Model Predictive Control of Modern High-Degree-of-Freedom Turbocharged Spark Ignited Engines with External Cooled EGR

Rohit Vishvanath Koli
Clemson University, koli.rohit@gmail.com

Follow this and additional works at: https://tigerprints.clemson.edu/all_dissertations

Recommended Citation
Koli, Rohit Vishvanath, "Model Predictive Control of Modern High-Degree-of-Freedom Turbocharged Spark Ignited Engines with External Cooled EGR" (2018). All Dissertations. 2211.
https://tigerprints.clemson.edu/all_dissertations/2211

This Dissertation is brought to you for free and open access by the Dissertations at TigerPrints. It has been accepted for inclusion in All Dissertations by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.
MODEL PREDICTIVE CONTROL OF MODERN HIGH-DEGREE-OF-FREEDOM TURBOCHARGED SPARK IGNITED ENGINES WITH EXTERNAL COOLED EGR

A Dissertation
Presented to
the Graduate School
of Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Automotive Engineering

by
Rohit Vishvanath Koli
August 2018

Accepted by:
Dr. Robert Prucka, Committee Chair
Dr. Zoran Filipi
Dr. Qilun Zhu
Dr. Srikanth Pilla
ABSTRACT

The efficiency of modern downsized SI engines has been significantly improved using cooled Low-Pressure Exhaust Gas Recirculation, Turbocharging and Variable Valve Timing actuation. Control of these sub-systems is challenging due to their interdependence and the increased number of actuators associated with engine control. Much research has been done on developing algorithms which improve the transient turbocharged engine response without affecting fuel-economy. With the addition of newer technologies like external cooled EGR the control complexity has increased exponentially.

This research proposes a methodology to evaluate the ability of a Model Predictive Controller to coordinate engine and air-path actuators simultaneously. A semi-physical engine model has been developed and analyzed for non-linearity. The computational burden of implementing this control law has been addressed by utilizing a semi-physical engine system model and basic analytical differentiation. The resulting linearization process requires less than 10% of the time required for widely used numerical linearization approach. Based on this approach a Nonlinear MPC-Quadratic Program has been formulated and solved with preliminary validation applied to a 1D Engine model followed by implementation on an experimental rapid prototyping control system.

The MPC based control demonstrates the ability to co-ordinate different engine and air-path actuators simultaneously for torque-tracking with minimal constraint violation. Avenues for further improvement have been identified and discussed.
DEDICATION

This dissertation is dedicated to my family in India. Without your support, this endeavor would not have been possible for me.
ACKNOWLEDGMENTS

I am extremely grateful to my advisor Dr. Robert Prucka for his valuable guidance and support. He is among the few academic advisors who proactively prioritize the students’ well-being. He has taught me that the easiest person to convince is the person in the mirror, which has inspired me to cultivate and maintain the rigor of fundamental research.

I would also like to thank my friend Dr. Qilun Zhu for his continual input and the several hour long discussions we have had regarding engine control, science and life. His work has been a great source of inspiration for my research. Thank you for having the patience and the will to help me to the best of your ability I would also like to thank Dr. Zoran Filipi for his valuable advice on turbocharged engine control on modeling. His class lecturers on advanced IC engine controls was a memorable experience for me.

The research work performed was in collaboration with industrial partners Robert Bosch LLC and Fiat Chrysler Automobiles. I would like to thank Jason Schwanke and Dr. Shyam Jade from Robert Bosch LLC for their valuable industrial insight and quick assistance in this research. I would also like to thank Michael Prucka and the folks at Fiat Chrysler Automobiles for their valuable inputs and collaboration.

Lastly, I would like to thank all my fellow CUICAR students and friends, Konstantinos Siokos, Shuonan Xu, Harikesh Arunachalam, Anirudh Allam, Dan Egan, Dennis Robertson, Tommy Powell, Ryan O’Donnell, and others for their help and assistance.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Down-sizing SI internal combustion engines</td>
<td>1</td>
</tr>
</tbody>
</table>
Control problems associated with high-degree of freedom engines .......... 2
Motivation for utilization of Model Predictive Control ............................. 7
Challenges in implementing MPC on turbocharged SI engines ............... 8
Research objectives and outline ............................................................. 11

II. EXPERIMENTAL TESTING ARCHITECTURE ...................................... 13

Test Engine Setup .................................................................................. 12
Data Acquisition .................................................................................... 15
Rapid Control Prototyping Setup ........................................................... 16

III. CONTROL ORIENTED ENGINE MODELING ...................................... 20

Parametrized Turbocharger model ......................................................... 19
Data driven parameter models for engine cylinders ............................. 38
Continuous time state space model of the engine ............................... 47

IV. NON-LINEARITY ANALYSIS OF CONTROL ORIENTED MODEL ... 53

Local linear approximation ................................................................. 51
Step and sinusoidal input transient testing .......................................... 51
<table>
<thead>
<tr>
<th>V. NON-LINEAR MODEL PREDICTIVE CONTROL DESIGN</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizon length and sample time determination</td>
<td>59</td>
</tr>
<tr>
<td>Non-linear system model linearization</td>
<td>61</td>
</tr>
<tr>
<td>Model discretization in time</td>
<td>70</td>
</tr>
<tr>
<td>MPC Cost function formulation</td>
<td>72</td>
</tr>
<tr>
<td>MPC Quadratic program problem formulation and optimization</td>
<td>73</td>
</tr>
<tr>
<td>Co-simulation with GT-Power</td>
<td>74</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VI. LOWER-LEVEL CONTROL AND ESTIMATION</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valve mass flow rate control</td>
<td>80</td>
</tr>
<tr>
<td>Low-pressure EGR ∆P control</td>
<td>82</td>
</tr>
<tr>
<td>Turbo speed estimation and state measurement</td>
<td>82</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VII. EXPERIMENTAL VALIDATION OF MPC</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental control system hardware configuration</td>
<td>87</td>
</tr>
<tr>
<td>Experimental test results: Aggressive tip-in and tip-out maneuvers</td>
<td>93</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>VIII. CONCLUSIONS AND FUTURE WORK</td>
<td>102</td>
</tr>
<tr>
<td>Summary of research work</td>
<td>97</td>
</tr>
<tr>
<td>Broader impacts and practical aspects</td>
<td>99</td>
</tr>
<tr>
<td>Future work and potential for improvement</td>
<td>100</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>107</td>
</tr>
<tr>
<td>A: A Control Algorithm for Low Pressure – EGR Systems using a Smith</td>
<td>102</td>
</tr>
<tr>
<td>B: Nonlinear Model Predictive Control of Dual Loop - Exhaust Gas</td>
<td>125</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>152</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table II-1. Engine specifications</td>
<td>14</td>
</tr>
<tr>
<td>Table III-1. Error metrics of the compressor mass flow rate models</td>
<td>25</td>
</tr>
<tr>
<td>Table III-2. Actuator combinations used for data collection for ANN model training</td>
<td>40</td>
</tr>
<tr>
<td>Table III-3. Inputs used for training ANNs</td>
<td>42</td>
</tr>
<tr>
<td>Table III-4. Accuracy of trained ANNs</td>
<td>48</td>
</tr>
<tr>
<td>Table III-5. Units used for normalization of state space model</td>
<td>49</td>
</tr>
<tr>
<td>Table III-6. Error metrics of the 0D model for four different transient engine trajectories</td>
<td>51</td>
</tr>
<tr>
<td>Table IV-1. Minimum number of model linearizations required in horizon to maintain IMEPn model accuracy for varying input frequency</td>
<td>57</td>
</tr>
<tr>
<td>Table IV-2. Minimum number of model linearizations required in horizon to maintain IMEPn model accuracy for varying input frequency</td>
<td>59</td>
</tr>
<tr>
<td>Table A-1. Conventional PID control vs Smith Predictor Control performance comparison</td>
<td>129</td>
</tr>
<tr>
<td>TABLE A-2. Performance summary of EGR % control using NMPC</td>
<td>150</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>Figure I-1. High-degree of freedom SI Turbocharged engine</td>
<td>2</td>
</tr>
<tr>
<td>Figure I-2. Load control using intake throttle and wastegate on downsized turbocharged engines</td>
<td>4</td>
</tr>
<tr>
<td>Figure II-1. The General Motors LTG engine [92]</td>
<td>13</td>
</tr>
<tr>
<td>Figure II-2. Exhaust side of the test engine</td>
<td>15</td>
</tr>
<tr>
<td>Figure II-3. Intake side of tests engine</td>
<td>15</td>
</tr>
<tr>
<td>Figure II-4. Engine and dynamometer interface</td>
<td>16</td>
</tr>
<tr>
<td>Figure II-5: The BOSCH MED17 Engine ECU</td>
<td>17</td>
</tr>
<tr>
<td>Figure II-6. Dspace rapid prototyping system</td>
<td>19</td>
</tr>
<tr>
<td>Figure II-7. ETAS ES910 rapid prototyping system</td>
<td>19</td>
</tr>
<tr>
<td>Figure III-1. Polynomial approximation for Jensen compressor model coefficients</td>
<td>24</td>
</tr>
<tr>
<td>Figure III-2. Modeled compressor iso-speed lines vs manufacturer provided map points</td>
<td>25</td>
</tr>
<tr>
<td>Figure III-3. Modeled compressor efficiency curve vs manufacturer provided compressor map points</td>
<td>28</td>
</tr>
<tr>
<td>Figure III-4. 0D Modeled vs GT-Power modeled compressor inlet temperature</td>
<td>29</td>
</tr>
<tr>
<td>Figure III-5. 0D Modeled vs GT-Power modeled compressor outlet temperature</td>
<td>30</td>
</tr>
<tr>
<td>Figure III-6. Modeled turbine mass flow rate vs manufacturer provided turbine map points</td>
<td>32</td>
</tr>
<tr>
<td>Figure III-7. Turbine pressure ratio as a function of turbine mass flow rate with map points for four different turbine inlet temperatures</td>
<td>34</td>
</tr>
<tr>
<td>Figure III-8. Modeled turbine efficiency curve vs map points</td>
<td>36</td>
</tr>
<tr>
<td>Figure III-9. Comparison of measured and GT-Power modeled transient turbocharger speed response with tuned inertia for 1500 RPM</td>
<td>38</td>
</tr>
<tr>
<td>Figure III-10. Comparison of measured and GT-Power modeled transient turbocharger speed response with tuned inertia for 2500 RPM</td>
<td>38</td>
</tr>
<tr>
<td>Figure III-11. Experimental setup to facilitate data acquisition for engine cylinder models</td>
<td>41</td>
</tr>
<tr>
<td>Figure III-12. Modeled vs measured volumetric efficiency [%] for all data with +/- 10% error bands represented by dashed red line</td>
<td>43</td>
</tr>
<tr>
<td>Figure III-13. Modeled vs measured IMEPn for all data with +/- 10% error bands represented by dashed red line</td>
<td>44</td>
</tr>
<tr>
<td>Figure III-14. Modeled vs measured $COVIMEPn$ for all data with +/- 10% error bands represented by dashed red line</td>
<td>45</td>
</tr>
<tr>
<td>Figure III-15. Modeled vs measured $KI2$ for all data with +/- 10% error bands represented by dashed red line</td>
<td>46</td>
</tr>
</tbody>
</table>
Figure III-16. Modeled vs measured steady state exhaust temperature for all data +/- 10% error bands represented by dashed red line. The accuracy of each of these neural networks is summarized in...

