_{ht} + \dot{Q}_{gen}}{c_v} + \sum_{i=1}^{n_{in}} \dot{m}_{in,i} (\gamma hr T_{in,i} - T_{cv}) - \sum_{j=1}^{n_{out}} \dot{m}_{out,j} (\gamma hr T_{out,j} - T_{cv}) \right) \tag{2.2}
\]

where \(p_{cv}\) is the mass averaged pressure within the control volume, \(R\) is the gas constant, \(V_{cv}\) is the volume of the control volume, \(\dot{Q}_{ht}\) is the heat transfer rate into the control volume, \(\dot{Q}_{gen}\) is the rate of heat generation within the control volume, \(n_{in}\) is the number of incoming fluid paths, \(n_{out}\) is the number of outgoing fluid paths, \(\dot{m}_{in,i}\) is the mass flow rate of the \(i^{th}\) stream entering the control volume, \(T_{in,i}\) is the gas temperature of the \(i^{th}\) stream entering the control volume, \(\dot{m}_{out,j}\) is the mass flow rate of the \(j^{th}\) stream exiting the control volume, \(T_{out,j}\) is the gas temperature of the \(j^{th}\) stream exiting the control volume and \(T_{cv}\) is the mass averaged gas temperature within
the control volume. Additionally, $\gamma_{hr}$ is the specific heat ratio of the gaseous mixture, $c_p$ is the constant pressure specific heat and $c_v$ is the constant volume specific heat. If the gases within the control volume are assumed to mix perfectly, then these equations can be simplified further. More specifically, the properties of the gases exiting the control volume can be assumed to be equal to the mass averaged properties of the control volume. This simplifies the above equations to

$$\frac{dp_{cv}}{dt} = \frac{\gamma_{hr} R}{V_{cv}} \left( \frac{\dot{Q}_{ht} + \dot{Q}_{gen}}{c_p} + \sum_{i=1}^{n_{in}} \dot{m}_{in,i} T_{in,i} - T_{cv} \sum_{j=1}^{n_{out}} \dot{m}_{out,j} \right)$$  \hspace{1cm} (2.3)$$

and

$$\frac{dT_{cv}}{dt} = \frac{RT_{cv}}{p_{cv} V_{cv}} \left( \frac{\dot{Q}_{ht} + \dot{Q}_{gen}}{c_v} + \sum_{i=1}^{n_{in}} \dot{m}_{in,i} (\gamma_{hr} T_{in,i} - T_{cv}) - T_{cv} \sum_{j=1}^{n_{out}} \dot{m}_{out,j} (\gamma_{hr} - 1) \right).$$  \hspace{1cm} (2.4)$$

For engines with EGR or lean burning engines, the oxygen concentration within the control volumes also varies dynamically. These variations affect not only the in-cylinder oxygen concentration but also the specific heat values. In addition to varying with temperature, the specific heat values of the gaseous mixtures within the air path depend on their compositions. The constant pressure and constant volume specific heat values can be predicted as a function of the oxygen concentration within the control volume using

$$c_p = \frac{[O_2]_{cv}}{[O_2]_{air}} c_{p_{air}} + \left( 1 - \frac{[O_2]_{cv}}{[O_2]_{air}} \right) c_{p_{comb}}$$  \hspace{1cm} (2.5)$$

and

$$c_v = \frac{[O_2]_{cv}}{[O_2]_{air}} c_{v_{air}} + \left( 1 - \frac{[O_2]_{cv}}{[O_2]_{air}} \right) c_{v_{comb}}.$$  \hspace{1cm} (2.6)$$

where $[O_2]_{cv}$ is the oxygen concentration within the control volume on a per mass basis, $[O_2]_{air}$ is the oxygen concentration of air on a per mass basis, $c_{p_{air}}$ is the constant pressure specific heat of air, $c_{v_{air}}$ is the constant volume specific heat of air,
$c_{p,\text{comb}}$ is the constant pressure specific heat of the products of complete combustion and $c_{v,\text{comb}}$ is the constant volume specific heat of the products of complete combustion. In this dissertation, all oxygen concentrations are defined on a per mass basis (i.e. the ratio of the mass of oxygen to the total mass).

To capture these effects, each control volume needs to consider one additional state, the oxygen concentration. The change in the oxygen concentration within a zero-dimensional control volume is governed by

$$\frac{d(m_{cv}[O_2]_{cv})}{dt} = \sum_{i=1}^{n_{in}} \dot{m}_{in,i}[O_2]_{in,i} - \sum_{j=1}^{n_{out}} \dot{m}_{out,j}[O_2]_{out,j} + \dot{R}_{O_2}$$

(2.7)

where $m_{cv}$ is the total mass within the control volume, $[O_2]_{in,i}$ is the oxygen concentration on a per mass basis of the $i^{th}$ stream entering the control volume, $[O_2]_{out,j}$ is the oxygen concentration on a per mass basis of the $j^{th}$ stream exiting the control volume and $\dot{R}_{O_2}$ is the rate at which oxygen is produced by chemical reactions.

Applying the chain rule,

$$m_{cv} \frac{d[O_2]_{cv}}{dt} + [O_2]_{cv} \frac{dm_{cv}}{dt} = \sum_{i=1}^{n_{in}} \dot{m}_{in,i}[O_2]_{in,i} - \sum_{j=1}^{n_{out}} \dot{m}_{out,j}[O_2]_{out,j} + \dot{R}_{O_2}. \quad (2.8)$$

From the conservation of mass law, the change control volume mass is known to be governed by

$$\frac{dm_{cv}}{dt} = \sum_{i=1}^{n_{in}} \dot{m}_{in,i} - \sum_{j=1}^{n_{out}} \dot{m}_{out,j}. \quad (2.9)$$

Combining these results produces

$$m_{cv} \frac{d[O_2]_{cv}}{dt} = \sum_{i=1}^{n_{in}} \dot{m}_{in,i}[O_2]_{in,i} - \sum_{j=1}^{n_{out}} \dot{m}_{out,j}[O_2]_{out,j}$$

$$- [O_2]_{cv} \left( \sum_{i=1}^{n_{in}} \dot{m}_{in,i} - \sum_{j=1}^{n_{out}} \dot{m}_{out,j} \right) + \dot{R}_{O_2}. \quad (2.10)$$

To simplify this equation, the rate of oxygen production can be ignored and the gases within the control volume can be assumed to mix perfectly thus allowing the
oxygen concentration of the gases exiting the control volume to be approximated with the mass averaged oxygen concentration of the control volume [11,61–63,65,69,72,73]. With this assumption, (2.10) reduces to

\[
\frac{d[O_2]_{cv}}{dt} = \frac{1}{m_{cv}} \left( \sum_{i=1}^{n_{in}} \dot{m}_{in,i}[O_2]_{in,i} - [O_2]_c (\sum_{i=1}^{n_{in}} \dot{m}_{in,i}) \right). \tag{2.11}
\]

or equivalently

\[
\frac{d[O_2]_{cv}}{dt} = \frac{RT_{cv}}{p_{cv} V_{cv}} \left( \sum_{i=1}^{n_{in}} \dot{m}_{in,i}[O_2]_{in,i} - [O_2]_c (\sum_{i=1}^{n_{in}} \dot{m}_{in,i}) \right). \tag{2.12}
\]

For an engine with EGR, the outputs of the lumped intake control volume are the mass flow rate, the temperature and the oxygen concentration of the air charge entering the engine. The inputs are the mass flow rate, temperature and oxygen concentrations of the fresh air flowing entering the manifold and the mass flow, the temperature and the oxygen concentration of the exhaust gases entering the manifold through the EGR system. Similarly, the outputs of the lumped EGR system are the mass flow rate, the temperature and the oxygen concentration of the exhaust gases entering the lumped intake system. The inputs are the mass flow rate, the temperature and the oxygen concentration of the exhaust gases entering the EGR system. For the lumped exhaust manifold system, the outputs are the mass flow rate, the temperature and the oxygen concentration of the exhaust gases entering the EGR system and the mass flow rate, the temperature and the oxygen concentration of the exhaust gases continuing downstream toward the tailpipe. The inputs are the mass flow rate and temperature of the combustion gases exiting the engine.

In a zero-dimensional model, the mass flow rates of the gases entering and exiting each control volume are typically calculated using algebraic equations. For example,
the flow through an orifice such as an EGR valve or an air throttle is normally predicted with the orifice flow equation based the assumption of an isentropic expansion, written as

$$\dot{m}_r = \frac{C_d A p_u}{\sqrt{RT_r}} \sqrt{\gamma hr} \ f \left( \frac{p_d}{p_u} \right)$$

(2.13)

where the correction factor $f()$ is governed by

$$f \left( \frac{p_d}{p_u} \right) = \begin{cases} \sqrt{\frac{2}{\gamma hr - 1}} \left( \frac{p_d}{p_u} \right)^{\frac{2}{\gamma hr}} - \left( \frac{p_d}{p_u} \right)^{\frac{\gamma hr + 1}{\gamma hr}} & \text{if } \frac{p_d}{p_u} \geq \left( \frac{2}{\gamma hr + 1} \right)^{\frac{\gamma hr}{\gamma hr - 1}} \\ \sqrt{\left( \frac{2}{\gamma hr - 1} \right)^{\gamma hr + 1/(\gamma hr - 1)}} & \text{otherwise} \end{cases}$$

(2.14)

and where $\dot{m}_r$ is the mass flow rate through the restriction, $C_d$ is the discharge coefficient of the restrictions, $A$ is the cross sectional area, $p_u$ is the upstream pressure, $p_d$ is the downstream pressure and $T_r$ is the gas temperature at the restriction. The flow into the engine through the intake valves is typically estimated using an empirically based concept known as volumetric efficiency. The volumetric efficiency of an engine is the ratio of the actual mass of air charge trapped per induction event to the maximum possible mass of air charge that could be inducted for the same intake manifold conditions (temperature and pressure). Using this definition of volumetric efficiency, the mass flow rate of air charge into the engine, $\dot{m}_{eng}$, can be predicted with

$$\dot{m}_{eng} = \frac{\eta_v p_{int,cv} NV_d}{RT_{int,cv} 120}$$

(2.15)

where $\eta_v$ is the volumetric efficiency, $p_{int,cv}$ is the pressure in the intake control volume, $N$ is the engine speed in revolutions per second, $V_d$ is the displacement of the engine and $T_{int,cv}$ is the temperature of the intake control volume.

A wide range of methods are used to model the turbocharger and compressor. The most common formulations either directly analyze the power transfer from the turbocharger to the compressor or they model the lumped rotational inertia of the
system and predict the response using Newton’s second law. When the power of
the compressor is chosen to be the dynamic turbocharger state as in [9, 11, 70], the
turbocharger is characterized by the following equation

\[ \frac{dP_c}{dt} = \frac{1}{\tau_{tc}} (P_t - P_c) \]  

(2.16)

where \( \tau_{tc} \) is the time constant of the turbocharger system, \( P_t \) is the power of the
turbocharger and \( P_c \) is the power of the compressor. The power of the turbocharger
can be approximated with

\[ P_t = \dot{m}_t \eta_t c_p T_{exh,cv} \left( 1 - \left( \frac{p_a}{p_{exh,cv}} \right)^{\gamma_{hr} - 1} \right) \]  

(2.17)

where \( \dot{m}_t \) is the mass flow rate through the turbocharger, \( \eta_t \) is the isentropic efficiency
of the turbocharger, \( T_{exh,cv} \) is the temperature of the exhaust control volume, \( p_{exh,cv} \)
is the pressure of the exhaust control volume and \( p_a \) is the ambient pressure. The
mass flow rate through the turbocharger can be predicted using

\[ \dot{m}_t = \frac{C_t A_t p_{exh,cv}}{\sqrt{RT_{exh,cv}}} \sqrt{\gamma_{hr}} f \left( \frac{p_a}{p_{exh,cv}} \right) \]  

(2.18)

where \( C_t \) is the discharge coefficient of the turbocharger and \( A_t \) is the area of the
turbocharger. To improve the model fit, the pressure ratio correction factor described
in (2.14) is sometimes modified slightly. Using the compressor power state, the mass
flow rate through the compressor \( \dot{m}_c \) can be calculated according to

\[ \dot{m}_c = \frac{\eta_c P_c}{c_p T_a \left( \frac{p_{int,cv}}{p_a} \right)^{\gamma_{hr} - 1} - 1} \]  

(2.19)

where \( \eta_c \) is the isentropic efficiency of the compressor and \( T_a \) is the ambient tem-
perature. A diagram of a complete zero-dimensional air path model for a typical
turbocharged engine with EGR is provided in Figure 2.3.
A zero-dimensional approach to modeling the air path system is reasonably well suited for predicting the mass averaged properties of each lumped control volume. Because the mass flow rate of air charge into the engine is predicted using a volumetric efficiency concept defined with respect to the mass averaged conditions within the intake manifold control volume, the inducted air mass can also be well predicted by these models. However, this modeling approach is not well suited for predicting the instantaneous fluid properties at specific locations within the fluid system, because the physical system never achieves perfect mixing. When the pipes which comprise
a control volume have a nontrivial characteristic length to volume ratio (which is most often the case), the gases exiting the control volume will not have the same properties as the mass averaged control volume properties during a transient. In particular, a zero-dimensional filling and emptying air path model predicts the oxygen concentration of the air charge inducted into the engine rather poorly, because the oxygen concentration of the inducted air charge is incorrectly assumed to be equal to the mass averaged oxygen concentration of the intake control volume. Even though this method of predicting the in-cylinder oxygen concentration is inherently flawed, it is the prevailing approach used in the open literature.

**One-Dimensional Air Path Models**

As with a zero-dimensional model, a one-dimensional air path model partitions the fluid paths into different control volumes (e.g. intake, EGR and exhaust). The properties within each of these control volumes, however, are no longer assumed to be uniform; variations along a characteristic length are permitted. Accounting for these spacial variations enables a one-dimensional air path model to capture the pressure wave dynamics, the oxygen transport dynamics and the oxygen mixing dynamics. As shown in [74], the governing one-dimensional equations can be derived from the fundamental conservation laws (i.e. mass, energy, momentum and species). The resulting partial differential equations are the continuity equation,

\[
\frac{\partial \rho}{\partial t} + \frac{\partial (\rho v)}{\partial z} + \frac{\rho v dA}{A \, dz} = 0, \quad (2.20)
\]

the energy equation,

\[
\frac{\partial (\rho Au)}{\partial t} + \frac{\partial (\rho Avh)}{\partial z} - (\dot{q}_{ht} + \dot{q}_{gen})\rho A = 0, \quad (2.21)
\]
the momentum equation,

$$\frac{\partial (\rho A v)}{\partial t} + \frac{\partial (\rho A v^2)}{\partial z} + A \frac{\partial \rho}{\partial t} = 0,$$  \hspace{1cm} (2.22)

and the convection-diffusion equation,

$$\frac{\partial [O_2]}{\partial t} - D \frac{\partial^2 [O_2]}{\partial z^2} + v \frac{\partial [O_2]}{\partial z} + \dot{r}_{O_2} = 0,$$  \hspace{1cm} (2.23)

where \( z \) is the characteristic length variable, \( \rho \) is the fluid density, \( v \) is the fluid velocity, \( A \) is the cross sectional area of the control volume, \( u \) is the specific internal energy, \( h \) is the specific enthalpy, \( \dot{q}_{ht} \) is the heat transfer rate, \( \dot{q}_{gen} \) is the heat generation rate, \([O_2]\) is the oxygen concentration, \( D \) is the diffusivity of oxygen in the gaseous mixture and \( \dot{r}_{O_2} \) is the rate of oxygen is produced by chemical reactions. A diagram of a complete one-dimensional air path model for a typical turbocharged engine with EGR is provided in Figure 2.3. The defining characteristic of a one-dimensional model is that the properties within each subsystem (i.e. density, internal energy, velocity and oxygen concentration) vary as function of a characteristic length variable.

To solve these partial differential equations, they are normally converted into ordinary differential equations by spatially discretizing them. This practice is akin to sub-partitioning the intake, EGR and exhaust systems into even smaller control volumes. The properties within these control volumes (i.e. density, internal energy, velocity and oxygen concentration) are assumed to be uniform. A typical approach might partition each subsystems into anywhere from five to one hundred control volumes. Since each one of these control volumes is defined by four states, the overall model can contain hundreds, thousands or even tens of thousands of states. Although a one-dimensional air path model provides a higher fidelity system description, it is too complex to be directly used in control applications.
2.4.2 Oxygen Storage Based TWC Model

Because of the complex concentration dependent and temperature dependent chemical kinetics involved in the catalytic reactions, modeling a TWC is a very difficult undertaking. One of the most common approaches is to model the catalyst as an oxygen integrator with saturation bounds \([15,33,34]\). The lower saturation bound represents the absence of stored oxygen, and the upper saturation bound represents the maximum oxygen storage capacity. In this method, the stored oxygen fraction within the TWC, \(\Theta_s\), is chosen to be the state variable and its dynamics are modeled...
by
\[
\dot{\Theta}_s = \begin{cases} 
\frac{0.23}{C_{TWC}} k_{rate} \dot{m}_{cat} (1 - \phi_{cat}) & 0 \leq \Theta_s \leq 1 \\
0 & \text{otherwise}
\end{cases}
\] (2.24)

where $C_{TWC}$ is the storage capacity of the TWC, $\phi_{cat}$ is the fuel/air equivalence ratio (EQR) of the mixture at the inlet of the catalyst, $k_{rate}$ is the oxygen storage/release rate and $\dot{m}_{cat}$ is the mass flow rate entering the catalyst. The EQR of a combustion mixture, $\phi$, can be calculated using
\[
\phi = \left( \frac{m_{fuel,\text{trap}}}{m_{air,\text{trapped}}} \right) \text{AFR}_s = \left( \frac{m_{fuel,\text{trap}}}{[O_2]_{\text{cyl}} [O_2]_{\text{air}} m_{\text{trapped}}} \right) \text{AFR}_s
\] (2.25)

where $m_{fuel,\text{trap}}$ is the trapped fuel mass, $m_{air,\text{trapped}}$ is the trapped unburned air mass, $m_{\text{trapped}}$ is the total trapped air mass, $[O_2]_{\text{cyl}}$ is the in-cylinder oxygen concentration, $[O_2]_{\text{air}}$ is the oxygen concentration of ideal air and AFR$_s$ is the stoichiometric air-to-fuel ratio. Although this type of model does not predict the behavior of a catalyst all that well, its performance is comparable to other more complex models such as [16].

### 2.4.3 Fuel Film Evaporation Dynamics Model

For a port fuel injected engine, the mass of fuel delivered by a fuel injector on a given cycle is not necessarily the mass of fuel that enters the engine on that cycle. A film of liquid fuel forms on the walls of the intake ports. This film is continuously evaporating, but periodically fuel is added during discrete injection events. In steady-state the cycle averaged mass of the fuel puddle stays constant, and the mass of fuel injected equals the mass of fuel entering the engine. During transients, however, the fuel film can remove or supply fuel to the nominal trapped fuel mass. Several fuel dynamics models have been developed to approximate this phenomenon including [75], [76] and [77]. A certain fraction of the injected fuel $\chi_{fuel}$ is assumed to add to
the fuel film while the remaining fraction directly follows the charge flowing into the engine. By using a first order response characterized by a time constant $\tau_{\text{evap}}$, the fuel film mass $m_{\text{film}}$ can be modeled as

$$\dot{m}_{\text{film}} = \dot{m}_{\text{fuel,in}} - \dot{m}_{\text{fuel,out}} = \chi_{\text{fuel}} \dot{m}_{\text{fuel,inj}} - \frac{m_{\text{film}}}{\tau_{\text{evap}}} \tag{2.26}$$

where $\dot{m}_{\text{fuel,inj}}$ is the fuel injection rate. The resulting flow of fuel into the engine $\dot{m}_{\text{fuel,eng}}$ is

$$\dot{m}_{\text{fuel,eng}} = (1 - \chi_{\text{fuel}}) \dot{m}_{\text{fuel,inj}} + \frac{m_{\text{film}}}{\tau_{\text{evap}}} \tag{2.27}$$

Because the evaporation rate depends on the factors such as the intake port temperature and the charge flow rate, this model is often implemented using gain scheduling in practice.

### 2.5 Summary

Because production fueling controllers undergo extensive calibration, the potential for a model based fueling control approach to outperform the current experimentally based approaches in terms of closed loop performance is minimal. When considering the calibration requirements, robustness, closed loop stability and engine platform to engine platform performance uniformity, a model based approach could still potentially be a significant improvement over the current production approaches as long as an equivalent level of closed loop performance is achieved. Current model based techniques have yet to satisfy these requirements and therefore reach the true potential of a model based approach largely because the oxygen mixing and transport dynamics are not properly captured. Existing air path truth models inaccurately assume perfect mixing within the intake, EGR and exhaust systems. This assumption is also carried down to the control-oriented models based on these truth models.
This dissertation improves upon current control-oriented engine modeling techniques by explicitly capturing the oxygen transport and mixing dynamics. This oxygen dynamics model can be integrated into almost any engine modeling structure without having to redesign the TWC model, the fuel dynamics model or the filling and emptying dynamics model because they do not depend on the oxygen concentration dynamics. The improvement to the overall plant model accuracy also facilitates the design of advanced model based controllers and estimators which are able to overcome many of the issues which plague current approaches.

The biggest problem with conventional fueling controllers in diesel engines is that they do not effectively account for dynamic behavior within the air path system. Depending on the system operating mode, different air path setpoints are needed to best satisfy the changing trade-offs between torque response, emissions and fuel economy. To account for these air path setpoint changes, a separate set of fueling parameter tables is used for each system operating mode. By scheduling the fueling parameters on the variables which directly influence the cylinder contents and/or the combustion process such as the in-cylinder oxygen concentration, the optimal fueling parameters can be selected regardless of the air path dynamics, providing improved robustness. Moreover, a single set of fueling parameter tables can be used for all of the system operating modes thereby significantly reducing the calibration effort.

For such a fueling controller to be effective, all of the scheduling variables including the in-cylinder oxygen concentration must be dynamically estimated. From an air path truth model which captures the transport and mixing dynamics, a real-time estimator can be designed using model reduction techniques and modern control theory. In this dissertation several different estimators including a linear parameter
varying (LPV) state observer and an extended Kalman filter are designed and compared to determine the most suitable approach. These results demonstrate that the in-cylinder oxygen concentration can be estimated effectively even with production grade sensors as long as the oxygen dynamics are properly modeled.

In the research field of AFR control for gasoline engines, one of the major weaknesses of current controllers is that the feedback control objective is not directly tied to the overall objective of emission reduction. Normally the feedback objective is to regulate the measured oxygen sensor voltage to a slowly varying setpoint value which nominally corresponds to stoichiometry. The best catalyst conversion efficiencies are achieved when the AFR is varied periodically in a precisely controlled manner. As a direct result, an AFR controller which strives to track the AFR signal and promotes the highest catalyst conversion efficiency is more capable of reducing the tailpipe emission. Tracking based AFR controllers are uncommon because the delay characterized relationship between delivering fueling and measuring the AFR of the exhaust stream is not well predicted with current models. This input/output relationship can be more effectively predicted by directly modeling the transport and mixing dynamics within the exhaust system.

The benefits of accounting for the transport and mixing delays within the exhaust system are also demonstrated through a model based cylinder imbalance estimator. This technique is a very effective means of identifying air and/or fuel imbalances which result in cylinder-to-cylinder AFR variations. With this model based approach, not only is the calibration effort significantly smaller than similar techniques but also the cylinder imbalance is predicted smoothly across the complete operating space. For a gasoline engine, this information could be used to adjust the mass of fuel delivered by
each fuel injector. Conversely for a diesel engine, this information could be used to adjust a cylinder averaged in-cylinder oxygen concentration estimate. The resulting cylinder specific in-cylinder oxygen concentration estimate could be used instead of the cylinder averaged value when the fueling parameters are being looked up. In this formulation the total fuel mass would not be changed; only the way in which fuel is delivered is changed.
CHAPTER 3

Plug Flow Based Oxygen Dynamics Modeling

The truth models derived in this chapter provide the basis for the rest of this dissertation. The lumped oxygen dynamics model presented in Section 3.2 is later simplified in Section 4.1 to facilitate real time in-cylinder oxygen concentration estimation. This real time estimator in turn enables the in-cylinder oxygen concentration based diesel fueling controller that is detailed in Chapter 5. Unlike the lumped oxygen dynamics model, the high fidelity exhaust oxygen dynamics model described in Section 3.3 models each one of the exhaust runners separately and accounts for the cylinder firing order. These additional considerations produce a more accurate prediction of the oxygen concentration at the pre-catalyst oxygen sensor thereby allowing the AFR switching period controller that is specified in Chapter 6 to be calibrated almost entirely in simulation. The final model described in Section 3.4 represents the input/output relationship between injecting fueling and measuring the exhaust oxygen concentration as a finite impulse response. Characterizing the system in this manner allows for cylinder imbalances to be estimated with the techniques outlined in Section 4.2.
3.1 Overview

The dynamics of a nonreactive gaseous species flowing in a gaseous medium are dictated by two factors, the convection rate and the diffusion rate. Since the convection rate depends on the linear velocity of the gaseous species and the diffusion rate depends on the diffusivity of the gaseous species, the relative importance of the gas transport dynamics can be determined by analyzing the ratio of the gas velocity to the diffusivity coefficient. In other words, the manner in which a change in the concentration at the inlet of a one-dimensional control volume influences the concentration at the outlet of that control volume strongly depends on this ratio.

Figure 3.1 demonstrates the response of the gas concentration at the outlet of a pipe to a step change in the inlet concentration as a function of the ratio of the gas velocity to the diffusivity. To non-dimensionalize the time axis, the time variable is scaled by the ratio of the volumetric flow rate divided by the volume of the control volume. For large values of this ratio, the system closely resembles a delayed step response. This delay is the transport delay. As the ratio is decreased, the response resembles that of a filter delayed in time. When the ratio approaches zero, the transport delay is completely negligible and the response resembles that of a first order system. Because the diffusion effects dominate, the gases within the control volume are mixed well enough to produce a uniform mixture. The perfect mixing assumption used in zero-dimensional models assumes exactly this scenario.

In an internal combustion engine, the ratio of the bulk velocity to the diffusivity of oxygen in the air path gases is typically very large. Therefore, the transport dynamics are much more significant than the mixing dynamics. A traditional zero-dimensional filling and emptying approach to oxygen concentration modeling is inherently flawed.
Figure 3.1: Effect of diffusion rate on the step response of one-dimensional gaseous system

because it is based on a perfect mixing assumption which ignores the transport dynamics. To capture the transport dynamics, a one-dimensional (or higher) oxygen dynamics model must be used.

To implement a one-dimensional air path model, the governing partial differential equations must be spatially discretized into small control volumes. As long as the number of control volumes is sufficiently large, a spatially discretized one-dimensional air path model can still effectively capture the time varying transport dynamics. Because each control volume is characterized by a density, an internal energy, a velocity and an oxygen concentration, partitioning each subsystem into too many control volumes can make solving the resulting set of the coupled nonlinear differential equations exceedingly difficult.
For a control-oriented air path model, the spatial variation of the gas density, internal energy and velocity are not of particular interest, so they can be ignored. The spatial variation in the oxygen concentration, however, cannot. This dissertation develops a new air path modeling scheme that uses a zero-dimensional model to capture the temperature and pressure dynamics and a one-dimensional model to capture the oxygen dynamics. The focus of this modeling approach is the development of a hierarchy of oxygen dynamics models beginning from the one-dimensional convection-diffusion equation which accurately captures both the oxygen transport and mixing dynamics in a computationally efficient manner.

The oxygen concentration at any point in the intake, EGR or exhaust system can be predicted by tracking small “packets” of gas (burned gas and/or fresh air) as they move through these systems. For stoichiometric gasoline engines, the oxygen concentration of the gas in the exhaust and EGR systems is always close to zero, but in lean burning engines the exhaust gas may contain a significant concentration of oxygen. To track the packets as they move through the exhaust, EGR and intake systems, the three representative one-dimensional subsystems are individually discretized into many cells of constant volume. Each of these constant volume cells holds a packet of fresh/burned gas with a uniform oxygen concentration. Every time a cylinder exhausts its combustion products, all of the gas packets within the exhaust system are pushed further downstream. Depending on the EGR flow rate, some fraction of the gas packets that leave the exhaust manifold will enter the EGR system. The gas packets that enter the EGR system displace the same number of packets into the intake system. This displaced EGR gas combines with fresh air and is eventually inducted
into the engine. Modeling the system in this type of plug flow manner inherently captures the time-varying transport delay as well as the diffusion dynamics.

The governing plug flow equations can be directly derived from the one-dimensional oxygen diffusion-convection equation described in (2.23). If all of the chemical reactions are assumed to occur inside the combustion chamber, then the oxygen production term can be eliminated. Next, this equation is spatially discretized using a discretization length of $\Delta z$ to give rise to

\[
\frac{d[O_2]_i(t)}{dt} = D \frac{[O_2]_{i-1}(t) - 2[O_2]_i(t) + [O_2]_{i+1}(t)}{\Delta z^2} - \frac{v}{\Delta z} \left( [O_2]_i(t) - [O_2]_{i-1}(t) \right)
\]  

(3.1)

where $[O_2]_i$ is oxygen concentration of the $i^{th}$ control volume cell. Assuming constant cross sectional area, the oxygen dynamics equation can be rewritten in terms of the volumetric flow rate in the manner

\[
\frac{d[O_2]_i(t)}{dt} = \left( D_1 + \frac{q_{in}(t)}{V_{cv}} \right) [O_2]_{i-1}(t) + \left( -2D_1 - \frac{q_{in}(t)}{V_{cv}} \right) [O_2]_i(t) + D_1 [O_2]_{i+1}(t)
\]

(3.2)

where

\[
D_1 = \frac{D}{\Delta z^2},
\]

(3.3)

$q_{in}$ is the volumetric flow rate entering the control volume and $V_{cv}$ is the volume of the control volume.

Next, this time domain oxygen concentration model can be converted into the engine event domain in which a combustion occurrence is defined as an event. In the event domain, the number of events per engine cycle is equal to the number of cylinders. One can convert between the event domain defined by variable $e$ and the
time domain defined by variable $t$ using

$$
t = \frac{120}{n_{cyl}N} e. \quad (3.4)
$$

The derivative of this equation can be approximated as

$$
dt = \frac{120}{n_{cyl}N} de, \quad (3.5)
$$

because the rate of change of engine speed is insignificant in this setting. The volumetric flow rates can also be converted from the time domain to the event domain. For any volumetric flow rate variable $q$, define $q^*(e)$ as the volumetric flow rate per event which satisfies

$$
q^*(e) = \frac{120}{n_{cyl}N} q(t). \quad (3.6)
$$

With these relationships, the continuous time domain oxygen concentration model described by (3.2) can be converted to the continuous event domain model

$$
\frac{n_{cyl}N}{120} \frac{d[O_2]_i(e)}{de} = \left( D_1 + q_{in}^* \frac{n_{cyl}N}{120V_{cv}} \right)[O_2]_{i-1}(e) +
\left( -2D_1 - q_{in}^* \frac{n_{cyl}N}{120V_{cv}} \right)[O_2]_i(e) + D_1[O_2]_{i+1}(e). \quad (3.7)
$$

By defining $D_2$ as

$$
D_2 = \frac{120}{n_{cyl}N} D_1 \quad (3.8)
$$

the event domain equation reduces to

$$
\frac{d[O_2]_i(e)}{de} = \left( D_2 + \frac{q_{in}^*}{V_{cv}} \right)[O_2]_{i-1}(e) + \left( -2D_2 - \frac{q_{in}^*}{V_{cv}} \right)[O_2]_i(e) + D_2[O_2]_{i+1}(e). \quad (3.9)
$$

Since each unit control volume has the same volume $V_{cv}$, the volumetric flow rate can be normalized according to

$$
K^*(e) = \frac{q_{in}^*(e)}{V_{cv}} \quad (3.10)
$$
where the units of $K^*(e)$ are control volumes per event. The oxygen concentration
dynamics can now be expressed as

$$
\frac{d[O_2]_i(e)}{de} = (D_2 + K^*)[O_2]_{i-1}(e) + (-2D_2 - K^*)[O_2]_i(e) + D_2[O_2]_{i+1}(e).
$$

(3.11)

Applying the discrete Euler approximation with an integration step (in event
space) of $\Delta e$, the changing rate of the oxygen concentration within a unit control
volume can be approximated as

$$
\frac{d[O_2]_{cv}(e)}{de} \approx \frac{[O_2]_{cv}(e + \Delta e) - [O_2]_{cv}(e)}{\Delta e}.
$$

(3.12)

With this approximation the oxygen dynamics equation becomes

$$
[O_2]_i(e + \Delta e) = \Delta e(D_2 + K^*)[O_2]_{i-1}(e) + (1 - 2\Delta e D_2 - \Delta e K^*)[O_2]_i(e) + \Delta e D_2[O_2]_{i+1}(e).
$$

(3.13)

If the diffusion effects were negligible, then this equation would reduce to

$$
[O_2]_i(e + \Delta e) = \Delta e K^*[O_2]_{i-1}(e) + (1 - \Delta e K^*)[O_2]_i(e).
$$

(3.14)

On the other hand, if the volumetric flow rate was zero, then (3.13) would simplify
to

$$
[O_2]_i(e + \Delta e) = \Delta e D_2[O_2]_{i-1}(e) + (1 - 2\Delta e D_2)[O_2]_i(e) + \Delta e D_2[O_2]_{i+1}(e).
$$

(3.15)

Since the time scales of the convention and diffusion processes are radically different,
they are modeled separately using different integration step sizes. In a plug flow oxy-
gen dynamics model, the transport dynamics due to convection are solved assuming
no diffusion and then the diffusion effects are solved assuming zero net flow.
Because the transport dynamics depend on the volumetric flow rate, the integration step size for the transport dynamics equation is chosen to be proportional to the volumetric flow rate according to

$$\Delta e = \frac{1}{\lceil K^*(e) \rceil}$$

where \( \lceil \cdot \rceil : \mathbb{R} \mapsto \mathbb{Z} \) is the standard ceiling function defined by

$$\lfloor y \rfloor = \min\{m \in \mathbb{Z} \mid m \geq y\}.$$  \hspace{1cm} (3.17)

With this choice of step size, the oxygen transport dynamics equation can be simplified substantially. Defining \( \beta \) as

$$\beta(e) = \frac{K^*(e)}{\lceil K^*(e) \rceil}$$

the oxygen transport dynamics model can be represented as

$$[O_2]_i(e + \Delta e) = \beta(e)[O_2]_{i-1}(e) + \left(1 - \beta(e)\right)[O_2]_i(e).$$  \hspace{1cm} (3.19)

When \( \beta \) is exactly unity, this model formulation perfectly captures the transport delays of the system. It is important to note that if a subsystem is discretized into a sufficiently large number of control volumes, then \( \beta \) will be approximately unity regardless of the volumetric flow rate. Also, by construction the step size is a unit fraction. Therefore, all of the oxygen concentration calculations are guaranteed to be updated every whole event number.

Correspondingly, the oxygen diffusion model is updated once per event by choosing an integration step size of one event. With this choice, the diffusion dynamics equation described in (3.15) reduces to

$$[O_2]_i(e + 1) = D_2[O_2]_{i-1}(e) + \left(1 - 2D_2\right)[O_2]_i(e) + D_2[O_2]_{i+1}(e).$$  \hspace{1cm} (3.20)
Depending on the volumetric flow rate and the number of control volumes, the diffusion integration step size could be as many as a hundred times larger than the integration step size used for the transport dynamics. Separating the diffusion and transport dynamics in this manner improves the computational efficiency of the model while preserving its accuracy. As seen by this derivation, a plug flow approach to oxygen dynamics modeling is simply a discrete approximation of the one-dimensional oxygen convection-diffusion equation which separately solves for the two effects.

The next section describes how a fundamental plug flow oxygen dynamics model of a single subsystem can be applied to the intake, exhaust and EGR subsystems to generate a complete oxygen dynamics model of an engine air path system. As will be shown, the inputs to a plug flow oxygen dynamics air path model are the volumetric flow rates within each subsystem, the mass flow rate into the engine, the trapped residual fraction and the injected fuel mass. The injected fuel mass is a known control input and the trapped residual fraction can be empirically modeled. Using the ideal gas law, the volumetric flow rates can be predicted with a zero-dimensional filling and emptying air path model. This zero-dimensional model need only model the temperature and pressure dynamics, because the variation in the specific heat values caused by the change in gas composition can be calculated based on the spatial average of the oxygen concentrations predicted by the plug flow oxygen dynamics model according to

\[ c_p = \frac{[O_2]_{cv}}{[O_2]_{air}} c_{p_{air}} + \left(1 - \frac{[O_2]_{cv}}{[O_2]_{air}}\right) c_{p_{comb}} \]  
(3.21)

and

\[ c_v = \frac{[O_2]_{cv}}{[O_2]_{air}} c_{v_{air}} + \left(1 - \frac{[O_2]_{cv}}{[O_2]_{air}}\right) c_{v_{comb}} \]  
(3.22)
Figure 3.2: Diagram of an air path model which models the filling and emptying dynamics with a zero-dimensional model and models the oxygen transport and mixing dynamics with a lumped plug flow oxygen dynamics model.
where

$$[O_2]_{cv} = \sum_{i=1}^{n_{cv}} [O_2]_i$$

(3.23)

and $n_{cv}$ is the total number of control volumes. Figure 3.2 illustrates how the complete set of dynamic air path states can be split between the two models. All of the mass flow rates, temperatures and pressures are predicted with the zero-dimensional model and all of the oxygen concentrations are predicted with plug flow model.

It is important to note that when coupled with a zero dimensional filling and emptying model, a one-dimensional plug flow oxygen dynamics model does not account for wave dynamics. Moreover, all spatial variation in temperature, pressure and gas velocity are ignored. Instead, a plug flow oxygen dynamics model assumes that the volumetric flow rate entering a given subsystem is equal to the volumetric flow rate leaving that subsystem. With this assumption, the volumetric flow rate variables used in the plug flow model can be calculated using either the mass flow rate entering the corresponding zero-dimensional control volume, the mass flow rate exiting the corresponding zero-dimensional control volume or a weighted average of the two mass flow rates. In this dissertation, the volumetric flow rate within the intake system is calculated based on the sum of the mass flow rate of fresh air entering the intake manifold (i.e. the mass flow rate through the compressor) and the mass flow rate through the EGR valve. The volumetric flow rate within the EGR system is calculated based on the mass flow rate through the EGR valve. Finally, the characteristic volumetric flow rate within the exhaust system is calculated based on the sum of the mass flow rate entering the engine and the mass flow rate of fuel.
3.2 Lumped Oxygen Dynamics Model

With a lumped plug flow based oxygen dynamics model, each of the three subsystems is approximated by a straight linear pipe which has the same total volume as the physical system. At each event, the volume that the newly exhausted gas mass will occupy is calculated using the ideal gas law. The ceiling function is used to convert this volume into a discrete number of unit control volumes (cells) according to

\[
\delta_{exh}(k) = \left\lceil \frac{\dot{m}_{exh}(k) T_{exh,cv}(k) R}{p_{exh,cv}(k)} \times \frac{n_{exh}}{V_{exh,cv}(k)} \right\rceil = \lceil \bar{\delta}_{exh}(k) \rceil
\]

where \(\delta_{exh}\) is the total number of cells that the new exhaust gas will occupy, \(\dot{m}_{exh}\) is the exhaust gas mass flow rate on a per event basis, \(n_{exh}\) is the total number of cells used to model the exhaust manifold and \(k\) is the discrete time index in engine events (i.e. \(\Delta e = 1\)). Any mass flow rate \(\dot{m}(t)\) can be converted from the time domain to the discrete event domain using

\[
\dot{m}^*(k) = \frac{120}{n_{cyl} N} \dot{m}(t).
\]

The newly exhausted cells enter the model at the cell corresponding to the exhaust ports of the engine, \(EXH_1\), one cell at a time according to

\[
EXH_1^{q_{exh}}(k) = [O_2]_{comb}(k - d_{fuel,exh})
\]

where \([O_2]_{comb}\) is the oxygen concentration on a per mass basis of the gases within the cylinders after combustion occurs, \(q_{exh}\) is the flow index of the exhaust system and \(d_{fuel,exh}\) is the fixed delay between the fueling and exhausting events. The general notation \(EXH_i^{q_{exh}}\) represents the oxygen concentration of the \(i^{th}\) cell of the exhaust system after the \(q_{exh}\)th cell has been exhausted.
To make room for this new cell, each cell in the fluid path must move downstream one cell. The propagation of the cells within the exhaust manifold is modeled as

$$EXH_{i}^{q_{exh}}(k) = (1 - \beta_{exh}(k))EXH_{i}^{\delta_{exh}^{-1}}(k) + \beta_{exh}(k)EXH_{i-1}^{\delta_{exh}^{-1}}(k)$$  \hspace{1cm} (3.27)$$

$$\forall \ i \in \{2, 3, ..., n_{exh}\}$$ where

$$\beta_{exh}(k) = \frac{\delta_{exh}(k)}{\delta_{exh}(k)}.$$  \hspace{1cm} (3.28)$$

This is equivalent to (3.19), derived in the previous section. To model the effect of each new cell entering the exhaust system, the calculations in (3.26) and (3.27) must be repeated for all integer values $q_{exh}$ from 1 to $\delta_{exh}(k)$.