Figure III-17. Open-loop system trajectory for 0D model vs experimentally collected data...

Figure IV-1. Deviation between the non-linear system model and the linearized system model as a function of time step N from the linearization. The deviation is induced by perturbing $m_{ITV}$ corresponding to commanded IMEPn...

Figure IV-2. Model error as a function of commanded open-loop change in IMEPn shows that the deviation between linearized and non-linear model varies as a function of operating condition and the time step after the point of linearization...

Figure IV-3. Model error as a function of frequency of commanded open-loop model input shows that the deviation between linearized and non-linear model generally peaks between 1.5 Hz and 0.5 Hz depending on operating condition...

Figure V-1. Approximate relationship between horizon length considered for MPC, size of the model and the computational expense of real-time execution...

Figure V-2. The selection of MPC sample time and prediction horizon time dictates the choice of appropriate engine dynamics in the engine model...

Figure V-3. Graphical layout of Artificial Neural Network...

Figure V-4. Comparison between execution times of IMEPn ANN model linearization using numerical and analytical approach...

Figure V-5. Comparison between execution times of complete system model linearization using numerical and analytical approach...

Figure V-6. Co-simulation setup between GT-Power and Simulink...

Figure V-7. IMEPn response of MPC co-simulated with GT-Power...

Figure V-8. Control actuator trajectories for MPC cosimulation with GT-Power, air-path actuators are on the left hand side and cylinder actuators are on the right hand side. Black dashed lines are the steady-state reference inputs and blue lines are the inputs generated by MPC...

Figure V-9. Compressor operation during co-simulation of GT-Power with MPC...

Figure VI-1. The gradient of the flow function becomes very steep as the pressure difference across the valve diminishes...

Figure VI-2. Real-time estimated vs measured turbocharger speed using fixed point iteration method...

Figure VII-1. Rapid control prototyping setup for experimental validation of MPC...

Figure VII-2. Tip-in IMEPn response for MPC and Open-loop control strategy...

Figure VII-3. Intake throttle mass flow rate trajectories for MPC and Open-loop control...

Figure VII-4. Wastegate mass flow rate trajectories for MPC and Open-loop control...

Figure VII-5. External cooled EGR mass flow rate trajectories for MPC and Open-loop control shows that MPC deliberately stops EGR flow on the aggressive tip-in...

Figure VII-6. Spark timing trajectories for MPC and Open-loop control...

Figure VII-7. Camshaft position timing trajectories for MPC and Open-loop control...
Figure VII-8 ANN Modeled volumetric efficiency using MPC inputs and open-loop inputs for tip-in maneuver ................................................................................................ 95
Figure VII-9 Measured turbine inlet temperature for MPC and Open-loop control shows that MPC commands late spark timing to increase specific exhaust gas enthalpy for turbocharger response .................................................................................................................. 95
Figure VII-10 Turbocharger speed for MPC increases much more rapidly than open-loop control ........................................................................................................................................ 96
Figure VII-11 Measured $K_I2$ for MPC does not exceed the limit of 1 despite advanced spark timing during tip-in ........................................................................................................................................ 97
Figure VII-12. Tip-out IMEPn response for MPC and Open-loop control strategy............ 98
Figure VII-13 Air-path trajectories for aggressive tip-out maneuver............................... 100
Figure VII-14 Cylinder actuator trajectories for aggressive tip-out maneuver.............. 100
Figure VII-15 Compressor operation during the aggressive tip-out............................... 101
Figure A-1 Control system setup .................................................................................. 110
Figure A-2 Dimensional Discharge coefficient map obtained from optimization routine ........................................................................................................................................ 113
Figure A-3 Measured vs modeled EGR valve mass flow rate.................................... 114
Figure A-4 Modeled low pressure EGR path ............................................................... 115
Figure A-5. Closed loop EGR fraction controller with Smith predictor....................... 118
Figure A-6 Modeled vs measured transport delay ....................................................... 120
Figure A-7. Comparison between modeled and measured EGR fraction at EGR sensor location; engine operated at 1700 RPM 3 Bar BMEP .................................................... 121
Figure A-8. Comparison between modeled and measured EGR fraction at EGR sensor location; engine operated at 2100 RPM 5 Bar BMEP .................................................... 122
Figure A-9. Comparison between modeled and measured EGR fraction at EGR sensor location; engine operated at 2100 RPM 2.5 Bar BMEP .................................................. 123
Figure A-10. Step response of EGR fraction controller with conventional PID controller at 1500RPM 7 Bar BMEP ................................................................. 124
Figure A-11. Step response of EGR fraction controller with Smith Predictor at 1500RPM 7 Bar BMEP ................................................................. 125
Figure A-12. Step response of EGR fraction controller with conventional PID controller at 1300RPM 4 Bar BMEP ................................................................. 126
Figure A-13. Step response of EGR fraction controller with Smith Predictor .......... 127
Figure A-14. Step response of EGR fraction controller with conventional PID controller at 2000RPM 9 Bar BMEP ................................................................. 128
Figure A-15. Step response of EGR fraction controller with Smith Predictor at 2000RPM 9 Bar BMEP ................................................................. 128
Figure A-16. Schematic of Engine and Air path model............................................. 133
Figure A-17. Step response of oxygen sensor based EGR measurement installed at compressor outlet location ................................................................. 135
Figure A-18. Comparison of step response of transport delay models with true dead time delay response .......................................................................................... 136
Figure A-19 Closed-loop control structure of DL-EGR NMPC................................. 139
Figure A-20 Pressure states and intake manifold EGR dilution trajectories for HP-EGR control only. ................................................................. 141
Figure A-21 Input trajectories for HP-EGR control only. ........................................ 142
Figure A-22 Pressure states and intake manifold EGR dilution trajectories for LP-EGR control only. ................................................................. 144
Figure A-23 Input trajectories for LP-EGR control only. ......................................... 145
Figure A-24 Pressure states and intake manifold EGR dilution trajectories for Dual-loop EGR control. ................................................................. 147
Figure A-25 Input trajectories for Dual-loop EGR control. ....................................... 148
Figure A-26 Transient intake manifold EGR dilution for multiple EGR architectures. 149
I. INTRODUCTION

1.1 Down-sizing SI internal combustion engines

The practice of reducing engine displacement whilst maintaining and surpassing performance requirements using turbocharging i.e. ‘downsizing’ has been widely followed by the Automotive industry [1][2][3][15]. The recently projected industrywide requirement of a fleet average of 54.5mpg for the model year 2025 requires radical improvements in fuel economy [4]. This can be facilitated partially by further downsizing with more sophisticated methods of turbocharging. Classically, the performance of turbocharged engines during transient engine operation is influenced strongly by the turbocharger behavior. Since turbochargers are mechanically independent from the engine, there are undesired transient phenomenon which severely affect performance and consequently driver perception during these situations. This phenomenon is called the ‘turbo-lag’ and is literally a time delay between the driver-demanded engine torque and the response of the engine. Dedicated systems have been widely researched and utilized to improve turbocharger response [5][8][9][10]. However, these systems usually result in transient fuel economy and emissions penalties.

Lighter turbocharger rotor assemblies have also been investigated by [6][7] which result in quicker turbocharger acceleration due to lower moment of inertia. Utilization of Variable Geometry Turbines (VGT) has also shown significant improvement in transient response on turbocharged diesel applications [19][22]. VNT application on SI engines has not only shown improvement in transient response but also improved low speed
boosting ability [23]-[26]. However, exhaust gas temperature, cost, and reliability limitations have impeded penetration of this technology into mainstream gasoline powered vehicles. Electrical turbocharger compounding concepts have been studied in [26][28] for improved turbocharger response and exhaust heat recovery. All of these solutions which focus on improved hardware characteristics and performance result in elevated cost and added system complexity which is unfavorable for mass production.

1.2 Control problems associated with high-degree of freedom engines

Modern engines are highly over-actuated with respect to torque control, wherein multiple actuator combinations exist for a single engine torque level at a given engine
speed. Usually a unique combination exists which optimizes the steady-state fuel economy of the engine. However, multiple optimum combinations with negligible fuel economy difference may exist [11][12]. The actuator combinations themselves can be very different from each other which further confounds the determination of the appropriate steady state engine-actuator combination.

The dynamics of the different subsystems on the engine are also vastly different in timescales. The impact of each actuator on its corresponding subsystem and coupling with other subsystems affects the overall engine’s dynamic response drastically. Because of the fluid coupling control of the turbocharger actuators is critical for engine performance and fuel economy. Early turbocharger control methods have been outlined [14] with introduction to rule based electronic control of waste-gate and blow-off valve. One of the most significant challenges associated with turbocharged engine control has been the coordination of the waste-gate and the intake throttle valve in turbocharged SI engines. The primary load control actuator for SI engines is the intake throttle valve. However, in order to increase the intake manifold pressure beyond atmospheric pressure, the turbocharger is utilized. Exhaust enthalpy drop across the turbine results in power transfer to the compressor. The compressor performs work on the inlet air, thereby increasing it’s temperature and pressure, also known as ‘boost’ pressure. The waste-gate valve, which controls the turbine bypass flow has been typically used to limit the maximum boost pressure using a pre-calibrated mechanical waste-gate actuator. However, with the advent of electronic waste-gate control this valve is kept wide open under throttled operation to minimize the exhaust pressure and hence minimize pumping
work performed by the engine. However, under transient operation, the control of the intake throttle and waste-gate valve requires complex strategies to simultaneously minimize fuel consumption and improve turbocharger response.

Figure I-2. Load control using intake throttle and wastegate on downsized turbocharged engines

Much research has been conducted in the control and coordination of the turbocharger and air-path actuators since then with an exhaustive list of papers and patents [13][17]. Karnik, et. al have evaluated the performance of decentralized Single Input Single Output (SISO) controllers vs full-state feedback and the reduced versions of full state feedback control in [16] using linearized engine models. The results of this work begin to show the benefit of utilizing Multi-Input-Multi-Output (MIMO) control laws for waste-gate and intake throttle coordination. Typically each of the subsystems
have been controlled with individual control loops targeting a particular set-point as outlined in [13]. Non-linear control approaches have been introduced in [17] using Parallel Distributed Compensation control. Turbocharger waste-gate control under boosted conditions is explored in [18], where singular perturbation is utilized to obtain a first order non-linear model. The feed-forward control law is derived based on inversion of the reduced order model, which is then augmented with an error-integral based feedback term. The potential of Variable Valve Timing (VVT) to improve turbocharger response has been mentioned in [59][60] by improving air-flow into the cylinders and in some conditions directly from the intake port to the exhaust port during the valve overlap phase. Colin et al. [57][58] have proposed a decentralized control system which consists of feed-forward intake throttle controller and NMPC based waste-gate controller for engine torque control. The control of Variable Valve Timing (VVT) into the engine torque controller for control of the trapped Residual Gas Fraction (RGF) has been introduced using a static neural network model. Interaction between the turbocharger response and VVT has been discussed but not exploited in the control strategy due to the decentralized structure of the controller which targets RGF and Torque independently.

Spark timing also has a significant impact on the combustion process, exhaust temperature and the cycle to cycle variability of combustion [54][71]. For best fuel-economy, the spark-timing is calibrated to position fifty percent of the in cylinder fuel–air mixture burnt at eight crank angle degrees after top dead center. However, on turbocharged SI engines the spark timing at higher loads is often limited by the onset of knocking in the cylinder due to end-gas auto-ignition [65] [68]. Such abnormal
combustion causes premature damage to the engine. The Spark timing is retarded to a safe level such that knocking is minimized to acceptable levels. This is called the knock-limited-spark-advance timing. Retarding spark timing for temporary increase of the specific enthalpy of the exhaust gas have been shown to reduce turbocharger lag in [72]and [9]. There is significant potential to improve transient response by utilizing this actuator. However, due to the knock and combustion stability constraints mentioned above, simultaneous control of spark timing along with other engine actuators is challenging.

Current generation production systems generally comprise multiple Single Input Single Output (SISO) type controllers with lookup-tables and maps. Although this approach is simple to adopt for OEMs, it requires significant calibration effort and tuning to ensure stability and robustness. This entails significant real-time testing on the dynamometer and in some cases the vehicle as well. The recent findings of implementing Exhaust Gas Recirculation (EGR) in SI turbocharged engines has also shown potential to improve fuel economy drastically [11][12][29]-[36]. However, this adds yet another control actuator which needs to be considered in the overall engine control system. Moreover, low-pressure cooled EGR also has known issues associated with the inability of feed-forward flow modeling through the valve and transport delay of the EGR Air mixture through the intake air-path [37]-[41]. This delay is more significant for SI engines compared to CI engines because the combustion process of SI engines is far more sensitive to external dilution, since external EGR has a significant effect on the combustion stability and knock propensity.
Because the number of ‘control knobs’ on engines is ever increasing OEMs are beginning to realize the potential of using model based control methodologies which could significantly reduce the required calibration and tuning efforts. Multi-Input-Multi-Output (MIMO) control systems have shown significant promise in being capable of coordinating multiple engine actuators based on multiple targeted outputs. One particular MIMO control approach which has gained attention in recent years is Model Predictive Control (MPC).

1.3 Motivation for utilization of Model Predictive Control

Model Predictive Control is a class of control algorithms in which the control actions for the system are generated as a result of an optimization problem over a fixed time period into the future from the current time. This period is called the prediction horizon. As the closed loop control system executes and proceeds in time, the horizon recedes further equally. MPC is hence also known as Receding Horizon Control (RHC) This optimization problem is formulated from a mathematical objective function, the current condition and a dynamic control oriented model of the system. The general MPC algorithm can be summarized as follows:

1. Generate a predicted future reference trajectory of the system for the prediction horizon based on a known model, current state of the system and reference inputs.

2. Formulate and solve an optimization problem based on a user-defined objective function related to the behavior of the system for the prediction horizon.

3. Apply the inputs that correspond to the current time to the system
4. Re-evaluate the state of the system at periodic intervals and repeat the process.

It is clearly visible that this strategy automatically considers the slow and the fast dynamics of the system simultaneously due to pre-evaluated knowledge of the system’s behavior for a specific time period in the future. The main advantage of applying MPC to engines is the significant reduction of calibration and tuning efforts required. MPC tuning can be achieved easily by simple manipulation of the constant terms within the objective function. This is greatly simplified compared to the conventional tuning and calibration approach used for conventional SISO and LUT based methods. Another major advantage of using MPC is the ability to handle constraints. These constraints can be simple saturation limit constraints for the control actions or more complex operational constraints. This attribute is particularly favorable for turbocharged engines. Operational constraints for turbocharger protection, combustion stability and knocking which are extremely detrimental to the mechanical integrity, fuel economy and exhaust emissions can be efficiently incorporated into the MPC framework.

1.4 Challenges in implementing MPC on turbocharged SI engines

The predominant challenge in widespread implementation of MPC in production is the real-time computational ability. In order to execute MPC in real-time a sufficiently powerful microcontroller with ample memory is required. In order to justify the usage of MPC for engine control in production, the MPC has to be capable of operating at any engine operating condition and any kind of engine transient, with consideration of fuel
economy, emissions, hardware protection, drivability and most importantly torque delivery. The dynamic engine model complexity and accuracy begins to become a concern when all of the above mentioned aspects have to be considered.

Internal combustion engine behavior is highly complicated and requires high-order models to capture high levels of accuracy and mimic relevant dynamic behavior. The addition of a turbocharger to the engine introduces additional coupled, non-minimum phase dynamics which complicates the model further. The turbocharger’s inertial dynamics have a time constant which is in the order of seconds. This is significantly slower compared to the cycle by cycle dynamics of the engine cylinders. Incorporating all of these dynamics into one prediction model and consideration of a time period long enough to account for the turbocharger’s inertial dynamics results in MPC based optimization problems which are impractical to execute in real-time even on research-grade high performance micro-controllers. Hence, choosing an appropriate prediction model for the MPC is the most important task.

A linear prediction model is the favorable choice for MPC formulation due to it’s simplicity in implementation. Using a fixed pre-determined linear model also eliminates the computational cost of obtaining a linearized model from a more accurate non-linear system model in real-time. The prediction model utilized in many of the past research articles is a local linear approximation of the engine dynamics [73][75]. Many of these research articles focus mainly on the control of the air-path of the diesel engines using predetermined equilibrium-point linearized system models stored in the memory of the micro-controller. Some of these articles utilize black-box models identified directly from
experimental or high-fidelity simulation data of the engine. The number of models used to approximate the engine dynamics is also an open question since utilizing lesser models results in larger inaccuracy in the prediction of system behavior during transients and complicates the transitioning between linear models. These approaches have been demonstrated to work well mainly because of the linear-like behavior of the air-path dynamics (intake manifold and compressor outlet pressure). However, this approach has not been demonstrated to track highly non-linear outputs associated with engine behavior like engine torque while simultaneously adhering to combustion constraints which are also complicated non-linear functions of the system states, inputs and parameters.

Application of MPC to gasoline SI engines has been recently explored by Wiese et al in [77] for evacuation of EGR trapped in the intake manifold using a Linear Time Varying MPC. This approach demonstrates the ability of the MPC to schedule the VVT actuation in order to short-circuit the flow from intake ports directly to the exhaust port using available preview of the future torque reference. Bemporad et al have explored the ability of utilizing this short-circuit flow in order to improve turbocharger boost response in [82] using a linear MPC with pre-stored linear black-box models to control the air-path actuators. Similar to the diesel MPC research, the air-path dynamics of the SI turbocharged engine are also linear-like, and hence the local-linear formulation has been shown to work.

Naturally aspirated SI engine torque control while simultaneously adhering to actuator and combustion constraints has been recently investigated in [79]-[81],[87] and [54]. Combustion variability and knock constraint have been considered in the
formulation. Sequential Quadratic Programming based approach is used to solve the MPC iteratively. The same approach has been further investigated in [83] for Air-path control of a SI turbocharged engine which has a continuous surge control valve.

1.5 Research objectives and outline

This research focuses on developing an MPC framework using a control oriented engine model of a SI turbocharged gasoline engine. The air-path model is largely physical and the cylinder model is purely data driven. The goal is to unify the control of air-path and engine cylinder actuators into a single control system which automatically co-ordinates these actuators to improve transient turbocharged engine response while respecting the engine’s operational constraints. In order to achieve this goal many sub-tasks are required to be performed.

1. Develop a control oriented model of the engine which captures the necessary and sufficient dynamic behavior of the engine
2. Develop an algorithm that utilizes the control oriented model and formulates an MPC based optimization problem with a user defined objective function comprising of multiple reference targets
3. Identify avenues and implement strategies to minimize the computational burden of formulating the MPC based optimization problem
4. Solve the MPC based optimization problem by utilizing a solver algorithm capable of executing in real-time on a research grade rapid-prototyping micro-controller
5. Build a Simulation framework to tune the MPC controller by coupling it with a virtual engine developed using a 1D-Engine simulation software.