Due to the pulsating nature of both the breathing and exhausting processes, the mass flow rates within each of the three subsystems are not constant even for a window of time as small as one event. To account for the slower diffusion effects, the following set of diffusion calculations are made once per engine cycle according to:

$$EXH_{i}^{\delta_{exh}+1}(k) = (1 - \mathbb{D})EXH_{i}^{\delta_{exh}}(k) + \mathbb{D}EXH_{i}^{\delta_{exh}+1}(k),$$  \hspace{1cm} (3.29)$$

$$EXH_{i}^{\delta_{exh}+1}(k) = \mathbb{D}EXH_{i-1}^{\delta_{exh}}(k) + (1 - 2\mathbb{D})EXH_{i}^{\delta_{exh}}(k) + \mathbb{D}EXH_{i+1}^{\delta_{exh}}(k)$$  \hspace{1cm} (3.30)$$

$$\forall \ i \in \{2, 3, ..., n_{exh} - 1\}$$ and

$$EXH_{n_{exh}}^{\delta_{exh}+1}(k) = \mathbb{D}EXH_{n_{exh}-1}^{\delta_{exh}}(k) + (1 - \mathbb{D})EXH_{n_{exh}}^{\delta_{exh}}(k)$$  \hspace{1cm} (3.31)$$

where $\mathbb{D}$ is the effective diffusion coefficient formerly represented by $D_2$. Because the complexities of fluid mixing within the physical three-dimensional system are not perfectly capture with a one-dimensional approximation, the best model performance is achieved with this coefficient is treated as calibration variable and identified experimentally.
Part of the exhaust gases that leave the exhaust manifold enter the EGR loop while most of the exhaust gases continue down the exhaust system through the turbocharger. The number of cells in the EGR system that will be occupied by these gases, $\delta_{egr}$, is governed by

$$\delta_{egr}(k) = \left\lceil \frac{\dot{m}^*_egr(k) T_{egr,cv}(k) R}{p_{egr,cv}(k)} \times \frac{n_{egr}}{V_{egr,cv}} \right\rceil = \left\lceil \bar{\delta}_{egr}(k) \right\rceil \quad (3.32)$$

where $\dot{m}^*_egr$ is the mass flow rate of EGR on a per event basis and $n_{egr}$ is the total number of cells used to model the EGR system. Define $r_{egr}$ as the ratio of the number of occupied exhaust cells to the number of occupied EGR cells described by

$$r_{egr}(k) = \frac{\delta_{exh}(k)}{\delta_{egr}(k)}. \quad (3.33)$$

To determine the oxygen concentration of the cells entering the EGR system, the incoming $\delta_{exh}$ cells are split into $\delta_{egr}$ groups of $r_{egr}$ exhaust cells. If $r_{egr}(k) > 1$, then each group will contain two fractional cells and may contain multiple whole cells. If instead $r_{egr}(k) \leq 1$, then each group will contain either two fractional cells or a single factional cell. The oxygen concentration of each new EGR cell entering the EGR system is the weighted average of the oxygen concentrations of the exhaust cells in that grouping.

For the case where $r_{egr}(k) > 1$, each grouping of exhaust cells can be split into the following three partitions: a first partition which contains a fractional cell, a middle partition which contains multiple whole cells and a last partition which contains a fractional cell. The contribution from the first partition of the $j^{th}$ group of exhaust cells, $F^j_{egr}$, is calculated according to

$$F^j_{egr} = EXH^{[r_{egr} \times (j-1)]\ast} \left( [r_{egr} \times (j-1)]\ast - r_{egr} \times (j-1) \right) \quad (3.34)$$

57
∀ \ j \in \{1, 2, \ldots, \delta_{egr}\} \text{ where } \lceil \ \rceil^* : \mathbb{R} \mapsto \mathbb{Z} \text{ represents the strictly ceiling function defined as }

\lceil y \rceil^* = \min\{m \in \mathbb{Z} \mid m > y\}. \quad (3.35)

Similarly, the contribution from the last partition of the \(j^{th}\) group of exhaust cells, \(L_{egr}^j\), is calculated with

\[ L_{egr}^j = EXH_{n_{exh}}^{\lceil r_{egr} \times j \rceil} \left( r_{egr} \times j - \lfloor r_{egr} \times j \rfloor^* \right) \] (3.36)

∀ \ j \in \{1, 2, \ldots, \delta_{egr}\} \text{ where } \lfloor \ \rfloor^* : \mathbb{R} \mapsto \mathbb{Z} \text{ represents the strictly floor function defined as }

\lfloor y \rfloor^* = \max\{m \in \mathbb{Z} \mid m < y\}. \quad (3.37)

Lastly, the contribution from the middle partition of the \(j^{th}\) group of exhaust cells, \(M_{egr}^j\), is defined as

\[ M_{egr}^j = \sum_{c = \lceil (j-1) \times r_{egr}(k) \rceil^*}^{\lceil j \times r_{egr}(k) \rceil^* - 1} EXH_{n_{exh}}^{c+1} (k) \] (3.38)

∀ \ j \in \{1, 2, \ldots, \delta_{egr}\}. The same definitions for \(F_{egr}^j\), \(M_{egr}^j\) and \(L_{egr}^j\) can be used for the case when \(r_{egr}(k) < 1\). In this situation, \(M_{egr}^j\) will always be identically zero. The diagram presented in Figure 3.3 demonstrates how cells are divided for the case when seven exhaust cells are converted into three EGR cells.

Once each of the three contributions has been calculated, the oxygen concentration of the cells entering the EGR system can be calculated. Define \(q_{egr}\) as the flow index of the EGR system in the same way that \(q_{exh}\) represents the flow index of the exhaust system. Let the general notation \(EGR_{i}^{q_{egr}}\) represent the oxygen concentration of the \(i^{th}\) cell of the EGR system after the \(q_{egr}\)th cell has entered the EGR system. If
Cells leaving the exhaust manifold

Cells entering the EGR system

Figure 3.3: Cell splitting diagram

\[
\left( (q_{egr} - 1) \times r_{egr}(k) \right)^* \leq \left[ q_{egr} \times r_{egr}(k) \right]^*, \text{ then}
\]

\[
EGR_{q_{egr}}^1(k) = \frac{1}{r_{egr}(k)} \left( F_{egr}^q + L_{egr}^q + M_{egr}^q \right),
\]

otherwise

\[
EGR_{q_{egr}}^1(k) = EXH_{n_{exh}}^{(q_{egr}-1) \times r_{egr}(k)]^*}(k).
\]

The second equation corresponds to the case when \( r_{egr}(k) < 1 \) and only a fraction of a single cell is contained within the \( q_{egr} \)th grouping of exhaust cells. The propagation of the downstream cells within the EGR loop is exactly the same as the exhaust manifold. This process is modeled as

\[
EGR_{q_{egr}}^i(k) = (1 - \beta_{egr}(k))EGR_{q_{egr}}^{i-1}(k) + \beta_{egr}(k)EGR_{q_{egr}}^{i-1}(k)
\]

\( \forall i \in \{2, 3, ..., n_{egr}\} \) and \( \forall q_{egr} \in \{1, 2, ..., \delta_{egr}\} \) where

\[
\beta_{egr}(k) = \frac{\delta_{egr}(k)}{\delta_{egr}(k)}.
\]
The diffusion process is also modeled in the same manner as the exhaust system as seen by the following equations:

\[
EGR_{i}^{\delta_{egr}+1}(k) = (1 - \mathbb{D})EGR_{i}^{\delta_{egr}}(k) + \mathbb{D}EGR_{i+1}^{\delta_{egr}}(k), \quad (3.43)
\]

\[
EGR_{i}^{\delta_{egr}+1}(k) = \mathbb{D}EGR_{i}^{\delta_{egr}}(k) + (1 - 2\mathbb{D})EGR_{i}^{\delta_{egr}}(k) + \mathbb{D}EGR_{i+1}^{\delta_{egr}}(k) \quad (3.44)
\]

\[
\forall \ i \in \{2, 3, ..., n_{egr} - 1\} \text{ and }
\]

\[
EGR_{n_{egr}}^{\delta_{egr}+1}(k) = \mathbb{D}EGR_{n_{egr} - 1}^{\delta_{egr}}(k) + (1 - \mathbb{D})EGR_{n_{egr}}^{\delta_{egr}}(k). \quad (3.45)
\]

Each cell that exits the EGR system mixes with fresh air and enters the intake system. The number of cells occupied this mixture of EGR and fresh air will, \(\delta_{int}\), is governed by

\[
\delta_{int}(k) = \left\lceil \frac{\dot{m}_{charge}(k)T_{int,cv}(k)R}{p_{int,cv}(k)} \times \frac{n_{int}}{V_{int,cv}} \right\rceil = \left\lceil \dot{\delta}_{int}(k) \right\rceil \quad (3.46)
\]

where \(\dot{m}_{charge}\) is the mass flow rate or air charge on a per event basis and \(n_{int}\) is the total number of cells used to model the intake system. Define \(r_{int}\) as the ratio of the number of occupied intake cells to the number of occupied intake cells described by

\[
r_{int}(k) = \frac{\delta_{egr}(k)}{\delta_{int}(k)}. \quad (3.47)
\]

As with the EGR system, this ratio will be used to separate the \(\delta_{egr}\) EGR cells into \(\delta_{int}\) groups of \(r_{int}\) intake cells. The contribution from the first partition of the \(j^{th}\) group of EGR cells, \(F_{int}^{j}\), is calculated according to

\[
F_{int}^{j} = EGR_{n_{int}}^{\lceil r_{int}(k) \times (j-1) \rceil} \left( [r_{int} \times (j-1)] - r_{int} \times (j-1) \right) \quad (3.48)
\]

\(\forall \ j \in \{1, 2, ..., \delta_{int}\}. \) Similarly, the contribution from the last partition of the \(j^{th}\) group of EGR cells, \(L_{int}^{j}\), can be calculated with

\[
L_{int}^{j} = EGR_{n_{int}}^{\lceil r_{int} \times j \rceil} \left( r_{int} \times j + \lceil r_{int} \times j \rceil \right) \quad (3.49)
\]
∀ \ j \in \{1, 2, \ldots, \ delta_{\text{int}}\}. Lastly, the contribution from the middle partition of the \( j \)th group of EGR cells, \( M_{\text{int}}^j \), is defined as

\[
M_{\text{int}}^j = \left\lfloor \frac{j \times r_{\text{int}}(k)}{c} \right\rfloor - 1 \sum_{e = \left\lfloor (j-1) \times r_{\text{int}}(k) \right\rfloor}^{\left\lfloor j \times r_{\text{int}}(k) \right\rfloor} EGR_{n_{\text{int}}}(k)
\]

∀ \ j \in \{1, 2, \ldots, \ delta_{\text{int}}\}.

Because both fresh air and EGR enter the intake system, the oxygen concentration of the cells entering the intake depends on both the oxygen concentration of the ambient air and the cells leaving the EGR system. Define \( q_{\text{int}} \) as the flow index of the intake system. If \( \left\lfloor (q_{\text{int}} - 1) \times r_{\text{int}}(k) \right\rfloor \leq \left\lfloor q_{\text{int}} \times r_{\text{int}}(k) \right\rfloor \), then the effect of the EGR gases on the oxygen concentration of the \( q_{\text{int}} \)th cell entering the intake system, \( G_{q_{\text{int}}}(k) \), is

\[
G_{q_{\text{int}}}(k) = \frac{1}{r_{\text{int}}(k)} \left( F_{q_{\text{int}}}^n + L_{n_{\text{int}}} + M_{q_{\text{int}}}^n \right),
\]

otherwise

\[
G_{q_{\text{int}}}(k) = EGR_{n_{\text{egr}}}^{(q_{\text{int}} - 1) \times r_{\text{int}}(k)}(k).
\]

The second equation corresponds to the case when \( r_{\text{int}}(k) < 1 \) and only a fraction of a single cell is contained within the \( q_{\text{int}} \)th grouping of EGR cells. The oxygen concentration of the first cell in the intake system after the \( q_{\text{int}} \)th cell has entered the intake system can be predicted using

\[
INT_{1}^{q_{\text{int}}}(k) = \frac{\dot{m}_{\text{egr}}(k)G_{n_{\text{int}}}(k) + \dot{m}_{\text{fresh}}(k)[O_2]_{\text{amb}}(k)}{\dot{m}_{\text{charge}}(k)}
\]

where \( \dot{m}_{\text{fresh}} \) is the mass flow rate of fresh air on a per event basis and the general notation \( INT_{i}^{q_{\text{int}}} \) refers to the oxygen concentration of the \( i \)th cell of the intake system after the \( q_{\text{int}} \)th cell has entered the intake system. The propagation of the downstream
cells within the intake manifold is modeled as

\[ INT_i^{q_{int}}(k) = (1 - \beta_{int}(k))INT_i^{q_{int}-1}(k) + \beta_{int}(k)INT_{i-1}^{q_{int}-1}(k) \]  

∀ \( i \in \{2, 3, ..., \ n_{int}\} \) and ∀ \( q_{int} \in \{1, 2, ..., \ \delta_{int}\} \)

where

\[ \beta_{int}(k) = \frac{\dot{\delta}_{int}(k)}{\delta_{int}(k)}. \]  

As with the other subsystems, the diffusion process is modeled according to

\[ INT_1^{\delta_{int}+1}(k) = (1 - \mathbb{D})INT_1^{\delta_{int}}(k) + \mathbb{D}INT_2^{\delta_{int}}(k), \]  

\[ INT_i^{\delta_{int}+1}(k) = \mathbb{D}INT_{i-1}^{\delta_{int}}(k) + (1 - 2\mathbb{D})INT_i^{\delta_{int}}(k) + \mathbb{D}INT_{i+1}^{\delta_{int}}(k), \]  

∀ \( i \in \{2, 3, ..., \ n_{int} - 1\} \) and

\[ INT_{n_{int}}^{\delta_{int}+1}(k) = \mathbb{D}INT_{n_{int}-1}^{\delta_{int}}(k) + (1 - \mathbb{D})INT_{n_{int}}^{\delta_{int}}(k). \]  

The in-cylinder oxygen concentration depends on both the air charge inducted from the intake system and the residual gases that remained in the cylinders. The contribution from the air charge, \([O_2]_{int}\), can be calculated using

\[ [O_2]_{int}(k) = \frac{1}{\delta_{int}(k)} \sum_{i=1}^{\delta_{int}(k)} INT_i^{\delta_{int}}(k). \]  

Define the trapped residual fraction \( \gamma_{res} \) as the ratio of the trapped residual mass \( \dot{m}_{res}^* \) to the total trapped mass \( \dot{m}_{trapped}^* \) as in

\[ \gamma_{res}(k) = \frac{\dot{m}_{res}^*(k)}{\dot{m}_{trapped}^*(k)}. \]  

Using this definition, the total trapped mass \( \dot{m}_{trapped}^* \) can be related to the flow rate entering the engine with

\[ \dot{m}_{trapped}^*(k) = \frac{\dot{m}_{eng}^*(k)}{1 - \gamma_{res}(k)}. \]
In steady-state, the in-cylinder oxygen concentration is the mass weighted average of the oxygen concentration of the inducted air charge and the oxygen concentration of the residual gases from the previous combustion cycle. In terms of the trapped residual fraction,

\[
[O_2]_{cyl}(k) = (1 - \gamma_{res}(k))[O_2]_{int}(k) + [O_2]_{comb}(k - d_{fuel,int})\gamma_{res}(k)
\]  

(3.62)

where \([O_2]_{cyl}\) is the in-cylinder oxygen concentration on a per mass basis and \(d_{fuel,int}\) is the fixed delay between the fueling and intake events. The mass of oxygen trapped in the cylinder \(\dot{m}_{[O_2]_{cyl}}^*\) is therefore

\[
\dot{m}_{[O_2]_{cyl}}^*(k) = [O_2]_{cyl}(k)\dot{m}_{trapped}(k).
\]  

(3.63)

During combustion much of the oxygen within the cylinders is expended. The remaining oxygen concentration after combustion \([O_2]_{comb}\) can be predicted using

\[
[O_2]_{comb}(k) = \frac{\dot{m}_{[O_2]_{cyl}}^*(k - d_{int.fuel}) - AFR_s[O_2]_{air}\dot{m}_{fuel}(k)}{\dot{m}_{trapped}(k - d_{int.fuel}) + \dot{m}_{fuel}(k)}
\]  

(3.64)

where \(d_{int.fuel}\) is the fixed delay between the induction and fueling events, \(AFR_s\) is the stoichiometric air-to-fuel ratio, \([O_2]_{air}\) is the concentration of oxygen in fresh air and \(\dot{m}_{fuel}^*\) is the total injected fuel mass per event. Before beginning the calculations for the next event, the initial conditions for each of the subsystems must be defined according to

\[
EXH_i^0(k+1) = EXH_i^{k_{exh}+1}(k) \quad \forall \ i \in \{1, 2, 3, ..., n_{exh}\};
\]

\[
EGR_i^0(k+1) = EGR_i^{k_{egr}+1}(k) \quad \forall \ i \in \{1, 2, 3, ..., n_{egr}\};
\]

\[
INT_i^0(k+1) = INT_i^{k_{int}+1}(k) \quad \forall \ i \in \{1, 2, 3, ..., n_{int}\}.
\]  

(3.65)

Unlike a complete air path model, this plug flow based oxygen dynamics model does not predict the thermodynamic conditions of the system. Instead, the temperature, pressure and mass flow rate within the exhaust, EGR and intake systems are
considered as inputs. All of these quantities can be predicted with a zero-dimensional filling and emptying model or alternatively direct measurements can be used. In addition to these inputs, the trapped residual fraction must also be estimated. A plug flow based oxygen dynamics model is essentially a large filter with an input consisting of three terms: the oxygen contribution from the ambient and the oxygen removed during combustion. Each one of these terms has a coefficient (gain) which is necessarily less than one, so although a plug flow based oxygen dynamics model relies on many measurements/estimates as inputs, this type of model is reasonably insensitive to measurement noise, sensor dynamics and estimation errors.

3.2.1 Oxygen Dynamics Modeling Via GT-Power

Experimentally measuring the in-cylinder oxygen concentration is very difficult and not feasible in a production setting. Using a high fidelity engine model such as GT-Power allows for a more accurate prediction of oxygen dynamics. Figure 3.4 shows the steady-state BSFC comparison between experimental data and GT-Power predictions with an $R^2$ value of 94.2%.

![Figure 3.4: Steady-state BSFC comparison: Experimental versus GT-Power](image-url)
as GT-Power, however, the in-cylinder oxygen concentration can be predicted with reasonable confidence. GT-Power is a commercially available software package capable of accurately predicting the one-dimensional gas dynamics within an engine. To quantify the ability of a lumped oxygen dynamics model to predict in-cylinder oxygen concentrations, a GT-Power model of a six cylinder heavy-duty diesel engine with a VGT was developed. This model was calibrated using steady-state data collected experimentally and then validated using transient experimental data.

![Graph showing measured versus predicted NO\textsubscript{x} concentrations with R\textsuperscript{2} = 82.9%](image)

**Figure 3.5:** Steady-state NO\textsubscript{x} comparison: Experimental versus GT-Power

For the steady-state calibration, each of the actuator models were tuned to ensure that the VGT and EGR position inputs affect the air path system the same as the experimental data in terms of quantities such as the fresh air flow, EGR flow, turbine speed, exhaust pressure and intake pressure. Because the specific fuel consumption
and NO\textsubscript{x} production predictions are so important to the selection of the fueling parameters, particular attention was given to the predictive combustion model. The correlations between the experimental and predicted specific fuel consumption and NO\textsubscript{x} production are shown in Figure 3.4 and Figure 3.5. The $R^2$ correlation defined by

\begin{equation}
R^2 = 1 - \frac{\sum_{k=1}^{n_{\text{data}}} (y_{\text{pred}} - y_{\text{meas}})^2}{\sum_{i=1}^{n_{\text{data}}} (y_{\text{meas}} - y_{\text{meas}})^2}.
\end{equation}

was 94.2% for the specific fuel consumption prediction and 82.9% for the NO\textsubscript{x} prediction, both reasonably good for this type of model.

![Figure 3.6: Transient fresh air flow rate comparison: Experimental versus GT-Power](image)

To validate the GT-Power model of the engine, the response of the model over a heavy duty FTP drive cycle was compared to experimental data. In the simulation, the trajectories for the engine speed, fueling parameters, EGR valve position and VGT position measured/commanded in the experimental drive cycle were imposed as inputs to the GT-Power model. For each of the relevant quantities, the predicted and experimentally measured values had very high correlations. The $R^2$ correlations were 91.1% for the fresh air mass flow rate, 91.3% for the EGR mass flow rate, 90.9% for
the oxygen concentration of the exhaust stream and 81.1% for the NO\textsubscript{x} production.

To account for the sensor dynamics of the NO\textsubscript{x} sensor, the GT-Power predictions were filtered with a first order filter and then compared to the measurement data. Figures 3.6-3.9 depicts how well the dynamic variation in each of the four quantities is captured by the GT-Power model for the first 100 seconds of the drive cycle. This same section of the drive cycle was also used to validate the lumped oxygen dynamics model.

Figure 3.7: Transient EGR flow rate comparison: Experimental versus GT-Power

Figure 3.8: Transient [O\textsubscript{2}]\textsubscript{exh} comparison: Experimental versus GT-Power
GT-Power was also used to generate a static lookup function which predicts the trapped residual fraction. The inputs to this lookup function are nonlinear functions of intake manifold pressure, exhaust manifold pressure, engine speed, EGR flow rate...
and fresh air flow rate. Although the performance of this function as shown in Figure 3.10 is rather unremarkable, the trapped residual fraction estimate only affects the in-cylinder oxygen concentration prediction and this effect is minor. As a result, the errors produced by the trapped residual fraction model do not significantly affect the performance of the lumped oxygen dynamics model.

### 3.2.2 GT-Power Validation of a Lumped Oxygen Dynamics Model

Because of its computational requirements, a lumped oxygen dynamics model is not suited for real time implementation in a production ECU. The lumped oxygen dynamics model developed in Section 3.2 was designed to help reduce the need for experimental data collection. These same reasons motivated the development of traditional filling and emptying models. With a traditional filling and emptying model, it is possible to predict the response of the complete air path system including temperatures, pressures, flow rates and oxygen concentrations without requiring an additional experimental data. This same goal with an even more accurate representation of the oxygen mixing and transport dynamic can be accomplished by replacing the oxygen concentration model within a traditional filling and emptying model with a lumped oxygen dynamics model.

Since both a traditional oxygen concentration model and a lumped oxygen dynamics model consider the same input variables, these two modeling approaches can be easily compared in a direct manner. Using the temperatures, pressures and flow rates predicted by GT-Power, the in-cylinder oxygen concentration over the first 100 seconds of a heavy duty FTP drive cycle was predicted with both models and compared to the high fidelity GT-Power model. The lumped oxygen dynamics model produces
excellent performance as shown in Figure 3.11. The top plot in Figure 3.11 compares the in-cylinder oxygen concentration predicted by GT-Power to the in-cylinder oxygen concentration predicted by the lumped oxygen dynamics model, whereas the bottom plot compares the predicted oxygen concentration at the entrance to the exhaust manifold.

Figure 3.11: Oxygen concentration prediction comparison: GT-Power versus lumped oxygen dynamics model

By dynamically accounting for the transport delay, the oxygen concentration predicted by a lumped oxygen dynamics model is always in phase with the actual oxygen concentration. Because a traditional oxygen concentration model does not account for transport delays, the oxygen concentrations predicted by this type of model always leads the actual concentrations. Whenever the oxygen concentration changes, a
Figure 3.12: Oxygen concentration prediction error comparison: Lumped oxygen dynamics model versus traditional oxygen concentration model.

The traditional oxygen concentration model produces significant errors. The in-cylinder oxygen concentrations produced by these two models are compared in Figure 3.12. For the span of events where the oxygen concentration stayed relatively constant, both models produced approximately the same errors. Anytime the injected fuel mass or EGR flow rate changes, however, a traditional oxygen concentration model produces markedly larger errors. During heavy transients when an accurate in-cylinder oxygen concentration prediction is most beneficial, a traditional model is least accurate. Over the complete test, the root mean squared (RMS) error in the in-cylinder oxygen concentration prediction was 0.109% for the lumped oxygen dynamics model and 0.192% for a traditional model.

### 3.3 High Fidelity Exhaust Oxygen Dynamics Model

The approach described in Section 3.2 is particularly well suited for predicting the in-cylinder oxygen concentration in a lean burning engine with an external EGR
loop. Regardless of the estimation method, however, the predicted trapped cylinder mass and the predicted in-cylinder oxygen concentration will have some degree of uncertainty. To correct for these estimation errors, the oxygen concentration of the exhaust stream as measured by an oxygen sensor can be used in feedback. This sensor is usually placed at the confluence junction of the exhaust manifold runners in naturally aspirated engines or downstream of the turbocharger in turbocharged engines. Although much smaller than the transport delays of a complete EGR loop, the transport delay between the exhaust valves and the confluence junction of the location of the oxygen sensor within the exhaust system is still significant. To reliably predict the oxygen concentration at the oxygen sensor located in the exhaust system, the fidelity of the plug flow model present in Section 3.2 must be increased to account for the shorter timescale. The result is a high fidelity exhaust oxygen dynamics (HFEOD) model.

For the purpose of predicting the in-cylinder oxygen concentration, the behavior of each cylinder is assumed to be identical, therefore allowing a single lumped pipe model for the exhaust system. To predict/identify cylinder imbalances, the physical differences in the exhaust runners of each cylinder must be taken into account. The variation in the transport delay caused by these differences varies with the firing order and the exhaust flow rate. To model these effects, each physical pipe in the exhaust manifold is modeled with a separate control volume subsystem.

As with the lumped oxygen dynamics model, the oxygen concentration at any point in the exhaust system can be predicted by tracking packets of combustion gases as they move through the exhaust system. Each control volume subsystem, which now corresponds to a physical linear section of pipe, is discretized into many cells of
constant volume. An example of a 4-2-1 exhaust manifold being discretized into 100 cells is shown in Figure 3.13.

Figure 3.13: Example 4-2-1 exhaust system separated into 100 cells

When a cylinder exhausts its combustion products, the gas packets downstream from the exhaust ports are pushed further downstream. However, only packets in the fluid path of the currently exhausting cylinder actually move. The volume occupied by the newly exhaust gas mass is still calculated by (3.24). Each cell now truly represents a physical location in the exhaust system, so these newly exhausted cells enter the model at the cell corresponding to the exhaust ports of the exhausting cylinder. If pipe $a$ represents the exhaust runner of this cylinder, then the oxygen concentration of the first cell in this pipe after packet $q_{exh}$ has been exhausted is governed by

$$S_{Eexc_{a,1}}(k) = [O_2]_{comb}(k - d_{fuel,exh})$$  \hspace{1cm} (3.67)
where as before $q_{exh}$ is the flow index. The general notation $Sec_{ξ,i}^{q_{exh}}(k)$ refers to the oxygen concentration of the $i^{th}$ cell in pipe $ξ$ after the $q_{exh}^{th}$ gas packet of event $k$ has been exhausted.

To make room for this new cell, each cell in the fluid path must move downstream one cell. The propagation of the cells within any pipe $ξ$ in the fluid path is modeled as

$$Sec_{ξ,i}^{q_{exh}}(k) = (1 - \beta_{exh}(k))Sec_{ξ,i}^{q_{exh}-1}(k) + \beta_{exh}(k)Sec_{ξ,i-1}^{q_{exh}-1}(k)$$  \hspace{1cm} (3.68)

∀ $i \in \{2, 3, ..., n_ξ\}$ where $n_ξ$ is the number of cells in pipe $ξ$ and $\beta_{exh}$ is still governed by (3.28). When a cell transitions between piping sections, the gases mix resulting in a new oxygen concentration. Consider the case when exhaust gases are pushed through pipe $a$ into pipe $c$ through a pipe junction where the outlets of pipes $a$ and $b$ converge into the inlet of pipe $c$. The oxygen concentration of the first cell in pipe $c$ after the exhausting of the $q_{exh}^{th}$ packet is predicted according to

$$Sec_{c,1}^{q_{exh}}(k) = (1 - \beta_{exh}(k))Sec_{c,1}^{q_{exh}-1}(k) + \beta_{exh}(k)M_{ab}Sec_{a,n_a}^{q_{exh}-1}(k)$$

$$+ \beta_{exh}(k)(1 - M_{ab})Sec_{b,n_b}^{q_{exh}-1}(k)$$  \hspace{1cm} (3.69)

where $M_{ab} \in (0.5, 1.0]$ models the mixing effects of the confluence junction of pipes $a$ and $b$. A portion of the gases leaving pipe $a$ is exchanged with an equal proportion of gases in the last cell of pipe $b$. The parameter $M_{ab}$ has a minimum bound of 0.5 to ensure that the pipe in the fluid path has a larger effect than the other pipe not in the fluid path. The oxygen concentration of the last element of the pipe not in the fluid path is defined by

$$Sec_{b,n_b}^{q_{exh}}(k) = \left(1 - \mathbb{B} - \beta_{exh}(k)(1 - M_{ab})\right)Sec_{b,n_b}^{q_{exh}-1}(k) + \mathbb{B}Sec_{b,n_b-1}^{q_{exh}-1}(k)$$

$$+ \beta_{exh}(k)(1 - M_{ab})Sec_{a,n_a}^{q_{exh}-1}(k).$$  \hspace{1cm} (3.70)
Although there is no net motion, the gases in each of the pipes not in the fluid path still experience mixing via diffusion, modeled by

$$Sec_{b,i}^{q_{exh}}(k) = \mathbb{D}Sec_{b,i+1}^{q_{exh}}(k) + (1 - 2\mathbb{D})Sec_{b,i}^{q_{exh}}(k) + \mathbb{D}Sec_{b,i-1}^{q_{exh}}(k)$$  \hspace{1cm} (3.71)

\(\forall \ i \in \{2, 3, \ldots, n_b - 1\}\). For the last element of a pipe, diffusion is modeled as

$$Sec_{b,1}^{q_{exh}}(k) = \mathbb{D}Sec_{b,2}^{q_{exh}}(k) + (1 - \mathbb{D})Sec_{b,1}^{q_{exh}}(k).$$  \hspace{1cm} (3.72)

To model the effect of each new cell entering the exhaust system, the calculations in (3.67) - (3.72) must be repeated for all integer values of \(q_{exh}\) between 1 and \(\delta_{exh}\). During each event, multiple packets of exhaust gas move past any given location in the exhaust system. Therefore, the event averaged oxygen concentration of the \(i^{th}\) cell of any pipe \(\xi\) can be calculated as

$$\overline{Sec_{\xi,i}}(k) = \frac{1}{\delta_{exh}(k)} \sum_{q_{exh}=1}^{\delta_{exh}(k)} Sec_{\xi,i}^{q_{exh}}(k).$$  \hspace{1cm} (3.73)

If cell \(Sec_{n_{pipes},n_{pipes}}\) corresponded to the location of the oxygen sensor, then \(\overline{Sec_{n_{pipes},n_{pipes}}}\) would represent the predicted oxygen concentration at the oxygen sensor. Because the dynamics of an oxygen sensor are nontrivial, they must be accounted for when predicting the measured oxygen concentration. An oxygen sensor can be approximated by a first order system with a fixed-in-time time constant. To implement this model in the discrete domain, the number of seconds per event must be calculated using

$$\Delta t(k) = \frac{120}{n_{cyl}N(k)}.$$  \hspace{1cm} (3.74)

The expected oxygen sensor response can be calculated by applying the Euler approximation to the first order system model which has a time constant \(\tau_{O_2}\) in the
manner

\[
\overline{\text{Sec}_{z,n}(k)} - [O_2]_{ego}(k) = \tau_{O_2} \frac{d([O_2]_{ego}(k))}{dt} \\
\approx \tau_{O_2} \frac{[O_2]_{ego}(k) - [O_2]_{ego}(k-1)}{\Delta t(k)}
\] (3.75)

which simplifies to

\[
[O_2]_{ego}(k) = \overline{\text{Sec}_{n_{pipe},n_{pipe}}(k)} \frac{\Delta t(k)}{\Delta t(k) + \tau_{O_2}} + [O_2]_{ego}(k-1) \tau_{O_2}
\] (3.76)

Before beginning the calculations for the next event, the initial conditions for each pipe \(\xi\) must be defined according to

\[\forall i \in \{1, 2, 3, ..., n_{\xi}\}.
\]

### 3.3.1 High Fidelity Exhaust Oxygen Dynamics Model Calibration

This section details the calibration procedure for the more difficult scenario when a high frequency UEGO sensor is unavailable and all measurements are made using a switching-type EGO sensor. The response of a production grade UEGO sensor is too slow to accurately resolve event-to-event fluctuations, and is therefore not fast enough to calibrate an HFEOD model. Because a UEGO sensor can measure a wider range of oxygen concentrations, calibration using a fast responding lab grade UEGO sensor would be more accurate. However, the primary objective of this section is to demonstrate the effectiveness of the exhaust oxygen concentration model under worst case conditions, thereby motivating the use of a simple sensor set for both the calibration process and on-line usage. Note that for the following calibration
procedure information from a UEGO sensor is required to calibrate an EGO sensor model.

For the case when only an EGO sensor is available, the relationship between voltage and oxygen concentration (or equivalently AFR) must be calibrated. To this end, AFR values from 0.96 to 1.04 can be commanded in increments of 0.002, then from 1.04 back down to 0.96 to capture any hysteresis effects. To generate the EGO sensor curve, the voltage should be averaged for each AFR set point, facilitating a two way conversion from voltage to AFR in a limited range. Two datasets are needed to tune the mixing and diffusion parameters which are the only true calibration parameters. For the engine with exhaust system shown in Figure 3.13, three mixing parameters ($M_{12}$, $M_{34}$, and $M_{56}$) and one diffusion parameter ($D$) require calibration. To ensure that the resulting model is conservative, these parameters must be constants and should not be scheduled as a function of the operating conditions. To isolate the specific response of each cylinder, steps and impulses must be commanded to each cylinder at randomly selected operating conditions. Then to capture the transient behavior, a drive cycle such as an FTP-72 drive cycle should be driven experimentally. Using standard optimization procedures, the globally optimal parameters can then be calculated. A suggested cost function is the 2-norm of the error between the model predicted EGO voltage and the measured EGO voltage over the two datasets.

### 3.3.2 High Fidelity Exhaust Oxygen Dynamics Model Validation

To demonstrate the effectiveness of a high fidelity exhaust oxygen dynamics model under the worst case conditions described in the previous section, a model was developed to represent the 2.4 liter gasoline engine with variable valve timing in a 2006
Saturn Ion. The firing order for this engine is the industry standard, 1-3-4-2. All of the measurements were made using the production sensor set except for the addition of a Bosch LSU 4.9 wide range UEGO sensor installed adjacent to the production pre-catalyst EGO sensor. The response times of the EGO and UEGO sensors are approximately 50 and 100 milliseconds, respectively. All testing was performed on a 150 horsepower twin roll chassis dynamometer.

For this model, the 4-2-1 exhaust configuration was discretized into 100 cells to provide the best compromise between accuracy and computational complexity. An illustration of the geometry of the model is given in Figure 3.13. This number can be thought of as typical for a 4-cylinder engine. Either the total number of cells or the window length could be increased, but the increased accuracy would not offset the increased computational burden. The four calibration parameters were calibrated according to the procedures outlined in the previous section using a genetic algorithm.

In order to validate the model performance, an additional set of experimental data was collected. At constant manifold pressure and engine speed, step and impulse signals were added to the fuel commands of individual cylinders. Because the exhaust geometry is symmetric, the response of cylinder-1 is nearly the same as cylinder-4 and the response of cylinder-2 is nearly the same as cylinder-3. For this reason only the responses of two cylinders are presented. Cylinder-1 and cylinder-3 were chosen because they are the first two cylinders in the firing order; the results for cylinder-4 and cylinder-2 are similar. Relative to the outer cylinders (1 and 4), the inner cylinders (2 and 3) have a much shorter delay primarily due to the firing order. Referencing Figure 3.13, the exhaust gas expelled from cylinder-1 for sufficiently low flow rates such as idle fill sections 1 of the exhaust system and part of section 5.
The exhaust gases in section 5 remain there until three events later when cylinder-2 exhausts and pushes the contents of section 5 into section 7. Under these same conditions, the exhaust gas expelled from cylinder-3 would fill sections 3 and 6. The very next event cylinder-4 exhausts and pushes the contents of section 6 into section 7, resulting in a shorter delay.

These significant cylinder-to-cylinder variations in the plant delay are well predicted by a HFEOD model. A comparison between the measured and predicted impulse responses for cylinder-1 at idle is shown in Figure 3.14. The impulse is commanded at event 0, but does not appear in the EGO voltage until 11 events later. As shown in Figure 3.15, an impulse commanded to cylinder-3 at idle appears after only 8 events. Although the amplitude was slightly under-predicted for cylinder-3, the predicted impulse responses show good agreement with the measured responses. Because this model is based on an EGO sensor, it is prone to small amplitude errors. If a fast responding UEGO sensor is used, an even better agreement in amplitudes would be expected.

Equivalent performance was also achieved under moderate loads (Figure 3.16 and Figure 3.17). For this data the mass flow rate was held constant at 210 milligrams per event and the engine speed was fixed at 1900 revolutions per second. At this higher flow rate the response speeds of cylinders 1 and 4 are much more comparable to cylinders 3 and 2, because the exhausted gases fill more of the exhaust system. When any cylinder exhausts at this flow rate, part of the exhausted gas reaches section 7 (referencing Figure 3.13). Because at every event the contents of section 7 move, some of the newly exhausted gases will reach the oxygen sensor one event later. Since each cylinder exhibits this same behavior, the effect of the firing order
Figure 3.14: Measured versus model predicted impulse response of cylinder-1 at idle which still determines when sections 5 and 6 move is less pronounced. As the flow rate increases, the plant delay decreases and variation in the response shape between cylinders also decreases.

With higher engine speeds, the relative response of the EGO sensor is slower. As a result, the EGO voltage is dependent on a wider range of inputs. This behavior, however, is also captured very well by a HFEOD model. To provide an objective performance metric, the $R^2$ correlation was calculated using (3.66). The $R^2$ correlations between the impulse response predicted by the HFEOD model and the measured impulse response for representative operating conditions are presented in Table 3.1. In all cases the $R^2$ value was greater than 75%; these are deemed to be favorable results considering the variability of the engine and EGO sensor.
For these load conditions, the impulse responses predicted by the HFEOD model have peaks and valleys that match up with the measured responses. Because the locations are aligned, the HFEOD model also accurately predicts the response of the engine under cylinder imbalance, which is essentially a cascade of several impulse responses. If the shape of the impulse response is shifted at all, then the cylinder imbalance response will be predicted poorly. Figure 3.18 compares the response of the HFEOD model (dashed line) to the measured response (solid line) to a cylinder-1 imbalance error during idle. For this figure the event in which cylinder-1 commands fuel is distinguished with a circle. Recall from Figure 3.14 that the impulse response peaks 11 events after the fuel is commanded; as a consequence, a cylinder-1 imbalance error should cause an increase in the EGO voltage the event before cylinder-1
Figure 3.16: Measured versus model predicted impulse response of cylinder-1 under moderate load

commands fuel. As expected, both the measured and predicted responses had a high EGO voltage during these events.

The order in which fuel is commanded is the same as the firing order, except for an event shift. As such, fuel is commanded to cylinder-3 the event after fuel is delivered to cylinder-1. Unlike cylinder-1 and cylinder-4, the impulse responses of cylinder-2 and cylinder-3 have two distinct peaks. These peaks occur 8 events and 12 events, respectively, after fuel is commanded to those cylinders. Based on this observation, a cylinder-3 imbalance error should have a peak when fuel is commanded to cylinder-3 or, equivalently, one event after fuel is commanded to cylinder-1. As shown in Figure 3.19, the peaks of both the measured and predicted responses did occur one event after fuel is commanded to cylinder-1 (data points denoted by circle).
Figure 3.17: Measured versus model predicted impulse response of cylinder-3 under moderate load

Again, the $R^2$ metric was used to determine the quality of the fit. This information is provided in Table 3.2 for cylinder-1 and Table 3.3 for cylinder-3. Like the previous results, the performance of cylinder-4 and cylinder-2 were approximately the same. To highlight the variability of the engine, these tables include a fit relative to the average imbalance response. The fit of the average response is essentially the best that any model could achieve and is therefore given the label “Ideal Prediction”. With this in mind, the HFEOD model was able to approximate the average response of the engine to a cylinder imbalance error exceptionally well. The $R^2$ fit is normally within 10% of the $R^2$ fit of the average response. The standard deviation between the EGO voltage predicted by the HFEOD model and the measured voltage appears slightly high. However, this is mostly a result of the variability of the system, confirmed by
Table 3.1: Cylinder impulse response $R^2$ comparison: HFEOD model versus measured

<table>
<thead>
<tr>
<th>Mass Flow Rate (mg/event)</th>
<th>Engine Speed (RPM)</th>
<th>Cylinders 1 and 4 $R^2$(%)</th>
<th>Cylinders 2 and 3 $R^2$(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
<td>800</td>
<td>93.7</td>
<td>81.8</td>
</tr>
<tr>
<td>210</td>
<td>1900</td>
<td>95.1</td>
<td>88.4</td>
</tr>
<tr>
<td>250</td>
<td>1400</td>
<td>77.9</td>
<td>86.2</td>
</tr>
</tbody>
</table>

Table 3.2: Cylinder-1 imbalance $R^2$ comparison: Models (HFEOD and ideal) versus measured

<table>
<thead>
<tr>
<th>Mass Flow Rate (mg/event)</th>
<th>Engine Speed (RPM)</th>
<th>HFEOD prediction $R^2$(%)</th>
<th>Std (mV)</th>
<th>Ideal prediction $R^2$(%)</th>
<th>Std (mV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
<td>800</td>
<td>81.3</td>
<td>117.8</td>
<td>95.8</td>
<td>62.9</td>
</tr>
<tr>
<td>210</td>
<td>1900</td>
<td>72.2</td>
<td>51.5</td>
<td>78.4</td>
<td>47.4</td>
</tr>
<tr>
<td>250</td>
<td>1400</td>
<td>74.3</td>
<td>108.8</td>
<td>83.3</td>
<td>90.9</td>
</tr>
</tbody>
</table>

comparing the standard deviation of the average response to the standard deviation of the HFEOD model.

Although a HFEOD model is able to accurately predict the effects of cylinder imbalance on the oxygen sensor measurement, the form of this model is not conducive to identifying cylinder imbalance. In the next section the high fidelity exhaust oxygen dynamics model will be developed further to dynamically identify the relationship between input fueling commands and output exhaust oxygen sensor measurements. This model will be the basis for a model based cylinder imbalance estimator.
Figure 3.18: Measured versus model predicted EGO response of a cylinder-1 imbalance error at idle

Figure 3.19: Measured versus model predicted EGO response of cylinder-3 imbalance error at idle
Table 3.3: Cylinder-1 imbalance $R^2$ comparison: Models (HFEOD and ideal) versus measured

<table>
<thead>
<tr>
<th>Mass Flow Rate (mg/event)</th>
<th>Engine Speed (RPM)</th>
<th>HFEOD prediction $R^2$(%) Std (mV)</th>
<th>Ideal prediction $R^2$(%) Std (mV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
<td>800</td>
<td>73.1 121.9</td>
<td>82.3 98.9</td>
</tr>
<tr>
<td>210</td>
<td>1900</td>
<td>58.6 81.8</td>
<td>61.7 78.8</td>
</tr>
<tr>
<td>250</td>
<td>1400</td>
<td>70.0 96.6</td>
<td>79.5 86.0</td>
</tr>
</tbody>
</table>

3.4 Oxygen Dynamics Modeling Through a Finite Impulse Response Characterization

The goal of the high fidelity exhaust oxygen dynamics model presented in Section 3.3 is to accurately predict the oxygen concentration at the oxygen sensor. By comparing the predicted oxygen concentration measurement to the actual oxygen concentration measurement, it is possible to identify and correct in-cylinder oxygen concentration estimation errors which uniformly affect all of the cylinders. With the representation presented in the previous section, however, it is not possible to directly identify and link an error between a measured and predicted oxygen concentration back to a specific cylinder fueling event for period cylinder imbalance error correction.

By expanding and recasting the high fidelity exhaust oxygen dynamics model described in Section 3.3, it is possible to characterize the system in terms of a finite impulse response (FIR) model. For convenience this formulation uses the EQR domain rather than oxygen concentration. The output of the new plant, the EQR at the exhaust oxygen sensor, is defined exclusively by recursions of the input EQR in
the manner

$$\phi_{O_2}(k) = IO(k)^T \begin{bmatrix} \phi_{in}(k) \\ \phi_{in}(k-1) \\ \vdots \\ \phi_{in}(k-n_{\text{window}} + 1) \end{bmatrix} = IO(k)^T u_{\text{window}}(k)$$

(3.78)

where $n_{\text{window}}$ is the size of the input window, $\phi_{O_2}(k) \in \mathbb{R}$ is the predicted EQR at the oxygen sensor, $\phi_{in}(k) \in \mathbb{R}$ is the EQR of the pre-combustion mixture, $IO(k) \in \mathbb{R}^{n_{\text{window}}}$ contains the coefficients describing the impulse response of the system and $u_{\text{window}}(k) \in \mathbb{R}^{n_{\text{window}}}$ is a moving window of the past EQR commands. Because the input/output relationship is linear, several properties can be exploited. The coefficients of the model vary with time (events), mass flow rate and the cylinder firing order. In this way, complexity has essentially been shifted from the input/output structure to the coefficient prediction model. The coefficients are calculated using a variation of the HFEOD model, referred to henceforth as the \textit{coefficient prediction model}.

To recast the high fidelity exhaust oxygen dynamics model into the coefficient prediction model, another dimension was added to each cell of the model. In the original model, both the location and the oxygen concentration of the combustion gas packets are tracked through the exhaust system. To develop an FIR model, a relative time stamp corresponding to the number of events since fuel was injected must also be considered. Each discretized cell of the exhaust system must now store information about its time stamped composition instead of just its oxygen concentration. As the exhaust gas packets move through the exhaust, the composition changes. By the time a packet reaches the oxygen sensor, it will contain contributions from multiple cylinder fueling events.
Each cell is defined by a unique set of weighting factors corresponding to a window of inputs. This set of factors is captured in a vector referred to as the *input dependency vector*. Each element in this vector represents the fractional composition of the cell with respect to a specific fueling input. By convention, the first element corresponds to the current input, the second element corresponds to the input commanded one event earlier, and so on. The EQR of any cell can be calculated by taking the inner product of the input dependency vector and a vector containing a window of the previous EQR inputs. The length of the input dependency vector is constant but depends on the engine being modeled. This vector must be long enough to capture the complete dependence under the worst case conditions (low flow rate and high engine speed). When the flow rate is low, the transport delay is largest. At high engine speeds, the time constant of the oxygen sensor increases in the event domain. For a four cylinder engine a window size of 30 events captures more than 99% of the impulse response under these worst case conditions.

Tracking the location and composition of the exhaust gas packets first requires prediction of the number of cells entering the exhaust system using (3.24). Because the cells entering the exhaust system from the cylinder have not yet mixed with any other gases, the input dependency vectors of these packets have only one nonzero term. This relationship is captured by

\[
SEC_{q_{exh}}(1, j) = \begin{cases} 
1 & \text{if } j = d_{fuel,exh} \\
0 & \text{otherwise}
\end{cases}
\]

\(\forall j \in \{1, 2, ..., n_{window}\}\) where \(SEC_{q_{exh}} \in \mathbb{R}^{n_a \times n_{window}}\) contains the input dependency vectors of the cells within pipe \(a\) after packet \(q_{exh}\) has been exhausted. The general notation \(SEC_{\xi q_{exh}}(i, j)\) represents the scalar value of the \(j^{th}\) element of the input
dependence vector of the \( i^{th} \) cell in pipe \( \xi \) after packet \( q_{exh} \) has been exhausted. Although not explicitly shown, each of these variables varies with time (events).

To make room for these new cells, an equal number of cells must exit the exhaust system. However, not every cell moves; only the cells in the fluid path between the exhaust valves and the oxygen sensor move. The propagation of the cells within the pipes in the fluid path is modeled with the vector equation

\[
SEC_q^{exh}(i,:) = (1 - \beta_{exh})SEC_q^{exh-1}(i,:) + \beta_{exh}SEC_q^{exh-1}(i-1,:)
\]

\( \forall \ i \in \{2, 3, ..., n_a\} \) where \( \beta_{exh} \) is governed by (3.28). The general notation \( SEC_{\xi}^{q_{exh}}(i,:) \) refers to the input dependency vector of the \( i^{th} \) cell in pipe \( \xi \) after packet \( q_{exh} \) has been exhausted. When a cell transitions from one pipe to another, the gases from all three of the pipes at the junction mix. The input dependency vector that gets pushed into the first cell of the downstream pipe is described by

\[
SEC_c^{q_{exh}}(1,:) = (1 - \beta_{exh})SEC_c^{q_{exh-1}}(1,:) + \beta_{exh}M_{ab}SEC_a^{q_{exh-1}}(n_a, :)
+ \beta_{exh}(1 - M_{ab})SEC_b^{q_{exh-1}}(n_b, :)
\]

\( 3.81 \)

where \( SEC_b^{q_{exh}} \in \mathbb{R}^{n_b \times n_{window}}, SEC_c^{q_{exh}} \in \mathbb{R}^{n_c \times n_{window}} \) and \( M_{ab} \in (0.5, 1.0] \) represents the mixing process. A portion of the input dependency vector is exchanged between the last cells of the pipe that is in the fluid path and the pipe not in the fluid path. The input dependency vector of the last element of the pipe not in the fluid path is defined by

\[
SEC_b^{q_{exh}}(n_b,:) = \left( 1 - \mathbb{D} - \beta_{exh}(1 - M_{ab}) \right)SEC_b^{q_{exh-1}}(n_b,:)
+ \mathbb{D}SEC_b^{q_{exh-1}}(n_b - 1,:) + \beta_{exh}(1 - M_{ab})SEC_a^{q_{exh-1}}(n_a,:).
\]

\( 3.82 \)
Although there is no net motion, the gases in each of the pipes not in the fluid path still experience mixing via diffusion, modeled by

\[
SEC_b^{q_{exh}}(i,:) = DSEC_b^{q_{exh}-1}(i + 1,:) + (1 - 2D)SEC_b^{q_{exh}-1}(i,:) + DSEC_b^{q_{exh}-1}(i - 1,:)
\]

\((3.83)\)

\(\forall i \in \{2, 3, ..., n_b - 1\}\) and

\[
SEC_b^{q_{exh}}(1,:) = DSEC_b^{q_{exh}-1}(2,:) + (1 - D)SEC_b^{q_{exh}-1}(1, :)
\]

\((3.84)\)

To model the effect of each new cell entering the exhaust system, the calculations in (3.79) - (3.84) must be repeated for all integer values of \(q_{exh}\) between 1 and \(\delta_{exh}\).

Multiple cells move past the EGO sensor during each event. Therefore, the input to the oxygen sensor model is the average input dependency vector of these cells. The oxygen sensor is approximated by a first order system with a fixed-in-time time constant. To implement this model in the discrete domain, again the Euler approximation is used to generate

\[
IO(k) = \frac{SEC_{n_{pipes}}(n_{pipes} ;:) \Delta t(k) + IO(k - 1)\tau_{O_2}}{\Delta t(k) + \tau_{O_2}}.
\]

\((3.85)\)

In this equation, \(SEC_{n_{pipes}}(n_{pipes} ;:) \in \mathbb{R}^{n_{window}}\) which represents the averaged composition flowing past the oxygen sensor can be calculated using

\[
SEC_{n_{pipes}}(n_{pipes} ;:) = \frac{1}{\Delta t} \sum_{q_{exh}=1}^{\delta_{exh}} SEC_{n_{pipes}}(n_{pipes}, q_{exh}).
\]

\((3.86)\)

where the \(n_{pipes}^{th}\) cell within pipe \(n_{pipes}\) is assumed to correspond to the location of the oxygen sensor. Because (3.85) is a stable filter, \(IO(k)\) will converge to the coefficients of the finite impulse response of the system regardless of the initial conditions.

Before the next set of FIR coefficients are predicted, another calculation must be made. Relative to the output of the next event, each input has an additional delay
of one event. Therefore, all of the input dependency vectors must be shifted one event when the new initial conditions corresponding to \( q_{exh} = 0 \) are defined. This is captured with

\[
SEC_0^\xi(\cdot, j) = \begin{cases} 
SEC_{\delta_{exh}}(\cdot, j - 1) & \forall \ j \in \{2, 3, \ldots, n_{window}\} \\
0 & \text{if } j = 1
\end{cases}
\]  
(3.87)

where \( \xi \in \{a, b, c, \ldots, n_{pipes}\} \). Although several calculations are required to find these coefficients, they define a very simple FIR model between fuel injection and exhaust oxygen sensor measurement.

Tracking the input dependency vector in this manner guarantees that the model satisfies the conservation constraints. Both the total mass and the mass of each species, where the mass of exhaust gas expelled each event is considered a separate species, must be conserved. As is commonly done in cylinder imbalance modeling, these equations can be converted into a matrix representation. In a matrix form, the conservation laws require that all of the row-wise and column-wise summations must be unity. Because all of the equations governing the input dependency vector are algebraic sums of input dependency vectors, these properties are automatically satisfied.

Figure 3.20 demonstrates how the input dependency vector changes for a simple system by showing snapshots of the vector for successive events. In this example, a two cylinder engine is modeled using one cell for each of the three exhaust pipes, and an output window of five events was chosen. The dimensions of \( SEC_1, SEC_2 \) and \( SEC_3 \) corresponding to these values are \( 1 \times 5 \). The two blocks on the left in each event represent the exhaust runners and the block on the right represents the confluence piping that contains the oxygen sensor. The output of this model is the input dependency vector of the block on the right. For simplicity, at each event only
Figure 3.20: Evolution of the input dependency vector
one packet is exhausted, and the mixing/diffusion processes are captured by a single
coefficient, $\beta \in (0.5, 1]$. To demonstrate the conservation of species law, a unique
variable identifier with a numerical value of unity is assigned to each new packet.

The first event represents an initial condition where each section of the exhaust
contains gases expelled from three different events ($a$, $b$ and $c$) corresponding to the
fueling events 3, 2 and 1 events prior, respectively. In between each event, each
element of the input dependency vectors gets shifted down by one to account for the
additional one event delay. In the second event, cylinder-2 exhausts and pushes the
contents of section 1, which is now $[0, 0, b, 0, 0]$, into section 3. As these contents are
being pushed, they mix with the contents of section 2. After mixing, the resulting
input dependency vector of section 3 becomes the weighted average of section 1 and
section 2 with weights $(1 - \beta)$ and $\beta$, respectively. This mixing also affects the input
dependency vector of section 1 equally and oppositely ($\beta$ multiplies section 1 and
$(1 - \beta)$ multiples section 2). The newly exhausted gases which occupy section 2 are
only dependent on a single fueling event represented by the new variable $d$.

In the next event cylinder-1 exhausts and pushes the contents of section 1 into
section 3. Again the input dependency vector of section 3 is the weighted average
of sections 1 and 2 but the weighting factor switches. The contents of section 2 are
also affected by this mixing process equally and oppositely ($\beta$ multiplies section 2
and $(1 - \beta)$ multiplies section 1). The newly exhausted gases which occupy section 1
are only dependent on a single fueling event represented by the new variable $e$. The
continuation of this alternating process is shown for two more events corresponding
to the fueling event variables $f$ and $g$. 

93
The values to the far right of Figure 3.20 indicate the cumulative influence of each exhaust gas packet on the output. Although it takes several events, the cumulative influence of each exhaust packet on the output is eventually unity. Notice that for every event the sum of the terms in each input dependency vector is also unity. These two summations verify that the species and mass conservation constraints are satisfied.

3.4.1 Input/Output Analysis: Fuel Delivery and EGO Measurement

The equations governing the coefficient prediction model within the FIR exhaust system model are nearly identical to the high fidelity exhaust oxygen dynamics model. The only difference is the dimension of the dynamic variables within the equations. For this reason, the same calibration procedures used in Section 3.3.1 are also applicable to a FIR exhaust system model. Given that the performance of the high fidelity exhaust oxygen dynamics model has already been demonstrated, this section explores the input/output relationship between fuel delivery and EQR measurement.

Recall from Section 3.3.2 that cylinder-3 imbalance errors during idle result in a peak EGO voltage one event after cylinder-1 commands fuel. Under the same conditions, the peak of the impulse response of cylinder-4 occurs 11 events after commanding fuel. Fuel is commanded to cylinder-4 two events after fuel is commanded to cylinder-1. Referencing cylinder-1, a cylinder-4 imbalance error results in a peak in the EGO voltage one event after fuel is commanded to cylinder-1. This is also the same event that cylinder-3 imbalance errors appear. Similarly, cylinder-2 imbalance errors appear in the EGO voltage predominately at the same events as a cylinder-1
imbalance error (one event before cylinder-1 commands fuel). These results only hold at idle, but they have significant implications for cylinder imbalance identification.

To visually represent these concepts, consider Figure 3.21 which shows how the output depends on the inputs at idle for the event prior to cylinder-1 commanding fuel. The values on the $x$-axis represent the number of events since that input was commanded. For example, the value at event 10 represents the fractional weight that the input, commanded 10 events previously, has on the current output. The largest contributions result from cylinder-1 and cylinder-2. The cumulative dependence of the output on an increasing window of inputs (beginning at event 0) is also included in the figure. From this figure, it is clear that the effects of inputs commanded more than 17 events prior are almost negligible.

The output measured at the event after cylinder-1 commands fuel depends on the inputs in a similar manner, except that the contributions from cylinder-1 switch with cylinder-4 and the contributions from cylinder-2 switch with cylinder-3. When cylinder-1 commands fuel, the relationship between the output and input changes profoundly. As shown in Figure 3.22, each cylinder affects the output nearly identically. Notice that the cumulative weight still reaches unity after about 18 events.

In addition to observing how previous inputs affect the current output, the relationship between the current input and future outputs can also be explored, particularly as it is connected to the impulse response of the cylinder. At idle cylinder-1 and cylinder-4 influence the output predominately after an 11 event delay. The effect at idle of an input applied to cylinder-1 on the output measurement is shown in Figure 3.23, with a cumulative curve. The values on the $x$-axis represent the number of events between commanding fuel and output measurement. For example, the value
Figure 3.21: Output dependencies: Event prior to commanding fuel to cylinder-1 at idle

at event 10 corresponds to the fractional weight that the current input will have on the output 10 events in the future. After approximately 18 events, nearly all of the input effect has passed through the output. Both of these results are consistent with the impulse and cylinder imbalance analyses given previously. Inputs commanded to either cylinder-2 or cylinder-3 reach the output sooner than inputs from cylinder-1 and cylinder-4. Figure 3.24 shows the effect that cylinder-3 has on the output at idle. The responses of cylinder-2 and cylinder-3 are bimodal. Therefore, even though the contributions from these cylinders initially reach the output quicker, the same number of events are required to completely push the combustion gases through the exhaust.
Figure 3.22: Output dependencies: Event when fuel is commanded to cylinder-1 at idle

Under all conditions, the contributions to the output must sum to unity. Excluding the effect of noise, the output of the sensor is clearly strongly influenced by the inputs. Therefore, the contribution from all of the inputs to each specific output must always sum to unity. This statement is not true for the contributions of each individual input; only in steady-state will the contributions from each specific input, to all of the outputs, sum to unity. Depending on the changing rate of the operating conditions, an input will have a larger or smaller total contribution to the outputs relative to the other inputs. To illustrate this point, a simple set of tests was performed. These tests are not meant to replicate the physical behavior of an engine; rather, they are used to isolate the effects of mass flow rate and engine speed.
In the first case considered here, the mass flow rate is changed while the engine speed is fixed at 800 RPM. As shown in Figure 3.25, the summation of the input dependencies (solid line) is not always unity, unlike the output dependencies (dashed line). At the first dashed vertical line corresponding to event 30, the trapped air mass is increased from 120 to 200 milligrams per event. Then at the second dashed vertical line corresponding to event 60, the trapped air mass is decreased back down to 120 milligrams per cylinder. Changing the mass flow rate affects the inputs just prior to the change. The number of cells an input will occupy in the exhaust is proportional to the mass flow rate. This is also the number of cells that pass by the oxygen sensor. Each cell has the same weight, but inputs with different mass flow rate occupy a different number of cells. The summation of the input dependencies for the event
just prior to the mass flow rate change can be calculated directly by dividing the old flow rate by the new flow rate. For the example shown in Figure 3.25, the input dependencies sum for event 29 is 0.6; similarly, the input dependencies sum for event 59 is 1.67. For the inputs commanded longer than one event prior to a mass flow rate change, the effect of the change is less significant. This result is intuitive, because these inputs see a combination of flow rates.

Engine speed also affects the input contribution summation, for different inputs. How a particular input influences the output is dependent on the mass flow rate at that event and the subsequent few events (number of events in the transport delay). Engine speed does not affect the input dependencies until one of the cells reaches the oxygen sensor. As a result, the inputs affected by a change in engine speed are

Figure 3.24: Influence on the output of a single fueling input from cylinder-3
those that have cells that reach the oxygen sensor both before and after the change in engine speed. Due to the total plant delay, the affected inputs are those commanded several events prior to the change in engine speed. This is readily apparent in the second experiment shown in Figure 3.26. In this case, mass flow rate is fixed at 120 mg per event while the engine speed is increased from 800 to 2000 RPM at event 40 (indicated by a dashed line). At event 80 (the second dashed line), the engine speed is decreased back down to 800 RPM. When the engine speed is increased, the inputs that contribute to the output at the event prior to this change will continue to the affect the output longer. These same inputs contribute to fewer output events when the engine speed decreases.

Figure 3.25: Summation of the input and output contributions during step changes in mass flow rate with a fixed engine speed
3.5 Summary

By applying a plug flow approximation, oxygen transport and mixing dynamics can be incorporated into an air path model. Three different models based on this approach have been developed: a lumped oxygen dynamics model, a high fidelity exhaust oxygen dynamics model and a finite impulse response exhaust oxygen dynamics model. A lumped oxygen dynamics model can be used to predict the oxygen concentration at any point in the air path system including inside the engine cylinders. Compared to a traditional filling and emptying oxygen concentration model, a lumped oxygen dynamics model is able to more effectively predict the in-cylinder oxygen concentration before and after combustion especially during transients. Due
to its computational complexity, a lumped oxygen dynamics model cannot be directly implemented in a production ECU. Section 4.1 is devoted to simplifying this model so that a real time in-cylinder oxygen concentration estimator can be designed.

By modeling each of the exhaust runners separately and accounting for the firing order, a high fidelity exhaust oxygen dynamics model captures how the transport delays of the exhaust system vary from cylinder to cylinder with remarkable accuracy. This model is able to predict the input/output relationship between injecting fueling and measuring the exhaust oxygen concentration exceptionally well as seen by the impulse and step response comparison results. This relationship also characterizes the plant dynamics which for a gasoline engine are enclosed by the feedback fueling control loop. Although a high fidelity exhaust oxygen dynamics model is not integrated into the switching frequency based gasoline fueling control presented in Chapter 6, a high fidelity exhaust oxygen dynamics model is used to calibrate the switching frequency control gains in simulation.

The relationship between injecting fuel and measuring the oxygen concentration of the exhaust stream is also key to identifying cylinder imbalances. A FIR exhaust oxygen dynamics model directly predicts how each of the past input fuel injection events affect a given output exhaust oxygen concentration measurement. With this model it is also possible to identify how a given fueling event will influence future exhaust oxygen concentration measurements. Both of these capabilities are exploited in Section 4.2 to produce a model based cylinder imbalance estimator.
CHAPTER 4

Model Based Estimation Applications

To effectively control an internal combustion engine, many time varying parameters must be dynamically estimated in real time. Two such parameters are the in-cylinder oxygen concentration and the relative cylinder imbalance of each cylinder. For all types of combustion, the in-cylinder oxygen concentration strongly impacts the combustion process, especially the heat release rate. As the in-cylinder oxygen concentration changes, the combustion strategy must be adapted to ensure quality combustion. In a spark ignited gasoline engine, this is achieved by adjusting the spark timing. In a diesel engine, combustion performance is maintained primarily by changing the fuel injection timings. Maximizing the combustion performance of an engine also requires that each cylinder behaves identically. To achieve uniform performance across each of the cylinders, imbalance errors must be estimated and then corrected.

Predicting the in-cylinder oxygen concentration and the imbalance between cylinders is very difficult because neither quantity can be directly measured in a production engine. For situations such as these when a desired variable cannot be directly measured, model based estimation techniques are particularly useful. The first section
in this chapter develops a closed loop estimator of the in-cylinder oxygen concentration within a diesel engine based on a simplified version of the lumped oxygen dynamics model developed in Section 3.2. To arrive at the final estimator design, several different approaches including the traditional filling and emptying approach were compared in simulation. The final model based design is able to robustly estimate the in-cylinder oxygen concentration thereby enabling an in-cylinder oxygen concentration based diesel fueling controller which is able to better control the diesel combustion process. The details of this diesel fueling controller are provided in Chapter 5. In the second section within this chapter, the design of a cylinder imbalance estimator based on the FIR exhaust oxygen dynamics model derived in Section 3.4 is presented. The performance of this design is validated on a gasoline engine.

4.1 Real Time In-cylinder Oxygen Concentration Estimation

In the open literature, most air path controllers and estimators are not directly based on the nonlinear filling and emptying type models described in Section 2.4.1, but rather on simplified versions of such models. High order filling and emptying models are simplified through a variety of techniques including linearization and/or model order reduction, often by making an isothermal assumption and replacing certain estimated quantities with measurements [11–13, 61, 63, 65–67, 69, 78]. The most common states/quantities which are replaced are the manifold pressure, fresh air mass flow rate and the exhaust oxygen concentration. Although the accuracy is often reduced significantly, a simplified air path model is able to be executed in real
time and is more conducive to observer design. The modeling trade-off between accuracy, simplicity and ease of observer design often significantly limits the achievable performance of a given estimation approach.

The following subsections develop and compare improved techniques for estimating in-cylinder oxygen concentration in real time. Unlike current designs, each of the proposed estimators are based on models which account for the transport delays of the air path system. More specifically the high order lumped oxygen dynamics model described in Section 3.2 is simplified to facilitate real time execution. Because the original truth model of the air path system is more accurate than a traditional filling and emptying model, the modeling trade-off between accuracy, simplicity and ease of observer design for a reduce order model based on the lumped oxygen dynamics model is more favorable than for a reduced order model based on a filling and emptying model.

The first estimator design is a simplified version of the lumped oxygen dynamics model called a delay based oxygen dynamics model which is run in an open loop manner. Although this reduced order model is simple enough to run in real time, the form in which it is realized makes estimator design extremely difficult. When this model is further simplified (at the expense of its open loop accuracy), the system can be represented in a form in which closed loop estimator design is feasible. The estimators designed based on this model (a LPV state estimator and an extended Kalman filter) are more robust to measurement uncertainty than an open loop oxygen dynamics model.

To combine the strengths of each estimator, a hybrid feed-forward/feedback estimation structure which incorporates both an open loop delay based oxygen dynamics
model and a closed loop state estimator is designed. The resulting design achieves a better trade-off between accuracy and robustness that any single estimator could. As will be demonstrated in Section 4.1.5, this hybrid feed-forward/feedback estimation approach is able to accurately predict the in-cylinder oxygen concentration during both steady-state and transient operation even in the presence of input and output disturbances.

4.1.1 Delay Based Oxygen Dynamics Model

The plug flow based air path model described in Section 3.2 is too complex to be run in real time; however, by separately and explicitly calculating the transport delays of the system it is possible to achieve approximately the same level of accuracy with a reduced order model. In steady-state, all of the plug flow calculations can be simplified thereby allowing the transport delay to be directly predicted with a single expression. Using this relationship, the nominal transport delays of the exhaust, EGR and intake systems are predicted. Depending on the nominal transport delay, different oxygen concentration values within a set of two time history buffers are used. With this formulation, the effects of the time-varying transport delays are accounted for by filtering the oxygen concentration obtained from the buffer. Although this reduced order approach is not strictly based on physics, it produces very good oxygen concentration estimates with a minimal number of calculations and dynamic states.
The nominal transport delays of the exhaust system \((d_{\text{exh}})\), the EGR system \((d_{\text{egr}})\) and intake system \((d_{\text{int}})\) can be estimated using
\[
d_{\text{exh}}(k) = \frac{V_{\text{exh},cv} p_{\text{exh},cv}(k)}{\dot{m}^*_{\text{exh}}(k) T_{\text{exh},cv}(k) R},
\]
\[
d_{\text{egr}}(k) = \frac{V_{\text{egr},cv} p_{\text{egr},cv}(k)}{\dot{m}^*_{\text{egr}}(k) T_{\text{egr},cv}(k) R},
\]
\[
d_{\text{int}}(k) = \frac{V_{\text{int},cv} p_{\text{int},cv}(k)}{(\dot{m}^*_{\text{egr}}(k) + \dot{m}^*_{\text{fresh}}(k)) T_{\text{int},cv}(k) R}.
\]
The memory buffers contain a fixed number of discrete values, so the nominal transport delays must be converted to an integer and bounded from above before they can be used. With the following three equations, the nominal exhaust, EGR and intake delays are converted into integer delays:
\[
d_{\text{exh}}(k) = \begin{cases} \lfloor d_{\text{exh}}(k) \rfloor & \text{if } d_{\text{exh}}(k) < d_{\text{exh},\text{max}} \\ d_{\text{exh},\text{max}} & \text{otherwise} \end{cases},
\]
\[
d_{\text{egr}}(k) = \begin{cases} \lfloor d_{\text{egr}}(k) \rfloor & \text{if } d_{\text{egr}}(k) < d_{\text{egr},\text{max}} \\ d_{\text{egr},\text{max}} & \text{otherwise} \end{cases}
\]
and
\[
d_{\text{int}}(k) = \begin{cases} \lfloor d_{\text{int}}(k) \rfloor & \text{if } d_{\text{int}}(k) < d_{\text{int},\text{max}} \\ d_{\text{int},\text{max}} & \text{otherwise} \end{cases}
\]
where \(d_{\text{exh},\text{max}}\) is the largest allowable exhaust system transport delay, \(d_{\text{egr},\text{max}}\) is the largest allowable EGR system transport delay and \(d_{\text{int},\text{max}}\) is the largest allowable intake system transport delay.

During each event new exhaust gas enters the EGR system. The resulting oxygen concentration of the EGR system depends on the oxygen concentration of the EGR system before the exhaust gas enters the system, and on the oxygen concentration of the exhaust gas that enters the EGR system. Because of transport delay, the oxygen concentration of the gases entering the system depends on the oxygen concentration of the exhaust gas that was expelled several events prior. Accordingly, the nominal
oxygen concentration at the confluence function where the EGR system meets the intake system, \([O_2]_{egr}^{\dagger}\), is approximated with

\[
[O_2]_{egr}^{\dagger}(k) = [O_2]_{exh}^{\dagger}(k - d_{exh}(k) - d_{egr}(k)) \frac{\dot{m}_{egr}(k)}{\epsilon \dot{m}_{egr,max}}
+ [O_2]_{egr}^{\dagger}(k - 1) \frac{\epsilon \dot{m}_{egr,max} - \dot{m}_{egr}(k)}{\epsilon \dot{m}_{egr,max}}
\]  

(4.7)

where \([O_2]_{exh}^{\dagger}\) is the nominal oxygen concentration at the exhaust ports, \(\dot{m}_{egr,max}\) is the maximum possible mass flow rate of EGR and \(\epsilon > 1\) captures the mixing and diffusion phenomena within the EGR system. In the same manner, the oxygen concentration of the lumped intake system depends on the oxygen concentrations of the EGR system and the ambient air time-shifted to account for the transport delay of the intake system. Using their respective mass flow rates as weighting factors, the nominal oxygen concentration at the intake ports, \([O_2]_{int}^{\dagger}\), is modeled according to

\[
[O_2]_{int}^{\dagger}(k) = \frac{[O_2]_{egr}^{\dagger}(k - d_{int}(k)) \dot{m}_{egr}(k - d_{int}(k)) + [O_2]_{amb} \dot{m}_{fresh}(k - d_{int}(k))}{\dot{m}_{egr}(k - d_{int}(k)) + \dot{m}_{fresh}(k - d_{int}(k))}.
\]  

(4.8)

The nominal in-cylinder oxygen concentration depends on the oxygen concentration of the air charge entering the cylinders and the residual gas left in the cylinders. Using the trapped residual fraction defined in (3.60), the total trapped mass can be related to the charge flow rate with

\[
m_{trapped}^*(k) = \frac{\dot{m}_{egr}^*(k - d_{int}(k)) + \dot{m}_{fresh}^*(k - d_{int}(k))}{1 - \gamma_{res}(k)}.
\]  

(4.9)

By assuming the trapped residual gases are equivalent to the gases that were just exhausted, the nominal in-cylinder oxygen concentration, \([O_2]_{cyl}^{\dagger}\), is approximated by

\[
[O_2]_{cyl}^{\dagger}(k) = (1 - \gamma_{res}(k))[O_2]_{int}^{\dagger}(k) + [O_2]_{exh}^{\dagger}(k)\gamma_{res}(k).
\]  

(4.10)
Define $d_{int,exh}$ as the fixed delay between the induction and exhaust strokes. After $d_{int,exh}$ events, the inducted cylinder contents are burned and expelled into the exhaust system; therefore, the oxygen concentration of the lumped exhaust system is governed by

$$[O_2]_{exh}^{\dagger}(k + d_{int,exh}) = \frac{[O_2]_{cyl}^{\dagger}(k)m_{trapped}^*(k) - AFR_a[O_2]_{air}m_{fuel}^*(k)}{m_{trapped}^*(k) + m_{fuel}^*(k)}.$$ 

(4.11)

To account for the mixing effects, the nominal in-cylinder oxygen concentration is filtered using

$$[O_2]_{cyl}(k) = \nu [O_2]_{cyl}^{\dagger}(k) + (1 - \nu)[O_2]_{cyl}(k - 1)$$ 

(4.12)

where $[O_2]_{cyl}$ is the final in-cylinder oxygen concentration estimate and $\nu$ is the filter constant.

As long as all of the temperatures, pressures and flow rate inputs can be well estimated/measured, however, a delay based oxygen dynamics model can be used to predict the in-cylinder oxygen concentration in a open loop manner. Although a post-turbocharger UEGO sensor could potentially enable some form of feedback correction, the variable time delay nature of this model makes designing a closed loop estimator extremely difficult. With a few additional assumptions, a delay based oxygen dynamics model can be realized in a more useful form for state estimator design. The model simplification and estimator design processes are detailed in the next subsection.

### 4.1.2 In-cylinder Oxygen Concentration Prediction via Linear Parameter Varying State Estimation

The most common approach to estimating the oxygen concentration within an engine (or equivalently the intake manifold if perfect mixing is assumed) is to design
a state observer based a simple low order system model. These models can be based on a reduced order filling and emptying model of the air path system [61,63,65,67,78–80] or they can be based on an empirical model which assume a desired linear structure [6,7,79–82]. The order of these systems is either second [67,78], third [61,65,80] or fourth [63]. Since their higher order counterparts do not account for the dynamic transport delays with in the air path, neither do these reduced order models thus significantly limiting their accuracy. In many cases these reduced order models must be gain scheduled on the operating conditions to achieve reasonable accuracy. By necessity, the state observers based on these models are also gain scheduled. Because the physical plant is not modeled particularly well with these simple models, the resulting state estimator must not only be asymptotically stable but also robust to a wide range of uncertainties to be practical. Such robustness properties are often lacking or at least not explicitly proven in current estimator designs. More specifically, most stability proofs assume fixed operating conditions and perfect knowledge all of the time varying inputs/parameters.

A delay based oxygen dynamics model performs well as an open loop in-cylinder oxygen concentration estimator in ideal situations where all of the mass flow rate, temperature and pressure inputs are exactly known. All of these inputs are measured or estimated in production engines, but not perfectly. Measurement and estimation errors propagate through the delay based oxygen dynamics model and into the output in-cylinder oxygen concentration estimate. Although the model remains stable, small errors appear in the estimated in-cylinder oxygen concentration. In its current form, additional sensor measurements cannot be easily incorporated back into a delay based oxygen dynamics model to close the estimation loop.
A delay based oxygen dynamic model can be alternatively represented as a time dependent LPV state space model. The order of this model is equal to \( d_{\text{exh,max}} + d_{\text{egr,max}} + d_{\text{int,max}} \). For the six cylinder diesel engine studied in this dissertation, the largest combined transport delay is 40 events, so in state space form the delay based oxygen dynamics model of this engine has a total of 40 states. Designing and implementing an LPV observer for a time varying system of this order would be impractical. By changing the time scale and applying some simplifying assumptions, the system can be reduced to a more manageable LPV state space model.

Removing the intermediate variables, the original delay based model can be represented with the following three dynamic equations:

\[
[O_2]_{\text{exh}}(k) = \alpha_1(k - d_{\text{int,exh}})[O_2]_{\text{cyl}}(k - d_{\text{int,exh}}) - \left(1 - \alpha_1(k - d_{\text{int,exh}})\right)AFR_s[O_2]_{\text{air}}, \tag{4.13}
\]

\[
[O_2]_{\text{egr}}(k) = \alpha_0(k)[O_2]_{\text{exh}}(k - d_{\text{exh}}(k) - d_{\text{egr}}(k)) + (1 - \alpha_0(k))[O_2]_{\text{egr}}(k - 1), \tag{4.14}
\]

and

\[
[O_2]_{\text{cyl}}(k) = \left(1 - \gamma_{\text{res}}(k)\right)\alpha_2(k - d_{\text{int}}(k))[O_2]_{\text{egr}}(k - d_{\text{int}}(k)) + \gamma_{\text{res}}(k)[O_2]_{\text{exh}}(k) + \left(1 - \alpha_2(k - d_{\text{int}}(k))\right)(1 - \gamma_{\text{res}})[O_2]_{\text{amb}}, \tag{4.15}
\]

where

\[
\alpha_0(k) = \frac{\dot{m}_{\text{egr}}(k)}{\dot{m}_{\text{egr}}(k) + \epsilon \dot{m}_{\text{egr,max}}}, \tag{4.16}
\]

\[
\alpha_1(k) = \frac{m_{\text{trapped}}(k)}{m_{\text{trapped}}(k) + m_{\text{fuel}}(k)}. \tag{4.17}
\]
and

$$\alpha_2(k) = \frac{\dot{m}_{egr}(k)}{\dot{m}_{egr}(k) + \dot{m}_{fresh}(k)}.$$  \hspace{1cm} (4.18)

If a larger time step corresponding to an engine cycle rather than an engine event is used and the transport delays for the exhaust, EGR and intake systems in terms of engine cycles are assumed to be constant, then these equations can be simplified even further. For the six cylinder diesel engine studied in this dissertation, the average combined delay of the exhaust and EGR system is approximately two engine cycles whereas the average transport delay of the intake system is approximately one engine cycle. The transport delay between the exhaust ports and the location of the post-turbocharger oxygen sensor \(d_{exh, eugo}\) is also approximately one engine cycle. The average fixed delay between the induction and exhaust strokes was found to be best approximated by a single engine cycle. Additionally, the trapped residual fraction was assumed to be a constant value of 0.089.

To account for the timescale change, the time constant of the EGR mixing model had to be modified. With the slower timescale, the mixing effects within the EGR system were found to be more uniform. Without significantly affecting the oxygen concentration predictions, the mixing effects can be approximated by a filter with a fixed time constant \(\zeta\). Through numerical simulations, the best constant value of \(\zeta\) for the engine in this study was found to be 0.67. To further simplify these equations, the nominal in-cylinder oxygen concentration \([O_2]_{cyl}^{\dagger}\) can be assumed to be the actual in-cylinder oxygen concentration \([O_2]_{cyl}\).

With all of the assumptions summarized in Table 4.1, the air path system can be modeled as a fourth order LPV system. Defining \(\kappa\) as the engine cycle index, the four
Table 4.1: Assumptions used to derive the LPV air path model

<table>
<thead>
<tr>
<th>Assumption</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_{int,exh}$</td>
<td>1 engine cycle</td>
</tr>
<tr>
<td>$d_{egr} + d_{exh}$</td>
<td>2 engine cycles</td>
</tr>
<tr>
<td>$d_{int}$</td>
<td>1 engine cycle</td>
</tr>
<tr>
<td>$d_{exh,uego}$</td>
<td>1 engine cycle</td>
</tr>
<tr>
<td>$\gamma_{res}$</td>
<td>0.089</td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>$\zeta = 0.67$</td>
</tr>
<tr>
<td>$[O_2]_{cyl}$</td>
<td>$[O_2]_{cyl}$</td>
</tr>
</tbody>
</table>

State variables correspond to

$$x_1(k+1) = \frac{1}{6} \sum_{i=1}^{6} [O_2]_{exh}^\dagger (6 k + i),$$  \hspace{1cm} (4.19)$$

$$x_2(k+1) = \frac{1}{6} \sum_{i=1}^{6} [O_2]_{exh}^\dagger (6 k - 6 + i),$$  \hspace{1cm} (4.20)$$

$$x_3(k+1) = \frac{1}{6} \sum_{i=1}^{6} [O_2]_{egr}^\dagger (6 k + i),$$  \hspace{1cm} (4.21)$$

$$x_4(k+1) = \frac{1}{6} \sum_{i=1}^{6} [O_2]_{cyl}^\dagger (6 k + i).$$  \hspace{1cm} (4.22)$$

The output of this system $y$ is the post-turbocharger exhaust oxygen concentration measurement. Define $w$ as the difference between the cycle averaged post-turbocharger oxygen concentration measurement and the cycle averaged true oxygen concentration. Under these conditions, the system can be represented with the following state space representation

$$x(k+1) = \left( A_0 + \alpha_1(k)A_1 + \alpha_2(k)A_2 \right) x(k) + \left( B_0 + \alpha_1(k)B_1 + \alpha_2(k)B_2 \right) u_{LPV}(k)$$

$$y(k) = C x(k) + w(k),$$  \hspace{1cm} (4.23)$$

where
Although both of the $\alpha_1$ and $\alpha_2$ parameters can be derived from measurements, these measurements contains some degree of uncertainty. Using feedback from an oxygen sensor located in the exhaust stream, it is possible to develop a state estimator for this system. For this estimator to be practical, however, it must be robust to uncertainties in the $\alpha_1$ and $\alpha_2$ parameters as well as in the presence of output measurement errors $w$. The robust stability of a similar class of system has been previously studied in [83–85]. The following theorem and subsequent proof are based on the arguments presented in those works but have been adapted for a system of the form represented by (4.23).