6. Develop an experimental validation framework for real-time testing of the MPC
II. EXPERIMENTAL METHODS AND RESEARCH ENGINE SETUP

2.1 Test engine description

The experimental test engine is a turbocharged spark-ignited gasoline engine designed by General Motors. It is codenamed as the GM-LTG engine and is available in several standard and premium production vehicles from General Motors.

![The General Motors LTG engine](image)

The LTG engine is a representative of modern downsized engines with multiple actuators. Additionally a low-pressure, cooled EGR circuit was added to the stock engine as well. A detailed list of engine specifications is outlined in Table II-1.
Table II-1. Engine specifications

<table>
<thead>
<tr>
<th>Specification</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Displacement</td>
<td>2.0L</td>
</tr>
<tr>
<td>Compression ratio</td>
<td>9.5:1</td>
</tr>
<tr>
<td>Air-EGR induction system</td>
<td>Turbocharged with Low-pressure EGR</td>
</tr>
<tr>
<td>Fuel System</td>
<td>Gasoline – Direct Injection</td>
</tr>
<tr>
<td>Bore</td>
<td>86</td>
</tr>
<tr>
<td>Stroke</td>
<td>86</td>
</tr>
<tr>
<td>Valve train</td>
<td>DOHC 4 valves per cylinder</td>
</tr>
<tr>
<td>Maximum engine speed</td>
<td>7000 RPM</td>
</tr>
</tbody>
</table>

An aftermarket stainless-steel EGR cooler for a Ford F250 is utilized as the LP-EGR system. The EGR valve from the Audi 1.8TDI is installed as the LP-EGR flow control valve. Engine coolant is circulated through the LP-EGR cooler to emulate real-world cooled EGR systems used in production passenger car diesel engines. Excessive EGR cooling has also been known to cause condensation in the compressor inlet location leading to water droplet impingement on the compressor vanes. An additional pre-compressor throttle is also installed upstream of the EGR mixing location. The original pneumatic waste-gate actuator has been replaced with an electronic linear actuator for improved position control of the waste-gate valve. Figure II-3, Figure II-2 show the hot side as well as the cold side of the engine. The actuators and systems most relevant to turbocharged SI engine control are shown in these figures.
Figure II-2. Exhaust side of the test engine

Figure II-3. Intake side of tests engine
2.2 Data acquisition and dynamometer interface

High and low speed data acquisition systems have been utilized on the test engine. The high speed data acquisition consists of the crank angle resolved pressure measurement of the cylinders, intake and exhaust ports and across the low pressure EGR valve. The low speed data acquisition systems mainly consists of time averaged pressure, temperature and mass flow rate measurements at various locations on the engine. The test engine crankshaft is connected to the dynamometer using a custom flywheel and driveshaft. FEV’s Test automation system is utilized to control the speed and torque setpoint of the dynamometer.

Figure II-4. Engine and dynamometer interface
2.3 Rapid control prototyping hardware

Since the base engine is a production engine, it has production intent sensors and actuators associated with basic engine functionality. An open Bosch ECU MED17 has been interfaced with the production engine sensors and actuators. This ECU is also capable of real time cylinder pressure feedback based diagnostics. The embedded application software deployed onto the ECU has CCP bypass functionality which allows run-time read-access to all of the ECU variables and write-access to the variables associated with spark timing, camshaft timing and throttle position.
Additional subsystem like continuous blow off valve, electronic wastegate control actuator and pre-compressor throttle are controlled using the ETAS ES930 and ES910 units. The ES910 is a rapid control prototyping ECU and calibration interface for the Bosch MED17 ECU. The control applications are initially developed in Simulink/Stateflow and compiled using the Intecrio Realtime system target file. This target file compiles the base level Simulink/Stateflow model into a Scoop IX file with the extension .six. This file is then imported into a software called ETAS Intecrio which facilitates interfacing with the MED17 ECU, ES930 and other systems via CAN. The execution rate and several other diagnostics can also be configured in Intecrio. The ES910’s real-time computational ability is ‘limited’ with respect to running complicated algorithms with large memory requirement and/or small execution time step. These algorithms include computationally intensive Model Predictive Control, Moving Horizon Estimation and complex plant-models for Hardware In the Loop simulation. For such cases the Dspace DS1006 is utilized. The DS1006 system used has four processor boards which can execute different applications independently. This system has high speed CAN controllers which can be interfaced with other hardware. One of these controllers is utilized to establish communication with the ES910, while the other is utilized to communicate with the dynamometer control system for vehicle level simulations with Engine In the Loop capability.
Figure II-6. Dspace rapid prototyping system

Figure II-7. ETAS ES910 rapid prototyping system
A modular approach has been adopted for modeling, wherein the individual components of the air path have been modeled separately and ultimately combined together to form a control oriented air path model.

### 3.1 Parametrized Turbocharger model

Selection of the turbocharger model is critical for model accuracy and execution speed of the engine model. The turbocharger models are split into the compressor models and the turbine models.

#### 3.1.1 Compressor mass flow rate models

To model the pressure dynamics in the boost manifold the compressor mass flow rate is modeled as a function of turbocharger speed and the pressure ratio across the compressor. There are several models available in literature with varying degrees of accuracy and targeted applications. Artificial Neural Network (ANN) based models have been proposed in [47]. Although the prediction of these models within the compressor-map is good, low speed extrapolation is an issue with this approach. Furthermore, the structure of the ANN required to ensure in-map and low-speed extrapolation accuracy changes as a function of the specific compressor chosen as mentioned in [48]. Hence, a physics based classical approach was chosen to model the compressor mass flow rate.

Two models were evaluated for the compressor used in this research. The first one is the ellipse model [44][46] The ellipse model is based on the ellipsoidal parametrization of the iso-speed lines obtained from steady state gas-stand tests of the compressor. These
results of these tests are called the compressor maps and are provided by the manufacturer for normal operating region of the compressor. The ellipse model supports compressor mass flow rate prediction in the mild and deep surge region as well as the choke region. However, the disadvantage of this approach is that the model changes mathematical structure when transitioning from the normal region to the surge region of the compressor map. Incorporating a variable structure model into MPC based control algorithms is not favorable. Hence only the normal region of the compressor map was modeled using the Ellipse model. The compressor map is modeled in the normal region as follows.

\[ \dot{m}_{\text{Comp}} = (\dot{m}_{\text{Max}} - \dot{m}_{\text{ZSL}}) \left[ 1 - \left( \frac{n}{n_{\text{ZSL}}} \right)^{\frac{1}{C_1}} \right] + \dot{m}_{\text{ZSL}} \]

1

\[ C_1 = f_{C_1}(\omega_T) = C_{1,0} + C_{1,1}\omega_T \]

2

\[ C_2 = f_{C_2}(\omega_T) = C_{2,0} + C_{2,1}\omega_T^{C_{2,2}} \]

3

\[ \dot{m}_{\text{Max}} = f_{\dot{m}_{\text{Max}}}(\omega_T) = C_{3,0} + C_{3,1}\omega_T \]

4

\[ \dot{m}_{\text{ZSL}} = f_{\dot{m}_{\text{ZSL}}}(\omega_T) = C_{4,1}\omega_T \]

5

\[ \Pi_{\text{ZSL}} = f_{\Pi_{\text{ZSL}}}(\omega_T) = 1 + C_{5,1}\omega_T^{C_{5,2}} \]

6

Where, \( \dot{m} \) is the mass flow rate and \( C_{l,j} \) have real values. \( \Pi \) is the pressure ratio across the compressor. \( \omega_T \) is the turbocharger’s rotational speed. Within the manufacturer’s compressor map, \( C_1 \) and \( C_2 \) are identified by utilizing global, least-
squares optimization to minimize the difference between the manufacturer’s and modeled iso-speed lines. The suffixes ZSL and Max correspond to two distinct locations on the iso-speed line. The former corresponds to the point on the surge line and the latter corresponds to the point on the $\Pi = 0$ line.

The second compressor model is proposed by Hadef et al. [49][50] which is an evolution of the Jensen and Kristensen empirical mean value model [51]. The latter has been widely used in publications relating to modeling, control and estimation of air path on turbocharged engines. At the heart of these models is the normalization of the nonlinear relationship between the compressor mass flow rate and the pressure ratio using a dimensionless head parameter $\psi$ and dimensionless flow parameter $\phi$. These parameters are defined as follows:

$$
\psi = \psi(\Pi, \omega_T) = \frac{C_{P_{Cin}} T_{Cin}}{0.5 U_C^2} \begin{bmatrix} K-1 \end{bmatrix} \frac{\Pi}{K-1} - 1
$$

$$
\phi = \frac{m_{Comp} R T_{Cin}}{(P_{Cin} d_c^2 \omega_T)}
$$

Where, $C_{P_{Cin}}$ and $T_{Cin}$ are the constant pressure specific heat capacity and temperature of the gas at the inlet of the compressor. $R$ and $K$ are the specific gas constant and the ratio of specific heat capacity respectively. $U_C$ is the compressor blade tip speed which is calculated as a function of the turbocharger speed $\omega_T$ and the compressor wheel diameter $d_c$ as follows.

$$
U_C = \frac{\pi}{60} d_c \omega_T
$$
The flow parameter \( \phi \) is not directly measurable and in fact it is the parameter which needs to be predicted in order to derive the compressor mass flow rate. Several functions have been explored to derive \( \phi \) from \( \psi \) depending on the shape of the iso-speed lines of the compressor [52][53]. The most widely used relationship is the inverse proportionality function which uses the Mach number across the ring orifice of the compressor. However, it requires additional determination of coefficients which are derived from the turbo speed dependent Mach number. Preserving the structure of the function \( \phi \) is modeled as a function of the head parameter as follows as described in [49][50].

\[
\phi = \frac{C(\omega_T)-A(\omega_T)\psi}{B(\omega_T)+\psi}
\]

Where, \( A, B, \) and \( C \) are modeled as second order polynomial functions of the turbocharger speed. The coefficients of these polynomial functions are obtained similar to the method used for obtaining \( C_1 \) and \( C_2 \) for the ellipse model.
Figure III-1. Polynomial approximation for Jensen compressor model coefficients

The modeled iso-speed lines of both of the models are shown in Figure III-3. The accuracy metrics for predicted mass flow rate as a function of compressor pressure ratio for both the models is shown in Table III-1. Error metrics of the compressor mass flow rate models. Although the Ellipse model accuracy is superior to the Jensen model, the ellipse model requires a different mathematical structure in the surge region. Determination of whether the compressor is operating in surge region is not possible without the prior knowledge of the mass flow rate across it. The Jensen model structure
can be utilized all the way till the zero compressor mass flow rate. However, the true compressor behavior in the deep surge region is not an accurate representation of the compressor behavior predicted by the Jensen model. Since the scope of this research is to not model these abnormal modes of operation, the Jensen model is used since it transitions continuously from the normal operation region to the surge region.