**Theorem 4.1.1** An observer of the form

\[
\hat{x}(k+1) = \left( A_0 + \alpha_1(k)A_1 + \alpha_2(k)A_2 \right) \hat{x}(k) + \left( B_0 + \alpha_1(k)B_1 + \alpha_2(k)B_2 \right) u_{LPV}(k)
+ \left( L_0 + \alpha_1(k)L_1 + \alpha_2(k)L_2 \right) \left( y(k) - \hat{y}(k) \right)
\]

\[
\hat{y}(k) = C\hat{x}(k)
\]

is an input to state stable (ISS) estimator of system (4.23) provided that there exists symmetric positive definite matrices $S_0, S_1, S_2 \in \mathbb{R}^{4 \times 4}$, square matrices $G_0, G_1, G_2 \in \mathbb{R}^{4 \times 4}$,
\[ \mathbb{R}^{4 \times 4}, \text{ row matrices } F_0, F_1, F_2 \in \mathbb{R}^{1 \times 4} \text{ and a constant } \varpi > 0 \text{ that satisfy} \]
\[ \begin{bmatrix}
G_i + G_i^T - S_j & G_i^T A_i - F_i^T C \\
A_i^T G_i - C^T F_i & \frac{1}{\varpi} S_i
\end{bmatrix} > 0 \quad (4.25) \]
\[ \forall i, j \in \{0, 1, 2\} \text{ and the observer gains are selected according to } L_0 = (G_0^T)^{-1} F_0^T, \]
\[ L_1 = (G_1^T)^{-1} F_1^T \text{ and } L_2 = (G_2^T)^{-1} F_2^T. \]

Proof: If block matrix shown in (4.25) is strictly positive definite, then the sub-matrices on the major diagonal must also be strictly positive definite. This means that
\[ G_i + G_i^T - S_j > 0 \quad (4.26) \]
and \( G_i \) must therefore be full rank. By construction \( S_j \) is strictly positive definite which also implies that \( S_j^{-1} \) is also strictly positive definite and there exists a full rank matrix \( W_j \in \mathbb{R}^{4 \times 4} \) that satisfies
\[ S_j^{-1} = W_j^T W_j > 0. \quad (4.27) \]

A matrix \( Z_{\text{psd}} \in \mathbb{R}^{n \times n} \) is said to be positive semidefinite if there exists a matrix \( V \in \mathbb{R}^{n \times n} \) such that \( Z_{\text{psd}} = V^T V \). By this definition we can observe that
\[ \left[ (S_j - G_i)^T W_j^T \right] W_j (S_j - G_i) \geq 0. \quad (4.28) \]
Substituting (4.27) into the above results in
\[ (S_j - G_i)^T S_j^{-1} (S_j - G_i) \geq 0, \quad (4.29) \]
which simplifies to
\[ S_j - G_i - G_i^T + G_i^T S_j^{-1} G_i \geq 0, \quad (4.30) \]
or equivalently
\[ G_i^T S_j^{-1} G_i \geq G_i + G_i^T - S_j. \quad (4.31) \]
This inequality can be combined with the original linear matrix inequality (LMI) (4.25) to generate

\[
\begin{bmatrix}
G_i^T S_j^{-1} G_i & G_i^T A_i - F_i^T C \\
A_i^T G_i - C^T F_i & \frac{1}{9 + \varpi} S_i
\end{bmatrix} > 0.
\] (4.32)

This matrix can be factored according to

\[
\begin{bmatrix}
G_i^T & 0 \\
0 & S_j
\end{bmatrix} M_1 \begin{bmatrix}
G_i & 0 \\
0 & S_j
\end{bmatrix} > 0,
\] (4.33)

where

\[
M_1 = \begin{bmatrix}
S_j^{-1} A_i^T - S_j^{-1} C^T F_i G_i^{-1} & A_i S_j^{-1} - (G_i^T)^{-1} F_i^T C S_j^{-1} \\
S_j^{-1} A_i^T - S_j^{-1} C^T F_i G_i^{-1} & \frac{1}{9 + \varpi} S_j^{-1} S_i S_j^{-1}
\end{bmatrix}.
\]

Because \(G_i\) is full rank the two block matrices multiplying \(M_1\) from the left and from the right are full rank and transposes of each other. If a matrix \(Z_{pd} \in \mathbb{R}^{n \times n}\) is strictly positive definite then for any full rank matrix \(W \in \mathbb{R}^{n \times n}\), \(W^T Z_{pd} W > 0\). Making use of this fact, the LMI in (4.33) can be simplified to

\[
\begin{bmatrix}
G_i^T & 0 \\
0 & S_j
\end{bmatrix}^{-1} \begin{bmatrix}
G_i^T & 0 \\
0 & S_j
\end{bmatrix} M_1 \begin{bmatrix}
G_i & 0 \\
0 & S_j
\end{bmatrix}^{-1} > 0
\]

\[
M_1 > 0.
\] (4.34)

Recalling that \(L_0 = (G_0^T)^{-1} F_0^T\), \(L_1 = (G_1^T)^{-1} F_1^T\) and \(L_2 = (G_2^T)^{-1} F_2^T\), this inequality can be rewritten as

\[
\begin{bmatrix}
S_j^{-1} \\
S_j^{-1} [A_i - L_i C]^T
\end{bmatrix} \begin{bmatrix}
A_i - L_i C \\
\frac{1}{9 + \varpi} S_j^{-1} S_i S_j^{-1}
\end{bmatrix} > 0.
\] (4.35)

Defining \(\overline{A}_i\) as

\[
\overline{A}_i = A_i - L_i C,
\] (4.36)
it follows that

\[
\begin{bmatrix}
S_j & 0 \\
0 & S_j
\end{bmatrix}
\begin{bmatrix}
S_j^{-1} \\
S_j^{-1} [A_i - L_i C]^T [A_i - L_i C] S_j^{-1}
\end{bmatrix}
\begin{bmatrix}
\frac{1}{9 + \infty} S_j^{-1} S_j \\
S_j 0
\end{bmatrix}
> 0
\]

\[
\begin{bmatrix}
S_j & S_j A_i \\
A_i^T S_j & \frac{1}{9 + \infty} S_i
\end{bmatrix}
> 0
\]

(4.37)

\( \forall i \in \{0, 1, 2\} \) and \( \forall j \in \{0, 1, 2\} \). For each \( i \), the corresponding \( j \) inequalities can be multiplied by 1, \( \hat{\alpha}_1(\bar{k} + 1) \) and \( \hat{\alpha}_2(\bar{k} + 1) \), respectively, and then summed to provide

\[
\begin{bmatrix}
S_0 + \hat{\alpha}_1 S_1 + \hat{\alpha}_2 S_2 \\
A_i^T \left( S_0 + \hat{\alpha}_1 S_1 + \hat{\alpha}_2 S_2 \right) \frac{1}{9 + \infty} \left( 1 + \hat{\alpha}_1 + \hat{\alpha}_2 S_i \right)
\end{bmatrix}
> 0
\]

(4.38)

where \( \hat{\alpha}_1 = \hat{\alpha}_1(\bar{k} + 1) \) and \( \hat{\alpha}_2 = \hat{\alpha}_2(\bar{k} + 1) \).

Define \( \mathcal{P}(\hat{\alpha}_1(\bar{k}), \hat{\alpha}_2(\bar{k})) \) as

\[
\mathcal{P}(\hat{\alpha}_1(\bar{k}), \hat{\alpha}_2(\bar{k})) = S_0 + \hat{\alpha}_1(\bar{k} + 1) S_1 + \hat{\alpha}_2(\bar{k} + 1) S_2.
\]

Because \( \mathcal{P}(\hat{\alpha}_1(\bar{k}), \hat{\alpha}_2(\bar{k})) \) is part of the major diagonal, it must also be positive definite. In terms of \( \mathcal{P}(\hat{\alpha}_1(\bar{k}), \hat{\alpha}_2(\bar{k})) \), (4.38) becomes

\[
\begin{bmatrix}
\mathcal{P}(\hat{\alpha}_1(\bar{k} + 1), \hat{\alpha}_2(\bar{k} + 1)) \\
A_i^T \mathcal{P}(\hat{\alpha}_1(\bar{k} + 1), \hat{\alpha}_2(\bar{k} + 1)) \\
\left( \mathcal{P}(\hat{\alpha}_1(\bar{k} + 1), \hat{\alpha}_2(\bar{k} + 1)) \right) \frac{1}{9 + \infty} \left( 1 + \hat{\alpha}_1(\bar{k} + 1) + \hat{\alpha}_2(\bar{k} + 1) \right) S_i
\end{bmatrix}
> 0
\]

(4.40)

\( \forall i \in \{0, 1, 2\} \). Multiplying each of these \( i \) inequalities with \( \hat{\alpha}_i(\bar{k}) \) and summing generates

\[
\begin{bmatrix}
\theta(\bar{k}) \mathcal{P}_2 \\
A^T \mathcal{P}_2 \\
\frac{\mathcal{P}_2 A}{9 + \infty} \theta(\bar{k} + 1) \mathcal{P}_1
\end{bmatrix}
> 0
\]

(4.41)

where \( \theta(\bar{k}) = 1 + \hat{\alpha}_1(\bar{k}) + \hat{\alpha}_2(\bar{k}) \), \( \mathcal{P}_1 = \mathcal{P}(\hat{\alpha}_1(\bar{k}), \hat{\alpha}_2(\bar{k})) \), \( \mathcal{P}_2 = \mathcal{P}(\hat{\alpha}_1(\bar{k} + 1), \hat{\alpha}_2(\bar{k} + 1)) \) and

\[
\mathcal{A} = \mathcal{A}(\hat{\alpha}_1(\bar{k}), \hat{\alpha}_2(\bar{k})) = (A_0 - L_0 C) + \hat{\alpha}_1(\bar{k})(A_1 - L_1 C) + \hat{\alpha}_2(\bar{k})(A_2 - L_2 C)
\]

\[
= \bar{A}_0 + \hat{\alpha}_1(\bar{k}) \bar{A}_1 + \hat{\alpha}_2(\bar{k}) \bar{A}_2.
\]

(4.42)
We can multiply both sides of this positive definite matrix with any full rank matrix to generate another positive definite matrix as in

\[
\begin{bmatrix}
I & 0 \\
-A^T & \theta(k) I
\end{bmatrix}
\begin{bmatrix}
\theta(k) P_2 & P_2 A \\
A^T P_2 & \frac{1}{9+\bar{\omega}} \theta(k+1) P_1
\end{bmatrix}
\begin{bmatrix}
I & -A \\
0 & \theta(k) I
\end{bmatrix} > 0,
\]

\[
\begin{bmatrix}
\theta(k) P_2 & P_2 A \\
0 & \frac{1}{9+\bar{\omega}} \theta(k) \theta(k+1) P_1 - A^T P_2 A
\end{bmatrix}
\begin{bmatrix}
I & -A \\
0 & \theta(k) I
\end{bmatrix} > 0,
\]

\[
\theta(k)
\begin{bmatrix}
P_2 & 0 \\
0 & \frac{1}{9+\bar{\omega}} \theta(k) \theta(k+1) P_1 - A^T P_2 A
\end{bmatrix} > 0.
\]

(4.43)

For this LMI to hold, the following inequality must necessarily be true

\[
\frac{\theta(k) \theta(k+1)}{9+\bar{\omega}} P_1 - A^T P_2 A > 0.
\]

(4.44)

Because \(\hat{\alpha}_1, \hat{\alpha}_2 \in (0, 1]\),

\[
1 \leq \theta(k) \theta(k+1) \leq 9
\]

(4.45)

and

\[
0 < \frac{\bar{\omega}}{9+\bar{\omega}} \leq 1 - \frac{\theta(k) \theta(k+1)}{9+\bar{\omega}} < 1.
\]

(4.46)

must hold. The positive definite matrix \(P_1\) scaled by \(1 - \frac{\theta(k) \theta(k+1)}{9+\bar{\omega}}\) remains positive definite. Adding \(1 - \frac{\theta(k) \theta(k+1)}{9+\bar{\omega}}\) \(P_1\) to each side of the inequality in (4.44) and simplifying generates

\[
P_1 - A^T P_2 A \geq \left(1 - \frac{\theta(k) \theta(k+1)}{9+\bar{\omega}}\right) P_1 \geq \frac{\bar{\omega}}{9+\bar{\omega}} P_1 > 0
\]

(4.47)

and in turn

\[
A^T P_2 A - P_1 \leq -\frac{\bar{\omega}}{9+\bar{\omega}} P_1 < 0.
\]

(4.48)

This inequality will be the basis of the following Lyapunov-based stability argument.
The dynamics of interest are the observer error dynamics. For an observer which satisfies the conditions described in Theorem 4.1.1, the state estimation error $\hat{x}(\bar{k}) = x(\bar{k}) - \hat{x}(\bar{k})$ is governed by

$$\begin{align*}
\hat{x}(\bar{k} + 1) &= \left[ (A_0 - L_0C) + \hat{\alpha}_1(\bar{k})(A_1 - L_1C) + \hat{\alpha}_2(\bar{k})(A_2 - L_2C) \right] \hat{x}(\bar{k}) \\
&\quad + \left[ \left( \alpha_1(\bar{k}) - \hat{\alpha}_1(\bar{k}) \right) A_1 + \left( \alpha_2(\bar{k}) - \hat{\alpha}_2(\bar{k}) \right) A_2 \right] x(\bar{k}) \\
&\quad + \left[ \left( \alpha_1(\bar{k}) - \hat{\alpha}_1(\bar{k}) \right) B_1 + \left( \alpha_2(\bar{k}) - \hat{\alpha}_2(\bar{k}) \right) B_2 \right] u_{LPV}(\bar{k}) \\
&\quad - \left[ L_0 + \hat{\alpha}_1(\bar{k}) L_1 + \hat{\alpha}_2(\bar{k}) L_2 \right] w(\bar{k}) \tag{4.49}
\end{align*}$$

where

$$v_d(\bar{k}) = \left[ \left( \alpha_1(\bar{k}) - \hat{\alpha}_1(\bar{k}) \right) A_1 + \left( \alpha_2(\bar{k}) - \hat{\alpha}_2(\bar{k}) \right) A_2 \right] x(\bar{k}) \\
+ \left[ \left( \alpha_1(\bar{k}) - \hat{\alpha}_1(\bar{k}) \right) B_1 + \left( \alpha_2(\bar{k}) - \hat{\alpha}_2(\bar{k}) \right) B_2 \right] u_{LPV}(\bar{k}) \\
- \left[ L_0 + \hat{\alpha}_1(\bar{k}) L_1 + \hat{\alpha}_2(\bar{k}) L_2 \right] w(\bar{k}). \tag{4.50}$$

Using the positive definite matrix defined in (4.39), we can define the following parameter dependent candidate ISS-Lyapunov function

$$V(\hat{\alpha}_1(\bar{k}), \hat{\alpha}_2(\bar{k}), \hat{x}(\bar{k})) = \hat{x}(\bar{k})^T P(\hat{\alpha}_1(\bar{k}), \hat{\alpha}_2(\bar{k})) \hat{x}(\bar{k}). \tag{4.51}$$

This ISS-Lyapunov function is bounded for $\forall \bar{k}$ by

$$\gamma_1 ||\hat{x}(\bar{k})||^2 \leq \hat{x}(\bar{k})^T P(\hat{\alpha}_1(\bar{k}), \hat{\alpha}_2(\bar{k})) \hat{x}(\bar{k}) \leq 3\gamma_2 ||\hat{x}(\bar{k})||^2 \tag{4.52}$$

where $\gamma_1 = \lambda_{\min}(S_0)$ and $\gamma_2 = \max_{i\in\{0, 1, 2\}} \lambda_{\max}(S_i)$. If the alpha parameters were perfectly known ($v(\bar{k}) = 0$), then the ISS-Lyapunov function difference defined as

$$\Delta V = V(\hat{\alpha}_1(\bar{k} + 1), \hat{\alpha}_2(\bar{k} + 1), \hat{x}(\bar{k} + 1)) - V(\hat{\alpha}_1(\bar{k}), \hat{\alpha}_2(\bar{k}), \hat{x}(\bar{k})) \tag{4.53}$$
is given by
\[
\Delta V = \tilde{x}(k)^T \left[ A(\hat{\alpha}_1(k), \hat{\alpha}_2(k))^T P(\hat{\alpha}_1(k+1), \hat{\alpha}_2(k+1)) A(\hat{\alpha}_1(k), \hat{\alpha}_2(k)) - P(\hat{\alpha}_1(k), \hat{\alpha}_2(k)) \right] \tilde{x}(k)
\]
\[
= \tilde{x}(k)^T \left[ A^T P_2 A - P_1 \right] \tilde{x}(k).
\]
(4.54)

From (4.48) we know that this difference is strictly negative and therefore there exists a constant \( \gamma_3 \geq \gamma_1 \frac{\rho}{\nu+\rho} > 0 \) such that \( \forall k \)
\[
V(\hat{\alpha}_1(k+1), \hat{\alpha}_2(k+1), \tilde{x}(k+1)) - V(\hat{\alpha}_1(k), \hat{\alpha}_2(k), \tilde{x}(k)) \leq -\gamma_3 \| \tilde{x}(k) \|^2.
\]
(4.55)

It follows that for the ideal case when the alpha parameters are exactly known the state estimator is globally uniformly asymptotically stable.

With uncertain parameter estimates the observer error dynamics as shown in (4.50) contain an additional disturbance term \( v_d(k) \). For the uncertain system, the Lyapunov function difference is
\[
\Delta V = (A\tilde{x}(k) + v_d(k))^T P_2 \left( A\tilde{x}(k) + v_d(k) \right) - \tilde{x}(k)^T P_1 \tilde{x}(k)
\]
\[
= \tilde{x}(k)^T \left[ A^T P_2 A - P_1 \right] \tilde{x}(k) + v_d(k)^T P_2 v_d(k) + 2v_d(k)^T P_2 A\tilde{x}(k).
\]
(4.56)

This equation is bounded by
\[
\Delta V \leq -\gamma_3 \| \tilde{x}(k) \|^2 + 3 \gamma_2 \| v_d(k) \|^2 + 2 \| v(k) \| \| P_2 \| \| A \| \| \tilde{x}(k) \|.
\]
(4.57)

Consider the constant \( \gamma_4 \) which satisfies
\[
\gamma_4 = 9 \left( \max_{i \in \{0, 1, 2\}} \| S_i \| \right) \left( \max_{j \in \{0, 1, 2\}} \| A_j \| \right) \geq \| P_2 \| \| A \|.
\]
(4.58)
Using the well known inequality

\[ 2ab \leq \kappa a^2 + \frac{1}{\kappa}b^2 \]  

\( \forall \kappa > 0 \) and \( \forall (a, b) \in \mathbb{R}^2 \), the final term in (4.57) can be bounded by

\[ 2\|v(k)\| \|P_2\| \|A\| \|\tilde{x}(k)\| \leq \kappa \|\tilde{x}(k)\|^2 + \frac{\gamma_4^2}{\kappa} \|v_d(k)\|^2. \]  \( \tag{4.60} \)

Substituting back into (4.57) generates the following inequality:

\[ \Delta V \leq -(\gamma_3 - \kappa) \|\tilde{x}(k)\|^2 + \left(3 \gamma_2 + \frac{\gamma_4^2}{\kappa}\right) \|v_d(k)\|^2. \]  \( \tag{4.61} \)

When considering \( \kappa \in (0, \gamma_3) \) it can readily be observed that the Lyapunov function \( V(\hat{\alpha}_1(k), \hat{\alpha}_2(k), \tilde{x}(k)) \) is an input to state stable (ISS) Lyapunov function. Although omitted here, a class \( KL \) function \( \beta_{ISS} \) and class \( K \) function \( \gamma_{ISS} \) which bound the estimation error according to

\[ \|\tilde{x}(k)\| \leq \beta_{ISS}(\|\tilde{x}(k)\|,k) + \gamma_{ISS}(\|v(k)\|) \]  \( \tag{4.62} \)

can be found explicitly. The derivation of this result is provided in [85].

Theorem 4.1.1 is very practical because the LMI conditions described in (4.25) can be directly used to design an observer. An LMI solver can be used to find a set of matrices which satisfy these LMI conditions (if they can be satisfied) and these matrices can be directly converted into the observer gains. The resulting observer is guaranteed to be ISS stable. Moreover by fixing \( \varpi > 0 \) and by adding the constraint that \( \gamma_1 = \lambda_{\text{min}}(S_0) \) should be maximized, the convergence rate of the observer can be maximized.
4.1.3 In-cylinder Oxygen Concentration Prediction via Extended Kalman Filter Theory

A Kalman filter is a recursive state estimator which minimizes the mean of the estimation error squared. Originally the Kalman filter was designed specifically for linear systems, but it has since been extended for certain types of nonlinear systems. To extend a Kalman filter to a nonlinear system, the nonlinear plant model must be linearized at each time step. Once linearized, the predictor-corrector equations which define a Kalman filter can be used directly. Unlike the linear case, the stability of an extended Kalman filter is not guaranteed.

For comparison, an extended Kalman filter was also designed for the simplified four state LPV oxygen dynamics model. To construct an extended Kalman filter, the system can be rewritten as the following time-varying system:

\[
\begin{bmatrix}
    x_1(k+1) \\
    x_2(k+1) \\
    x_3(k+1) \\
    x_4(k+1)
\end{bmatrix}
= 
\begin{bmatrix}
    0 & 0 & 0 & \alpha_1(k) \\
    1 & 0 & 0 & 0 \\
    0 & \zeta & 1-\zeta & 0 \\
    \gamma_{res} & 0 & (1-\gamma_{res})\alpha_2(k) & 0
\end{bmatrix}
\begin{bmatrix}
    x_1(k) \\
    x_2(k) \\
    x_3(k) \\
    x_4(k)
\end{bmatrix}
+ 
\begin{bmatrix}
    -(1-\alpha_1(k))AFR_s & 0 \\
    0 & 0 \\
    0 & 0 \\
    0 & (1-\gamma_{res})(1-\alpha_2(k))
\end{bmatrix}
\begin{bmatrix}
    [O_2]_{air} \\
    [O_2]_{amb}
\end{bmatrix}
\]
\[
y(k) = \begin{bmatrix}
    0 & 1 & 0 & 0
\end{bmatrix} x(k) + w(k). \quad (4.63)
\]

When the difference between the estimated and actual alpha parameter \( \hat{\alpha}_i(k) = \alpha_i(k) - \alpha_i(k) \) is thought of as process noise, the system can be put into the following form which resembles the standard linear system form for which the Kalman filter was originally developed:

\[
x(k+1) = \Phi_{KF}(k)x(k) + \Gamma_{KF}(x(k))\hat{\alpha}(k) + \Psi_{KF}(k)u_{KF}(k)
\]
\[
y(k) = H_{KF}x(k) + w(k) \quad (4.64)
\]

122
where $\Phi_{KF}(k) = \begin{bmatrix} 0 & 0 & 0 & \hat{\alpha}_1(k) \\ 1 & 0 & 0 & 0 \\ 0 & \zeta & 1-\zeta & 0 \\ \gamma_{res} & 0 & (1-\gamma_{res})\hat{\alpha}_2(k) & 0 \end{bmatrix},$

$\Gamma_{KF}(x(k)) = \begin{bmatrix} \Gamma_{KF,1}(x(k)), & \Gamma_{KF,2}(x(k)) \end{bmatrix},$

$\Gamma_{KF,1}(x(k)) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} + \begin{bmatrix} AFR_s \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} [O_2]_{air} \\ [O_2]_{amb} \end{bmatrix},$

$\Gamma_{KF,2}(x(k)) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1-\gamma_{res} & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \gamma_{res} - 1 \end{bmatrix} \begin{bmatrix} [O_2]_{air} \\ [O_2]_{amb} \end{bmatrix},$

$\Psi_{KF}(k) = \begin{bmatrix} -(1-\hat{\alpha}_1(k))AFR_s & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & (1-\gamma_{res})(1-\hat{\alpha}_2(k)) \end{bmatrix},$

$H_{KF} = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}, \quad \hat{\alpha}(k) = [\hat{\alpha}_1(k), \hat{\alpha}_2(k)]^T$ and

$u_{KF}(k) = \begin{bmatrix} [O_2]_{air}, & [O_2]_{amb} \end{bmatrix}^T.$

This system does not conform perfectly to the traditional linear form, however, because the process noise matrix $\Gamma_{KF}$ depends on the unknown state vector.

The standard Kalman filter design can be easily extended to this system by replacing $\Gamma_{KF}(x(k))$ with $\Gamma_{KF}(\hat{x}(k))$ in the estimation equations. The resulting Kalman filter predictor and corrector equations [86] are provided below:

$$\hat{x}(k+1|k) = \Phi_{KF}(k)\hat{x}(k) + \Psi_{KF}(k)u_{KF}(k), \quad (4.65)$$

$$\tilde{y}(k+1|k) = y(k+1) - H_{KF}\hat{x}(k+1|k), \quad (4.66)$$
\[ P_{KF}(\overline{k}+1|\overline{k}) = \Phi_{KF}(\overline{k})P_{KF}(\overline{k}|\overline{k})\Phi_{KF}(\overline{k})^T + \Gamma_{KF}(\hat{x}(\overline{k}|\overline{k}))Q_{KF}\Gamma_{KF}(\hat{x}(\overline{k}|\overline{k}))^T, \quad (4.67) \]

\[ K_{KF}(\overline{k}+1) = P_{KF}(\overline{k}+1|\overline{k})H_{KF}^T(H_{KF}P_{KF}(\overline{k}+1|\overline{k})H_{KF}^T + R_{KF})^{-1}, \quad (4.68) \]

\[ P_{KF}(\overline{k}+1|\overline{k}+1) = \left( I - K_{KF}(\overline{k}+1)H_{KF} \right)P_{KF}(\overline{k}+1|\overline{k}), \quad (4.69) \]

\[ \hat{x}(\overline{k}+1|\overline{k}+1) = \hat{x}(\overline{k}+1|\overline{k})K_{KF}(\overline{k}+1)\tilde{y}(\overline{k}+1), \quad (4.70) \]

where \( P_{KF} \) is the prediction error covariance matrix, \( K_{KF} \) is the Kalman gain, \( Q_{KF} \) is the process noise covariance matrix and \( R_{KF} \) is the measurement variance. Both \( Q_{KF} \) and \( R_{KF} \) are assumed to be stationary and have been identified experimentally.

Due to the additional complexity of the predictor-corrector equations, analyzing the robust stability of this extended Kalman filter is much more difficult than for the LPV estimator. The simulated closed loop performance suggests that the estimator is robust to measurement and output uncertainty, but these properties have not been proven analytically. It should be noted that the Kalman filter equations do not contain tuning parameters and therefore the stability properties cannot be altered.

### 4.1.4 Hybrid Estimation Scheme

Because a delay based oxygen dynamic model considers the time varying nature of the transport delays while the LPV state estimator and extended Kalman filter designs assume fixed transport delays, an open loop delay based oxygen model predicts the response of the system much better than an LPV state estimator or an extended Kalman filter when all of the model inputs are known. The same modeling structure that enables a delay based in-cylinder oxygen concentration estimator to capture time varying transport delays also inhibits state estimator design. Since production measurement systems are far from ideal, an effective in-cylinder oxygen concentration estimator must be robust to measurement uncertainty. Due to its open loop nature,
input errors affect the in-cylinder oxygen concentration predicted by a delay based oxygen dynamic model more significantly than a LPV state estimator or an extended Kalman filter.

In general, a delay based oxygen dynamics model better predicts the open loop response of the system, whereas a LPV state estimator or an extended Kalman filter are better able to identify and compensate for input errors. Using a hybrid feed-forward/feedback estimation structure, the strengths of each design can be simultaneously realized. In this hybrid structure, the delay based oxygen concentration estimator generates a feed-forward estimate of the oxygen concentration. Running in parallel to this estimator are two versions of the four state LPV oxygen dynamics model. One of these models is run in an open loop manner whereas the other incorporates feedback according to the LPV state estimator design. The open loop estimator based on the four state LPV oxygen dynamics model is governed by

\[
\hat{x}^\dagger(\bar{k} + 1) = \left( A_0 + \hat{\alpha}_1(\bar{k})A_1 + \hat{\alpha}_2(\bar{k})A_2 \right)\hat{x}^\dagger(\bar{k}) \\
+ \left( B_0 + \hat{\alpha}_1(\bar{k})B_1 + \hat{\alpha}_2(\bar{k})B_2 \right)u(\bar{k})
\] (4.71)

where \( \hat{x}^\dagger \) is the open loop estimate of the state vector \( x \). By comparing the closed loop in-cylinder oxygen concentration estimate to the open loop in-cylinder oxygen concentration estimator, a closed loop in-cylinder oxygen concentration correction factor \( f_{corr} \) can be determined according to

\[
f_{corr}(\bar{k}) = \kappa C_{in-cyl} \left( \hat{x}(\bar{k}) - \hat{x}^\dagger(\bar{k}) \right)
\] (4.72)

where \( C_{in-cyl} = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \), \( \kappa \in \mathbb{R} \) is the correction weighting factor and the state estimate \( \hat{x} \) is governed by (4.24). This correction factor can then be applied to the feed-forward in-cylinder oxygen concentration estimate produced by the delay
based oxygen dynamics model $[O_2]_{cyl}$ to provide a final corrected in-cylinder oxygen concentration estimate $[O_2]_{cyl,corr}$ using

$$[O_2]_{cyl,corr}(k) = [O_2]_{cyl}(k) + f_{corr}\left(\lfloor k/6 \rfloor\right)$$

(4.73)

where $\lfloor k/6 \rfloor$ indicates that the correction factor is updated only once per engine cycle. The value of $\kappa$ which produced the best performance for the engine studied in this dissertation is 0.97.

An overall control diagram of this hybrid estimation control structure is provided in Figure 4.1. Because this estimation structure is modular, the LPV state estimator could be replaced by an extend Kalman filter. For an extended Kalman filter based hybrid estimator, the correction factor is calculated in same manner as before except that (4.65 - 4.70) are used to predict the state estimate $\hat{x}$ rather than (4.24). The control diagram for this estimator would appear similar to Figure 4.1 except that the extended Kalman filter equations would replace the LPV state estimator equations.

### 4.1.5 Performance Comparison

To validate the performance of the lumped oxygen dynamics model, the in-cylinder oxygen concentration and exhaust manifold oxygen concentration were compared to the oxygen concentrations predicted by a high fidelity GT-Power model. The results show that a lumped oxygen dynamics model estimates the oxygen concentrations throughout the air path much better than a traditional filling and emptying oxygen concentration model, especially during transients. A similar procedure has been used to quantify the performance of each estimation scheme. The true in-cylinder and exhaust oxygen concentrations were again predicted by GT-Power and the same test
cycle consisting of the first one hundred seconds of a FTP heavy duty drive cycle was used.

As emphasized previously, an effective real time in-cylinder oxygen concentration estimator must remain accurate even in the presence of measurement uncertainty.
To test the robustness of each estimator, different input disturbance profiles were superimposed on the fresh air mass flow rate, EGR mass flow rate, fuel mass flow rate, temperature and pressure input measurement trajectories. Similarly, different output disturbances trajectories are superimposed on the exhaust oxygen concentration output measurement used in feedback. A total of one thousand different sets of disturbance profiles were generated by scaling and time shifting a nominal set of disturbance trajectories modeled from experimental data.

The disturbance profiles for each input contained both long term bias errors and short term dynamic errors. The bias error for both the nominal fresh air mass flow rate and the nominal EGR mass flow rate disturbance profiles was 5.0%; the bias error for the nominal fuel mass flow rate disturbance profile was 2.0%. All of nominal disturbance profiles for the temperature and pressure input measurements nominal disturbance profiles as well as the exhaust oxygen concentration output measurement were characterized by a bias error of 3.0%.

The peak dynamic error represented in the nominal fresh air mass flow rate and EGR mass flow rate disturbance profiles was 20.0%; the peak dynamic error represented in the nominal fuel mass flow rate disturbance profile was 8.0%. A peak dynamic error of 9.0% was selected for the nominal temperature and pressure disturbance profiles and 12.0% for the nominal exhaust oxygen concentration disturbance profile. For reference, these values are presented in Table 4.2. The scaling applied to these nominal trajectories was a randomly selected value between -1.0 and 1.0 to ensure both positive and negative bias and dynamic errors were represented. Figure 4.2 shows the nominal EGR mass flow rate disturbance trajectory and an example of
the type of shifting and scaling which varied randomly for each of the one thousand simulation cases.

Table 4.2: Characteristics of the nominal input and output disturbance trajectories

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Bias</th>
<th>Peak Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fresh air mass flow rate</td>
<td>5.0 %</td>
<td>20.0 %</td>
</tr>
<tr>
<td>EGR mass flow rate</td>
<td>5.0 %</td>
<td>20.0 %</td>
</tr>
<tr>
<td>Fuel mass flow rate</td>
<td>2.0 %</td>
<td>8.0 %</td>
</tr>
<tr>
<td>Temperature</td>
<td>3.0 %</td>
<td>9.0%</td>
</tr>
<tr>
<td>Pressure</td>
<td>3.0 %</td>
<td>9.0%</td>
</tr>
<tr>
<td>UEGO measurement</td>
<td>3.0 %</td>
<td>12.0 %</td>
</tr>
</tbody>
</table>

Figure 4.2: Illustration of the shifting and scaling applied to the nominal EGR disturbance trajectory

The one thousand different sets of disturbance profiles were simulated under two different conditions. In the first set of simulations, only the input disturbances were
applied. These conditions represent an ideal scenario for using the exhaust oxygen concentration measurement in a feedback estimation loop. After each induction event, the in-cylinder oxygen concentration was predicted using a filling and emptying oxygen concentration model (open loop), a four state LPV oxygen dynamics model (open loop), a delay based oxygen dynamics model (open loop), a LPV state estimator (closed loop), an extended Kalman filter (closed loop), a LPV based hybrid feed-forward/feedback estimator and an extended Kalman filter based hybrid feed-forward/feedback estimator. For each drive cycle simulation case, a single value corresponding to the RMS error in estimating the in-cylinder oxygen concentration over the drive cycle was used to quantify the performance of each estimator.

Table 4.3: RMS in-cylinder oxygen concentration estimation errors in percent oxygen (input disturbances only)

<table>
<thead>
<tr>
<th>Estimation Method</th>
<th>Ideal Case</th>
<th>Average</th>
<th>Worse Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filling &amp; Emptying Model</td>
<td>0.199</td>
<td>0.283</td>
<td>0.589</td>
</tr>
<tr>
<td>Four State LPV Model</td>
<td>0.156</td>
<td>0.248</td>
<td>0.480</td>
</tr>
<tr>
<td>Delay Based Model</td>
<td>0.102</td>
<td>0.219</td>
<td>0.525</td>
</tr>
<tr>
<td>LPV State Estimator</td>
<td>0.161</td>
<td>0.190</td>
<td>0.283</td>
</tr>
<tr>
<td>Extended Kalman Filter</td>
<td>0.162</td>
<td>0.182</td>
<td>0.257</td>
</tr>
<tr>
<td>Hybrid: Delay + LPV</td>
<td>0.101</td>
<td>0.138</td>
<td>0.257</td>
</tr>
<tr>
<td>Hybrid: Delay + Kalman</td>
<td>0.097</td>
<td>0.133</td>
<td>0.256</td>
</tr>
</tbody>
</table>

The results for each estimator in terms of the RMS error when all disturbances are zero, the average RMS error over the one thousand cases and the largest RMS error out of all of the cases tested are presented in Table 4.3. In the absence of input errors,
the estimators which account for the time varying transport delays produce the most accurate predictions. The delay based oxygen dynamics model and the two hybrid estimators all produced RMS errors around 0.10% in this ideal case. Recall that the four state LPV oxygen dynamics model was derived from the delay based oxygen dynamics model by assuming constant delays, a constant trapped residual fraction and adopting a slower timescale. These approximations which facilitated estimator design result in additional estimation errors. For the ideal case, the four state LPV oxygen dynamics model, the LPV state estimator and the extended Kalman filter produced RMS errors of approximately 0.16%. All of these estimator designs outperform a traditional filling and emptying oxygen concentration model which produced an RMS error of 0.199% when the inputs are known perfectly.

Because open loop estimators are more susceptible to input errors, the RMS estimation errors produced by the filling and emptying oxygen concentration model, the four state LPV oxygen dynamics model and the delay based oxygen dynamics model are widely distributed. Conversely, the RMS error distributions for the LPV state estimator, the extended Kalman filter and the hybrid feed-forward/feedback estimators are much narrower. To better illustrate this point, probability density and cumulative probability functions were created for each estimation approach using the one thousand RMS estimation error data points. It should be noted that these input errors are not simply random noise but rather the output of a dynamical system whose input is random noise. Figure 4.3 compares the probability density function of each estimator and Figure 4.4 compares the cumulative probability functions.

For sufficiently small input errors, the delay based oxygen dynamics model continues to produce lower RMS errors than both the LPV state estimator and the extended
Kalman filter. As the size of the input errors increase, the delay based oxygen dynamics model produces larger errors than the LPV state estimator and the extended Kalman filter. With the hybrid feed-forward/feedback estimator design, the left end
of the RMS error distribution (minimum RMS error) starts from nearly the same point as the delay based oxygen dynamics model but the distribution is almost as narrow as the LPV state estimator and the extend Kalman filter. Overall, the two hybrid feed-forward/feedback estimators consistently produce the smallest RMS errors and the filling and emptying oxygen concentration model consistently produces the largest RMS errors.

All of the estimator comparisons presented so far only considered the effect of input disturbances. When an inaccurate measurement is used in feedback, the resulting closed loop estimate can potentially produce more errors than an open loop estimate. To determine the effect of UEGO measurement errors, the same set of one thousand different input disturbance profiles were simulated over the same drive cycles but with the inclusion of the output disturbance profiles. This testing approach allows for the exact effect of output measurement errors to be determined. The performance metrics for this set of simulations are presented in Table 4.4. Since output disturbances only impact the LPV state estimator, the Kalman filter and the hybrid estimators, the affected performance metrics for these estimators have been emphasized with a bold font.

Compared to the case with only input disturbances, the average RMS error increased from 0.190% to 0.204% for the LPV estimators and from 0.182% to 0.204% for the extended Kalman filter. Likewise, the LPV based hybrid estimator produced an average RMS error of 0.138% up from 0.161% and the extended Kalman filter based hybrid estimator produced an average RMS error of 0.133% up from 0.159%. Using these one thousand RMS estimation error data points, new probability density and cumulative probability functions were created. As seen in Figures 4.5 and
Table 4.4: RMS in-cylinder oxygen concentration estimation errors in percent oxygen (input and output disturbances)

<table>
<thead>
<tr>
<th>Estimation Method</th>
<th>Ideal Case</th>
<th>Average</th>
<th>Worse Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filling &amp; Emptying Model</td>
<td>0.199</td>
<td>0.283</td>
<td>0.589</td>
</tr>
<tr>
<td>Four State LPV Model</td>
<td>0.156</td>
<td>0.248</td>
<td>0.480</td>
</tr>
<tr>
<td>Delay Based Model</td>
<td>0.102</td>
<td>0.219</td>
<td>0.525</td>
</tr>
<tr>
<td>LPV State Estimator</td>
<td>0.161</td>
<td>0.204</td>
<td>0.346</td>
</tr>
<tr>
<td>Extended Kalman Filter</td>
<td>0.162</td>
<td>0.204</td>
<td>0.351</td>
</tr>
<tr>
<td>Hybrid: Delay + LPV</td>
<td>0.102</td>
<td>0.161</td>
<td>0.350</td>
</tr>
<tr>
<td>Hybrid: Delay + Kalman</td>
<td>0.097</td>
<td>0.159</td>
<td>0.350</td>
</tr>
</tbody>
</table>

Figure 4.6, output disturbances widen the error distributions for the LPV state estimator, the extended Kalman filter and the hybrid estimators. Compared to the open loop estimators, the RMS errors are still significantly less scattered. Even when output disturbances are considered, the two hybrid feed-forward/feedback estimators consistently produce the best performance and the filling and emptying oxygen concentration model consistently produces the largest RMS errors.

The differences between the open loop, closed loop and hybrid estimator designs are even more apparent when the in-cylinder oxygen concentration estimation traces are directly compared as in Figure 4.7 and Figure 4.8. This data corresponds to one of the drive cycle simulations in which the combined effect of all of the input disturbances tends to produce lean bias errors. These errors are well corrected in the LPV state observer, the extended Kalman filter and the hybrid estimators. The extended Kalman filter and the extended Kalman filter based hybrid estimator produce in-cylinder oxygen concentration estimates which are nearly identical to the LPV state
estimator and the LPV state estimator based hybrid estimator respectively; thus, these two estimates were omitted from Figure 4.7 and Figure 4.8.
Figure 4.7: Estimated in-cylinder oxygen concentration comparison

Aside from the disturbance rejection performance, the other difference that distinguishes the estimators is their ability to capture the time varying transport delays within the air path system. During a transient such as the one depicted in Figure 4.9, the in-cylinder oxygen concentration predicted by the filling and emptying model significantly leads the actual in-cylinder oxygen concentration. The four state LPV model assumes a fixed transport delay; thus, the transient in-cylinder oxygen concentration will sometimes lead and sometimes lag the actual oxygen concentration, but not by a significant margin. The time varying transport delay compensation within the delay based estimator, however, produces by far the best transient performance. It should be noted that since the four state LPV model and the LPV state estimate are updated only once per event, their estimates appear jagged.
Figure 4.8: In-cylinder oxygen concentration estimation error comparison

For completeness, the performances of each estimator has also been quantified in terms of its instantaneous estimation errors. Table 4.5 contains the estimation errors corresponding to the 50\textsuperscript{th}, the 75\textsuperscript{th}, the 90\textsuperscript{th}, the 95\textsuperscript{th} and the 99\textsuperscript{th} percentiles for each estimator. Considering the size of the input and output disturbance magnitudes tested, the performance of the two hybrid feed-forward/feedback estimator designs are remarkable. In terms of percent error, 95\% of the in-cylinder oxygen concentration predictions produced by the two hybrid feed-forward/feedback estimators are within 2.0\% of the actual oxygen concentration and 99\% of their estimates are within 3.0\% of the actual oxygen concentration.

As demonstrated in Figure 4.3 and Figure 4.5, closing the estimation loop on the four state LPV oxygen dynamics model narrows the RMS error distribution but does
Engine Events
In-cylinder Oxygen Concentration (%)

Figure 4.9: Estimated in-cylinder oxygen concentration comparison zoomed to show the effects of modeling the transport delays

Table 4.5: Instantaneous in-cylinder oxygen concentration estimation error percentile statistics (input and output disturbances)

<table>
<thead>
<tr>
<th>Estimation Method</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filling &amp; Emptying Model</td>
<td>0.113</td>
<td>0.272</td>
<td>0.485</td>
<td>0.664</td>
<td>1.023</td>
</tr>
<tr>
<td>Four State LPV Model</td>
<td>0.132</td>
<td>0.268</td>
<td>0.434</td>
<td>0.566</td>
<td>0.797</td>
</tr>
<tr>
<td>Delay Based Model</td>
<td>0.088</td>
<td>0.217</td>
<td>0.396</td>
<td>0.529</td>
<td>0.821</td>
</tr>
<tr>
<td>LPV State Estimator</td>
<td>0.113</td>
<td>0.224</td>
<td>0.342</td>
<td>0.423</td>
<td>0.612</td>
</tr>
<tr>
<td>Extended Kalman Filter</td>
<td>0.110</td>
<td>0.220</td>
<td>0.338</td>
<td>0.421</td>
<td>0.618</td>
</tr>
<tr>
<td>Hybrid: Delay + LPV</td>
<td>0.071</td>
<td>0.161</td>
<td>0.276</td>
<td>0.360</td>
<td>0.541</td>
</tr>
<tr>
<td>Hybrid: Delay + Kalman</td>
<td>0.074</td>
<td>0.158</td>
<td>0.272</td>
<td>0.353</td>
<td>0.541</td>
</tr>
</tbody>
</table>

not improve the maximum performance achieved for the ideal case without any disturbances. Applying state estimation techniques to a filling and emptying model would produce approximately the same effect. Even with feedback from a UEGO sensor, a
filling and emptying based in-cylinder oxygen concentration estimator would not be very effective during transients because the open loop estimator errors generated by ignoring the transport delays within the air path system would still remain.

By a wide margin, the hybrid feed-forward/feedback estimators produce the most accurate in-cylinder oxygen concentration predictions. The performance of the extended Kalman filter and the LPV state estimator are nearly the same, so the two hybrid estimators produce approximately the same performance. Unlike the extended Kalman filter, however, the LPV state estimator has been analytically guaranteed to be robustly input to state stable with respect to both input and output disturbances. When also considering that an extended Kalman filter requires several additional dynamic states and significantly more calculations, a LPV state estimator is clearly more practical than an extended Kalman filter. In turn, the hybrid estimator based on the LPV state estimator is the most desirable solution. Because the open loop four state model and the delay based oxygen dynamics model are both stable, the overall hybrid estimator is also stable. The LPV state estimator based hybrid estimator could feasibly be implemented and run in real time in a production ECU, and it outperforms all other estimation strategies, especially a filling and emptying oxygen concentration model.

4.2 Model Based Cylinder Imbalance Estimation

This section develops a cylinder imbalance estimator based on the FIR exhaust oxygen dynamics model from Section 3.4 and validates its performance on the 2.4 liter gasoline engine for the case when feedback is provided by an EGO sensor. Identifying cylinder imbalance using a single pre-catalyst EGO sensor is very difficult,
because when the voltage output saturates, it is almost impossible to determine the magnitude of an imbalance error. When the EGO voltage leaves the linear range, the gain between voltage and oxygen concentration (or equivalently EQR) changes appreciably. To accurately identify the magnitude, the EGO voltage would have to stay within the linear range of the sensor; even under controlled steady-state conditions, this may not happen. If the cylinder imbalance estimate is used in a feedback control loop, the uncertainty in the magnitude of the error can be compensated. However, because of the saturation effect of an EGO sensor, the gain between the imbalance estimate and the actual imbalance decreases as the imbalance amplitude increases. In a feedback structure, this effect causes the controller to respond to large imbalances more slowly than for small imbalances. In most cases, this would not be a problem because it prevents the controller from making large, possibly destabilizing corrections. Because accurate estimation of the error magnitude is not possible with a switching EGO sensor, only deviations from some mean value are considered. Thus, to reduce the nonlinear effects of an EGO sensor, the EGO voltage is first converted into an estimated AFR. The mean in this case is calculated by taking a window average, over two engine cycles, of the estimated EQR. With this idea as a starting point, two different techniques for estimating cylinder imbalance are presented in the following sections.

### 4.2.1 Cylinder Imbalance Estimation Using Direct Matrix Inversion

Prediction of cylinder imbalance requires a modification of the FIR exhaust oxygen dynamics model. As other studies have proposed [51–54], the system can be represented with a square matrix of dimension $n_{cyl} \times n_{cyl}$. For this realization, the input
to the model is a vector of the relative cylinder imbalances in terms of EQR and the output is a vector of the measured EGO voltage deviations from the mean. This representation can be thought of as a zero-input, multi-output model. The FIR exhaust oxygen dynamics model predicts when and how input fuel injection commands affect the oxygen sensor measurement. A single fueling event command contributes to the output for a small cluster of events, so that, depending on the operating conditions, the size and location of this grouping relative to the commanding event changes.

For the cylinder imbalance prediction model, a new output reference is considered. Instead of referencing the output back to the event in which the fuel was commanded, the output is referenced to its location relative to the periodic cylinder fueling (firing) order. A systematic cylinder imbalance error will produce the same output every engine cycle (four events for the gasoline engine in this study) in steady-state. The contribution from each cylinder to the output relative to the fueling order is constant for each cylinder, but varies between cylinders. Consider the simplest case of a 4-cylinder engine where all cylinders have a fixed integer delay of eight, and there are no mixing or sensor dynamics. In this hypothetical case, the commanded fuel for each cylinder will appear in the EGO measurement eight events after it was commanded. At each event the exhaust gas that passes the oxygen sensor originated from the same cylinder that is currently fueling. Because the fueling order is fixed, it is a unique identifier.

In their original form, the input and output are vectors of length \( n_{\text{window}} \). For cylinder imbalance prediction, the system is transformed into

\[
\Phi_{\text{imb}}(k + 1) = \Phi_{\text{imb}}(k) \quad (4.74)
\]

\[
\Phi_{\text{err}}(k) = C_{\text{imb}}(k)\Phi_{\text{imb}}(k) \quad (4.75)
\]
where $\Phi_{\text{err}} \in \mathbb{R}^{n_{\text{cyl}}}$, $C_{\text{imb}} \in \mathbb{R}^{n_{\text{cyl}} \times n_{\text{cyl}}}$ and $\Phi_{\text{imb}} \in \mathbb{R}^{n_{\text{cyl}}}$ (time dependency suppressed).

The states of the system, $\Phi_{\text{imb}}$, are the individual cylinder imbalances, assumed to be slowly varying. The output, $\Phi_{\text{err}}$, is a vector of the deviation of measured EQR from the mean EQR. The first element of the output vector is the measurement deviation when the first cylinder is fueling. The second element corresponds to the measurement deviation when the second cylinder fuels, and so on. The first column of the system matrix contains the summations of all of the coefficients in the FIR model that correspond to cylinder-1. The second column contains the summations for the second cylinder, and so on. At each event only one row of the matrix and one element of the output are updated.

As an example consider the case of a four cylinder engine with a 1-3-4-2 firing order like the 2.4 liter gasoline engine studied in this dissertation. Define $\Sigma_1$, $\Sigma_2$, $\Sigma_3$ and $\Sigma_4$ as

\[
\Sigma_1(k) = \text{io}_1(k) + \text{io}_5(k) + \text{io}_9(k) + \ldots + \text{io}_{\left[\frac{n_{\text{window}}-1}{4}\right]+1}(k) \quad (4.76)
\]

\[
\Sigma_2(k) = \text{io}_2(k) + \text{io}_6(k) + \text{io}_10(k) + \ldots + \text{io}_{\left[\frac{n_{\text{window}}-2}{4}\right]+2}(k) \quad (4.77)
\]

\[
\Sigma_3(k) = \text{io}_3(k) + \text{io}_7(k) + \text{io}_11(k) + \ldots + \text{io}_{\left[\frac{n_{\text{window}}-3}{4}\right]+3}(k) \quad (4.78)
\]

\[
\Sigma_4(k) = \text{io}_4(k) + \text{io}_8(k) + \text{io}_12(k) + \ldots + \text{io}_{\left[\frac{n_{\text{window}}}{4}\right]}(k) \quad (4.79)
\]

where $\text{IO} = \left[ \text{io}_1 \quad \text{io}_2 \quad \text{io}_3 \ldots \quad \text{io}_{n_{\text{window}}} \right]$. When cylinder one is fueling, $C_{\text{imb}}(k)$ is defined by

\[
C_{\text{imb}}(k) = \begin{bmatrix}
\Sigma_4(k) & \Sigma_1(k) & \Sigma_3(k) & \Sigma_2(k) \\
\Sigma_1(k-3) & \Sigma_2(k-3) & \Sigma_4(k-3) & \Sigma_3(k-3) \\
\Sigma_2(k-2) & \Sigma_3(k-2) & \Sigma_1(k-2) & \Sigma_4(k-2) \\
\Sigma_3(k-1) & \Sigma_4(k-1) & \Sigma_2(k-1) & \Sigma_1(k-1)
\end{bmatrix} \quad (4.80)
\]
The next event when cylinder three fuels this matrix is updated to

\[
C_{imb}(k) = \begin{bmatrix}
\Sigma_4(k-1) & \Sigma_1(k-1) & \Sigma_3(k-1) & \Sigma_2(k-1) \\
\Sigma_1(k) & \Sigma_2(k) & \Sigma_4(k) & \Sigma_3(k) \\
\Sigma_2(k-3) & \Sigma_3(k-3) & \Sigma_1(k-3) & \Sigma_4(k-3) \\
\Sigma_3(k-2) & \Sigma_4(k-2) & \Sigma_2(k-2) & \Sigma_1(k-2)
\end{bmatrix}.
\]  
(4.81)

The next event when cylinder four fuels this matrix is updated to

\[
C_{imb}(k) = \begin{bmatrix}
\Sigma_4(k-2) & \Sigma_1(k-2) & \Sigma_3(k-2) & \Sigma_2(k-2) \\
\Sigma_1(k-1) & \Sigma_2(k-1) & \Sigma_4(k-1) & \Sigma_3(k-1) \\
\Sigma_2(k) & \Sigma_3(k) & \Sigma_1(k) & \Sigma_4(k) \\
\Sigma_3(k-3) & \Sigma_4(k-3) & \Sigma_2(k-3) & \Sigma_1(k-3)
\end{bmatrix}.
\]  
(4.82)

The next event when cylinder two fuels this matrix is updated to

\[
C_{imb}(k) = \begin{bmatrix}
\Sigma_4(k-3) & \Sigma_1(k-3) & \Sigma_3(k-3) & \Sigma_2(k-3) \\
\Sigma_1(k-2) & \Sigma_2(k-2) & \Sigma_4(k-2) & \Sigma_3(k-2) \\
\Sigma_2(k-1) & \Sigma_3(k-1) & \Sigma_1(k-1) & \Sigma_4(k-1) \\
\Sigma_3(k) & \Sigma_4(k) & \Sigma_2(k) & \Sigma_1(k)
\end{bmatrix}.
\]  
(4.83)

As the pattern indicates, only one row needs to be updated each event.

By inverting the $C_{imb}$ matrix, the imbalance of each cylinder can be estimated using

\[
\Phi_{imb}(k) = [C_{imb}(k)]^{-1}\Phi_{err}(k).
\]  
(4.84)

Calculating this inverse on-line would be too taxing. By partitioning the operating space into zones based on flow rate and engine speed and assuming that within each zone the $C_{imb}$ matrix is constant, this problem can be overcome by calculating the inverse of these matrices off-line and then storing them in the ECU for direct use. For an operating condition in the interior of a zone, good performance can be achieved. However, at the boundary between zones, stable performance cannot be guaranteed, due to the switching nature of the estimation scheme. Although this type of system realization could be used to predict cylinder imbalance, a more robust technique based on recursive least squares (RLS) can alternatively be applied, addressed next.
4.2.2 Cylinder Imbalance Identification using RLS

Instead of converting the FIR model into the matrix form of (4.75), the formulation
\[ \phi_{\text{err}} = (c_{\text{imb}})^T \Phi_{\text{imb}} \] (4.85)
can be used, where \( \phi_{\text{err}} \in \mathbb{R}, c_{\text{imb}} \in \mathbb{R}^{n_{\text{cyl}}} \) and \( \Phi_{\text{imb}} \in \mathbb{R}^{n_{\text{cyl}}} \). In fact, note that
\[
\Phi_{\text{err}} = \begin{bmatrix}
\phi_{\text{err},1} \\
\phi_{\text{err},2} \\
\vdots \\
\phi_{\text{err},n_{\text{cyl}}}
\end{bmatrix} = C_{\text{imb}} \Phi_{\text{imb}} = \begin{bmatrix}
(c_{\text{imb},1})^T \\
(c_{\text{imb},2})^T \\
\vdots \\
(c_{\text{imb},n_{\text{cyl}}})^T
\end{bmatrix} \Phi_{\text{imb}} \tag{4.86}
\]
which is a single row of (4.75). Without matrix inversion, the cylinder imbalance can be calculated using RLS [86] via an iterative procedure defined by
\[
K_{\text{rls}}(k+1) = P_{\text{rls}}(k)c_{\text{imb}}(k+1)\left[c_{\text{imb}}(k+1)^T P_{\text{rls}}(k)c_{\text{imb}}(k+1) + \lambda_f \right]^{-1}, \tag{4.87}
\]
\[
P_{\text{rls}}(k+1) = \left[I - K_{\text{rls}}(k+1)c_{\text{imb}}(k+1)^T \right] \frac{P_{\text{rls}}(k)}{\lambda_f}, \tag{4.88}
\]
and
\[
\hat{\Phi}_{\text{imb}}(k+1) = K_{\text{rls}}(k+1) \left[ \phi_{\text{err}}(k+1) - c_{\text{imb}}(k+1)^T \hat{\Phi}_{\text{imb}}(k) \right] \\
+ \hat{\Phi}_{\text{imb}}(k), \tag{4.89}
\]
where \( K_{\text{rls}}(k) \in \mathbb{R}^{n_{\text{cyl}}}, P_{\text{rls}}(k) \in \mathbb{R}^{n_{\text{cyl}} \times n_{\text{cyl}}}, \lambda_f \in \mathbb{R} \) and \( \hat{\Phi}_{\text{imb}}(k) \in \mathbb{R}^{n_{\text{cyl}}}. \)

4.2.3 Experimental Validation

Because an RLS approach to cylinder imbalance estimation is more desirable than a matrix inversion approach, the experimental validation focused on the RLS technique. For the experimental validation, the four cylinder gasoline engine was operated in steady-state at the same three representative operating conditions shown in Table 3.1. For each set of testing, an imbalance error was applied to only one of the cylinders.
Even without tuning the forgetting factor, $\lambda_f$, an RLS estimator is able to provide a reasonable estimate of the cylinder imbalance. Figure 4.10 shows the RLS imbalance predicted at idle under a 10% rich imbalance applied to cylinder-1 at event 300. Because the voltage output of the EGO sensor saturates, the AFR signal cannot be completely reconstructed, leading to a significant amount of uncertainty in the measurement. As a result, the exact magnitude of the AFR imbalance cannot be determined. Considering the nonlinear response of a typical EGO sensor, the performance of this baseline RLS cylinder imbalance estimator is deemed acceptable.

By increasing the weight of newer measurements (decreasing the forgetting factor below unity), the speed of convergence can be increased, with the inherent trade-off being that the resulting estimate is more susceptible to noise. To reduce the noise, a discrete filter can be applied. With a forgetting factor of 0.989 and a first order filter time constant of 0.02, the estimate converges rapidly and the noise amplitude is small, as seen in Figure 4.11. Imbalance errors in the other cylinders were estimated as effectively and this same level of performance was achieved at the other operating conditions.

4.2.4 Practical Detection of Cylinder Characteristics

Cylinder imbalance estimation is particularly challenging at idle for this engine, because the exhaust gases from multiple cylinders reach the oxygen sensor at the same time. At idle, the largest contributions from cylinder-1 and cylinder-2 reach the oxygen sensor simultaneously. Similarly, the largest contributions from cylinder-3 and cylinder-4 reach the oxygen sensor at the same time. Although the cylinder imbalance responses of each cylinder are unique, they are difficult to distinguish.
Figure 4.10: Identification of a cylinder-1 imbalance error applied at event 300 during idle (baseline)

For some operating conditions, therefore, important characteristics affecting cylinder imbalance may not be detected. However, an FIR exhaust oxygen dynamics model can be simulated to determine under what conditions, if any, the system exhibits this tendency.

For a given engine, only the exhaust flow rate and engine speed affect the output dependencies. A simple metric to indicate some degree of detection ability is the condition number of the output matrix; that is, when the FIR model is put in the matrix form described by (4.75), the condition number of $C_{imb}$ provides a good measure of detection capability. If the condition number is too high, then cylinder imbalance errors cannot be identified with confidence, especially when considering the level of measurement uncertainty. Figure 4.12 shows how the condition number varies as a function of the operating conditions for the engine used in this study. The condition
Figure 4.11: Identification of a cylinder-1 imbalance error applied at event 300 during idle, with first-order filter (tuned)

number is very high at idle (noted with a star in the figure). Using an analysis such as this, guidelines for when to use the cylinder imbalance estimator can be constructed. Each engine will have different characteristics depending on the number of cylinders and exhaust configuration. As the number of cylinders increases, it becomes more difficult to identify cylinder imbalance because the frequency content of the measurement signal increases. It is conceivable that the exhaust mixing within a system is so uniform that cylinder imbalance cannot be detected at all. For engines that use EGO sensors, additional constraints must be applied, because cylinder imbalance errors can only be observed when the nominal AFR is near stoichiometry; useful guidelines are provided in [87].
4.2.5 Throughput Considerations

An FIR exhaust oxygen dynamics model can be used in an offline capacity to generate sets of matrix based cylinder imbalance models. For the RLS based cylinder imbalance prediction method, however, the FIR model is intended for real time implementation, requiring an analysis of throughput for a typical production ECU. Fortunately, the calculations required to execute an FIR model are very simple; that is, excluding one division, all of the operations are either multiplications or additions. The computational intensity is not derived from the complexity of the equations, but rather from the number of calculations. As a first run, the number of cells in an FIR model and the input window length could be decreased. Even with these simplifications, a significant number of operations, on the order of thousands, is required to execute the model. Current ECUs would not be able to handle this number of calculations every event. Steady-state cylinder imbalance errors, however, could be
corrected without executing the FIR model at every event. The coefficient prediction calculations could be distributed over several engine events and then stored until the next set of calculations are completed. In steady-state, the FIR model is constant, so decreasing the updating frequency of the model would not affect the accuracy of the cylinder imbalance estimate. Therefore, the RLS based cylinder imbalance estimator could be used for on-line estimation of steady-state cylinder imbalance errors.

4.3 Summary

Estimating the in-cylinder oxygen concentration in a turbocharged diesel engine is much more difficult than in a stoichiometric gasoline engine because the in-cylinder oxygen concentration does not scale linearly with the trapped EGR fraction. In this difficult situation, accounting for the oxygen transport and mixing dynamics is paramount to accurately estimating the in-cylinder oxygen concentration. With a hybrid feed-forward/feedback estimation approach, the transport delay is effectively modeled and input disturbances are adequately rejected, producing remarkable performance. The hybrid oxygen concentration estimator developed in this chapter could also be applied to a turbocharged or naturally aspirated stoichiometric gasoline engine with an external EGR loop. Although the in-cylinder oxygen concentration would most certainly be estimated with a higher accuracy, the relative benefit of this estimation scheme would not be as high because the effect of the transport delay is not as significant.

Likewise, the cylinder imbalance estimator developed in this chapter could also be applied to a lean burning engine or a turbocharged engine. For a turbocharged engine with an oxygen sensor located downstream from the turbocharger, the transport
delays are larger and the exhaust gases mixing effects could potentially be more significant. In certain situations, the mixing effects could prevent cylinder imbalance from being estimated; however, these cases can be identified a priori with a FIR model using the techniques described in Section 4.2.4. As for lean burning engines (both gasoline and diesel), a UEGO sensor would necessarily be required. Given the slow sensor dynamics which characterize current production type UEGO sensors, cylinder imbalance estimation is bound to be more difficult. The effects of a slower response rate on the imbalance detection potential can also be identified with a FIR model though. In the future, faster responding UEGO sensors could be adopted making the estimation problem easier.

The proposed in-cylinder oxygen concentration and cylinder imbalance estimators both exhibit excellent performance and enable tighter overall engine control. With a RLS cylinder imbalance estimator, imbalance errors resulting from either breathing or fueling differences can be estimated. In a stoichiometric gasoline engine, this information could be used to adjust the fueling of the unbalanced cylinder(s) to ensure the combustion in each cylinder is stoichiometric. In a diesel engine, balancing the torque produced by each cylinder is more important than balancing the EQR. Although a crank shaft speed fluctuation based cylinder imbalance estimator is probably better suited for the application of torque balancing within a diesel engine, an exhaust oxygen concentration based RLS imbalance estimator can be utilized in an alternative fashion to enhance the performance of an in-cylinder oxygen concentration based fueling controller. The EQR imbalance identified by the proposed RLS estimator can be converted into an in-cylinder oxygen concentration imbalance. Applying this imbalance estimate to a nominal in-cylinder oxygen concentration estimate (such as the
one produced by an LPV hybrid estimator) before selecting the fueling parameters will ensure that optimal combustion is maintained even in the presence of cylinder imbalance.

With the level of accuracy in predicting the in-cylinder oxygen concentration achieved by the LPV hybrid estimator, the fueling parameters of a diesel engine can be dynamically selected based on the actual cylinder conditions rather than some reference target. Because the LPV hybrid estimator is robust to both input and output disturbances, the optimal fueling parameters can be selected for each combustion event. Not only does this improve the transient combustion performance, but also the overall robustness of the fuel control system. With a fueling controller scheduling on the in-cylinder oxygen concentration predicted by a LPV hybrid estimator, the fueling parameters will remain optimal even if the air path controller misses its setpoint due to measurement errors, the air path system is physically unable to reach a setpoint or the air path actuation behavior is intentionally altered to prevent turbocharger surge/choke or rich combustion. Chapter 5 further describes and demonstrates the performance and robustness benefits of an in-cylinder oxygen concentration based diesel fueling control architecture.
A typical diesel engine controller determines both the desired air path setpoints and the fueling parameters based on a feed-forward mapping of the engine speed and the desired fuel injection mass. Because the timescale of the air path system is significantly slower than the fuel delivery system, the fueling parameters are not perfectly matched to the cylinder conditions; instead, the fueling parameters are matched to the desired air path setpoints. By mapping the fueling parameters based on the in-cylinder conditions, it is possible to reduce the fuel consumption rate and/or emission production rate during transient operation thereby reducing the total fuel consumed and/or the total emissions produced during real world driving patterns. Although the fueling parameters depend on many of the properties which characterize the cylinder contents, the in-cylinder oxygen concentration was found to be one of the most significant in specifying the “optimal” values. Using the feed-forward/feedback hybrid estimation structure described in Section 4.1, it is possible to accurately and robustly predict the in-cylinder oxygen concentration for design of such an in-cylinder oxygen concentration based fueling controller. The advantages of an oxygen concentration
based fueling controller over a conventional controller are demonstrated in the follow-
ing sections.

5.1 Controller Description

As interest in alternative diesel combustion modes has grown, the research in this field has begun to emphasize the development of methods for accurately predicting the dynamically changing cylinder conditions. Because alternative combustion modes such as low temperature combustion (LTC) and homogeneous charge compression ignition (HCCI) are prone to combustion instability, the fueling commands must be precisely matched to the contents of the cylinders to achieve desirable combustion. As demonstrated in [61], an estimate of the cylinder contents which has a 5% error can cause misfires in an HCCI engine. The chemical kinetics and computational fluid dynamics studies in [88] and [89] showed that the in-cylinder oxygen concentration has a strong effect on HCCI combustion. In particular, decreasing the in-cylinder oxygen concentration was found to significantly increase the ignition delay as well as decrease the heat release rate.

Although less researched, matching the fueling commands to the exact cylinder conditions is also important in conventional diesel combustion. In a conventional control structure, fueling parameters are determined based on a feed-forward mapping of the engine speed, the desired fuel mass and the system operating mode (e.g. steady-state, transient, altitude, etc.). In parallel to the fueling controller, an air path controller strives to regulate the air path system to a set of desired setpoints which are also scheduled on engine speed, the desired fuel mass, and the operating mode [5]. A simple schematic of this type of approach is provided in Figure 5.1. Because of the
storage dynamics and the transport delays within the air path system, the desired air path setpoints cannot be reached instantaneously. The variations in the cylinder contents caused by these dynamics are not typically taken into accounted in current fuel control systems.

![Diagram](image)

Figure 5.1: Conventional air path and fueling control approach

The combustion process depends on two factors: the in-cylinder conditions before fuel is injected and the fuel delivery characteristics. Scheduling the fueling parameters on the in-cylinder conditions rather than the air path setpoints is therefore a more logical approach. Figure 5.2 illustrates how such a controller could be integrated into the overall diesel engine control structure. As with a traditional controller, the fueling parameters are still selected based on sets of lookup tables, but in this architecture the indexing variables include the properties of the cylinder contents.
With this structure, the fueling controller is no longer autonomous with respect to the air path system. The interconnection between the fueling controller and the air path system through the cylinder contents estimator significantly improves the robustness of the fuel control system. Even though the air path setpoints still vary as a function of the system operating mode, this interconnection eliminates the necessity of having different sets of fueling tables for each different system operating mode. By expanding the dimensions of the fueling tables by one degree (three dimensions instead of two), the effect of the different air path setpoints can be explicitly captured. This approach can potentially reduce the calibration time and effort required for the fueling system if the number of table entries does not increase significantly. To determine how a set of three-dimensional fueling tables with coarse grid spacing compares to a set of conventional two-dimensional fueling tables with a finer grid spacing, and to determine the cylinder properties that have the strongest influence on the combustion process (and thus the best candidates for table indexing variables), a data driven approach described in the next section was undertaken.
5.2 Table Based Fueling Parameter Representation Comparisons

In a traditional engine control structure, both the air path setpoints and the fueling parameters are directly scheduled on the engine speed, the total fuel injection mass and the system operating mode. Often this structure is realized by using different sets of two-dimensional tables to represent the air path setpoints and the fueling parameters for each system operating mode. The most common air path setpoint variables are either flow rate variables (e.g. fresh air flow rate) or variables related to the intake manifold conditions (e.g. manifold pressure) because these variables can be directly measured. However, these types of variables cannot be linked to specific cylinder conditions in a one-to-one manner, so the effectiveness of a conventional fueling control structure is limited even when the air path setpoints are tracked well.
At a minimum three different system operating modes are required to characterize
the different modes of normal engine operation: steady-state, transient and altitude.
Fundamentally, the only difference between the transient, steady-state and altitude
fueling tables within a conventional controller is that they correspond to different air
path setpoints. Instead of using multiple sets of two-dimensional tables, the fueling
parameters for these same conditions can be predicted using a single set of three-
dimensional tables. For the engine studied in this dissertation, a total of five different
fueling parameters must be predicted: main SOI, pilot SOI, post SOI, pilot quantity
and post quantity.

To determine the effectiveness of a conventional fueling parameter prediction ap-
proach on this diesel engine, a set of optimized fueling parameters corresponding
to transient operation were fit using a set of five (one per fueling parameter) two-
dimensional tables indexed by engine speed and total fuel mass. To approximate
current industry practices, the size of these tables was selected to be 20 x 20. The
fit potential of the proposed approach involving three-dimensional tables was also
determined. Considering that a single set of three-dimensional fueling tables would
replace at least three sets of conventional fueling tables (i.e. steady-state, transient
and altitude), the size of the three-dimensional tables was selected to be 10 x 10 x 10
to keep the total number of entries comparable.

Before the effectiveness of any set of scheduling variables was determined, the
optimal fueling parameters were first identified. To rapidly identify the optimal fueling
parameters for a wide range of cylinder conditions, the GT-Power model of the engine
was utilized. The air path setpoints and fueling parameters chosen by the production
engine controller were used as a starting point. Even in simulation, simultaneously
optimizing both the air path and fueling systems is a significant undertaking. To simplify the optimization process, the VGT position was fixed to reduce the air path setpoint calibration problem from a two degree of freedom optimization (charge flow rate and EGR fraction) to a single degree of freedom optimization (EGR fraction). Again the evaluation cycle was chosen to be the first 100 seconds of a heavy duty FTP drive cycle. Using a variation of the standard design of experiments (DoE) approach, the optimal air path setpoints and the fueling parameters for each combustion event of this transient drive cycle were identified directly. Instead of performing the DoE at individual fixed steady-state operating conditions, a DoE was performed across the transient drive cycle.

The transient DoE optimization was performed by fixing the EGR valve position trajectory and simulating the same drive cycle repeatedly with different sets of fueling parameters. Once all of the drive cycles were simulated, the optimal fueling parameters were determined on a per event basis by comparing the indicated specific fuel consumption rate (ISFC) and the brake specific NO\textsubscript{x} production rate (BSNO\textsubscript{x}) achieved by each of the different fueling parameter combinations. More specifically, the fueling parameters which minimized the ISFC while limiting the BSNO\textsubscript{x} to a maximum level of 1.0 grams per horsepower-hour (g/hp-hr) were considered to be optimal with respect to the in-cylinder conditions corresponding to that combustion event. If a BSNO\textsubscript{x} of 1.0 g/hp-hr could not be achieved, then the fueling parameters which produced the lowest BSNO\textsubscript{x} level while still meeting the torque request were chosen.

Figure 5.3 compares the BSNO\textsubscript{x} versus ISFC trade-off produced by the optimal set of fueling parameters to the trade-offs of alternative sets of fueling parameters for a
single operating condition encountered during the drive cycle (1000 RPM, 65 mg fuel). For confidentiality reasons, all of the ISFC values presented in this section have been normalized with respect to the full range of achievable ISFC values (i.e. percentage of full scale). Since this drive cycle contains 4,000 combustion events, a total of 4,000 optimal sets of fueling parameters were generated. This procedure was repeated for five different EGR actuation trajectories to produce a total of 20,000 sets of optimal fueling parameters. To determine the optimal air path setpoints for this drive cycle, the ISFC and BSNO$_x$ values achieved by the optimal fueling parameters of each of the five different EGR actuation trajectories were compared on a per event basis using the same optimization constraints as before. A total of 5,500 different instantaneous air path setpoints were optimized and their corresponding fueling parameters were recorded.

![Figure 5.3: Fueling parameter calibration example](image_url)
The transient fueling tables within a conventional fueling controller would be used to predict these fueling parameters. For each of the fueling parameters, a 20 x 20 table indexed by engine speed and fuel mass was fit to this dataset. To quantify the effectiveness of this representation method, the $R^2$ correlations between the table predicted values and the optimal values of each fueling parameter were calculated and the results are presented in Table 5.1.

Once the intake valves close, only the fuel injection strategy can control the combustion process. The goal of the proposed fueling controller is therefore to produce optimal combustion under all possible in-cylinder conditions, not just the in-cylinder conditions corresponding to the optimal air path setpoints. The effectiveness of different three-dimensional table representations was evaluated using all of the optimal fueling parameter data (20,000 points). After the tables were fit to this data, the $R^2$ correlations between the optimal and the table-predicted fueling parameters were calculated for two different datasets. The first dataset contained only the 4,000 data points corresponding to the optimal air path setpoint, whereas the second dataset contained all 20,000 data points.

Table 5.1: Fueling table $R^2$ fit comparison (only optimal in-cylinder conditions)

<table>
<thead>
<tr>
<th>All values in (%)</th>
<th>Main SOI</th>
<th>Pilot SOI</th>
<th>Post SOI</th>
<th>Pilot Qty</th>
<th>Post Qty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional (RPM and Fuel)</td>
<td>91.2</td>
<td>93.0</td>
<td>96.3</td>
<td>98.4</td>
<td>93.1</td>
</tr>
<tr>
<td>Variable set 1</td>
<td>93.7</td>
<td>92.6</td>
<td>97.6</td>
<td>94.7</td>
<td>86.3</td>
</tr>
<tr>
<td>Variable set 2</td>
<td>91.5</td>
<td>91.4</td>
<td>96.6</td>
<td>93.3</td>
<td>79.0</td>
</tr>
<tr>
<td>Variable set 3</td>
<td>89.8</td>
<td>89.4</td>
<td>96.3</td>
<td>94.3</td>
<td>79.8</td>
</tr>
<tr>
<td>Average</td>
<td>78.9</td>
<td>77.2</td>
<td>90.4</td>
<td>84.4</td>
<td>43.3</td>
</tr>
<tr>
<td>Minimum</td>
<td>26.0</td>
<td>31.7</td>
<td>68.5</td>
<td>65.2</td>
<td>-44.6</td>
</tr>
</tbody>
</table>
The combustion process depends on many factors to varying degrees of importance. To identify the best method of representing the optimal fueling parameters using three-dimensional tables (i.e. the best table indexing variables), the $R^2$ fit correlations were determined for all possible three variable combinations of the following list of parameters: total trapped charge mass, trapped oxygen mass, trapped burned gas mass, in-cylinder oxygen concentration, total fuel mass, AFR, engine speed, charge pressure, charge temperature and charge density. For many of the variable combinations, the variation in the fueling parameters was not well predicted. Overall, three sets of variables outperformed all other combinations.

The $R^2$ correlation results corresponding to the dataset containing only the optimal air path setpoints are presented in Table 5.1 and the results corresponding to the full dataset are presented in Table 5.2. For each of the five fueling parameters, these tables show the $R^2$ fits of the three top performing variable sets as well as the average and minimum $R^2$ fits out of all of the variable combinations. The three variable sets in the order they appear in the table are \{engine speed, total fuel mass, in-cylinder oxygen concentration\}, \{engine speed, trapped oxygen mass, AFR\} and \{engine speed, total fuel mass, trapped oxygen mass\}. At least one of the variables in each of these sets depends on the in-cylinder oxygen concentration, motivating the use of an in-cylinder oxygen concentration based fueling controller. Of the three sets of variables, the trio of engine speed, total fuel mass and in-cylinder oxygen concentration are the most desirable because they predict the optimal fueling parameters most effectively, and each of the variables can be predicted with a high degree of certainty.

161
Table 5.2: Fueling table $R^2$ fit comparison (all in-cylinder conditions)

<table>
<thead>
<tr>
<th>All values in (%)</th>
<th>Main SOI</th>
<th>Pilot SOI</th>
<th>Post SOI</th>
<th>Pilot Qty</th>
<th>Post Qty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable set 1</td>
<td>95.0</td>
<td>94.3</td>
<td>96.9</td>
<td>90.0</td>
<td>90.8</td>
</tr>
<tr>
<td>Variable set 2</td>
<td>92.0</td>
<td>92.2</td>
<td>95.1</td>
<td>88.1</td>
<td>80.3</td>
</tr>
<tr>
<td>Variable set 3</td>
<td>89.5</td>
<td>90.0</td>
<td>94.4</td>
<td>88.3</td>
<td>80.3</td>
</tr>
<tr>
<td>Average</td>
<td>81.8</td>
<td>80.0</td>
<td>89.7</td>
<td>81.1</td>
<td>74.8</td>
</tr>
<tr>
<td>Minimum</td>
<td>41.4</td>
<td>46.1</td>
<td>70.0</td>
<td>67.1</td>
<td>43.2</td>
</tr>
</tbody>
</table>

As long as the air path controller perfectly tracks the desired air path setpoints, the optimal fueling parameters can be predicted about as effectively with a set of conventional two-dimensional tables indexed by engine speed and total fuel mass, as with a set of three-dimensional tables indexed by engine speed, total fuel mass and in-cylinder oxygen concentration. This is evident in Table 5.1. Unlike conventional two-dimensional fueling tables, however, three-dimensional tables are able to predict the fueling parameters regardless of whether or not the air path controller achieves its setpoint targets. When all of the fueling parameter data (that is, not just the data corresponding to the optimal in-cylinder conditions) was used to determine the $R^2$ correlations, the relative fitness improved; this indicates that the variations in the optimal fueling parameters caused by changes in the in-cylinder conditions are well predicted by adding another dimension to the fueling tables.

The biggest limiting factor in fitting the fueling parameters using the cylinder contents variables are the size constraints of the tables. As the number of table entries is increased, the $R^2$ fit continues to increase. Conversely, the most prominent limiting factor in fitting the fueling parameters using a conventional two-dimensional
approach is that the same engine speed and fueling conditions can have different cylinder conditions even when the air path setpoints are perfectly achieved. In practice, the instantaneous values for the air path variables tend to lag their desired values thereby limiting the performance even further. For the heavy duty FTP drive cycle run experimentally, the $R^2$ correlation between the achieved values and desired values of the two air path setpoint variables, the charge flow rate into the intake manifold and fraction of EGR gases within the intake manifold, were 85.0% and 89.5%, respectively. The next section compares how air path tracking errors affect the performance of a conventional fueling controller versus an in-cylinder oxygen concentration based fueling controller.

5.3 Transient Performance Comparison

Following the control structure shown in Figure 5.2, an in-cylinder oxygen concentration based fueling controller has been implemented in GT-Power. Within this controller, the five fueling parameters (main SOI, pilot SOI, post SOI, pilot quantity and post quantity) are predicted with five sets of 10 x 10 x 10 tables indexed by the total fueling mass, the engine speed and the in-cylinder oxygen concentration. The in-cylinder oxygen concentration is dynamically predicted using a delay based oxygen dynamics model section whereas the exact values for the total fueling and engine speed were used directly. Although a LPV hybrid estimator was shown to be more effective, a delay based oxygen dynamics model was selected because it is much easier to implement into the GT-Power modeling environment. A conventional fueling controller based on five sets of 20 x 20 tables indexed by the total fueling mass and the engine speed has also been implemented into a separate GT-Power model.
Ideally, an air path controller would perfectly regulate the air path variables to their desired setpoints. To represent this scenario, the optimal air path actuator trajectories were directly specified and both fueling controllers were simulated in GT-Power for the first 100 seconds of a heavy duty FTP drive cycle. Under these conditions, both of the fueling control structures produced similar results. Figure 5.4 compares the instantaneous ISFC produced by the optimal fueling parameters to the instantaneous ISFC produced by the fueling parameters selected by the oxygen concentration based fueling controller and the conventional fueling controller. Similarly, Figure 5.5 compares the instantaneous BSNO\textsubscript{x} performance. The spans of events without data correspond to decelerations where no fuel was injected. Neither controller was able to perfectly predict the optimal fueling parameters, thus resulting in small difference in the instantaneous values of the ISFC and BSNO\textsubscript{x}. However, these differences do not significantly affect the overall drive cycle performance.