![Graph showing modeled compressor iso-speed lines vs manufacturer provided map points.](image)

**Figure III-2.** Modeled compressor iso-speed lines vs manufacturer provided map points

<table>
<thead>
<tr>
<th>Model</th>
<th>RMS error [kg/min]</th>
<th>Max error [kg/min]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ellipse model</td>
<td>0.04</td>
<td>0.1</td>
</tr>
<tr>
<td>Jensen model</td>
<td>0.09</td>
<td>1.14</td>
</tr>
</tbody>
</table>
The compressor model is a quasi-steady state model [49]. During highly transient engine operation $\Pi$ can be calculated to have a value which is much higher than the maximum possible pressure ratio that the compressor model can deliver at the corresponding turbocharger speed and zero mass flow rate. This leads to an erroneous computation of $\psi$ and hence, to limit the operation of the compressor model to have realistic $\phi$ values even during the transients, the $\psi$ parameter is saturated as follows.

$$\psi_{\text{Max}} = \frac{C(\omega_T)}{A(\omega_T)}$$

$$\psi_{\text{Final}} = \min(\psi_{\text{Max}}, \psi)$$

### 3.1.2 Compressor efficiency model

The compressor efficiency can vary significantly depending upon the operating region on the compressor map. Assumption of a constant compressor efficiency has a significant impact with regards to turbocharger inertial dynamics since the turbocharger speed dynamics are directly coupled with this parameter. The compressor isentropic efficiency $\eta_{\text{Comp}}$ is defined using the following equation.

$$\eta_{\text{Comp}} = \frac{P_{C,\text{Ideal}}}{P_{C,\text{Actual}}}$$

Where, $P_{C,\text{Ideal}}$ is the power required to compress the gas from inlet to the outlet pressure, defined by the isentropic process. $P_{C,\text{Actual}}$ is the actual power required by the compressor to pressurize the gas. $P_{C,\text{Actual}}$ is always lesser than $P_{C,\text{Ideal}}$ due to the heat
transfer and other losses associated with the pumping of inlet gas. Since $P_{C \text{Actual}}$ is not measured on the engine, determination of $\eta_{\text{Comp}}$ cannot be made directly from this equation. [53] has shown that it can be identified as a function of $\phi$. However, $\phi$ is the parameter which is derived from $\psi$. Hence, it is identified directly as a polynomial function of the $\psi$ as follows.

$$\eta_{\text{comp}} = b_1 \psi^3 + b_2 \psi^2 + b_3 \psi + b_4$$

Where, $b_1, b_2, b_3$ and $b_4$ are identified to minimize the error between the manufacturer’s data and the modeled $\eta_{\text{comp}}$. Figure III-3 shows the fit between the modeled and the manufacturer’s compressor efficiency data. Majority of the modeled compressor efficiency curve is within 5% of the manufacturer’s data with some points exceeding the 5% error band marginally.
3.1.3 Compressor outlet temperature model

The compressor outlet temperature is simply calculated from the computed compressor efficiency and the pressure ratio across the compressor as follows.

$$T_{\text{CompOut}} = T_{\text{Cin}} \left( \frac{P_{\text{Bst}}}{P_{\text{Cin}}} \frac{K - 1}{\eta_{\text{Comp}}} - 1 \right) + 1$$

Where, $T_{\text{Cin}}$ is the compressor inlet temperature calculated using the adiabatic mixing equation as follows.
\[ T_{cin} = \frac{m_{Egr}T_{Egr} + m_{Air}T_{Air}}{m_{Egr} + m_{Air}} \]

Where, \( m_{Egr} \) is assumed to be modeled and \( m_{Air} \) is directly measured from the engine air mass flow rate sensor which is a few inches upstream of the EGR mixing location. \( T_{Egr} \) is the temperature of the EGR and it is assumed to be available as a measurement. Preliminary validation against GT-Power simulation of the compressor inlet and outlet temperature models is shown in Figure III-4 and Figure III-5.

![Figure III-4. 0D Modeled vs GT-Power modeled compressor inlet temperature](image-url)
3.1.3 Turbine mass flow rate and pressure ratio model

Two turbine mass flow rate models have been developed. One model which predicts the mass flow rate across the turbine as a function of the pressure ratio across it and another model which predicts the pressure ratio as a function of known mass flow rate across the turbine. The turbine mass flow rate is represented as a reduced Turbine Flow Parameter as follows.
\[ TFP = \frac{\dot{m}_{Turb} \sqrt{T_{Tin}}}{P_{Tin}} \times 10^5 \]

The manufacturer’s turbine data is represented in this format and hence conversion to mass flow rate is required in order to fit a model to the data. The model used to predict turbine mass flow rate as a function of pressure ratio is a modified version of the standard restriction model described by Eriksson in [10] has been utilized to model the relationship between the pressure ratio and mass flow rate across the turbine. The structure of this model is very similar to the simplified orifice flow model and is defined as follows.

\[ \dot{m}_{Turb} = A_{Turb} \frac{P_{Tin}}{\sqrt{R \cdot T_{Tin}}} \cdot \psi_{PR} \left( \frac{P_{Tin}}{P_{Tout}} \right) \]

Where, the flow function \( \psi_{PR} \) is given by,

\[ \psi_{PR} \left( \frac{P_{Tin}}{P_{Tout}} \right) = \left( 2 \sqrt{\frac{P_{Tin}}{P_{Tout}}} \left[ 1 - \sqrt{\frac{P_{Tin}}{P_{Tout}}} \right] \right)^{1/2} \]

The area parameter \( A_{Turb} \) is modeled as a linear function of turbocharger speed

\[ A_{Turb} = j_1 \omega_T + j_2 \]

The parameters \( j_1 \) and \( j_2 \) are identified using the manufacturer’s data. The main difference between this flow function and the one used in the orifice model is that \( \psi_{PR} \) is applied to the square root of the pressure ratio across the turbine as mentioned in [53]. The flow function of the standard orifice model changes its structure when the pressure
ratio reduces below the critical pressure ratio. This causes the standard orifice model to operate under the choked flow regime where the flow function $\psi_{PR}$ loses coupling with the turbine outlet pressure. Utilizing the square root of the pressure ratio instead of the pressure ratio itself, significantly improves the accuracy of the restriction based model. Figure III-6 shows the modeled and the manufacturer’s turbine TFP points.

![Modeled turbine mass flow rate vs manufacturer provided turbine map points](image)

Due to it’s non-linear nature and the square root terms present in the model, the inversion of this restriction based model is very difficult. Due to this concern, a simple $3^{rd}$ order polynomial model is defined which predicts turbine pressure ratio as a function of mass flow rate as follows.
\[ \frac{p_{\text{tin}}}{p_{\text{tout}}} = \alpha_1 (\dot{m}_{\text{turb}})^2 + \alpha_2 (\dot{m}_{\text{turb}}) + \alpha_3 \]  

Where, \( \alpha_3 \) is a constant with value 1. The rest of the \( \alpha_i \) terms are not constant since there is a dependency on turbine inlet temperature and the mass flow rate. Based on the turbine map, the \( \alpha_i \) terms were identified to be linear and second order polynomial terms of turbine inlet temperature as follows.

\[ \alpha_1 = \beta_1 T_{\text{tin}} + \beta_2 \]  

\[ \alpha_2 = \beta_3 T_{\text{tin}}^2 + \beta_4 T_{\text{tin}} + \beta_5 \]  

Where, the \( \beta_i \) terms are constants identified by minimizing the error between the model and the manufacturer data. The fit of the model vs the manufacturer’s data is shown for four different turbine inlet temperatures in Figure III-7.
Figure III-7. Turbine pressure ratio as a function of turbine mass flow rate with map points for four different turbine inlet temperatures.
3.1.4 Turbine efficiency model

In order to derive turbine isentropic efficiency, the normalized parameter blade speed ratio [65] is defined as follows.

\[
BSR = \frac{\omega_{Tr_t}}{\sqrt{2C_{PEXh}T_{Tin}(1-(\frac{P_{Tin}}{P_{Tout}})^{K_{EXh}-1})}}
\]

Where, \(c_{p_{EXh}}\) and \(K_{EXh}\) are the constant pressure specific heat capacity and the ratio of specific heat capacities of the exhaust gas, modeled as polynomial functions of exhaust temperature as follows.

\[
\eta_{Turb} = d_1 BSR^2 + d_2 BSR + d_3
\]

The parameters \(d_i\) are obtained similarly to the previous polynomial approximations in the compressor and turbine models. The turbine efficiency is shown in Figure III-8 as a function of \(BSR\). The model is accurate within five percent for majority of the turbine map points except for the six points at the lowest turbo speed.
3.1.5 Turbocharger torque models

The compressor and turbine torque models are given by the following equations [42]

\[
\tau_{\text{Comp}} = \dot{m}_{\text{Comp}} \cdot c_p A T_{\text{cin}} \left[ \frac{P_{\text{bst}}}{P_{\text{cin}}} \right]^{K-1} \left[ \frac{1}{\eta_{\text{Comp}} \omega_T} \right] \tag{26}
\]

\[
\tau_{\text{Turb}} = \dot{m}_{\text{Turb}} \cdot c_p A T_{\text{exh}} \left[ 1 - \left( \frac{P_{\text{tin}}}{P_{\text{out}}} \right)^{K_{\text{exh}}-1} \right] \left[ \frac{\eta_{\text{Turb}}}{\omega_T} \right] \tag{27}
\]

These equations are used in the turbocharger rotational dynamics equation as follows
\[ \frac{d\omega_T}{dt} = \frac{1}{J_T} (\tau_{Turb} - \tau_{Comp} - \tau_{Fric}) \]

Where, \( \tau_{Fric} \) is the friction torque which is assumed to be zero, since this parameter is associated with the mechanical efficiency term which is lumped into the turbine efficiency as mentioned in [53]. \( J_T \) is the moment of inertia of the rotor of the turbocharger. This parameter can be measured based on techniques outlined in [43]. Alternatively it can be modeled if the exact geometry and material composition of the turbocharger rotor assembly is available. These parameters can be used on the part file in any commercial computer aided design software to estimate the moment of inertia of the entire rotor assembly. For this research, this parameter has been derived experimentally by comparing the turbocharger speed behavior of the 1D Engine simulation model to the experimental data for a transient engine response. The moment of inertia utilized in the 1D Model is adjusted till the 1D Engine model mimics the turbo speed behavior of the real engine. The final value was determined to be \( 1.7 \times 10^{-5} kg/m^2 \) which is within the range of values for similar sized turbochargers. The turbocharger speed response for the tuned \( J_T \) can be seen in Figure III-9 and Figure III-10. The time constants of the 1D-Engine simulation as well as the experimentally measured turbocharger speed appear very similar to each other.
Figure III-9. Comparison of measured and GT-Power modeled transient turbocharger speed response with tuned inertia for 1500 RPM

Figure III-10. Comparison of measured and GT-Power modeled transient turbocharger speed response with tuned inertia for 2500 RPM
3.2 Data driven parameter models for engine cylinders

Modeling of the cylinder gas-exchange and combustion processes has been an active research area since the advent of internal combustion engines. Parametric models based on physics, empirical relations and measured data have been widely developed and investigated. Since, engine modeling is not the focus of this research and the means to obtain measurement data is easily available, a data driven approach has been adopted to model these processes. The data was obtained by operating the engine at multiple actuator combinations at 2000 and 3000 RPM. Each combination was allowed to settle for some time before transitioning to the next combination. This time period is critical because lesser time results in some of the slower dynamic behavior like exhaust temperature to be transient even at the point of transition to the next combination. More time per operating point is desirable but would lead to increased time required to obtain the entire data set. A time period of five seconds was chosen per operating combination which resulted in a total of forty hours of dynamometer test time. The actuator combinations are shown in Table III-2. The actuator combinations are programmed into an automated test cycle and deployed onto the ES910’s real-time controller.

In the control oriented model of the engine, there are four models that correspond to the gas-exchange and in-cylinder combustion process of the engine. The gas-exchange model is the volumetric efficiency ($\eta_{VE}$) model. The remaining cylinder process models are used to predict the Net Indicated Mean Effective Pressure ($IMEP_n$), Coefficient of
Variation of $IMEP_n$ ($COV_{IMEP_n}$), Turbine inlet gas temperature ($T_{Exh}$), and the square of the Knock Intensity ($KI^2$).

Table III-2. Actuator combinations used for data collection for ANN model training

<table>
<thead>
<tr>
<th>Variable</th>
<th>Points</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPM</td>
<td>2000, 3000</td>
<td>RPM</td>
</tr>
<tr>
<td>Load (WG)</td>
<td>0, 50, 100</td>
<td>%</td>
</tr>
<tr>
<td>EGR</td>
<td>0, 5, 10, 15, 20</td>
<td>%</td>
</tr>
<tr>
<td>Spark</td>
<td>10, 5, 0, -5, -10</td>
<td>deg from base</td>
</tr>
<tr>
<td>P_int (throttle)</td>
<td>25, 31, 37, 44</td>
<td>%</td>
</tr>
<tr>
<td>ICL</td>
<td>40, 30, 20, 10, 0</td>
<td>(bTDC)</td>
</tr>
<tr>
<td>ECL</td>
<td>5, 15, 25, 35, 45</td>
<td>(aTDC)</td>
</tr>
</tbody>
</table>

The data acquisition is done conveniently using ETAS INCA’s Open Hardware Integration (OHI) interface. Cyclic, crank-angle resolved cylinder pressure measurement is required for the calculation of $IMEP_n$ and $KI^2$. AVL Indicom’s calc-graf tool is utilized to calculate $KI^2$ in real time. $IMEP_n$ is automatically calculated by AVL Indicom per cycle. The INCA OHI interface is configured to communicate in real-time with AVL Indicom to receive a configured list of parameters computed from the cylinder pressure. This list of parameters is then available in the INCA workspace which also consists of the control actuator signals sent to the engine. This test structure shown in Figure III-11.
facilitates collection of data in one file per test using only one software (INCA) which makes post processing of this data much simpler.

Figure III-11. Experimental setup to facilitate data acquisition for engine cylinder models

Once data collection is complete, it is post-processed to identify the steady state portion of the five seconds per combination of input actuators. This is done because the measurement of $\eta_{VE}$ requires measurement of mass flow rate across the intake valves which is not done directly on the engine. Instead, the mass air flow rate sensor which is far upstream of the cylinders is utilized, which during transient operation is not representative of the engine mass flow rate.

Similarly, $COV_{IMEP_n}$ calculation requires an ensemble of $IMEP_n$ measurements at a steady state. Typically, this calculation is done based on hundreds of engine cycles of steady state data. For the un-boosted test conditions upto 600 consecutive steady-state
cycles could be identified from the data. However, for some of the high-load boosted conditions, the $IMEP_n$ would still be in a slow transient at the end of the five seconds due to the slow dynamics (exhaust manifold temperature, intercooler outlet gas temperature) of the engine. For such conditions, the last twenty cycles of the engine operation before the actuator combination change were considered to be steady-state operating conditions.

The inputs to the ANN models are given in Table III-3

Table III-3 Inputs used for training ANNs

<table>
<thead>
<tr>
<th>Input</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spark</td>
<td>deg bTDC</td>
</tr>
<tr>
<td>Intake manifold pressure</td>
<td>Bar</td>
</tr>
<tr>
<td>Intake temperature</td>
<td>K</td>
</tr>
<tr>
<td>Intake valve phasing</td>
<td>deg aTDC</td>
</tr>
<tr>
<td>Exhaust valve phasing</td>
<td>deg aTDC</td>
</tr>
<tr>
<td>RPM</td>
<td>rpm</td>
</tr>
<tr>
<td>Partial pressure of EGR</td>
<td>bar</td>
</tr>
</tbody>
</table>
3.2.1 $\eta_{VE}$ model

Neural network based methods have been investigated for engine air-flow estimation by Nicolau et al in [84] and Wu et al in [85]. Neural networks training is performed using Matlab’s neural network toolbox [86]. The accuracy for the model for the collected data can be seen in Figure III-12. This network is modeled using a fifteen neuron ANN with a single hidden layer. Majority of the test data points fall within the +/- 10% band. As seen in Figure III-12

![Graph showing modeled vs measured volumetric efficiency](image)

Figure III-12. Modeled vs measured volumetric efficiency [%] for all data with +/- 10% error bands represented by dashed red line
3.2.2 \( IMEP_n \) model

Similar to the volumetric efficiency model the IMEPn is modeled using a fifteen neuron, single hidden layer ANN. The model seems to lose accuracy at the lower load points due to the high combustion variability noticed at these conditions resulting in noisy \( IMEP_n \) data.

![Figure III-13. Modeled vs measured IMEPn for all data with +/- 10% error bands represented by dashed red line](image_url)
3.2.3 $COV_{IMEP_n}$ model

The $COV_{IMEP_n}$ is computed as follows for steady state ensembles of data.

$$COV_{IMEP_n} = \frac{stddev(IMEP_n)}{mean(IMEP_n)}$$

Where, $stddev(IMEP_n)$ is the standard deviation of the $IMEP_n$ and $mean(IMEP_n)$ is the average $IMEP_n$. The model is able to capture trends but lacks accuracy.

Figure III-14. Modeled vs measured $COV_{IMEP_n}$ for all data with +/- 10% error bands represented by dashed red line
3.2.4 $KI^2$ model

The $KI^2$ model has the worst accuracy of all the models. This is primarily due to the lack of training data in the heavy knocking region. Since knocking is very detrimental to the engine, collection of data at heavy knock was limited.

Figure III-15. Modeled vs measured $KI^2$ for all data with +/- 10% error bands represented by dashed red line
3.2. $T_{Exh}$ model

Similar to the volumetric efficiency and $IMEP_n$ models, the steady state exhaust temperature model is able to predict most of the points within the +/- 10% region. The model was fitted to steady state exhaust temperature because the time constant associated with the transient exhaust temperature is much longer than the turbocharger time constant. Including these dynamics in the control model would require lengthening of the horizon and consequently increasing the sampling time of the MPC to maintain real-time execution feasibility, which would then require removing some of the pressure dynamics to keep number of states in the model low.

Figure III-16. Modeled vs measured steady state exhaust temperature for all data +/- 10% error bands represented by dashed red line. The accuracy of each of these neural networks is summarized in Table III-4
Table III-4. Accuracy of trained ANNs

<table>
<thead>
<tr>
<th>Parameter</th>
<th>RMS Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$IMEP_n$</td>
<td>0.45 Bar</td>
</tr>
<tr>
<td>$COV_{IMEP_n}$</td>
<td>1.4%</td>
</tr>
<tr>
<td>$K_1^2$</td>
<td>0.297 Bar$^2$</td>
</tr>
<tr>
<td>$\eta_{VE}$</td>
<td>6.7%</td>
</tr>
<tr>
<td>$T_{Exh}$</td>
<td>14 Deg K</td>
</tr>
</tbody>
</table>
3.3 Continuous time state space model of the system

Based on the models described in the previous sections, a continuous time state-space model of the system is formed. The inputs and outputs are also normalized in magnitude to have single digit values during operation. The units which are used for normalization are shown in Table III-5

Table III-5 Units used for normalization of state space model

<table>
<thead>
<tr>
<th>State/input/parameter</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spark timing</td>
<td>deg bTDC/10</td>
</tr>
<tr>
<td>Mass flow rate</td>
<td>kg/min</td>
</tr>
<tr>
<td>Intake valve phasing</td>
<td>deg aTDC/10</td>
</tr>
<tr>
<td>Exhaust valve phasing</td>
<td>deg aTDC/10</td>
</tr>
<tr>
<td>Pressure</td>
<td>Bar</td>
</tr>
<tr>
<td>IMEPn</td>
<td>Bar</td>
</tr>
<tr>
<td>COV</td>
<td>[Ratio]</td>
</tr>
<tr>
<td>Turbo speed</td>
<td>kRad/sec</td>
</tr>
<tr>
<td>Temperature</td>
<td>Deg K</td>
</tr>
<tr>
<td>Volume</td>
<td>m^3</td>
</tr>
</tbody>
</table>