Figure 5.4: ISFC Comparison: Ideal air path response
Figure 5.5: BSNOx Comparison: Ideal air path response

The drive cycle averaged ISFC and BSNO\textsubscript{x} results are reported in Table 5.3. With respect to the benchmark set by the optimal fueling parameters, the average ISFC in percentage of full scale produced by the oxygen concentration based fueling controller was 0.44% larger and the average ISFC produced by the conventional fueling controller was 0.69% larger. Because of the inverse relationship between ISFC and BSNO\textsubscript{x}, both the oxygen concentration based fueling controller and the conventional fueling controller produced on average less NO\textsubscript{x} than the optimal fueling parameters. Recall that the optimization objective of the fueling parameter calibration was to minimize the ISFC under the constraint that the BSNO\textsubscript{x} is less than 1.0 g/hp-hr. Even though the two fueling controllers produced less NO\textsubscript{x} on average, the optimal fueling parameters met the BSNO\textsubscript{x} target more consistently. For the oxygen concentration based fueling controller, 1.07% fewer combustion events achieved a BSNO\textsubscript{x} value less
than 1.2 g/hp-hr, as compared to the optimal fuel parameters. For the conventional controller, 1.59% fewer combustion events met this criteria.

Table 5.3: Drive cycle averaged combustion performance results: Ideal air path response

<table>
<thead>
<tr>
<th></th>
<th>ISFC (% of full scale)</th>
<th>BSNOx (g/hp-hr)</th>
<th>% of data with BSNOx &lt; 1.2 g/hp-hr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal</td>
<td>29.75</td>
<td>0.970</td>
<td>83.6%</td>
</tr>
<tr>
<td>$[O_2]_{cyl}$ Based</td>
<td>30.19</td>
<td>0.942</td>
<td>82.5%</td>
</tr>
<tr>
<td>Conventional</td>
<td>30.44</td>
<td>0.939</td>
<td>82.0%</td>
</tr>
</tbody>
</table>

Although both the oxygen concentration based fueling controller and the conventional fueling control produced on average higher values of ISFC and exceeded a BSNO$_x$ value of 1.2 g/hp-hr more frequently, the magnitude of these differences was very small for both controllers. For the in-cylinder oxygen concentration based fueling controller, these results demonstrate that the sets of three-dimensional tables indexed by in-cylinder oxygen concentration, engine speed and total fuel mass effectively predict the fueling parameters and that the in-cylinder oxygen concentration prediction errors are practically negligible. In these types of ideal conditions where the trajectories of the air path setpoint variables are perfectly achieved, a conventional fueling controller with two-dimensional tables also effectively predicts the fueling parameters.

In practice, the instantaneous values of the air path setpoint variables will often differ from their desired values. Although a closed loop air path controller is designed to minimize this difference, many factors can limit the convergence rate or altogether prevent a desired value from being achieved. These factors include component failure, turbocharger surge/choke behavior, imposed AFR limits and, most importantly, the
dynamics of the air path system. To compare the robustness of an in-cylinder oxygen concentration based fueling controller to a conventional fueling controller with respect to these types of errors, the same drive cycle was simulated but with slightly different values air path responses. Uniform bias shifts were applied to the actuator position trajectories to produces these differences. A total of eight additional drive cycles were simulated for each controller.

![In-cylinder oxygen concentration variations: Cases A - E](image)

**Figure 5.6: In-cylinder oxygen concentration variations: Cases A - E**

The actuator shifts applied to the first four drive cycles labeled “A”, “B”, “C” and “D” decreased the EGR flow rate and increased the fresh air flow rate. The resulting effects on the in-cylinder oxygen concentration are shown in Figure 5.6. The ideal case which was tested previously is labeled as case “E”. For the remaining four tests labeled “F”, “G”, “H” and “I,” the shifts applied to the actuator commands caused
a relative increase in the EGR flow rate. Figure 5.7 depicts the in-cylinder oxygen concentrations resulting from these shifts. The types of uniform bias errors simulated in these tests are not meant to represent typical air path errors; rather, they were selected to better identify the robust behavior of the controllers.

![Figure 5.7: In-cylinder oxygen concentration variations: Cases E - I](image)

All of the drive cycle averaged results are compiled in Table 5.4. This same data is also visually depicted in Figure 5.8. These results clearly show that the two fueling controllers respond to actuator bias errors differently. Because a conventional fueling controller does not consider the state of the air path system, the conventional fueling controller responded exactly the same for each drive cycle. Although this response produced excellent performance when the air path setpoint trajectories are perfectly...
tracked, this same response produced poor performance when the air path setpoints were not achieved.

Table 5.4: Drive cycle averaged combustion performance results: Shifted air path actuator commands

<table>
<thead>
<tr>
<th>Case</th>
<th>$[O_2]_{cyl}$ (%)</th>
<th>ISFC ($%$ of full scale)</th>
<th>BSNO$_X$ (g/hp·hr) (%) of full scale</th>
<th>ISFC (g/hp·hr)</th>
<th>BSNO$_X$ (g/hp·hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>20.57</td>
<td>41.38</td>
<td>1.758</td>
<td>32.31</td>
<td>2.788</td>
</tr>
<tr>
<td>B</td>
<td>19.95</td>
<td>36.06</td>
<td>1.367</td>
<td>31.00</td>
<td>1.907</td>
</tr>
<tr>
<td>C</td>
<td>19.43</td>
<td>32.56</td>
<td>1.138</td>
<td>30.44</td>
<td>1.369</td>
</tr>
<tr>
<td>D</td>
<td>19.07</td>
<td>31.06</td>
<td>1.020</td>
<td>30.31</td>
<td>1.102</td>
</tr>
<tr>
<td>E</td>
<td>18.74</td>
<td>30.19</td>
<td>0.942</td>
<td>30.44</td>
<td>0.939</td>
</tr>
<tr>
<td>F</td>
<td>18.53</td>
<td>30.32</td>
<td>0.875</td>
<td>31.13</td>
<td>0.839</td>
</tr>
<tr>
<td>G</td>
<td>17.72</td>
<td>30.94</td>
<td>0.826</td>
<td>33.75</td>
<td>0.631</td>
</tr>
<tr>
<td>H</td>
<td>17.36</td>
<td>32.06</td>
<td>0.778</td>
<td>36.06</td>
<td>0.562</td>
</tr>
<tr>
<td>I</td>
<td>16.82</td>
<td>34.75</td>
<td>0.711</td>
<td>41.06</td>
<td>0.493</td>
</tr>
</tbody>
</table>

In general as the in-cylinder oxygen concentration decreases (i.e. diluent concentration increases), the heat release rate decreases. To offset this effect, the fuel injection timing should be advanced. Conversely, as the oxygen concentration increases (i.e. diluent concentration decreases), the combustion temperature tends to increase, thus increasing the NO$_X$ production rate. To prevent this increase, the fuel injection timing should be retarded. These types of corrective actions are not taken by a conventional fueling controller. For drive cycles (A, B, C, D) with actuator biases which cause an increase in the in-cylinder oxygen concentration, the conventional fueling controller consistently injected fuel sooner than it should resulting in higher ISFC values. On the other hand, for drive cycles (F, G, H, I) with actuator biases
which cause a decrease in the in-cylinder oxygen concentration, the conventional fueling controller consistently injected fuel later than it should, resulting in drastically larger BSNO\textsubscript{x} values.

With an in-cylinder oxygen concentration based fueling controller, the changes to the oxygen concentration caused by the air path actuator biases are correctly compensated. Based on the objective function used to calibrate the controller, the in-cylinder oxygen concentration based fueling controller produces a nearly constant NO\textsubscript{x} production rate of about 1 g/hp-hr. Under certain conditions, however, this target NO\textsubscript{x} production rate cannot be achieved through fuel injection adjustments alone. In these cases, the oxygen concentration based fueling controller selects the fueling parameters that minimize the BSNO\textsubscript{x} to as close as possible to 1.0 g/hp-hr.
For the drive cycles in which the EGR flow rate was decreased (A, B, C, D), these
types of conditions were encountered more frequently.

![Figure 5.9: Trade-off between cycle averaged BSNOₓ and cycle averaged ISFC](image)

The differences between a conventional fueling controller and an oxygen concentra-
tion based fueling controller become even more apparent when they are directly
compared as in Figure 5.9. In this figure, the drive cycle averaged (ISFC, BSNOₓ) data
pairs corresponding to the same actuator bias errors are joined together. When
the EGR flow rates were below their setpoint values (A, B, C, D), the in-cylinder
oxygen concentration based fueling controller produced less NOₓ each combustion
event relative to the conventional fueling controller, thus resulting in lower drive cy-
cle averaged BSNOₓ values. These lower instantaneous NOₓ production rates were
accompanied by slightly higher specific fuel consumption rates and thus higher cycle averaged ISFC values. When the EGR flow rates are above their setpoint values (F, G, H, I), each milligram of fuel injected by the oxygen concentration based fueling controller generated more torque than the conventional fueling controller, leading to lower drive cycle averaged ISFC values.

To further illustrate these effects, the ISFC and BSNO$_x$ values produced by the two controllers have been directly compared on an event-by-event basis to generate families of probability density functions. The magnitude of the difference in the ISFC and the BSNO$_x$ values achieved by the two fueling controllers depends on the magnitude of the air path setpoint errors. Figure 5.10 depicts how the distribution of the ISFC difference between the conventional fueling controller and the oxygen concentration based fueling controller varied as a function of the in-cylinder oxygen concentration difference (achieved value minus the value corresponding to the air path setpoints). The median ISFC difference is also included in this figure on the bottom plane (i.e. 0% probability). Each of the probability density functions corresponds to a particular the oxygen concentration difference. The random variable responsible for these variations corresponds to the combined effects of variation in the engine speed, the fuel injection mass and the cylinder conditions.

When the oxygen concentration difference for a given combustion event was approximately -2%, for example, the ISFC produced by the conventional controller was between 0% and 15% larger than the ISFC produced by the oxygen concentration based fueling controller for 91.9% of the occurrences. This same information can extracted from Figure 5.10 by integrating the probability density function curve corresponding to an oxygen concentration difference of -2% from 0% to 15%. Each of
Figure 5.10: Instantaneous ISFC difference (Conventional - \([O_2]_{\text{cyl}}\) based) probability density functions

The distributions have approximately the same degree of narrowness indicating that the compensatory adjustments made by an in-cylinder oxygen concentration based fueling controller influence the combustion performance in a consistent manner.

A similar set of probability density functions shown in Figure 5.11 were created to illustrate how the BSNO\(_x\) difference between the conventional fueling controller and the oxygen concentration based fueling controller varies as function of the in-cylinder oxygen concentration difference. These probability density functions are also characterized by very narrow distributions. For negative oxygen concentration differences, the two controllers tended to produce approximately the same BSNO\(_x\).
the oxygen concentration difference is positive, an in-cylinder oxygen concentration based fueling control produced significantly less NO\textsubscript{x}. In fact, the magnitude of these changes was frequently larger than 1.0 g/hp-hr. These conditions correspond to combustion events where the oxygen concentration based fueling controller remained near the BSNO\textsubscript{x} target of 1.0 g/hp-hr while the conventional fueling controller produced more than double that same target.

Consider Figures 5.12 and 5.13 which characterize for the oxygen concentration based fueling controller and the conventional fueling controller, respectively, with the occurrence frequency on a logarithmic scale of the (ISFC, BSNO\textsubscript{x}) pair encountered.
across all of the drive cycle simulations. To put this data in perspective, the drive cycle averaged (ISFC, BSNO\textsubscript{X}) pairs originally presented in Figure 5.9 have also been superimposed into these distribution charts as white curves. The combustion produced by the in-cylinder oxygen concentration based controller most frequently resulted in (ISFC, BSNO\textsubscript{X}) pairs within the range \([18, 35] \times [0.4, 1.4]\). Conversely, the combustion produced by the conventional controller most frequently resulted in a wider range of (ISFC, BSNO\textsubscript{X}) pairs contained within the box outlined by \([18, 40] \times [0.1, 1.7]\). Moreover, the NO\textsubscript{X} produced by the oxygen concentration based fueling controller for almost all of the combustion events was below 4.5 g/hp-hr, whereas this BSNO\textsubscript{X} exceeded 4.5 g/hp-hr on numerous occasions with the conventional fueling controller.

Table 5.5: Drive cycle averaged combustion performance results: Dynamic air path actuator errors

<table>
<thead>
<tr>
<th>Case</th>
<th>([O_2]\text{cyl} ) Based ISFC (g/hp-hr)</th>
<th>([O_2]\text{cyl} ) Based BSNO\textsubscript{X} (g/hp-hr)</th>
<th>Conventional ISFC (g/hp-hr)</th>
<th>Conventional BSNO\textsubscript{X} (g/hp-hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>J</td>
<td>18.55 32.75 0.925 33.38 1.007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>18.67 33.44 0.972 33.50 1.115</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>18.12 34.38 0.982 36.88 1.040</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>18.70 35.69 1.133 36.06 1.745</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As opposed to the constant bias errors tested thus far, most air path errors are transient in nature. The relative performance of the two fueling controllers depends on the air path trajectory, so that an infinite number of air path error trajectories is possible. To demonstrate some of the possible outcomes, a few hypothetical dynamic air path error trajectories have been simulated. The first two cases labeled “J” and...
Figure 5.12: Distribution of (ISFC, BSNO$_x$) pairs for the oxygen concentration based fueling controller

“K” had dynamic errors which are approximately of the same order as those seen experimentally. The second two cases labeled “L” and “M” contained larger dynamic errors (both in amplitude and duration). The drive cycle averaged performance results for these drive cycles is assembled in Table 5.5. The cycle averaged (ISFC, BSNO$_x$) pairs are also plotted in Figure 5.14. Included for reference are the cycle averaged data pairs for drive cycles A through I.
For each of the four drive cycles, the in-cylinder oxygen concentration based fueling controller produced more efficient combustion than the conventional fueling controller in terms of both fuel consumption and NO\textsubscript{x} production. Between the two cases which simulate moderate dynamic air path errors, the cycle averaged ISFC achieved by the oxygen concentration fueling controller relative to the conventional fueling controller in percentage of full scale was as large as 0.63\% and the largest decrease in the cycle averaged BSNO\textsubscript{x} was 0.143 g/hp-hr. The larger dynamic errors simulated in cases L and M produced even larger relative changes. Under these conditions, the cycle averaged ISFC produced by the oxygen concentration based fueling controller

Figure 5.13: Distribution of (ISFC, BSNO\textsubscript{x}) pairs for the conventional fueling controller
Figure 5.14: Trade-off between cycle averaged BSNO$_x$ and cycle averaged ISFC: Dynamic air path errors in percentage of full scale was as much as 2.50% lower than the conventional fueling controller and the average BSNO$_x$ was decreased by much as 0.612 g/hp-hr.
5.4 Summary

A fueling controller should be scheduled on the variables which most directly affect the combustion process. When only two scheduling variables are used, variations in air path conditions are neglected. During normal engine operation, the mass storage and transport dynamics can be significant. Factoring in the air path actuator dynamics and the controller response rate, the instantaneous errors between the desired values of the air path setpoint variables and their actual values can be considerable. With a conventional fueling controller, these errors produce combustion which has an undesirable trade-off between BSNO$_x$ and ISFC. In some cases, the torque production decreases, whereas in others the NO$_x$ production rate increases.

One of the most significant factors influencing the combustion process is the in-cylinder oxygen concentration. Enabled by the dynamic estimators presented in Section 4.1, an in-cylinder oxygen concentration based fueling controller is able to produce combustion with a desirable BSNO$_x$ versus ISFC trade-off for all cylinder conditions. In steady-state, the performance of an oxygen concentration based fueling controller is equivalent to a conventional controller. During a transient, however, the robustness provided by an oxygen concentration based fueling controller enables a smoother torque delivery while maintaining an acceptable BSNO$_x$ level. Even when a component within the air path system fails or when exceptional control action is taken to prevent the turbocharger system from reaching its surge or choke limits, an oxygen concentration based fueling controller maintains a high level of performance.

Calibrating a conventional diesel engine controller requires a significant amount of data because for every system operating mode both the air path setpoints and the fueling parameters must be optimized. The in-cylinder oxygen concentration based
fueling controller described in this chapter uses the same type of air path control structure as a conventional diesel engine controller. Although separate sets of air path setpoint tables are still required for each system operating mode, only a single set of fueling tables is required.

To calibrate the transient air path setpoint and accompanying two-dimensional transient fueling tables of a conventional diesel engine controller, a wide range of air path setpoints and fueling parameter combinations must be tested. With this same data, the transient air path setpoint tables and all of the three-dimensional fueling parameter tables of an in-cylinder oxygen concentration based fueling controller could be calibrated. After the three-dimensional fueling tables have been calibrated, the remaining air path setpoint tables can be calibrated with relative ease. Since the optimal fueling parameters are already known, the air path tables can be chosen by directly optimizing the air path setpoints. Conversely with a conventional engine control structure, both the air path setpoints and the fueling parameters must still be jointly optimized. Because an in-cylinder oxygen concentration based fueling controller more effectively uses the calibration data, the overall calibration burden is smaller than a conventional diesel engine controller.
CHAPTER 6

AFR Switching Period Control in a Gasoline Engine

Several research studies including [19], [20], [21] and [22] have found that for constant flow rates the conversion efficiency of a TWC catalyst depends on the frequency, amplitude and shape of the inlet exhaust gas AFR trajectory. These findings motivated the design of a unified pre-catalyst/post-catalyst AFR control architecture which strives to track a periodic waveform that has a desired amplitude, frequency, shape and offset instead of trying to regulate the AFR to a constant value. This control architecture uses two switching type oxygen sensors to measure the pre-catalyst and post-catalyst AFR. EGO sensors were chosen over UEGO sensors because they have a faster response time, higher sensitivity near stoichiometry and a lower cost. For the pre-catalyst (inner) loop, a novel switching frequency regulation design is utilized and a setpoint regulation design is used for the post-catalyst (outer) loop.

By actively controlling the AFR fluctuations around stoichiometry, typically un-controlled in regulation controllers, the stored oxygen mass within the TWC can be maintained at its optimal level, thus maximizing the catalyst conversion efficiency. To complement the steady-state studies cited earlier, the proposed control architecture was used to study how different dynamic AFR trajectories affect the conversion efficiency of a TWC under realistic driving conditions where the exhaust composition and
exhaust flow rate are continuously changing. These effects were determined by changing the amplitude, frequency and the shape of the periodic AFR waveform which the controller strives to track. The results of this study show that two equivalent controllers which track different waveforms can produce significantly different tailpipe emissions distributions even when the two controllers achieve nearly the same precatalyst tracking performance. From these studies the waveform which best promotes high TWC conversion efficiency was identified. Since the proposed AFR controller can track any periodic signal, this same waveform was chosen for the final controller design.

To effectively track this waveform, the proposed switching frequency based precatalyst AFR controller must be able to reject a wide range of disturbances. Within a switching frequency based control structure, disturbances are detected by comparing the number of events between output switches to some desired value. Even though magnitude information is not provided by a switching EGO sensor, the average AFR can be inferred from the switching period. Because the feedback signal is actually the switching period, not the switching frequency, this controller will henceforth be referred to as an AFR switching period controller.

In addition to having a control objective that better corresponds to high catalyst conversion efficiency, this control architecture is more robust than current model based controllers and requires significantly less calibration effort than the direct tuning methods used by production manufacturers. Without an accurate model of the delay between injecting fuel and measuring the oxygen concentration of the exhaust stream with a pre-catalyst oxygen sensor, traditional model based techniques are only marginally effective. The high fidelity exhaust oxygen dynamics model described in
Section 3.3 accurately predicts this delay but like the lumped oxygen dynamics model, it is too complex to run in real time. By casting the AFR control problem in terms of a switching period target, the time varying plant delay is accounted for without the need for a traditional estimator thereby making its performance invariant to the transport delay.

Even though a version of the HFEOD model is not incorporated into this control structure, this model enables better performance through other means. Using a HFEOD model and other subsystem level models of an engine, all of the control gains for the inner and outer loop controllers are optimized in simulation. Because the oxygen mixing and transport dynamics are captured so well by a HFEOD model, the performance predicted in simulation closely resembles the observed experimental performance. With this model based calibration approach, the experimental testing requirements are at least an order of magnitude smaller than current production approaches.

Unlike current controllers, an AFR switching period controller is guaranteed to remain stable for all bounded disturbances. Additionally, the maximum number of events between output switches is at worst bounded by an affine function of the maximum disturbance magnitude. These claims will be proven analytically and demonstrated experimentally in this chapter. Compared to the existing production AFR controller, the AFR switching period controller rejects disturbances more effectively and produces fewer tailpipe emissions.
6.1 Additional Nomenclature

Characterizing an AFR switching period controller requires definition of numerous variables. The symbols which are typically used to define the inputs, outputs and gains of such a model have already been utilized by the models and estimators described in the previous chapters. Avoiding all of the possible variable redundancies is very challenging; therefore, to avoid confusion, a new nomenclature section specific to the variables which are used in chapter is now presented. All previously defined variables which are still relevant are also included for completeness.

**GENERAL VARIABLES**

Δ\text{switch},e \quad \text{Relative earliness of an output switching event.}

Δ\text{switch},l \quad \text{Relative lateness of an output switching event.}

Ω \quad \text{Closed set.}

Υ\text{e} \quad \text{Function which dictates when early switching control action is applied.}

Υ\text{l} \quad \text{Function which dictates when late switching control action is applied.}

Υ\text{square} \quad \text{Function which dictates when the once per late switch correction is applied (square wave only).}

α\text{switch} \quad \text{Output switching interpolation factor.}

δ \quad \text{Size of the switching deadzone.}

ε\text{switch} \quad \text{Number of events between output switches (only calculated when switch occurs).}
$\epsilon_{\text{switch},m}$ Number of events between output switches for the case of monotonic convergence.

$\epsilon_{\text{switch},nm}$ Number of events between output switches for the case of non-monotonic convergence.

$\gamma$ Early switching compensator gain.

$\kappa_{\text{min}}$ Minimum monotonic convergence rate.

$\lambda_{\text{min}}$ Minimum eigenvalue.

$\lambda_{\text{max}}$ Maximum eigenvalue.

$\phi_{\text{corr}}$ Closed loop corrected EQR.

$\phi_{\text{cat,in}}$ EQR entering the catalyst.

$\phi_{\text{cat,out}}$ EQR exiting the catalyst.

$\phi_{O_2}$ Predicted EQR at the oxygen sensor.

$\phi_{\text{post}}$ Predicted post-catalyst EQR measurement.

$\rho$ Maximum accumulated error based on the disturbance changing rate.

$\theta_m$ Maximum number of events before an output switch for the case of monotonic convergence.

$\theta_{nm}$ Maximum number of events before an output switch for the case of non-monotonic convergence.

$\zeta_{EQR}$ Variable which indicates whether the output switches from rich-to-lean or lean-to-rich.

$A$ EQR amplitude of the imposed periodic dither waveform.

$\text{AFR}_s$ Stoichiometric air-to-fuel ratio.

$K_e$ Integral control gain of the early switching compensator.
$K_{fast}$ Fast filter constant which corresponds to an oxygen storage mass which is at one its maximum bounds.

$K_i$ Integral control gain of the late switching compensator.

$K_{i,\text{max}}$ Upper bound on the integral control gain of the late switching compensator.

$K_{i,\text{min}}$ Lower bound on the integral control gain of the late switching compensator.

$K_{post}$ Post-catalyst filter gain.

$K_{slow}$ Slow filter constant which corresponds to an oxygen storage mass which is within its maximum bounds.

$K_{V,\text{switch}}$ Switching voltage control gain

$M$ EQR offset magnitude.

$M_\delta$ Effect of the deazone on the EQR offset magnitude.

$O$ Sustained EQR offset.

$O_s$ Mass of oxygen stored within the TWC.

$O_{\text{max}}$ Maximum mass of oxygen that can be stored within the TWC.

$[O_2]_{\text{air}}$ Oxygen concentration of air on a mass basis.

$P_l$ Proportional control gain of the late switching compensator.

$P_{l,\text{max}}$ Upper bound on the proportional control gain of the late switching compensator.

$P_{\text{square}}$ Gain applied to the once per late switch correction factor (square wave dither only).

$\Delta T$ Switching period adjustment.

$\Delta T_{\text{lower}}$ Switching period adjustment lower bound.
$\Delta T_{\text{upper}}$ Switching period adjustment upper bound.

$T_{\text{dither}}$ Period of the imposed periodic dither waveform.

$T_{\text{dither, nom}}$ Nominal period of the imposed periodic dither waveform.

$T_{\text{lean}}$ Period of the lean portion of the imposed periodic dither waveform.

$T_{\text{rich}}$ Period of the rich portion of the imposed periodic dither waveform.

$\Delta V$ Switching voltage adjustment.

$\Delta V_{\text{lower}}$ Switching voltage adjustment lower bound.

$\Delta V_{\text{upper}}$ Switching voltage adjustment upper bound.

$V_{\text{switch}}$ Pre-catalyst switching voltage threshold.

$V_{\text{post}}$ Measured post-catalyst EGO voltage.

$V_{\text{post, des}}$ Desired post-catalyst EGO voltage.

$V_{\text{pre}}$ Measured pre-catalyst EGO voltage.

$V_{\text{pre, nom}}$ Nominal pre-catalyst switching voltage.

$a$ Closed set magnitude bound.

$c$ True EQR of the combustion mixture.

$d$ Lumped EQR disturbance variable.

$d_{\text{max}}$ Upper amplitude bound on the lumped EQR disturbance variable.

$d_{\text{exh}}$ Transport delay of the exhaust manifold system.

$e_{\text{acc, m}}$ Largest accumulated error before late switching controller acts for the monotonic convergence case.
\( e_{acc,nm} \) Largest accumulated error before late switching controller acts for the non-monotonic convergence case.

\( e_c \) Maximum error caused by the early switching controller.

\( e_{max} \) Maximum EQR error.

\( e_{max,m} \) Maximum EQR error for the monotonic convergence case.

\( e_{max,nm} \) Maximum EQR error for the non-monotonic convergence case.

\( e_{o,p} \) Peak overshoot caused by the late switching controller.

\( e_{o,p,max} \) Maximum peak overshoot caused by the late switching controller.

\( e_{o,s} \) Maximum sustained overshoot caused by the late switching controller.

\( f \) Maximum offset error function.

\( i o_j \) \( j^{th} \) coefficient of the FIR model of the exhaust system.

\( k_{\text{switch}} \) Event number in which the last output switch occurred.

\( m_{\text{fuel,des}} \) Desired trapped fuel mass.

\( m_{\text{air,trapped}} \) Estimated trapped unburned air mass.

\( \dot{m}^{*}_{\text{exh}} \) Exhaust mass flow rate.

\( n_{\text{first}} \) Number of events before the effect of a given fueling event first appears in the output EGO measurement.

\( n_{last} \) Number of events before the complete effect of a given fueling event appears in the output EGO measurement.

\( n_{\text{switch}} \) Number of event since the last switching event.

\( n_{\text{window}} \) Size of the output window used by the FIR model.

\( o \) Desired EQR offset.
\( p \)  
Proportional contribution from the late switching compensator.

\( r_{\text{max}} \)  
Maximum amplitude change in the lumped EQR disturbance variable over one engine event.

\( u \)  
EQR contribution from the feedback controller.

\( v \)  
Combined integral contributions from the early and late switching compensators.

\( w \)  
EQR of the imposed periodic waveform.

\( x \)  
Difference between true EQR of the combustion mixture and the desired EQR.

\( x_{\text{avg}} \)  
Average EQR offset error.

\( y \)  
EQR measured by the pre-catalyst EGO sensor (either rich or lean).

\( z \)  
Closed loop corrected EQR.

**INDEX VARIABLES**

\( \theta \)  
Generic event offset.

\( i \)  
Generic index.

\( j \)  
Generic index.

\( k \)  
Engine event index.

\( s \)  
EGO switching event index.

**SPECIAL MODIFIERS**

\( \ast \)  
Special value of variable.

\( \_ \)  
Variable in the switching event domain.
6.2 Pre-Catalyst AFR Switching Period Controller Description

If the plant delay was known perfectly and the pre-catalyst EQR was measured perfectly, the AFR control problem posed in this dissertation could be formulated into a tracking control problem. In this ideal case, the measured AFR would be compared to a predicted AFR to generate an error signal. Many well known techniques such as the internal modeling principal could be used to develop a controller that asymptotically drives this error to zero. This type of control is only effective if the plant delay is predicted exactly, because phasing errors brought on by delay estimation error cannot be corrected. In direct contrast to the control objective, these phase errors result in uncontrolled limit cycle behavior. As shown in Section 3.4.1, the plant delay varies markedly and is very difficult to predict, especially during transients.

Wide range UEGO sensors have larger time constants, compared to switching EGO sensors, limiting the response speed that a controller can achieve. Although a switching EGO sensor has a quicker response, only a small range of AFR values near stoichiometry can be uniquely correlated to a measurement voltage. This limited range only further complicates the traditional tracking problem. By recasting the problem into a switching period control problem, it is possible to track a desired AFR trajectory in terms of amplitude, frequency and offset. In such a scheme, the measured AFR is not directly compared to a predicted AFR, so the phasing and delay estimation issues are eliminated.

The switching period based AFR controller proposed in this chapter makes use of the fact that the average AFR (or equivalently the AFR offset error) can be estimated by observing the number of consecutive events before the measured AFR switches
from rich to lean or lean to rich. The control architecture consists of a feed-forward controller, a feedback controller based on a pre-catalyst EGO sensor, and setpoint generator based on a post-catalyst EGO sensor. All of the AFR control laws are developed in the fuel/air equivalence ratio (EQR) domain as defined by (2.25). The output of the AFR switching period controller is a corrected EQR signal. This signal is the desired EQR corrected to overcome air and fuel estimation errors. To generate a fuel command for the injection system, the corrected EQR is converted into a desired trapped fuel mass using

\[
m_{\text{fuel,des}} = \frac{\dot{m}_{\text{air, trapped}} \phi_{\text{corr}}}{\text{AFR}_s}
\]

where \(m_{\text{fuel,des}}\) is the desired trapped fuel mass, \(\dot{m}_{\text{air, trapped}}\) is the estimated trapped air mass and \(\phi_{\text{corr}}\) is the corrected EQR. Following this, a fuel dynamics compensator adjust the injected fuel mass to ensure that the desired fuel mass is indeed the trapped fuel mass. Because of its modular design, any type of air estimation algorithm and any type of fuel dynamics compensator can be coupled with an AFR switching period controller.

The overall control scheme in block diagram form is shown in Figure 6.1. At the top of this diagram, a feed-forward controller determines the desired EQR which is composed of two components: a setpoint EQR and a periodic dither signal. In the middle of the diagram, a feedback controller monitors the pre-catalyst EGO sensor measurement and determines when the output switches from rich to lean and lean to rich. The richness or leanness of the output is determined in the switching detector block by comparing the measured pre-catalyst oxygen sensor voltage to the switching voltage threshold. If the measured voltage is larger than the threshold, then the mixture is considered rich; otherwise, the mixture is considered lean.
Figure 6.1: Pre-catalyst switching period and post-catalyst setpoint controllers.
Corrections are made when the output switches either faster than desired via the early switching compensator or slower than desired via the late switching compensator. The early switching compensator makes a one-time integral correction to the EQR offset whenever the output switches earlier than the desired switching period. The late switching compensator makes both proportional and integral corrections if more than the desired number of events have passed since the last output switching event. The late switching controller will continue to make corrections until the output switches. The desired switching period directly depends on the period of the reference dither signal.

Two different post-catalyst setpoint controllers shown on the left portion of the control diagram are compared in this chapter. In the first controller, the average EQR is directly controlled by concurrently adjusting both the EQR setpoint of the feed-forward controller and the rich/lean voltage threshold of the feedback controller. Alternatively, the second post-catalyst controller adjusts the average EQR by changing the switching periods of the rich and lean portions of the dither signal.

By developing the AFR switching period controller in the EQR domain, the complex air and fuel dynamics can be pulled outside of the control model. Obviously the accuracy of the fuel and air estimations are fundamental to the performance of any AFR controller. With this structure, however, these errors are viewed as additive input disturbances. The corrected EQR signal fed to the fuel dynamics compensator can be described by

\[ z(k) = o(k) + w(k) + u(k) \]  

(6.2)

where \( z \) is the corrected EQR previously defined as \( \phi_{corr} \), \( o \) is the desired EQR offset, \( w \) is the periodic dither waveform, \( u \) is the contribution from the feedback controller.
and $k$ is the time index in engine events. An engine event represents a subset of
the crank angle domain where the number of events per engine cycle is equal to the
number of cylinders. Although any periodic signal can be tracked by switching period
control, this section focuses on case where the dither is a sinusoid. When the periodic
dither waveform is a sinusoid, it can be described by

$$w(k) = A \sin \left( \frac{2\pi}{T_{\text{dither}}} k \right)$$ (6.3)

where $T_{\text{dither}}$ is the dither period in events and $A$ is the dither amplitude in EQR.
Assuming that fuel is injected one event prior to the induction event, the true EQR
of the combustion mixture, $c$, can be represented as

$$c(k + 1) = z(k) + d(k)$$ (6.4)

where $d$ is a lumped disturbance variable which represents the combined effect of
air estimation errors, fuel estimation errors, sensor errors and unknown contributions
from the evaporative canister system.

For simplicity, the feedback control law defines the state $x$ as the difference between
the true EQR and the desired EQR setpoint as in

$$x(k + 1) = c(k + 1) - o(k) = d(k) + w(k) + u(k).$$ (6.5)

As derived in Section 3.4, the measured pre-catalyst EQR can be modeled using the
FIR relationship described by (3.78). This same relationship can be equivalently
described in terms of $x$ as

$$\phi_{O_2}(k) = \sum_{i=1}^{n_{\text{window}}} i o_i(k) x(k - i).$$ (6.6)
The relative composition of the exhaust stream as measured by the pre-catalyst EGO sensor can therefore be defined as

\[ y(k) = \text{sign} \left( \sum_{i=1}^{n_{\text{last}}} io_i(k)x(k-i) \right) \]  

(6.7)

where \( n_{\text{last}} < n_{\text{window}} \) is the largest nonzero FIR coefficient. It should be noted that the FIR relationship described above is only used to prove the stability of the control system; the proposed AFR controller does contain any form of this model. The stability analysis presented in this paper is based on the fact that the plant delay is bounded by a maximum value of \( n_{\text{last}} \) events.

The switching period controller can be represented with the following:

\[ u(k) = v(k) + p(k) \]  

(6.8)

where \( p \) is the proportional contribution from the late switching compensator and \( v \) is the combined integrator contributions from the early and late switching compensator. The proportional correction is defined as

\[ p(k) = P_l \Upsilon_l(n_{\text{switch}}(k))y(k) \]  

(6.9)

and the integral correction is defined as

\[ v(k) = v(k-1) - K_l \Upsilon_l(n_{\text{switch}}(k)) y(k) - K_e(\epsilon_{\text{switch}}(k)) \Upsilon_e(\epsilon_{\text{switch}}(k)). \]  

(6.10)

In the above equations, both the late switching proportional gain \( P_l \) and the late switching integral gain \( K_l \) are smooth functions of the operating conditions (intake manifold pressure and engine speed). The function \( \Upsilon_l : \mathbb{Z}^+ \mapsto \{0, 1\} \) dictates when late switching control action is applied according to

\[ \Upsilon_l(n_{\text{switch}}(k)) = \begin{cases} 1 & \text{if } n_{\text{switch}}(k) > \left( \frac{T_{\text{dither}}}{2} + \delta \right) \\ 0 & \text{otherwise} \end{cases} \]  

(6.11)
where \( n_{\text{switch}} \) is the number of events since the last switching event and \( \delta \) is the size of the deadzone. This deadzone, which also appears in the early switching compensator, allows a small range of switching periods to produce zero control error, and must satisfy \( 0 \leq \delta < \frac{T_{\text{dither}}}{2} \). This improves the convergence rate of the controller at the potential cost of small steady-state errors. The contribution from the late switching compensator is only adjusted if more than \( \frac{T_{\text{dither}}}{2} + \delta \) events have passed since the last output switching event. Once activated, the late switching controller continuously adjusts the EQR command until the output finally switches.

The function \( \Upsilon_e : \mathbb{Z}^+ \mapsto \{0, 1\} \) seen in (6.10) dictates when early switching control action is applied according to

\[
\Upsilon_e(\epsilon_{\text{switch}}(k)) = \begin{cases} 
1 & \text{if } \epsilon_{\text{switch}}(k) < \left( \frac{T_{\text{dither}}}{2} - \delta \right) \\
0 & \text{otherwise}
\end{cases}
\]  

(6.12)

where \( \epsilon_{\text{switch}} \) is a variable that is nonzero only when the output switches; when the output switches,

\[
\epsilon_{\text{switch}}(k) = \begin{cases} 
k - k_{\text{switch}} & \text{if EGO switches} \\
0 & \text{otherwise}
\end{cases}
\]  

(6.13)

where \( k_{\text{switch}} \) is the event number of the previous output switch. The early switching compensator acts only if the number of events between output switches is less than \( \frac{T_{\text{dither}}}{2} - \delta \). Unlike the late switching compensator, the early switching compensator makes a maximum of only one correction per early switching event. Let \( x_{\text{avg}} \) be the average EQR offset error defined by

\[
x_{\text{avg}}(k) = \frac{1}{T_{\text{dither}}} \sum_{i=1}^{T_{\text{dither}}} x(k - i).
\]  

(6.14)

In the presence of a constant offset error that has a magnitude less than the dither amplitude, the number of events between EGO switchings for the open loop system
can be directly related to the average EQR offset error using

\[ x_{\text{avg}}(k) = \zeta A \cos \left( \frac{\pi \epsilon_{\text{switch}}(k)}{T_{\text{dither}}} \right) \]  

(6.15)

where

\[ \zeta = \begin{cases} -1 & \text{if rich-to-lean switch} \\ 1 & \text{if lean-to-rich switch} \end{cases} \]  

(6.16)

Exploiting this relationship, the early switching control gain is defined by

\[ K_e(\epsilon_{\text{switch}}(k)) = \gamma \zeta A \cos \left( \frac{\pi \epsilon_{\text{switch}}(k)}{T_{\text{dither}}} \right) \]  

(6.17)

where \( 0 < \gamma < 1 \).

### 6.3 Post-Catalyst Setpoint Controller Schemes

An AFR switching controller strives to track an oscillating pre-catalyst EQR signal which has a nominal EQR of unity (stoichiometry). To correct for sensor bias errors and regulate the oxygen storage of the catalyst, feedback from a post-catalyst EGO sensor must be used. Recall that the conversion efficiency of a TWC depends on the mass of oxygen stored within the catalyst. As long as the stored oxygen mass is less than the maximum oxygen storage capacity level and more than zero (i.e. completely depleted of oxygen), the effective conversion efficiency remains high.

Defining the oxygen storage ratio as the ratio of the current stored oxygen mass to the maximum stored oxygen mass, a post-catalyst controller indirectly attempts to regulate the oxygen storage ratio to approximately one half. With an average value of one half, the potential of a TWC to convert the harmful emissions produced from lean operation is equal to the potential to convert the harmful emissions produced from rich operation. The only way to observe the state of the catalyst is through the post-catalyst oxygen sensor which effectively reads stoichiometry whenever the
oxygen storage ratio is not zero or one. All storage ratios between zero and one produce approximately the same output and thus make the oxygen storage state very difficult to observe.

Although the oxygen storage capacity is effectively unobservable, it can be regulated albeit indirectly using feedback from a post-catalyst EGO sensor. When the oxygen storage of the catalyst is within its minimum and maximum bounds, a post-catalyst EGO sensor nominally reads the voltage that corresponds to stoichiometry. If the EGO sensor reads higher than this voltage, then the nominal pre-catalyst EQR needs to be reduced to increase the oxygen stored within the catalyst. Conversely if the EGO sensor reads lower than this voltage, then the nominal pre-catalyst EQR needs to be increased to reduce the oxygen storage level of the catalyst. In the proposed architecture, this is accomplished with a post-catalyst controller that adjusts the setpoints of the pre-catalyst switching period controller based on a voltage error between the measured post-catalyst EGO voltage and the desired voltage. In general the desired voltage should be approximately equal to the stoichiometric EGO voltage. Post-catalyst control via two different setpoints, switching voltage and rich/lean switching period ratio, has been tested.

6.3.1 Switching Voltage Setpoint Controller

For a very small range near stoichiometry, the voltage output of an EGO sensor is proportional to the EQR of the exhaust mixture. Therefore by adjusting the voltage threshold which defines a switch, the average pre-catalyst EQR can be controlled. As the switching voltage is increased, the average pre-catalyst EQR increases; the average pre-catalyst EQR decreases as the switching voltage is decreased. With this
control structure, the switching voltage can be described using

\[ V_{\text{switch}}(k) = V_{\text{pre,nom}} + \Delta V(k) \]  

(6.18)

where \( V_{\text{switch}}(k) \) is the current switching voltage, \( V_{\text{pre,nom}} \) is the nominal pre-catalyst switching voltage and \( \Delta V(k) \) is the adjustment made by the post-catalyst controller. The adjustment made by the post-catalyst switching voltage controller is governed by

\[ \Delta V(k) = \Delta V(k - 1) + K_{V_{\text{switch}}}(V_{\text{post,des}} - V_{\text{post}}(k - 1)) \]  

(6.19)

where \( K_{V_{\text{switch}}} \) is the switching voltage control gain, \( V_{\text{post,des}} \) is the desired post-catalyst EGO voltage and \( V_{\text{post}} \) is the measured post-catalyst EGO voltage. The control authority of the post-catalyst controller must be limited, so in addition to choosing a small gain for \( K_{V_{\text{switch}}} \) the post-catalyst voltage adjustment is saturated according to

\[ \Delta V(k) = \begin{cases} 
\Delta V_{\text{upper}} & \text{if } \Delta V(k) > \Delta V_{\text{upper}} \\
-\Delta V_{\text{lower}} & \text{if } \Delta V(k) < -\Delta V_{\text{lower}}
\end{cases} \]  

(6.20)

where \( \Delta V_{\text{upper}} > 0 \) is the largest allowable positive switching voltage change and \( \Delta V_{\text{lower}} > 0 \) is the largest allowable negative switching voltage change.