Based on the physical unit based normalization scheme above the state dynamics and the outputs are represented as follows.
\[
\dot{P}_{\text{Im}} = - \frac{V_d \eta_{\text{VE-RPM}} P_{\text{Im}}}{V_{\text{Im},120}} + \frac{R.T_{\text{Im}}}{V_{\text{Im},6 \times 10^6}} \dot{m}_{\text{ITV}} \\
\dot{P}_{\text{Bst}} = \left( \frac{R(T_{\text{Comp,out}} + T_{\text{IC,out}})}{2V_{\text{Bst}}} \right) (\dot{m}_{\text{Comp}} - \dot{m}_{\text{ITV}} - \beta_{\text{Bov}} (P_{\text{Bst}} - P_{\text{Cln}})) \\
\dot{P}_{\text{Egr}} = - \frac{V_d \eta_{\text{VE-RPM}} P_{\text{Egr}}}{V_{\text{Im},120}} + \frac{R.T_{\text{Im}}}{V_{\text{Im},6 \times 10^6}} \dot{m}_{\text{Egr}} \\
\dot{\omega}_T = \frac{J_T^{-1}}{1000} (\tau_{\text{Turb}} - \tau_{\text{Comp}}) \\
x = \begin{bmatrix} P_{\text{Im}} & P_{\text{Bst}} & P_{\text{Egr}} & \omega_T \end{bmatrix}^T, x \in \mathbb{R}^{4 \times 1}
\]

Where, suffixes \( \text{Im}, \text{Bst} \) and \( \text{Egr} \) mean Intake manifold, post-compressor boost manifold, and the EGR in the intake manifold. The input vector to the model is defined as follows.

\[
u = \begin{bmatrix} \dot{m}_{\text{ITV}} & \beta_{\text{Bov}} & \dot{m}_{\text{WG}} & \dot{m}_{\text{Egr}} & \text{SPK} & \text{ICL} & \text{ECL} \end{bmatrix}^T, u \in \mathbb{R}^{7 \times 1}
\]

Where, suffixes \( \text{ITV}, \text{Bov} \) and \( \text{WG} \) denote Intake throttle valve, Blow-off valve, and Turbine waste-gate respectively. The output that is to be tracked and the constraints are given by the following equations.

\[
y = \text{IMEP}_n, y \in \mathbb{R}
\]

\[
z = \begin{bmatrix} (\text{COV}_{\text{IMEP}_n}) & (K_1^2) & (P_{\text{Im}} - P_{\text{Bst}}) & (\dot{m}_{\text{SL}} - \dot{m}_{\text{comp}}) & (\dot{m}_{\text{WG}} - \dot{m}_{\text{Eng}}) \end{bmatrix}^T, z \in \mathbb{R}^{5 \times 1}
\]
Where, the constraint $P_{im} - P_{bst} < 0$ and $\\dot{m}_{WG} - \dot{m}_{Eng} < 0$ ensures that the flow across the intake throttle valve and the waste-gate is physically meaningful. The constraint $\\dot{m}_{SL} - \dot{m}_{Comp} < 0$ ensures that the compressor mass flow rate is always on the right hand side of the surge line modeled as a linear function of $P_{bst}$ as follows similar to [88].

$$\\dot{m}_{SL} = S_{G1}.P_{bst} + S_{G2}$$

Where, $S_{G1}$ and $S_{G2}$ are the coefficients of the surge-line of the compressor.

Open-loop validation of the state space model against experimentally measured engine data has been shown in Figure III-17 for 3000RPM. The complete error metrics for the 0D engine model validated against four different engine transients at 2000 and 3000 RPM are shown in Table III-6. The model is within +/- 10% accuracy for the states and the IMEPn output.

Table III-6 Error metrics of the 0D model for four different transient engine trajectories.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>RMS</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMEPn</td>
<td>1.07</td>
<td>4.9%</td>
</tr>
<tr>
<td>Pim</td>
<td>0.01</td>
<td>8.25%</td>
</tr>
<tr>
<td>Pbst</td>
<td>0.029</td>
<td>7.1%</td>
</tr>
<tr>
<td>Pegr</td>
<td>0.00008</td>
<td>7.6%</td>
</tr>
<tr>
<td>Wt</td>
<td>2.39</td>
<td>9.15%</td>
</tr>
</tbody>
</table>
Figure III-17 Open-loop system trajectory for 0D model vs experimentally collected data
IV. NON-LINEARITY ANALYSIS OF CONTROL ORIENTED MODEL

4.1 Local linear approximation

Utilization of the non-linear state-space model in MPC requires local linear approximation. Previous research has used approximated linear state-space models for utilization in MPC. The underlying assumption of these approaches is that the behavior of the non-linear engine system is linear-like for the entirety of the prediction horizon of the MPC. In order to understand the effect of this assumption on the performance of MPC, the following analysis has been performed.

- The deviation of the locally linearized model from the non-linear model is quantified as the Linear Approximation Error (LAE)
- The LAE has been evaluated for a sinusoidal as well as a step change in inputs

The analysis has been done for six different steady state operating points which are combinations of 2000 and 3000 RPM and IMEPn load levels of 5, 10 and 15 Bar.

4.2 Step input transient testing

The assumption that the linearized model is a sufficiently accurate representative of the non-linear model 30 to 37 of the system is often made while designing MPC algorithms. The IMEPn accuracy of the linearized model is of specific importance for MPC designed for torque reference tracking. The LAE for the IMEPn output has been quantified by first linearizing the model at a steady state operating point. Then the open
loop inputs which correspond to a deviation from the steady state IMEPn are applied to both the models. The LAE for IMEPn is then examined as a function of time-step N from the linearization point as shown in Figure IV-1.

![Figure IV-1 Deviation between the non-linear system model and the linearized system model as a function of time step N from the linearization. The deviation is induced by perturbing $\dot{m}_{ITV}$ corresponding to commanded IMEPn](image)

The IMEPn LAE is quantified as the absolute error between the IMEPn from the non-linear and the linearized model as follows

$$\delta IMEPn = |IMEPn_{NL} - IMEPn_L|$$
The $\delta IMEP_n$ is computed for different open loop step input $\dot{m}_{ITV}$ for which the final values correspond to the steady-state deviation from the linearization point IMEPn. This is denoted in general as $\Delta IMEPn$. Figure IV-2 shows the $\delta IMEPn$ as a function of the time-step after linearization $N$ and $\Delta IMEPn$ for all six operating conditions. As expected for $\Delta IMEPn = 0$, all of the conditions show $\delta IMEPn = 0$ since there is no change in input. Also, as expected, $\delta IMEPn = 0$ for $N = 0$ since it corresponds to the time at which the model is linearized. As $\Delta IMEPn$ is deviated from 0 and $N$ is increased, $\delta IMEPn$ begins to grow larger. The highest deviation is seen at the 3000 RPM 15 Bar case for $\Delta IMEPn = -4$. This implies that the non-linear $IMEPn$ output model varies more with time at this operating condition compared to the others. Similarly, the least deviation is seen at 3000 RPM 5 Bar case. Implementing MPC using such approximated models would result in possibly reduced optimality and constraint violation. Based on the analysis done on the six operating conditions shown in the errors induced by linear approximation are significantly different as a function of operating conditions. Additional linearization within the horizon causes $\delta IMEPn$ to naturally reduce due to improved local accuracy of the linearized model at the additional linearization points. However, since model linearization is a computationally expensive process, it is necessary to identify the number of times the non-linear model needs to be linearized in order to maintain a prescribed model accuracy. Table IV-1. shows the minimum number of linearizations required to maintain different minimum accuracy levels of modeled $IMEPn$. 
Figure IV-2 Model error as a function of commanded open-loop change in IMEPn shows that the deviation between linearized and non-linear model varies as a function of operating condition and the time–step after the point of linearization.

Table IV-1 Minimum number of model linearizations required in horizon to maintain IMEPn model accuracy for varying input frequency

<table>
<thead>
<tr>
<th>Pressure</th>
<th>2000 RPM</th>
<th>3000 RPM</th>
<th>Max IMEPn error [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 Bar</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>10 Bar</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>5 Bar</td>
<td>10</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>15 Bar</td>
<td>1</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>10 Bar</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5 Bar</td>
<td>&gt;10</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>15 Bar</td>
<td>5</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>10 Bar</td>
<td>5</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>5 Bar</td>
<td>&gt;10</td>
<td>10</td>
<td>1</td>
</tr>
</tbody>
</table>
4.3 Sinusoidal input transient testing

Similar to the step-input tests, sinusoidal inputs of frequency increasing from 0 to 2 Hz was applied to the model to understand the LAE as a function of input frequency. The significance of this test is to analyze the dependence between the LAE and the rate at which the engine transient occurs. Similar to the step input test, Table IV-2 shows the $\delta IMEP_n$ as a function of the time-step after linearization N and frequency of the input that was applied to the models. At every operating condition a certain frequency results in the maximum LAE. This frequency ranges between 0.5 and 1.5 Hz depending on the operating condition. However, similar to the step-input test the minimum $\delta IMEP_n$ is seen at the 3000 RPM 5 Bar case. The maximum $\delta IMEP_n$ is seen at the 3000 RPM 10 Bar case for 0.9 Hz input frequency. Similar to the step-input case the minimum number of linearizations required for different prescribed accuracy levels have been quantified. In both of the tests we can see that some of the operating conditions require the model to be linearized more than 10 times within the horizon to meet the desired accuracy level. This is due to the inaccuracy associated with discretization of the continuous model. Since the sample time chosen here is 50ms, for a prediction horizon of 500ms, the maximum number of linearizations that can be performed is 10.
Table IV-2 Minimum number of model linearizations required in horizon to maintain IMEPn model accuracy for varying input frequency

<table>
<thead>
<tr>
<th>Pressure</th>
<th>2000 RPM</th>
<th>3000 RPM</th>
<th>Max IMEPn error [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 Bar</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>10 Bar</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>5 Bar</td>
<td>&gt;10</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>15 Bar</td>
<td>2</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>10 Bar</td>
<td>5</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>5 Bar</td>
<td>&gt;10</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>15 Bar</td>
<td>10</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>10 Bar</td>
<td>10</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>5 Bar</td>
<td>&gt;10</td>
<td>&gt;10</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure IV-3 Model error as a function of frequency of commanded open-loop model input shows that the deviation between linearized and non-linear model generally peaks between 1.5 Hz and 0.5 Hz depending on operating condition.
The choice of prediction model utilized for MPC is critical for the stability, optimality and robustness of the closed loop controller performance. This research presents a method of evaluating the impact of local linear approximation error on the modeled IMEPn by manipulation of the throttle mass flow rate input. The same method could be applied to every other input of the model to quantify the effect on the rest of the outputs of the model.