### 6.3.2 Rich/Lean Switching Period Ratio Controller

The average pre-catalyst EQR can also be controlled by adjusting the rich/lean switching period ratio of the dither waveform. For the nominal case, the waveform spends one half of its period rich of stoichiometry and one half of its period lean. By increasing the number of rich events and correspondingly decreasing the number of lean events, the average pre-catalyst EQR can be increased. Similarly, the pre-catalyst EQR can be decreased by decreasing the number of rich events and correspondingly
Figure 6.2: Effect of the rich/lean switching period ratio setpoint on the dither signal increasing the number of lean events. The waveform still retains the same nominal dither period, $T_{dither,nom}$, but the switching point is either shifted sooner or later. Each half of the dither signal still maintains the same shape but the period of each section changes. The effect of the rich/lean switching period ratio setpoint for a sinusoidal dither with a 20 event period is shown in Figure 6.2.

With this control architecture, the dither period can be adjusted according to

$$T_{dither}(k) = \begin{cases} T_{rich}(k) & \text{if } V_{pre}(k) > V_{switch} \\ T_{lean}(k) & \text{if } V_{pre}(k) < V_{switch} \end{cases}$$ (6.21)

where $T_{rich}$ is the period of the rich portion of the dither signal and $T_{lean}$ is the period of the lean portion of the dither signal. The rich portion of the dither signal is directly controlled by the post-catalyst controller by

$$T_{rich}(k) = T_{dither,nom} + \Delta T(k)$$ (6.22)
where $\Delta T$ is the contribution from the post-catalyst rich/lean switching period ratio setpoint controller. The corresponding control law is

$$
\Delta T(k) = \Delta T(k-1) + K_{T_{\text{post}}} (V_{\text{post,nom}} - V_{\text{post}(k-1)})
$$

(6.23)

where $K_{T_{\text{post}}}$ is the rich/lean switching period ratio setpoint control gain. To limit the control authority, the rich/lean switching period ratio setpoint correction is saturated according to

$$
\Delta T(k) = \begin{cases} 
\Delta T_{\text{upper}} & \text{if } \Delta T(k) > \Delta T_{\text{upper}} \\
-\Delta T_{\text{lower}} & \text{if } \Delta T(k) < -\Delta T_{\text{lower}} 
\end{cases}
$$

(6.24)

where $\Delta T_{\text{upper}} > 0$ is the largest positive change in the rich period and $\Delta T_{\text{lower}} > 0$ is the largest negative change in the rich period. To ensure the resulting signal has a constant period, the dither period of the lean portion of the dither cycle is constrained by

$$
T_{\text{lean}}(k) = 2T_{\text{dither,nom}} - T_{\text{rich}}(k)
$$

(6.25)

or equivalently

$$
T_{\text{lean}}(k) = T_{\text{dither,nom}} - \Delta T(k).
$$

(6.26)

### 6.4 Stability Study

In this control design, air estimation errors and fueling errors, including those resulting from the fuel dynamics compensator, are treated as a single additive EQR disturbance. Thus the closed inner loop control system has three internal states: the air estimation error, the net fuel air and the integral term of the controller, $v$. Physically, both types of errors are bounded, so if the pre-catalyst EQR trajectory of the inner loop control system is bounded, then the internal state $v(k)$ must also be bounded. The analysis presented in this section focuses on the behavior of the pre-catalyst EQR trajectory when the dither period and switching voltage setpoints are
constant. The analysis in the first subsection considers the stability of the closed-loop system affected by a constant disturbance (or equivalently a non-zero initial condition for the state $x(k)$), whereas the second subsection characterizes the robust stability of the system with respect to time varying disturbances.

### 6.4.1 Global Stability and Convergence

When only the sign of an error signal is known, magnitude information must be inferred using other means. With switching period control, offset errors are deduced from the number of events between output switches. Magnitude information can only be estimated, however, for sufficiently small offset errors. When an offset error is too large, the controller can only make corrections proportional to the sign of the offset error. This type of control law inherently causes trajectories to overshoot their targets. For AFR systems which have time varying delays, this overshoot behavior becomes even more pronounced. The goal of switching period control is to minimize the average offset error between the EQR of the combustion mixture and the desired EQR offset while also matching the amplitude and frequency of the dither signal.

The above-mentioned goal is achieved by forcing all trajectories into a closed, invariant set which is centered at the origin. This set also has the property that all trajectories within this set cause the pre-catalyst oxygen sensor to switch regularly. Inside this set, the average offset error is small enough that its magnitude can be inferred. With an estimate of the offset magnitude, the controller is then able to reduce the offset error to a desired level. In this chapter, let the $\star$ designation represent a special value of that variable. In what follows, define a closed set $\Omega(a)$ as

$$\Omega(a) = \left\{ x \mid |x| \leq a \right\}. \quad (6.27)$$
Theorem 6.4.1 For the discrete time AFR switching period controller governed by the equations presented in Section 6.2, consisting of a periodic dither, a late switching compensator and an early switching compensator, there exists a scalar $a^* \in \mathbb{R}$ such that for any constant disturbance the corresponding AFR trajectory of the closed-loop system remains bounded and converges to an invariant set $\Omega(a^*)$ in finite time provided the gain of the late switching controller $K_l$ is bounded from above by

$$K_{l,\text{max}} < \frac{A}{n_{\text{last}}}.$$ 

(6.28)

Additionally, the set, $\Omega(a^*)$, has the special property that all trajectories within this set cross the origin (which physically corresponds to an oxygen sensor switching event) at least every $T_{\text{dither}}$ events.

Proof: For a system to be bounded and converge to an invariant set in finite time the following three conditions must be satisfied:

1. Boundedness condition: $\exists M^* < \infty$ such that $\forall k \in \mathbb{Z}^+$

$$|x(k)| \leq M^*,$$ 

(6.29)

2. Reaching condition: $\exists k^* \in \mathbb{Z}^+ < \infty$ such that

$$|x(k^*)| \leq a^*.$$ 

(6.30)

3. Invariance condition: $\forall x(k) \in \Omega(a^*)$

$$x(k + 1) \in \Omega(a^*).$$ 

(6.31)
Consider the EQR system model described in (6.5). The change in $x$ over the span of $T_{\text{dither}}$ events can be described with

$$
x(k + T_{\text{dither}}) - x(k) = w(k + T_{\text{dither}} - 1) - w(k - 1) + u(k + T_{\text{dither}} - 1) - u(k - 1) + d(k + T_{\text{dither}} - 1) - d(k - 1).
$$

(6.32)

Because the dither signal is periodic,

$$
w(k + T_{\text{dither}}) - w(k) = 0 \quad (6.33)
$$

∀ $k \in \mathbb{Z}^+$. Therefore in the ideal case with a constant disturbance, the change in $x$ simplifies to

$$
x(k + T_{\text{dither}}) - x(k) = u(k + T_{\text{dither}} - 1) - u(k - 1).
$$

(6.34)

When the controller is first activated, after the system has reached its nominal operating temperature, the late switching compensator takes action until the output switches. During this period the control law can be represented with

$$
u(k + 1) - u(k) = -K_I y(k) \quad (6.35)
$$

∀ $k \in \mathbb{Z}^+ \leq \epsilon_{\text{switch}}^*$ where $\epsilon_{\text{switch}}^*$ is the event when the first switch occurs. This control law bounds the change in the state trajectory with

$$
\left( x(k + T_{\text{dither}} + 1) - x(k + 1) \right) y(k) \leq -K_{I,\min} T_{\text{dither}} \quad (6.36)
$$

∀ $k \in \mathbb{Z}^+$ such that $k \leq \epsilon_{\text{switch}}^* - T_{\text{dither}}$, which guarantees an upper bound on the number of events needed to cause the output to switch as described by

$$
\epsilon_{\text{switch}}^* \leq T_{\text{dither}} + \frac{|x(0)|}{K_{I,\min}} + n_{\text{last}} \quad (6.37)
$$
where \( x(0) = d(0) \).

The largest difference in the dither signal between any two events is described by

\[
|w(k) - w(j)| \leq 2A \tag{6.38}
\]

\( \forall k \in \mathbb{Z}^+ \) and \( \forall j \in \mathbb{Z}^+ \). Assume that the output does not switch within the first \( T_{\text{dither}} \) events, then the state trajectory is bounded for the first \( T_{\text{dither}} \) events as described by the following: \( \forall k \in \mathbb{T} = \{0, 1, \ldots (T_{\text{dither}} - 1)\} \),

\[
|x(k)| \leq 2A + |x(0)|. \tag{6.39}
\]

After one dither period, the system monotonically converges to the origin when a time step of \( T_{\text{dither}} \) is considered.

Because of the delays within the system, the late switching controller will overshoot stoichiometry \( (x = 0) \). Under the worse case scenario, the effect of the applied control action will appear in the output within \( n_{\text{last}} \) events. The peak overshoot caused by the late switching controller must, therefore, satisfy

\[
|e_{o,p}| \leq n_{\text{last}}K_{l,max} + P_{l,max}. \tag{6.40}
\]

When the initial conditions are sufficiently far away \( (|x(0)| \geq P_{l,max} + K_{l,max}n_{\text{last}}) \), the peak overshoot is contracted to

\[
|e_{o,p}^\dagger| \leq n_{\text{last}}K_{l,max}. \tag{6.41}
\]

Note that once the output switches, the effect of the proportional gain is removed leaving a maximum sustained offset error of

\[
|e_{o,s}| \leq n_{\text{last}}K_{l,max} \tag{6.42}
\]
for the first case and
\[ |e_{o,s}^t| \leq n_{last} K_{i,max} - P_{i,min} \quad (6.43) \]

for the second case. From these results it is clear that having a proportional gain does not always improve the performance of the system. In general, a large proportional gain improves the disturbance rejection performance but at the expense of steady-state tracking.

If after this first switching event the closed loop controller were deactivated, then \( \exists \theta \in [0, T_{dither}] \) such that
\[ x(k + 1) - x(k) = w(k + \theta) - w(k - 1 + \theta) + e_{o,s}. \quad (6.44) \]

If the peak overshoot caused by the switching period controller is less than the dither amplitude, then the periodic dither component alone will cause another switching event and this switching event would occur after no more than \( T_{dither} \) events. This condition is satisfied when
\[ K_{i,max} < \frac{A}{n_{last}}. \quad (6.45) \]

The overall governing control law satisfies
\[ 0 \leq \text{sign}\left((u(k) - u(k - 1))y(k)\right), \quad (6.46) \]

so that \( \forall k \in \mathbb{Z}^+ \exists d_{exh}^* \in [1, n_{last}] \) such that
\[ 0 \leq \text{sign}\left((u(k) - u(k - 1))x(k - d_{exh}^*)\right). \quad (6.47) \]

This means that when active the switching period controller will force the system trajectory toward the origin unless \( x \) changes sign (the origin was reached) between events \( k - d_{exh}^* \) and \( k \). One can conclude that with the controller active, the next
switching event will still occur within $T_{dither}$ events. Note that by construction both the early compensator and late compensator cannot both be active at the same time. By choosing $0 < \gamma < 1$, the one time correction applied by the early switching controller helps remove the average offset without causing overshoot.

If $T_{dither} \leq n_{last}$ holds, then once the origin is reached, the maximum peak overshoot caused by the late switching compensator is

$$e_{o,p,max} \leq T_{dither} K_{l,max} + P_{l,max}.$$  \hfill (6.48)

Conversely, if $T_{dither} > n_{last}$, then once the origin is reached, the maximum peak overshoot caused by the late switching compensator is

$$e_{o,p,max} \leq n_{last} K_{l,max} + P_{l,max}.$$  \hfill (6.49)

If (6.45) holds, then the periodic dither will automatically cause another switch regardless of the dither period and the EQR error trajectory will be bounded by some $e^{\ast}_{max} > 0$ in the manner

$$|x(k)| < 2A + P_{l,max} \triangleq e^{\ast}_{max}$$ \hfill (6.50)

$\forall k \in \mathbb{Z}^+$ such that $k \geq e^{\ast}_{switch}$. This means that once a single EGO switching event has occurred all trajectories remain in the set $\Omega(e_{max}^{\ast})$. Given that (6.36), (6.39) and (6.50) hold true,

$$|x(k)| < \max \left( |x(0)| + 2A, e_{max} \right)$$  \hfill (6.51)

$\forall k \in \mathbb{Z}^+$. The system is globally bounded and contains an attractive invariant set $\Omega(a^{\ast})$ that satisfies (6.29), (6.30) and (6.31) for $M^{\ast} = |x(0)| + 2A$, $a^{\ast} = e_{max}^{\ast}$ and $k^{\ast} = e_{switch}^{\ast}$. Therefore, all trajectories converge to $\Omega(a^{\ast})$ within finite time and subsequently remain inside this set. \hfill \qed
6.4.2 Local Asymptotic Stability

Now that the trajectories have been shown to be well behaved, the cycle averaged behavior can be analyzed. Recall that the overall control objective is not to drive the state trajectory $x$ to zero; rather, the state trajectory should oscillate around zero with amplitude and period equal to those of the dither signal. By controlling only the dither offset, a switching period controller is able to accurately reproduce the dither signal. It will be shown that if the average EQR offset error $\bar{x}$ converges to zero, then the dither signal is perfectly reproduced.

The asymptotic stability proof presented in this section relies on the assumption that every late switch is followed by an early switch and every early switch is followed by a late switch. This assumption is met if the following four conditions hold:

1. The complete effect of the control action taken by the late switching compensator is seen in the output before the next late switching event.

2. The control action taken by the early switching compensator influences the output before the following late switching event.

3. The control action taken by the late switching compensator first appears in the output before the following early switching event occurs.

4. The proportional gain of the late switching compensator is not larger than the minimum sustained overshoot caused by the late switching compensator.

The first two conditions guarantee that when a late switching event occurs the instantaneous EQR offset at any late switching event is equal to the EQR offset of the late switching event of the previous dither cycle plus the change in the integral
contributions made by the early and late switching compensators. The third condition prevents the late switching compensator from inducing two additional early switching events (for a total of three). Figures 6.3 and 6.4 illustrate the importance of this condition by demonstrating the effects of properly and improperly selected dither periods. In the first case the plant delay is five events, whereas in the second case the plant delay is ten events. For the ten event delay case, the output switches from rich to lean before the late switching control action even reaches the output. Once it does reach the output, the control action causes a rich to lean switch shortly followed by a lean to rich switch. The first three conditions all hold if the dither period satisfies

\[ T_{\text{dither}} > 2n_{\text{last}} + \delta. \]  

(6.52)

Finally, the fourth condition ensures that unwanted early switches are not caused by the proportional gain of the late switching compensator. Figure 6.5 illustrates what can happen if the proportional gain of the late switching controller is too large compared to the integral gain. Here the dither period is sufficiently large compared to the plant delay (twenty events compared to three events), but the proportional gain can still cause unwanted early switching events. The proportional gain must be small enough relative to the product of the plant delay and the integral gain so that the total control action taken by the late switching compensator to achieve a switching event is still maintained when the proportional gain is removed. This condition is satisfied if the proportional control gain \( P_l \) is bounded by

\[ P_{l,\text{max}} \leq (n_{\text{first}} - 1)K_{i,\text{min}} \]  

(6.53)

where \( n_{\text{first}} \) is the minimum plant delay. With these results, the asymptotic stability of the switching period controller can be proved.
Figure 6.3: Response of the late switching compensator when dither period is properly selected ($T_{dither} = 20$, $d_{exh} = 5$, $\delta = 0$, $A = 0.01$).

Figure 6.4: Response of the late switching compensator when dither period is too short ($T_{dither} = 20$, $d_{exh} = 10$, $\delta = 0$, $A = 0.01$).
Figure 6.5: Response of the late switching controller with proper and improper proportional gain selection ($T_{dither} = 20$, $d_{exh} = 3$, $\delta = 0$, $A = 0.015$).

**Theorem 6.4.2** For all AFR switching period controllers that satisfy Theorem 6.4.1, each EQR trajectory of the closed-loop system corresponding to any constant disturbance asymptotically converges to a sinusoidal signal with a frequency and amplitude equal to the frequency and amplitude of the imposed periodic signal with an offset error (distance to the EQR setpoint) that is bounded by a local class $K$ function $f(.) : \mathbb{R} \mapsto \mathbb{R}$ of the deadzone size $\delta$ provided (6.52), (6.53) and

$$K_{t,\text{max}} < \frac{2A}{2n_{\text{last}} + \delta} \quad (6.54)$$

are all satisfied.

*Proof:* The first two conditions of Theorem 6.4.2 ensure that for every dither cycle, two corrections (one early and one late) are made. Consider a new time scale represented
by the time index $s$ that increments every early switching event. With this reference, the order of these corrections is late, then early. For readability, each variable in this time scale is distinguished with an overline. The state variable $\overline{x}$ represents the sustained error which is equal to the sum of the constant disturbance and the value held in the integrator of the controller. This analysis is only valid for trajectories inside the set $\Omega(a^*)$.

In the new time scale, the state equation is simply

$$\overline{x}(s + 1) = \overline{u}(s), \quad (6.55)$$

where $\overline{u}(s)$ is the combined integral contributions from the early and late switching compensators. The initial conditions for this time scale are dependent on the sustained overshoot caused by the late switching controller according to

$$\overline{x}(0) = \overline{u}(-1) = \zeta e_{o,s} < A. \quad (6.56)$$

Therefore, the set of all initial conditions for $\overline{x}$ is

$$\overline{x}(0) \in \{x | |x| \leq A\}. \quad (6.57)$$

Because only the integral contributions affect the sustained offset error, the control law in this time scale can be represented as

$$\overline{u}(s) = \overline{u}(s - 1) - \text{sign} (\overline{x}(s)) \Delta_{\text{switch,l}} K_l - \gamma e_s(\overline{x}(s)^* \quad (6.58)$$

where

$$\Delta_{\text{switch,l}} = \max \left(0, \epsilon_{\text{switch}} - \left(\frac{T_{\text{dither}}}{2} + \delta\right) \right) \quad (6.59)$$

and $\overline{x}(s)^*$ is the value of the sustained offset error when the early switch occurs. Because of the plant delay, not all of the late switching compensation affects the
measurement before the early switching occurs. Therefore the value of $\bar{x}(s)^\dagger$ is between $\bar{x}(s)$ and $\bar{x}(s+1)$ as described by

$$\bar{x}(s)^\dagger = \bar{x}(s) + \alpha_{\text{switch}}(\bar{x}(s+1) - \bar{x}(s)), \quad (6.60)$$

where $\alpha_{\text{switch}} \in [0, 1]$.

The closed-loop system can also be represented by

$$\bar{x}(s+1) = \bar{x}(s) + \bar{u}(s) - \bar{u}(s-1) \quad (6.61)$$

which simplifies to

$$\bar{x}(s+1) = \bar{x}(s) - \text{sign}(\bar{x}(s))\Delta_{\text{switch,l}}K_l - \gamma \Upsilon_e(s)\bar{x}(s)^\dagger. \quad (6.62)$$

Neither the late nor the early switching compensator will take control action unless the offset error causes a switching period error larger than $\delta$ events. Recall that (6.15) relates the number of events between output switches to the average error offset. For the closed-loop system, the instantaneous sustained error, rather than the average error, dictates the number of events between output switches. The relative lateness of a late switching event can be related to the instantaneous sustained error using

$$\bar{x}(s) = \zeta A \cos \left( \frac{\pi}{2} + \delta + \Delta_{\text{switch,l}} \right) \quad (6.63)$$

which simplifies to

$$\frac{\bar{x}(s)}{-\zeta} = |\bar{x}(s)| = A \sin \left( \frac{\pi}{2} + \Delta_{\text{switch,l}} \right). \quad (6.64)$$

Define the relative earliness of a switching event as

$$\Delta_{\text{switch,e}} = \max \left( 0, \left( \frac{T_{\text{dither}}}{2} + \delta \right) - \epsilon_{\text{switch}} \right). \quad (6.65)$$
The instantaneous sustained error can also be related to the relative earliness of an early switching event according to

\[
\overline{\pi}(s) = \zeta A \cos \left(\frac{\pi \left(\frac{T_{\text{dither}}}{2} - \delta - \Delta_{\text{switch},e}\right)}{T_{\text{dither}}}\right)
\]

(6.66)

which simplifies to

\[
\frac{\overline{\pi}(s)}{-\zeta} = |\pi(s)| = A \sin \left(\frac{-\pi(\Delta_{\text{switch},e} + \delta)}{T_{\text{dither}}}\right).
\]

(6.67)

If

\[
|\overline{\pi}(s)| > A \sin \left(\frac{\pi \delta}{T_{\text{dither}}}\right) \equiv M^*_0,
\]

(6.68)

then either the late or the early switching compensator will take action depending on the situation; otherwise, neither controller will act. The system will converge to the invariant set \(\overline{\Omega}(M^*_0)\) defined as

\[
\overline{\Omega}(M^*_0) = \left\{ \overline{\pi} \mid |\overline{\pi}| \leq M^*_0 \right\},
\]

(6.69)

if \(\exists \lambda_{\min}, \lambda_{\max} \in (-1, 1)\) such that \(\forall s \in \mathbb{Z}^+\) that satisfy \(\overline{\pi}(s) \in \left\{ \overline{\pi} \mid M^*_0 < |\overline{\pi}| \leq A \right\}\)

\[
\lambda_{\min} \leq \frac{\overline{\pi}(s + 1)}{\overline{\pi}(s)} \leq \lambda_{\max}
\]

(6.70)

and if \(\forall s \in \mathbb{Z}^+\) that satisfy \(\overline{\pi}(s) \in \overline{\Omega}(M^*_0)\), it holds that

\[
\overline{\pi}(s + 1) = \overline{\pi}(s).
\]

(6.71)

Clearly (6.71) holds, because neither controller makes any corrections when the offset error is this small.

Assuming that the control authority is split equally between the two compensators, the magnitude of the control action taken by each controller must be less than \(|\overline{\pi}(s)|\).

For the late switching controller this means the following must be true:

\[
\Delta_{\text{switch},l} K_l < |\overline{\pi}(s)|
\]

(6.72)
∀ s ∈ Z^+. The trigonometric relationship containing |x\(s)\) shown in (6.64) can be bounded by
\[
\frac{2A}{T_{\text{dither}}} \Delta_{\text{switch},l} \leq |x\(s)\| \leq \frac{4A}{T_{\text{dither}}} \Delta_{\text{switch},l}
\]
(6.73)

∀ \Delta_{\text{switch},l} ∈ [0, \frac{T_{\text{dither}}}{2} - \delta]. If \(K_l\) is chosen such that
\[
K_{l,max} < \frac{2A}{T_{\text{dither}}},
\]
(6.74)
then (6.72) is guaranteed to hold.

To ensure the dither signal automatically causes switching once the origin is reached, \(K_l\) must also satisfy (6.45). The overall requirement for \(K_l\) is therefore
\[
K_{l,max} < \min \left( \frac{A}{n_{last}} \frac{2A}{T_{\text{dither}}} \right).
\]
(6.75)

Also recall that the dither period is bounded by (6.52). Therefore the bound for \(K_l\) is simply
\[
K_{l,max} < \frac{2A}{2n_{last} + \delta}.
\]
(6.76)

By selecting \(K_l\) in this manner, the late switching controller will decrease the magnitude of the offset every time step without changing the sign of the error. As a result \(x\(s)\| can be bounded by
\[
0 < \frac{x\(s)\|}{\bar{x}\(s)\|} < 1.
\]
(6.77)
The magnitude of the correction made by the early compensator is \(\gamma |\bar{x}\(s)\|\). As a result, system eigenvalues have a lower bound of
\[
\lambda_{min} = -\gamma > -1
\]
(6.78)
and an upper bound of
\[
\lambda_{max} = 1 - \frac{T_{\text{dither}}}{4A} K_{l,min} < 1.
\]
(6.79)
∀ \pi \in \left\{ \pi \mid M^*_\delta < |\pi| \leq A \right\}.

In conclusion if \( K_{l,max} \) satisfies the constraint described in (6.76), then all trajectories of the original system converge to the invariant set \( \Omega(e^{*}_{max}) \) within a finite number of events, which is bounded above by \( e^{*}_{\text{switch}} \). Once inside the invariant set \( \Omega(e^{*}_{max}) \), the oxygen sensor will switch twice every dither period (one late switch and one early switch) as long as the period of the dither signal satisfies (6.52) and the late switching proportional gain satisfies (6.53). The combined effect of the disturbance plus the integral compensation one event prior to the late switching event strictly decreases as long as the EQR offset has a magnitude larger than \( \delta \). Additionally, when the offset is equal to or less than \( \delta \), the offset error remains constant (\( \lambda = 1 \)). Therefore the offset error converges to the invariant set \( \overline{\Omega}(M^*_\delta) \). Inside this set, the system trajectory simplifies from (6.5) to

\[ x(k+1) = p(k) + O. \] (6.80)

where \(|O| \leq M^*_\delta \). Therefore, the EQR of the combustion mixture reproduces the dither signal in frequency and amplitude but the offset may have an error which is bounded by a function \( \overline{f} : \mathbb{R} \mapsto \mathbb{R} \) of the deadzone size \( \delta \). This function, defined as

\[ \overline{f}(\delta) = A \sin \left( \frac{\pi \delta}{T_{\text{dither}}} \right), \] (6.81)

is strictly increasing on the set \( \delta \in [0, \frac{T_{\text{dither}}}{2}) \), so \( \overline{f} \) is a local class \( K \) function.

Having a finite value for \( \delta \) ensures that the system converges to the set \( \overline{\Omega}(M^*_\delta) \) in finite time, rendering the controller more practical. Even in the presence of low frequency disturbances, it is still possible for the system to reach this set. Alternatively, if \( \delta = 0 \), then \( \pi \) will asymptotically approach the origin. However, because the origin
may never be reached, the controllers will continue to make corrections. Because of
the plant delay, this action may distort the shape of the dither signal.

6.4.3 Stability Robustness

For the stability discussion in the previous subsection, only constant disturbances
were considered. However, in practical applications, neither air nor fuel estimates are
exactly correct and their errors vary with time. In this section the stability of the
system with respect to a lumped disturbance parameter $d$, representing the combined
effect of the estimation errors, sensor errors and unknown/external disturbances, is
analyzed.

**Theorem 6.4.3** For the AFR switching period controller, there exists a scalar $a \in \mathbb{R}$
such that all trajectories of the perturbed system remain bounded and converge to an
invariant set $\Omega(a)$ in finite time provided the disturbance trajectory satisfies

$$|d(k + 1) - d(k)| < r_{max} \quad (6.82)$$

and

$$|d(k)| < d_{max} \quad (6.83)$$

$\forall \; k \in \mathbb{Z}^+$. Additionally, the set $\Omega(a)$ has the property that there exists an integer
$\theta \in \mathbb{Z}^+$ such that all trajectories within this set cross the origin at least every $\theta$
events. Lastly, the convergence to set $\Omega(a)$ is monotonic with respect to a time step
of $T_{\text{dither}}$ provided that

$$K_{l,\text{min}} > d_{max}. \quad (6.84)$$
Proof: The two disturbance trajectory bounds are reasonable and would be satisfied for any properly calibrated set of air and fuel estimation algorithms. Considering these bounds on the disturbance the EQR error can be bounded by

\[ x(k + T_{dither}) \leq x(k) + T_{dither}r_{\text{max}} + u(k + T_{dither}) - u(k) \] (6.85)

and

\[ x(k + T_{dither}) \geq x(k) - T_{dither}r_{\text{max}} + u(k + T_{dither}) - u(k) \] (6.86)

\[ \forall \ k \leq \epsilon^*_{\text{switch}} - T_{\text{dither}}. \] If \( \exists \ \kappa_{\text{min}} > 0 \) such that

\[ (x(k + T_{dither}) - x(k))y(k - 1) \leq -\kappa_{\text{min}} \] (6.87)

\[ \forall \ k \leq \epsilon^*_{\text{switch}} - T_{\text{dither}}, \] then the system will converge to the origin monotonically (with respect to a time step of \( T_{\text{dither}} \)) during this time frame. For this to occur, the late switching control action must overpower the effect of any disturbance (high gain feedback). The control input must satisfy

\[ u(k + T_{dither}) - u(k) < -\text{sign}(x(k))T_{dither}r_{\text{max}} \] (6.88)

\[ \forall \ k \leq \epsilon^*_{\text{switch}} - T_{\text{dither}}. \] Comparing this condition to the control law described in (6.36), it is clear that the monotonic convergence condition is satisfied when

\[ K_{l,\text{min}} > r_{\text{max}}. \] (6.89)

With this monotonic convergence property, the number of events before the first output switching event is bounded by

\[ \epsilon^*_{\text{switch},m} \leq T_{\text{dither}} + \frac{|x(0)|}{K_{l,\text{min}} - r_{\text{max}}} \] (6.90)

where \( x(0) = d(0) \).
The condition in (6.89) does not affect the dynamics of the controller. The overshoot caused by the late switching compensator still satisfies the bounds described in (6.40-6.43). Because of the time varying nature of the disturbance, the corrective action made by the early switching controller could be in the wrong direction. The sustained error made by the early switching controller is therefore bounded by

$$e_e \leq \gamma A.$$  \hfill (6.91)

The maximum control error must now include the accrued error from the time varying disturbance. After reaching the origin, the late switching compensator will start acting if the output has not switched after \( (T_{dither}^2 + \delta) \) events. As a result, the largest error that can accumulate before the late switching controller activates is

$$e_{acc,m} = \min \left( r_{max} \left( \frac{T_{dither}}{2} + \delta + n_{last} \right), 2d_{max} \right) + 2A. \hfill (6.92)$$

Because of the integral structure of the controller, the maximum accrued error can take a maximum value equal to twice the maximum disturbance magnitude. If the disturbance magnitude is maintained at one bound, then eventually the controller will reject the error by storing the opposite value in the integrator. Then if the disturbance changes sign, the new error is twice that of the maximum disturbance magnitude.

Combining the maximum accrued disturbance error with the maximum overshoot caused by the controller results in a maximum total error of

$$e_{max,m}^* = \max(e_{o,p,max}, e_e) + e_{acc,m}.$$  \hfill (6.93)

An error this large can only be accumulated after \( T_{dither}^2 + \delta \) events without switching; subsequently, the event that follows the late switching compensator will begin acting
and the error will necessarily be reduced. For the monotonic convergence case, the number of events required to return to the origin, $\theta_m$, is bounded by

$$\theta_m \leq \frac{T_{dither}}{2} + \delta + n_{last} + \frac{e_{max,m}}{K_{l,min} - r_{max}}. \quad (6.94)$$

If $r_{max} \geq K_{l,min}$, then (6.92) can no longer be used to calculate the maximum accumulated error. When the disturbance can change faster than the controller can compensate, the EQR error magnitude can increase even when the late switching controller is active. Because the magnitude of the disturbance is bounded by $d_{max}$, the disturbance magnitude cannot increase indefinitely. Define $\rho$ as the first of the two comparison terms in (6.92) in the manner

$$\rho = r_{max} \left( \frac{T_{dither}}{2} + \delta + n_{last} \right). \quad (6.95)$$

Before the late switching controller activates, the system can accumulate an error as large as $\rho + 2A$ provided $\rho \leq 2d_{max}$. Once the late switching controller is activated, the disturbance is not more than $2d_{max} - \rho$ away from its maximum magnitude bound and the system can head away from the origin at a maximum rate of $r_{max} - K_{l,min}$.

Combining these observations, the maximum accumulated error must be bounded by

$$e_{acc,nm} = \min \left( \rho + (2d_{max} - \rho) \frac{r_{max} - K_{l,min}}{r_{max}}, \ 2d_{max} \right) + 2A \quad (6.96)$$

when $r_{max} > K_{l,min}$. The maximum total error follows the same definition as the uniform convergence case as described by

$$e_{max,nm}^* = \max(e_{o,p,max}, e_e) + e_{acc,nm}. \quad (6.97)$$

Even though the system does not converge monotonically to the origin, the controller ensures that the system reaches the origin within $e_{switch,nm}^*$ events where

$$e_{switch,nm}^* \leq T_{dither} + \frac{d_{max}}{K_{l,min}}. \quad (6.98)$$
Once the origin is reached, the output will still switch regularly. Between switching events, the magnitude of the disturbance can change as much as $2d_{\text{max}}$. Therefore the maximum number of events between output switches for the non-monotonic convergence case, $\theta_{nm}$, must now be proportional to the maximum disturbance magnitude as described by

$$\theta_{nm} = \frac{2d_{\text{max}} + 2A}{K_{l,\text{min}}} + \frac{T_{\text{dither}}}{2} + \delta + n_{\text{last}}.$$  \hspace{1cm} (6.99)

Regardless of the magnitude of $K_{l,\text{min}}$, the desired convergence and invariance properties described in (6.29), (6.30) and (6.31) still hold. When $K_{l,\text{min}} > r_{\text{max}}$, the constants that satisfy these conditions are $a^* = e^*_{\text{max},m}$, $M^* = |x(0)| + 2A$ and $k^* = \epsilon^*_{\text{switch},m}$. Therefore, the system converges monotonically to the invariant set $\Omega_r(e^*_{\text{max},m})$ and inside this set the origin is crossed at least every $\theta_m$ events. When $K_{l,\text{min}} \leq r_{\text{max}}$, the constants that satisfy the conditions are $a = e_{\text{max},nm}$,

$$M^* = \max \left( e^*_{\text{max},nm}, 2A + \min \left( d_{\text{max}}, |x(0)| + \frac{d_{\text{max}} - |x(0)|}{r_{\text{max}}}(r_{\text{max}} - K_{l,\text{min}}) \right) \right)$$ \hspace{1cm} (6.100)

and $k^* = \epsilon^*_{\text{switch},nm}$. Under these conditions, the system converges non-monotonically to the invariant set $\Omega_r(e^*_{\text{max},nm})$ and the output switches at least every $\theta_{nm}$ events.

Under all conditions the EQR error remains bounded, and this bound depends on the size of the invariant set $\Omega(a^*)$. Air and fuel estimation errors are constant for a given operating condition. Therefore as the engine speed and manifold pressure equilibrate, the lumped disturbance variable converges to a constant value. Once this occurs, the system behaves like the system would in the presence of a constant disturbance. Therefore, the same local convergence properties presented in Section 6.4.2
still apply. If the conditions in both Theorem 6.4.2 and Theorem 6.4.3 are satisfied, then the closed-loop AFR control system is globally stable and the AFR will asymptotically approach the reference trajectory whenever the disturbance remains constant for a sufficient number of events.

6.4.4 Implications on Control Gain Selection

Selecting control gain $K_l$ represents a trade off between robustness and steady-state performance. To reject disturbances and monotonically converge to the invariant set $\Omega(\epsilon_{\text{max}}^*)$, $K_l$ must always be larger than the maximum rate of change of the disturbance (6.89). However, as $K_l$ is increased, the size of the invariant set is also increased. To reproduce the sinusoidal dither in the presence of a constant disturbance, the control gain must satisfy the upper bound described in (6.76). For the system to monotonically reject disturbances and have desirable steady-state performance, the following inequalities must hold:

$$r_{\text{max}} < K_{l,\text{min}} < K_{l,\text{max}} < \frac{2A}{2n_{\text{last}} + \delta}.$$  \hspace{1cm} (6.101)

In most cases, $r_{\text{max}}$ is so large that selecting $K_l$ and $A$ to satisfy these conditions would produce unacceptable results, because the dither amplitude is limited by both driveability and emissions performance. As the dither amplitude is increased, the AFR tracking performance increases but the conversion efficiency of the TWC decreases. One of the purposes of dithering the AFR is to maintain the TWC conversion efficiency at a high level, but this benefit is lost if the dither amplitude is too large compared to the oxygen storage capacity of the TWC. Unlike the other calibration parameters, the selection of the dither amplitude requires more experimental tuning.
The dependence of the emission performance on the dither amplitude is the topic of ongoing research.

Note that it was shown in the previous sections that even if $r_{max} > K_{t,min}$, the system is stable, facilitating a discussion of the practical aspects of gain selection. The late switching control gain is scheduled as a function of the operating conditions. For most driving patterns, steady-state operation is confined to specific regions of the operating space, whereas most operating conditions are only encountered during transients. Disturbances are most pronounced during transients, so large gains for $K_l$ and $P_l$ should be chosen in these regions. For operating conditions that typically correspond to steady-state, small gains should be chosen to improve the dither tracking performance.

The $\gamma$ factor in the early switching controller strongly affects the convergence rate of the system in steady-state. Although the maximum overshoot caused by the early switching compensator is directly proportional to $\gamma$, the magnitude of this overshoot is very small compared to that of the late switching controller. Therefore, a value near unity should be chosen for $\gamma$. A switching period controller does not need a deadzone, but a small deadzone is beneficial because it reduces the convergence time. Unfortunately, the size of the deadzone also affects the maximum allowable steady-state offset error. To minimize tailpipe emissions, this offset must be very small. Therefore the deadzone variable, $\delta$, should not be more than about 10% of the dither period. Considering this constraint and (6.52), the size of the dither period should be approximately 2.5 times larger than the plant delay at idle to allow steady-state tracking while limiting the size of the maximum possible overshoot error.
6.5 Calibration Procedures

Although the guidelines presented in the previous subsection provide a very useful starting point, some form of testing is required to fine tune each of the control gains. Finding the optimal trade-off between disturbance rejection performance and steady-state tracking is difficult, because the frequency of steady-state operation and thus its importance varies with the operating conditions. In a traditional calibration procedure, hundreds of drive cycles would be experimentally tested in a sequential manner with refinements made in between each test. This type of process is extremely costly and often relies on subjective decisions.

Using models to represent the system, all of the control gains can be numerically optimized in simulation. With a minimal amount of experimental data, millions of different sets of control gains can be simulated. In a simulation environment, any number of fictitious disturbances can be applied to directly test the robustness of each set of control gains. Not only does this systematic procedure produce excellent performance, but also the optimization process can be constrained to ensure that the resulting controller possesses desirable stability properties.

6.5.1 Engine Model Components

To calibrate a switching period controller in simulation, the following components are required: a controller model, a trapped air estimator, a fuel dynamics compensator, a fuel dynamics model, a controller model, an oxygen dynamics model, a sensor model and a catalyst model. The first three components are actually embedded within the AFR controller. A switching period controller can be implemented using any air estimation model and any fuel dynamics compensation model. As such these models
are not the primary focus of this dissertation. One of the primary means of validating the AFR switching period controller is by direct comparison to the existing production controller. To provide the fairest comparison, the same volumetric efficiency based trapped air estimation algorithm used by the production controller was also used for the AFR switching period controller. A third order auto regressive moving average (ARMA) structure was selected for the fuel dynamics model and the fuel dynamics compensator. The coefficients of these models were fit to linear spline functions of the manifold pressure, engine speed, intake cam position and exhaust cam position in the same manner as [90]. A HFEOD model was used to capture the time varying oxygen transport and mixing dynamics.

One of the most difficult components to model is the TWC. In the most basic sense, a catalyst adsorbs oxygen gas during lean operation and expends it during rich operation. As long as the maximum oxygen storage capacity of a catalyst is not reached, a catalyst continues to store oxygen during lean operation. Similarly, the products of rich combustion are oxidized until the stored oxygen is completely depleted. When the oxygen storage is between these bounds, the EGO voltage tends to slowly approach a stoichiometric voltage. Outside of these bounds, the EGO voltage is closely correlated to the EQR of mixture entering the catalyst.

For the purpose of calibrating an AFR switching period controller, a catalyst can be modeled using an integrator to represent the oxygen process and a set of filters to represent the combined effect of the oxidation/reduction reactions and the EGO sensor dynamics. The oxygen mass stored within the catalyst can be predicted according to

\[ O_s(k) = O_s(k - 1) + (1 - \phi_{cat,in}(k)) \tilde{m}_{exh}^*(k)[O_2]_{air} \]  

\[(6.102)\]
where $O_s$ is the mass of oxygen stored in the catalyst, $\phi_{cat,in}$ is the EQR of the mixture at the inlet of the catalyst, $\dot{m}^*_{exh}$ is the exhaust mass flow rate and $[O_2]_{air}$ is the mass fraction of oxygen gas in air. Using this oxygen storage mass, the nominal catalyst outlet EQR, $\phi_{cat,out}$, can be calculated as

$$\phi_{cat,out}(k) = \begin{cases} 
1 & \text{if } 0 < O_s(k) < O_{\max} \\
\phi_{cat,in}(k) & \text{otherwise}
\end{cases}.$$  \hspace{1cm} (6.103)

To approximate the chemical kinetics and sensor dynamics, the nominal catalyst outlet EQR is filtered according to

$$\phi_{post}(k) = (1 - K_{post})\phi_{post}(k-1) + K_{post}\phi_{cat,out}(k)$$ \hspace{1cm} (6.104)

where

$$K_{post}(k) = \begin{cases} 
K_{slow} & \text{if } 0 < O_s(k) < O_{\max} \\
K_{fast} & \text{otherwise}
\end{cases}.$$  \hspace{1cm} (6.105)

$\phi_{post}$ is the EQR at the post-catalyst EGO sensor and the two filter constants $K_{slow}$ and $K_{fast}$ satisfy

$$0 < K_{slow} < K_{fast} < 1.$$  \hspace{1cm} (6.106)

This last condition captures the very slow relaxation behavior which is observed when the stored oxygen mass is within its upper and lower bounds. The final component in the catalyst model is an EGO sensor calibration curve which converts EQR values into voltages. Even though the proposed catalyst model has some deficiencies, it accurately captures the low frequency behavior, which is all that is required to effectively calibrate the post-catalyst controller.

With these models any type of drive cycle can be simulated given the appropriate inputs. These inputs include the intake pressure, intake temperature, exhaust pressure, exhaust temperature, intake cam position, exhaust cam position and engine
speed. Instead of predicting these inputs with an air path model, a cam timing control model and a vehicle model, their trajectories can be measured by experimentally driving the desired test cycle. This approach is an efficient and accurate means of generating the required input trajectories with a very small amount of experimental data.

### 6.5.2 Optimization Procedure

The performance of each set of control gains was evaluated over a series of simulated FTP-72 drive cycles. The input trajectories were generated by first experimentally running a FTP-72 drive cycle using the production open loop fueling controller. To ensure the optimal controller is robust to modeling and prediction errors, a wide range of disturbances were applied throughout the simulation. Disturbances in the form of steps, ramps, impulses and/or white noise were applied to each submodel to represent modeling/estimation errors. In addition to these artificial disturbances, a realistic net EQR disturbance profile was also used. This disturbance which lumps all of the air and fuel errors into a single value consists of the time shifted difference between the measured EQR and the desired EQR resulting from the open loop estimation errors. This measured disturbance is exactly what the closed-loop AFR controller must reject.

Because the pre-catalyst switching period controller is independent of the post-catalyst switching voltage controller, the optimization was performed into two parts. First the pre-catalyst switching period control gains were optimized with the post-catalyst controller disabled. Using the cumulative squared error between the measured pre-catalyst EQR and the desired pre-catalyst EQR setpoint as the cost function
produced both good steady-state tracking performance and disturbance rejection performance. Once the pre-catalyst gains had been found, the post-catalyst switching voltage control gains were optimized in a second optimization routine. The goal of this optimization was to calibrate the post-catalyst control gains so that the post-catalyst EGO voltage can be regulated to a desired valued. The post-catalyst setpoint voltage directly affects the tailpipe emissions which are very difficult to accurately predict. For the optimization routine multiple FTP-72 cycles were simulated and a different setpoint voltage was chosen for each cycle. For this optimization, the cumulative squared error between the measured post-catalyst EQR and stoichiometry was used as the cost function. With this cost function and procedure, the optimal post-catalyst setpoint controller is able to regulate the post-catalyst EGO voltage very well regardless of the setpoint voltage selected. The same optimization procedure was also repeated to calibrate the rich/lean switching period ratio control gains.

For any optimization technique, the convergence time is dependent on the initial seeding of the optimization. Rapid convergence was achieved by choosing the initial gain values to be constant across the operating conditions according to

\[ P_l = (n_{first} - 1)K_{l,min} \]  \hspace{1cm} (6.107)

and

\[ K_l = \frac{2A}{2n_{last} + \delta} - \varepsilon \]  \hspace{1cm} (6.108)

where \( \varepsilon \) is some small positive constant. With these control gains, the performance of the system is initially biased toward steady-state tracking performance. Once the optimization had converged, the operating regions which infrequently encounter
steady-state had control gains significantly higher than the initial guesses, resulting in an improvement in disturbance rejection performance.

These model aided calibration procedures optimize all of the control gains for both the pre-catalyst and post-catalyst controllers. Because the model used to represent the TWC does not predict tailpipe emissions, the post-catalyst setpoint voltage must be identified experimentally. To select the post-catalyst setpoint voltage, a set of FTP-72 or FTP-75 was performed experimentally. A different post-catalyst setpoint voltage was used for each drive cycle and the setpoint voltage which produced the best emissions performance was chosen.

6.6 Experimental Setup

To analyze the effects of the dither signal parameters (waveform, amplitude and period) and to quantify the performance of the switching period control architecture, experimental testing was performed on a 2006 Saturn Ion powered by a 2.4 liter 4-cylinder gasoline engine with variable intake and exhaust valve timing. The switching period controller including the trapped air mass estimator and the fuel dynamics compensator used the same sensor set as the production vehicle. To quantify the AFR tracking performance, two Bosch LSU 4.9 wide range UEGO sensors were installed in the same locations as the pre-catalyst and post-catalyst EGO sensors. All of the emission data was collected using a Horiba MEXA 7500 and all testing was performed on a 150 horsepower twin roll chassis dynamometer.

In the experimental investigation that follows, many different dither signals are compared. Because the control gains are dependent on the choice of the dither signal, a new set of gains were optimized using a genetic algorithm for each dither signal.
A genetic algorithm was used because it was able to rapidly converge to the globally optimal solution.

6.7 Post-Catalyst Setpoint Controller Comparison

For the overall control architecture to be effective, transient disturbances must be rejected quickly and the dither signal must be accurately reproduced in steady-state regardless of the post-catalyst setpoint. Changes in the post-catalyst setpoint should influence the mean of the EQR distribution over the drive cycle without affecting the spread of the distribution. Before different dither signals were compared, a preliminary set of tests was performed to determine which setpoint is most appropriate for post-catalyst AFR control.

First the range of EQR biases which can be achieved by varying each setpoint was identified by driving a FTP-72 drive cycle under five different constant post-catalyst setpoint combinations. In order to establish a baseline comparison, a nominal case defined by a 50% rich/50% lean switching period ratio with a pre-catalyst switching voltage of 450 mV was run. To demonstrate the potential of a switching period setpoint controller, the most extreme rich / lean setpoint values that still produced relatively unbiased disturbance rejection performance (60% rich/40% lean and 40% rich/60% lean) were also tested. When the switching period is split equally between periods of rich and lean operation, rich disturbances are rejected in the exact same manor as lean disturbances. For a switching period that favors rich operation, rich disturbances may not be rejected as fast as lean disturbances and vice versa for lean bias switching periods.
Table 6.1: EQR error statistics for a simulated FTP cycle

<table>
<thead>
<tr>
<th>Switching Voltage (mv)</th>
<th>Rich Period (%)</th>
<th>Lean Period (%)</th>
<th>EQR Std (%)</th>
<th>Average EQR Offset (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>450</td>
<td>50</td>
<td>50</td>
<td>3.41</td>
<td>0.12</td>
</tr>
<tr>
<td>700</td>
<td>50</td>
<td>50</td>
<td>3.50</td>
<td>0.29</td>
</tr>
<tr>
<td>300</td>
<td>50</td>
<td>50</td>
<td>3.51</td>
<td>-0.41</td>
</tr>
<tr>
<td>450</td>
<td>60</td>
<td>40</td>
<td>3.41</td>
<td>0.12</td>
</tr>
<tr>
<td>450</td>
<td>40</td>
<td>60</td>
<td>4.10</td>
<td>-0.04</td>
</tr>
</tbody>
</table>

Consider the case when the switching period is rich biased, for example. The minimum number of consecutive rich events required for the late switching compensator to act increases. Similarly, the minimum number of consecutive lean events required for the late switching compensator to act decreases. As a result, the largest rich error that can be accrued before taking control action is larger than the largest lean error. This effect becomes readily apparent when the rich/lean or lean/rich switching period ratios are larger than 1.5 to 1.

Switching voltage setpoints of 300 and 700 mV were also tested to demonstrate the potential of a switching voltage setpoint controller. These voltages were chosen because they lay on the edge of the linear region of an EGO sensor. Outside of the linear region, the signal to noise ratio is unacceptable. Additionally the calibration curve varies more from sensor to sensor outside of the linear range. The viability of each post-catalyst control setpoint control method was evaluated based on how much the setpoint affected both the mean and standard deviation of the pre-catalyst EQR distributions. The testing results are shown in Table 6.1.
Adjusting the switching period ratio setpoint from 40%/60% to 60%/40% changed the average EQR offset from -0.04% to 0.12%. Compared to the nominal case, the rich/lean switching period ratio of 60%/40% did not change the average EQR offset. As shown in Figure 6.6, the switching period tracking performance at this setpoint could be improved. Specifically, the number of measured rich events was nearly equal to the number of measured lean events. For the rich/lean switching period ratio setpoint of 40%/60%, the measured EQR tracked the reference better (Figure 6.7), but the average steady-state EQR was barely lean at 0.998. These results suggest that the rich / lean switching period ratio setpoint is not a good candidate for post-catalyst control.

The switching voltage setpoint produced more favorable performance. For setpoint voltage of 300 and 700 mV, the sinusoidal shape of the dither signal was maintained and the EQR offsets were closely approximated. Figure 6.8 demonstrates the steady-state tracking performance of the switching period control when the switching voltage setpoint was 700 mV. At this switching voltage, the average steady-state EQR was approximately 1.0075. Over the complete drive cycle, however, the average was 1.0029. The steady-state tracking performance at a switching voltage of 300 mV was also very good as demonstrated in Figure 6.9. The average steady-state EQR was 0.990, whereas the overall average for the cycle was 0.9959. Adjusting the switching voltage did not affect the tracking performance as evidenced by the approximately constant EQR standard deviation; only the EQR offset was affected. These results show that a post-catalyst switching voltage setpoint controller has the potential to regulate the oxygen storage level of the catalyst and correct for small sensor bias errors.
Figure 6.6: EQR tracking performance: 15/10 rich/lean switching period ratio (reference - dashed, measured - solid)

Figure 6.7: EQR tracking performance: 10/15 rich/lean switching period ratio (reference - dashed, measured - solid)
Figure 6.8: EQR tracking performance: 700 mV switching voltage (reference - dashed, measured - solid)

Figure 6.9: EQR tracking performance: 300 mV switching voltage (reference - dashed, measured - solid)
Based on these experiments a post-catalyst switching voltage setpoint controller was calibrated and implemented. To determine the optimal post-catalyst EGO voltage setpoint, a series of steady-state emissions tests were conducted. For setpoints ranging from 300 to 800 mV, emission data was collected at idle for a variety of dither signals. As expected, low setpoint voltages produced significant nitrogen oxide (NO$_x$) emissions and high setpoint voltages produced more hydrocarbon (HC) and carbon monoxide (CO) emissions. Overall, it was determined that a setpoint of 680 mV produce the best compromise between the three emissions species.

For a switching voltage setpoint controller to be effective, the pre-catalyst controller must be able to track a time varying EQR offset. To test this, the operating conditions were fixed at idle and the switching voltage setpoint was increased by 20 mV every 250 events starting from a switching voltage which corresponds to an EQR which is slightly lean of stoichiometry. As shown in the top plot of Figure 6.10, the post-catalyst EGO voltage increased as the pre-catalyst switching voltage setpoint increased. When considering the window averaged post-catalyst EGO voltage (dashed line), the increasing trend, which was a direct result of an increased average pre-catalyst EQR, is even more pronounced. The lower plot shows that the average pre-catalyst EQR increased, but the standard deviation did not. This demonstrates not only that the pre-catalyst switching period tracking performance is not affected by a dynamically changing switching voltage setpoint, but also that the switching voltage setpoint has enough control authority to be an effective post-catalyst controller. To decrease the time constant of the post-catalyst controller, the control gain $K_{V_{\text{switch}}}$ can be increased. If this control gain is too large, however, the changing rate
Figure 6.10: Effect of the pre-catalyst switching voltage on the post-catalyst EGO voltage

of the switching voltage may interfere with the tracking ability of the pre-catalyst switching period controller.

Without any post-catalyst control, the pre-catalyst EQR will remain near stoichiometry but breakthroughs in the catalyst emissions are still possible because the oxygen storage level of the catalyst may not be optimal. To demonstrate the effect of the post-catalyst switching voltage controller, a FTP-72 drive cycle was run with and without the post-catalyst controller. The measured post-catalyst EGO voltage over a representative portion of the drive cycle for the uncontrolled case is presented in Figure 6.11. As evidenced by the sections of events where the post catalyst EGO sensor
maintained a voltage of around 100 mV, the oxygen storage level frequently reached its maximum storage capacity. During these periods, the NO$_x$ emissions increased significantly.

For the case when post-catalyst control was applied, this type of behavior was largely eliminated. Under similar conditions, the post-catalyst controller maintains a EGO voltage near the desired voltage. Figure 6.12 shows the post-catalyst voltage when the post-catalyst voltage setpoint is 680 mV. Although the average voltage of 650 mV is slightly lower than the target, the post-catalyst controller does a good job of preventing extended excursions from stoichiometry. The switching voltage setpoint commanded by the post-catalyst controller during this time interval changes very slowly as shown in Figure 6.13.
Figure 6.11: Evolution of the post-catalyst voltage without post-catalyst control

Figure 6.12: Post-catalyst voltage regulation performance: Switching voltage setpoint of 680 mV
Figure 6.13: Pre-catalyst switching voltage setpoint commanded by the post-catalyst controller

6.8 Selecting the Periodic Waveform Signal

The performance of a AFR switching period controller is strongly dependent on the choice of the dither signal especially with respect to disturbance rejection performance and steady-state emissions performance. The shape of the dither signal, the amplitude of the dither and the dither period all influence the performance of a switching period controller. Increasing the size of the dither signal improves the accuracy of the AFR error estimator but also makes the TWC more susceptible to emission breakthroughs. In general a dither signal designed to promote disturbance rejection performance suffers from poor steady-state emissions performance.
As previously described, a post-catalyst controller indirectly attempts to regulate the average oxygen storage ratio to near one half to facilitate the rejection of both rich and lean disturbances. If the average storage ratio could be regulated to exactly one half, then the imposed dither signal could take any symmetric shape as long as the mass of oxygen stored during lean operation is not larger than to the maximum storage capacity. Because of its observability issues, the oxygen storage ratio cannot be regulated to exactly one half. Therefore, the oxygen mass stored during the lean portion of the dither signal should be only a small fraction of the maximum storage capacity of the TWC. A catalyst can be thought of as an integrator with saturation bounds at zero and the maximum storage capacity. The oxygen stored during the lean portion of a dither cycle can be estimated by integrating the lean portion of the waveform. Regardless of the waveform, this integral is proportional to the dither period and the dither amplitude.

In addition to the TWC requirements, the dither signal must also satisfy drive-ability requirements. As the EQR of the combustion mixture increases, the torque production also increases. Torque fluctuations caused by rapid lean EQR changes are more undesirable than rapid rich changes. At a minimum, the vehicle operator should not be able to detect the torque fluctuations caused by the dither signal. As a result the derivative of the dither signal must be bounded and this bound is proportional to the dither amplitude and inversely proportional to the dither period.

To determine the real world implications of the trade-off between steady-state tracking performance and disturbance rejection performance, emission testing was conducted at idle and the first 500 seconds of an FTP-72 cycle were driven. The two most important performance metrics for an AFR controller are the pre-catalyst EQR
distribution and the tailpipe emissions. Although only the tailpipe emissions are regulated by the Environmental Protection Agency, emissions tests are costly and require a tuned post-catalyst controller. By design, the pre-catalyst switching period controller is independent of the post-catalyst setpoint controller. Clearly the performance of the post-catalyst controller is dependent on the pre-catalyst controller performance, however the pre-catalyst performance is not dependent on the post-catalyst controller performance. Therefore, the effectiveness of the pre-catalyst portion of the switching period controller can be determined exclusively by examining the distribution of the pre-catalyst EQR measurements.

For the steady-state emissions tests, the engine was run at idle for two minutes and emission data was collected. A post-catalyst voltage setpoint of 680 mV for the post-catalyst switching voltage setpoint controller was used to maintain the oxygen storage during this set of testing. Conversely, no post-catalyst controller was used for the drive cycle testing. To provide a comparison metric, all of the data was normalized with respect to the performance of the production AFR control. The production AFR controller which has been certified to meet the Environmental Protection Agency and California Air Resource Board emission standards is the final product of a laborious and expensive experimentally based calibration process, whereas the switching period controller has been calibrated to this point entirely in simulation.

6.8.1 Waveform Shape Selection

In the open literature the predominant waveform used in tracking controllers is a sinusoid, yet the motivation behind this choice is rarely discussed. The shape of the
dither waveform has a strong impact on the emission performance of an AFR controller. Oscillating the AFR around stoichiometry improves the conversion efficiency of a three-way catalyst because the exhaust gas switches between a net oxidizer and a net reducer. The fluctuations in the emission concentrations produced by different waveforms are unique and therefore produce different conversion efficiencies. Specifically for a switching period controller, the shape of the waveform also impacts the design of the early switching compensator.

The purpose of the early switching control is to reject offset errors faster than the late switching compensator could alone and to ensure that offset errors asymptotically approach zero. As long as the offset is smaller than the dither amplitude, the oxygen sensor will automatically switch. When the period between switches is shorter than one half of the dither period, it is assumed to be caused by a constant offset error. For a sinusoidal dither signal, the switching period can be directly related to the offset error, \( \tau \), using

\[
\tau = \zeta A \sin \left( \frac{\pi \Delta_{\text{switch}}}{2T_{\text{dither}}} \right).
\]  

(6.109)

This relationship is used by the early switching compensator to make a correction that is proportional to the magnitude of the offset error as described by

\[
K_e(\epsilon_{\text{switch}}) = \gamma \tau.
\]  

(6.110)

For other waveforms such as a triangle or sawtooth wave, the offset error can also be related to the switching period. For a triangle, the offset error can be estimated using the following function of the switching period,

\[
\tau = \zeta \left( A - \frac{2A}{T_{\text{dither}}} \epsilon_{\text{switch}} \right).
\]  

(6.111)
and

\[ \bar{x} = -\zeta \frac{2A}{T_{\text{dither}} \Delta_{\text{switch}}} \]

(6.112)

for a sawtooth wave. Like the sinusoidal case, the early switching compensator can be designed to completely remove offset errors thereby allowing the dither signal to be perfectly reproduced.

Conversely, the measured switching period cannot be related to the offset error when a square wave is used. As long as the offset error is smaller than the dither amplitude, the output will switch at a fixed rate equal to one half of the dither period. Even though premature switching cannot be directly related to an offset error, it does provide evidence of a disturbance. If the oxygen sensor switches faster than one half of the dither period, then it must be the result of a disturbance. Although the magnitude of the disturbance cannot be determined, the sign of the disturbance is known. Therefore, the early switching compensator is no longer a function of the switching period, and is described by

\[ K_{e,\text{square}} = K_{\text{square}} \zeta. \]

(6.113)

Without the ability to completely remove constant offset errors, a square wave signal is prone to emissions breakthroughs during steady-state operation. However, the disturbance rejection potential of a square wave is higher than the other waveforms. If the output does not switch after one half of the dither period, then the amplitude of the current offset error must be larger than the dither amplitude. For the square wave switching period controller, the proportional gain of the late switching compensator is converted into another integrator that satisfies

\[ p(k) = P_{\text{square}} A \bar{y}_{\text{square}}(k) y(k) + p(k - 1) \]

(6.114)
where

\[ \Upsilon_{\text{square}}(k) = \begin{cases} 1 & \text{if } n_{\text{switch}}(k) = \left( T_{\text{dither}} + \delta \right) / 2 \\ 0 & \text{otherwise} \end{cases} \] (6.115)

and \( P_{\text{square}} \in (0, 1) \). With this modification, which is only applicable for a square wave, disturbances can be rejected faster.

The sensitivity between an offset error and the resulting change in the switching period affects both disturbance rejection and steady-state tracking performance. Waveforms which have a high sensitivity (small offset errors produce large switching period changes) near the nominal switching period of one half of the dither period enable a switching period controller to better regulate the offset error to zero; however, disturbance rejection performance must be sacrificed. Because small offset errors can cause large switching period changes, the gain of the late switching compensator must be relatively small. Conversely, waveforms which have a low sensitivity (large offset errors produce small switching period changes) near the nominal switching period of one half of the dither period enable better disturbance rejection performance at the expense of steady-state performance. For the low sensitivity case, if the oxygen sensor does not switch within one half of the dither period, then the offset error must be significant. Therefore, larger gains for the late switching compensator can be selected.

For the steady-state and drive cycle testing, each waveform had a dither period of 25 events and a dither amplitude of 1.5%. To reduce the driveability issues, the sawtooth wave was oriented such that the AFR jumps from lean to rich. The results of the drive cycle and steady-state tests are shown in Figure 6.14 and Figure 6.15, respectively. Over a drive cycle, the EQR distributions for the sine, triangle and square waves were very similar. Even though the sawtooth wave produced a slightly wider distribution, it was still narrower compared to the production controller. At idle
the sine wave produced the fewest emissions and the square wave produced the most. Each of the four waveforms produced fewer overall emissions than the production controller. The sine wave in particular produced fewer NO\textsubscript{x}, HC and CO emissions making it the best performing waveform overall.

6.8.2 Dither Period

Recall from Section 6.4.2 that the switching period must be twice as large as the plant delay to ensure that the early switching compensator performs properly. When this condition is not satisfied, the control action taken in response to a late switch could unintentionally cause two consecutive early switching events and thus potentially produce unwanted oscillations. If the early switching compensator was removed, then shorter dither periods could be considered. Under these conditions an
AFR switching period controller will still reproduce the desired dither waveform but with a slower rate of convergence.

Seven sine waves with different dither periods corresponding to 15, 20, 25, 30, 40, 50 and 75 events were tested. A constant dither amplitude of 1.5% was used in all cases. Since the plant delay at idle is approximately 11 events, the early switching compensator was excluded for the 15 and 20 event dither period cases. During steady-state idle operation, these cases performed about as well as the 25 event dither cases (Figure 6.16), because the disturbances were so small. During an FTP-72 drive cycle, the performance of the switching period controllers which used dither signal with

Figure 6.15: Effect of the dither waveform on idle emissions
periods less than 25 events showed significant performance degradation (Figure 6.17). This demonstrates the importance of the early switching compensator.

![Figure 6.16: Effect of the dither period on idle emissions](image)

It can take up to one half of the dither period to recognize a disturbance. The size of the error that can accumulate during this time is proportional to the dither period. When considering both steady-state and drive cycle performance, the optimal dither period for this engine is 25 events. This value corresponds very well to the analytical solution \(2n_{\text{delay}} + \delta + 1\) which is 24 events. When the dither period is increased from this value, both the disturbance rejection performance and emission performance should decrease. The experimental results shown in Figure 6.16 and Figure 6.17 support this claim; however, the EQR distribution during a drive cycle did not increase significantly until the dither period was increased to 75 events. It should
be noted that all of the dither periods tested performed better than the production controller.

### 6.8.3 Dither Amplitude

Of all the waveform characteristics, the dither amplitude has the dominant effect on tailpipe emissions. Because the oxygen storage capacity of the catalyst is finite, the frequency of catalyst breakthroughs decreases as the dither amplitude is decreased. However, the dither amplitude cannot be decreased without limit. At some point the pre-catalyst switching period controller can no longer track the dither signal, because the signal to noise ratio of the EGO sensor becomes too small. This effect is clearly seen in the emission results presented in Figure 6.18. For each test a sinusoidal dither with a period of 25 events was used.

![Figure 6.17: Effect of the dither period on the EQR distribution over a drive cycle](image)
Compared to the idle emissions results, the drive cycle results in Figure 6.19 are not as conclusive. The dither amplitude did not have a significant effect on the disturbance rejection performance. According to (6.76), the largest late switching control gain which still guarantees dither tracking is proportional to the dither amplitude. Because the operating regions marked by the largest disturbances are not maintained for very long, the optimal control gains selected by the genetic algorithm tended to ignore this condition. Moreover, in the regions which most often corresponded to steady-state operation, the optimal control gains are much smaller than this bound. As a result, the optimal gains corresponding to each different late switching controller were nearly identical, leading to good disturbance rejection performance.

For all of the dither amplitudes less than 2.5%, the standard deviation of the precatalyst EQR was nearly constant (the variation between cases was less than 7%). Conversely, the total tailpipe emissions varied by as much as 100% within this same range. This clearly demonstrates that EQR regulation performance is not equivalent to tailpipe emissions performance. At a dither amplitude of 3.5%, the standard deviation of the EQR distribution began to increase. This increase is partially due to the larger variation in the imposed dither signal.

### 6.8.4 Signal Selection Summary

Firstly, this testing confirms that different waveforms have different affects on the conversion efficiency of a TWC and thus produce different tailpipe emissions. Secondly, the disturbance rejection performance of an AFR switching period controller is also dependent on the choice of the dither waveform. For the engine used in this study, the best trade-off between steady-state tracking performance and robustness
Figure 6.18: Effect of the dither amplitude on idle emissions

Figure 6.19: Effect of the dither amplitude on the EQR distribution over a drive cycle
was achieved with a sinusoidal dither with a period of 25 events and an amplitude of 1.5%. A dither period of 25 events is very close to the analytically solution defined by $2n_{last} + \delta + 1$. Based on these results, that optimal dither signal for other engines would be expected to also be a sinusoid with a dither period of approximately $2n_{last} + \delta + 1$ events and a dither amplitude between 1.5% and 2%.

6.9 Performance Comparison: Switching Period Controller Versus Production Controller

To best quantify the performance of the tuned switching period controller, it has been directly compared to the performance of the production AFR controller. To meet the Environmental Protection Agency and California Air Resource Board emission standards, the production controller relies on extensive experimental calibration procedures. The motivation behind this comparison is to show that the switching period controller matches or exceeds the performance of the production controller in all three emissions species yet requires only a fraction of the calibration effort. As a result, a switching period controller gives a significant improvement over the current practice.

The EQR regulation and emissions performance of the two controllers were evaluated over a complete FTP-72 cycle. For the switching period controller three post-catalyst EGO voltage setpoints were tested. These setpoints correspond to the lowest practical setpoint of 300 mV, the highest practical setpoint of 800 mV and a setpoint of 680 mV which was found to produce the best trade-off between the three emissions species. In each case, a 25 event sinusoidal dither signal with a 1.5% amplitude was used. The results for this set of testing are presented in Table 6.2. Adjusting the
Table 6.2: EQR and emissions statistics for a FTP-72 cycle (all values relate to stock and are in percentages)

<table>
<thead>
<tr>
<th>Control Setpoint</th>
<th>Pre-cat EQR Std</th>
<th>Post-cat EQR Std</th>
<th>NO\textsubscript{x} CO + HC</th>
<th>HC CO + CO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stock</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>300 mV</td>
<td>90.8</td>
<td>69.9</td>
<td>291.0</td>
<td>59.7</td>
</tr>
<tr>
<td>680 mV</td>
<td>84.6</td>
<td>78.5</td>
<td>97.4</td>
<td>68.1</td>
</tr>
<tr>
<td>800 mV</td>
<td>90.6</td>
<td>74.6</td>
<td>69.9</td>
<td>70.8</td>
</tr>
</tbody>
</table>

post-catalyst EGO voltage setpoint shifted the mean of the pre-catalyst and post-catalyst EQR distributions thereby profoundly impacting the emission results. For all three cases, however, the standard deviations of the EQR distributions were nearly constant and smaller than the production controller. The standard deviation of the post-catalyst EQR distribution for a setpoint of 680 mV was slightly higher because that voltage is located at the knee of the sensor curve making it more difficult to regulate to that voltage. Nevertheless with a setpoint of 680 mV, the switching period controller produced less of each of the three emission species resulting in a 16.5% reduction in the total emissions.

When the response of the two controllers are compared side-by-side, the performance advantage of the switching period controller is readily apparent. Figure 6.20 compares the performance of the switching period controller with a setpoint of 680 mV to the production AFR controller over a portion of the FTP-72 cycle (750 to 950 seconds). The data was collected in the event domain, so the datasets do not perfectly align, though. As shown in the top plot, the engine speed trajectory matched well between the two drive cycles; however, it can be observed that the production
trace fluctuated slightly more than the switching period controller. This increased variation is mainly due to the larger torque fluctuation caused by larger EQR fluctuations. In other words, the switching period controller was better able to regulate the AFR during transients which resulted in smoother transitions between operation conditions. Although driveability is difficult to quantify, these results suggest that the driveability of the switching period controller is better than the production controller.

The middle plot of Figure 6.20 compares the pre-catalyst EQR traces during this portion of the FTP cycle. It is difficult to distinguish the two trajectories, so certain regions of data are highlighted with boxes. In these regions, the production controller produced significant EQR errors and the switching period controller produced minimal EQR errors. At no point in this section of data did the reverse scenario occur. Overall the production AFR controller produced EQR errors that are more
frequent and of a significantly higher magnitude. On three occasions corresponding to the second, third and fifth boxes, a lean EQR was followed by a rich EQR error. This overshooting behavior tends to result in NO\textsubscript{x}, HC and/or CO emission breakthroughs. The lower plot compares the cumulative EQR error with respect to stoichiometry. By the end of this cycle, the production controller produced a cumulative EQR error which was about 60% higher that the switching period controller.

To achieve desirable EQR performance over a drive cycle, an AFR controller must be able to reject disturbances rapidly. The disturbance rejection performance of each controller was isolated by applying a positive 10% EQR step disturbance and a negative 10% EQR step disturbance. As demonstrated in Figure 6.21, both controllers have approximately the same rate of convergence, but the production controller produces significant overshoot when correcting lean errors. The magnitude of the overshoot produced by the production controller was almost as large as the original disturbance. The elimination of overshoot is one of the reasons why the switching period controller produces better performance than the production controller.
The switching period control architecture presented in this dissertation extracts amplitude information out of a binary EGO sensor by analyzing the switching period. With this amplitude information, both steady-state and transient AFR errors can be rejected. For all bounded disturbances, the proposed control system remains stable. Additionally, the maximum number of events between output switches is at worst bounded by an affine function of the maximum disturbance magnitude, and the results are invariant to the oxygen transport delay. To regulate the oxygen stored in the TWC, the switching voltage threshold setpoint can be adjusted by a post-catalyst controller. With this structure, the pre-catalyst tracking performance is isolated from the oxygen storage and post-catalyst control dynamics. The tracking
performance does not depend on any plant delay estimate and this architecture can
be applied to any engine with any exhaust configuration.

The bulk of model based AFR controllers which have appeared in the open litera-
ture control the average AFR and allow for uncontrolled oscillations. With a switching
period controller, however, the actual AFR trajectory is controlled. The experimental
results presented in this chapter demonstrate that the features of an imposed dither
signal (waveform, amplitude and period) strongly affect the conversion efficiency of
a TWC and thus the tailpipe emissions. From this experimental testing, the AFR
trajectory which results in the fewest tailpipe emissions was determined. One of the
most notable results was that tailpipe emissions increase when the dither amplitude
is decreased beyond a certain point (1.5% EQR for the engine tested). When noting
that a regulation controller is equivalent to a tracking controller with zero amplitude
dither signal, these results indicate that a controller which tries to track an oscillating
AFR trajectory produces fewer emissions than a controller which tries to regulate the
AFR to a constant value.

Compared to the production controller, this control architecture produces smaller
pre-catalyst and post-catalyst AFR variations as well as lower emissions for all three
crucial species. Because an AFR switching period controller can be systematically
calibrated almost entirely in simulation, the calibration time and effort is more than
an order of magnitude smaller than the direct tuning methods currently employed.
All these benefits are possible using the same low cost EGO sensors. In terms of
calibration ease, performance, stability and sensor costs, an AFR switching period
controller outperforms the current production methods as well as other model based
controllers.
CHAPTER 7

Conclusions

The main contributions of this dissertation can be grouped into three broad categories: oxygen dynamics modeling, model based in-cylinder oxygen concentration estimation and model based fueling control. Using a physically based plug flow modeling approach, the oxygen transport and mixing dynamics within the air path of an engine can be predicted remarkably well. The lumped oxygen dynamics model of Section 3.2 is most useful in applications that have very large transport delays such as the air path loop from the exhaust ports to the intake ports through the EGR system. In situations where the transport delay is relatively small or the transport delay varies from cylinder to cylinder, each section of piping can be modeled as a separate control volume subsystem. Predicting the oxygen concentration at any point within the exhaust manifold, for example, can be achieved with the type of high fidelity exhaust oxygen dynamics modeling presented in Section 3.3. As exemplified in Section 3.4, a plug flow based oxygen dynamics model can also be modified to predict the FIR relationship between an input fuel injection event and an output oxygen concentration measurement.

Traditionally, the lumped filling and emptying approach used to model the thermal and fluid dynamics is also used to model the oxygen dynamics. This type of mass
averaged approach does not account for the transport delays within the air path system. Compared to these existing models, the three types of oxygen dynamics models presented in this dissertation predict the transient response of an air path system much better. Replacing the oxygen dynamics portion of a traditional filling and emptying air path model with a high fidelity oxygen dynamics model produces a much improved air path model which is capable of simulating the response of the air path system (temperatures, pressures, flow rates and oxygen concentrations) without relying on experimental data. The benefits of these models, however, extends well beyond direct simulation applications. Through the application of modern control theory, the true potential of these models can be realized.

Although significantly more accurate than alternative models, the model orders of the proposed oxygen dynamics models are also much higher. Even after simplification, the resulting oxygen dynamics models are more complex than conventional reduced order oxygen dynamics models. To accommodate these additional layers of complexity, more advanced estimation/control techniques must be utilized along with more sophisticated analyses. For these reasons, the model based estimators and model based controllers built around the reduced order versions of the proposed oxygen dynamics models are also significant contributions.

The model based estimation portion of this dissertation focuses on in-cylinder oxygen concentration prediction. To cope with the limited computational capabilities of a production ECU, a lumped oxygen dynamics model can be simplified to generate two complementary reduced order models (a delay based oxygen dynamics model and a LPV oxygen dynamics model) using the procedures outlined in Section 4.1. Unlike a LPV oxygen dynamic model, a delay based oxygen dynamics model accounts for
the time varying nature of the transport delays. The time-varying nature of this model makes state estimation impractical. On the other hand, the simpler structure of a LPV oxygen dynamics model is more amenable to state estimation. Using a set of LMI design constraints, a LPV state estimator based on a measurement of the exhaust oxygen concentration can be designed. These LMI conditions are the result of a comprehensive stability analysis that explicitly defined a set of sufficient conditions for stable state estimation. Selecting the observer gains in a manner that satisfies these LMIs ensures that the resulting in-cylinder oxygen concentration estimator is robust to both input and output measurement uncertainties.

Predicting the in-cylinder oxygen concentration using either a delay based oxygen dynamics model in an open loop manner or a LPV state estimator provides a considerable improvement over current techniques. When combined in the hybrid feed-forward/feedback estimation structure of Section 4.1.4, the in-cylinder oxygen concentration can be predicted even more accurately. A hybrid feed-forward/feedback in-cylinder oxygen concentration estimator can even account for cylinder-to-cylinder variations in the trapped air mass provided that these variation can be predicted through some external means.

Using a FIR exhaust oxygen dynamics model, cylinder imbalances causing variations in the post combustion oxygen concentration can be identified. Because each cylinder exhausts at different times, periodic variations in the measured exhaust oxygen concentration can be used to identify cylinder imbalances. Before the exhaust oxygen concentration measurement can be linked to the imbalance of a cylinder, the mixing and transport dynamics within the exhaust system must be identified. This relationship which varies as a function of the engine speed and the exhaust flow rate
can be smoothly predicted with a FIR exhaust oxygen dynamics model. Because this type of model is physically based, the conservation of mass and conservation of species requirements that are paramount to accurate cylinder imbalance estimation are automatically satisfied. On its own, an exhaust oxygen concentration based cylinder imbalance estimator cannot distinguish between EGR mixing imbalances and fueling imbalances. Conversely, a crankshaft fluctuation based or cylinder pressure based cylinder imbalance estimator can only identify torque imbalances. Although not developed in this dissertation, these two types of cylinder imbalance estimators could be paired to distinguish air imbalances from fuel imbalances.

With the accuracy provided by a hybrid feed-forward/feedback in-cylinder oxygen concentration estimator, the diesel fueling control problem can be reformulated to directly account for the effect of diluents on the combustion process thereby improving the overall robustness of the fuel control system. As shown in Section 5.2, the fueling parameters that produce the optimal trade-off between emission production and fuel consumption can be well predicted using sets of three-dimensional lookup tables indexed by in-cylinder oxygen concentration, engine speed, and total fuel injection mass. Scheduling the fueling parameters in this manner enables optimal combustion to be achieved for all instantaneous cylinder conditions. Regardless of the rate at which an air path setpoint controller reaches it desired setpoint or even if the desired setpoint is ever reached, an in-cylinder oxygen concentration based fueling controller will continue to meet its emissions targets in a manner that minimizes the fuel consumption rate.
Oxygen transport and mixing dynamics are also a fundamental part of the AFR control problem in gasoline engines. The experimental testing in Section 6.8.1 demonstrates that the average conversion efficiency of a three way catalyst during real world operation can be improved by intentionally dithering the AFR of the combustion mixture in a precisely controlled manner. This experimental insight combined with a better understanding of the oxygen transport and mixing dynamics motivated the design of a tracking based AFR control structure. Within this control architecture, a pre-catalyst AFR switching period controller tracks the periodic AFR trajectory that maximizes the conversion efficiency of the catalyst during nominally stoichiometric operation and a post-catalyst switching voltage setpoint controller maintains the oxygen storage level within the three way catalyst. Compared to a typical production AFR controller, an AFR switching period control structure produces smaller pre-catalyst and post-catalyst AFR variations as well as lower emissions for all three crucial species. Results were provided to show that this controller is also robustly stable for all bounded disturbances.

Throughout this dissertation, the models, estimators and controls have been constructed to meet the typical production constraints of an engine manufacturer. In particular, each component has been designed so that they require a minimal amount of experimental data to calibrate. All three of the proposed oxygen dynamics models are physically based and therefore require very little experimental data to calibrate; a brief transient dataset and a steady-state dataset in which steps and impulses are applied to the fueling commands of individual cylinders are sufficient. An oxygen dynamics model of just the exhaust system requires an exhaust oxygen concentration measurement, whereas an oxygen dynamics model of the complete air path system
requires measurements of both intake and exhaust oxygen concentrations as well as a model of the trapped residual fraction. Given that the trapped residual fraction only affects the in-cylinder oxygen concentration prediction and its effect is relatively minor, the trapped residual fraction model can be less accurate, yet still highly effective.

To simplify a lumped oxygen dynamics model to produce a delay based oxygen dynamics model, only two additional parameters corresponding to an EGR mixing constant and an output filter constant require calibration. These parameters can be identified using a simulated dataset produced by the original lumped oxygen dynamics model. Similarly, a LPV oxygen dynamics model requires only a single calibration constant and it, too, can be identified from a simulated dataset. When a solution exists, a LPV state estimator can be generated automatically using a LMI solver. When combining a LPV state estimator with a delay based oxygen dynamic model to form a hybrid feed-forward/feedback in-cylinder oxygen concentration estimator, one additional constant requires tuning. Again, this constant can be calibrated using a simulated dataset.

With a calibrated FIR exhaust oxygen dynamics model, a recursive least squares based cylinder imbalance estimator can be constructed directly. For applications that use switching type EGO sensors, the forgetting factor may need to be tuned and a first order filter may need to be designed. If necessary, these parameters can be calibrated using the portion of the original experimental dataset in which step changes are commanded to individual cylinders. All of calibration parameters within the oxygen dynamics models and estimators are constants. Comparable techniques use
calibration parameters which are scheduled on the operating conditions and therefore require much more experimental data to calibrate.

Because of the nonlinear interdependencies between the cylinder contents and the parameters which define how fuel is injected in a diesel engine, it is nearly impossible to avoid scheduling the fueling parameters and air path setpoints. In the proposed in-cylinder oxygen concentration based fueling controller, the overall calibration effort is reduced by consolidating multiple sets of two-dimensional fueling parameter lookup tables into a single set of three-dimensional lookup tables without modifying the air path setpoint lookup table structure. Although the total number of fueling table entries is approximately the same, experimental data can be incorporated into a three-dimensional fueling table more effectively. A typical dataset consisting of fueling parameter and air path setpoint sweeps that would be used to calibrate a single table entry of a conventional fueling parameter table could calibrate as many as ten table entries of an in-cylinder oxygen concentration based fueling parameter table.

Controlling the AFR within a gasoline engine is another problem in which parameter scheduling is often necessary. The proportional and integral control gains of an AFR switching period controller are scheduled on engine speed and mass air flow. Using a high fidelity exhaust oxygen dynamics models, however, these control gains can be calibrated entirely in simulation. In fact, only one parameter, the post catalyst setpoint voltage, needs to be calibrated experimentally. Compared to a production AFR controller, the calibration requirements of an AFR switching period controller are more than an order of magnitude smaller.
In summary, each of the modeling, estimation and control techniques presented in this dissertation provides several advantages over existing techniques. The models and estimators are more accurate, and the controllers more effectively meet their control objectives. Moreover, the estimators and controllers are robustly stable with respect to external disturbances. All of these improvements are realized even though the calibration requirements are markedly reduced, thus clearly making these designs better solutions.
BIBLIOGRAPHY