The transient tests conducted in this study show clear deviation between the non-linear model and its linearized version for different operating conditions. Based on this analysis it is clear that this deviation varies significantly with the type of transient induced in the system. The step transients are generally more forgiving with respect to the number of model linearizations required, compared to the sinusoidal transients. The sinusoidal transients clearly show a narrow frequency range where the deviation in the models is highest.

In order to maintain a specific minimum accuracy, it is possible to identify the minimum number of linearizations required within the horizon as a function of operating condition and the commanded transient to the MPC. However, for some operating conditions, linearization even at every stage within the prediction horizon is not sufficient to ensure that Max IMEPn error stays within 5%.
5.1 Horizon length and sample time determination

As mentioned in [54] by Zhu, a cascaded control structure is suitable for engine control. Complex cycle-cycle behavior and fast dynamics can be handled effectively by simpler, computationally cheaper controllers. However, the determination of the MPC sample time $\delta t$ is dependent on the engine dynamics considered in the MPC control model, the prediction/control horizon considered and the computational ability of the micro-controller on which the MPC is designed to execute on as shown in Figure V-1. Closed loop stability of MPC requires prediction horizon to be long enough to accommodate the slowest dynamics in the system [55][56]. The turbocharger speed has the slowest dynamics in the system with time constant in the order of seconds (ignoring the exhaust manifold wall temperature and the intercooler outlet temperature dynamics).

![Figure V-1 Approximate relationship between horizon length considered for MPC, size of the model and the computational expense of real-time execution.](image-url)
Figure V-2 The selection of MPC sample time and prediction horizon time dictates the choice of appropriate engine dynamics in the engine model.

The sample time is chosen to be 50 ms in order to have a MPC prediction/control horizon length of 10 steps which is 0.5 s. This allows consideration of the turbocharger dynamics, the filling and emptying dynamics of the boost manifold and the intake manifold. The exhaust manifold and the other control volumes on the engine are too small and hence have almost instantaneous dynamics which can then be handled by lower level controllers.

Based on the study shown in the previous chapter, it is evident that utilizing one linearized model to predict the behavior of the engine over a period of time similar to the prediction horizon of the MPC can result in significant errors in the modeled outputs. Ideally, linearization at every stage within the prediction horizon would be desirable to conserve accuracy. However, the process of obtaining the linearized model requires
computational resources. The following section will explore the methods used to minimize computational burden associated with linearization.

### 5.2 Non-linear system model linearization

The system model described in (30) to (37) can be linearized as follows using the Taylor’s series expansion (ignoring higher terms) in continuous time domain as follows.

\[
\dot{x} = f(x_0, u_0) + \delta \dot{x} = \dot{x}_0 + \frac{\partial f}{\partial x}|_{x_0,u_0} \delta x + \frac{\partial f}{\partial u}|_{x_0,u_0} \delta u
\]

\[
[\ddot{y} \dot{z}] = [\ddot{y}_0 \dot{z}_0] + \begin{bmatrix} \frac{\partial g}{\partial x}|_{x_0,u_0} \\ \frac{\partial h}{\partial x}|_{x_0,u_0} \end{bmatrix} \delta x + \begin{bmatrix} \frac{\partial g}{\partial u}|_{x_0,u_0} \\ \frac{\partial h}{\partial u}|_{x_0,u_0} \end{bmatrix} \delta u
\]

Where, \(x_0\) and \(u_0\) are the nominal state vectors and input vectors respectively. \(y_0\) and \(z_0\) are the corresponding nominal tracking output and constraint output vectors. \(\dot{x}_0\) is the vector of derivatives of the states evaluated at the nominal point. \(\delta x\) is the deviation of the states from the nominal point. \(\frac{\partial f}{\partial x}|_{x_0,u_0}\) and \(\frac{\partial f}{\partial u}|_{x_0,u_0}\) are the Jacobian matrices evaluated at \(x_0\) and \(u_0\) by partial differentiation of the state dynamics 3-30 to 3-33 with respect to \(x\) and \(u\). Similarly, \(\frac{\partial g}{\partial x}|_{x_0,u_0}\), \(\frac{\partial g}{\partial u}|_{x_0,u_0}\), \(\frac{\partial h}{\partial x}|_{x_0,u_0}\) and \(\frac{\partial h}{\partial u}|_{x_0,u_0}\) are the Jacobian matrices evaluated at \(x_0\) and \(u_0\) by partial differentiation of the output functions 3-36 to 3-37 with respect to \(x\) and \(u\).
5.2.1 Drawbacks of numerical linearization

These partial differential functions described above can be approximated by using the forward finite difference approach as follows.

\[
\frac{\partial c(x)}{\partial x} \approx \frac{(c(x+\Delta x) - c(x))}{\Delta x}
\]

Where, \(\Delta x\) is the perturbation magnitude which is chosen to be sufficiently small to approximate the slope of the arbitrary non-linear function \(c\). Utilizing this method is simple since it only requires evaluations of \(c\) at different values of the argument. However, since the engine system model has four states and seven inputs, utilizing this method to obtain the linearized system model at one nominal operating condition will require thirty eight evaluations of the system dynamics 3-30 to 3-33. Similarly, in order to obtain the linearized output models, the output functions 3-36 to 3-37 would have to be evaluated fifty six times. This would lead to ninety four evaluations of the system model in order to obtain the linearized system dynamics. This process would have to be repeated at every stage in the prediction horizon in order to minimize the linear approximation error, resulting in nine hundred and forty evaluations of the system model at every time-step of execution. This process takes a large amount of memory and processor throughput and leaves no resources left for the optimization algorithm. In the worst case, it will cause a task overrun wherein the computation fails to execute within the given time-step.
5.2.2 Hybrid linearization of the system model

The advantage of utilizing analytical differentiation is that unlike the numerical approach described in the previous section, the analytical solution is exact and it requires significantly less number of executions of the system model. A series of steps taken in order to maximize analytical differentiation within the system model is described in the following sections.

5.2.2.1 Fixed zero identification

The system model is first analyzed for potential to apply analytical partial differentiation. In this the first process is to identify the fixed zeros in the system model.

\[
\frac{\partial f}{\partial x} = \begin{bmatrix}
\frac{\partial f_{P_{lm}}}{\partial P_{lm}} & 0 & \frac{\partial f_{P_{lm}}}{\partial P_{E_{gr}}} & 0 \\
0 & \frac{\partial f_{P_{bst}}}{\partial P_{bst}} & 0 & \frac{\partial f_{P_{bst}}}{\partial \omega_{T}} \\
\frac{\partial f_{\omega_{T}}}{\partial P_{lm}} & \frac{\partial f_{\omega_{T}}}{\partial P_{bst}} & \frac{\partial f_{\omega_{T}}}{\partial P_{E_{gr}}} & \frac{\partial f_{\omega_{T}}}{\partial \omega_{T}}
\end{bmatrix}
\]
\[
\frac{\partial f}{\partial u} = \begin{bmatrix}
\frac{\partial f_{P_{Im}}}{\partial m_{ITV}} & 0 & 0 & 0 & \frac{\partial f_{P_{Im}}}{\partial IC_{L}} & \frac{\partial f_{P_{Im}}}{\partial EC_{L}} \\
\frac{\partial f_{P_{Bst}}}{\partial m_{ITV}} & \frac{\partial f_{P_{Bst}}}{\partial \beta_{BOV}} & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{\partial f_{\omega_{T}}}{\partial m_{WG}} & 0 & \frac{\partial f_{\omega_{T}}}{\partial SP_{K}} & \frac{\partial f_{\omega_{T}}}{\partial IC_{L}} \\
0 & 0 & \frac{\partial f_{\omega_{T}}}{\partial \beta_{BOV}} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{\partial f_{\omega_{T}}}{\partial IC_{L}} & 0 & \frac{\partial f_{\omega_{T}}}{\partial EC_{L}} & 0 \\
\end{bmatrix}
\]

\[
\frac{\partial g}{\partial x} = \begin{bmatrix}
\frac{\partial (IMEP_{n})}{\partial P_{Im}} & 0 & \frac{\partial (IMEP_{n})}{\partial P_{Egr}} & 0 \\
\end{bmatrix}
\]

\[
\frac{\partial h}{\partial u} = \begin{bmatrix}
\frac{\partial (COV_{IMEP_{n}})}{\partial P_{Im}} & 0 & \frac{\partial (COV_{IMEP_{n}})}{\partial P_{Egr}} & 0 \\
\frac{\partial (K1^2)}{\partial P_{Im}} & 0 & \frac{\partial (K2)}{\partial P_{Egr}} & 0 \\
0 & 0 & 0 & 0 \\
1 & 0 & \frac{\partial (\xi_{Surge})}{\partial P_{Bst}} & 0 \\
0 & \frac{\partial (\xi_{Surge})}{\partial \omega_{T}} & 0 & 0 \\
\end{bmatrix}
\]

The fixed zeros are identified by simply inspecting the differential equations in state dynamics and the output equations. A fixed zero would simply denote the absence of the corresponding state or output within that equation/model. It is evident that there is a significant number of fixed zeros in the model which substantially alleviates the computational burden of model linearization. Since the forward throttle flow constraint \(\xi_{ITV}\) is a linear function of the states the corresponding elements of \(\frac{\partial h}{\partial u}\) are 1 and -1 respectively.
5.2.2.2 Analytical and hybrid differentiation of system model

The analytical differentiation of the system dynamics for $P_{Im}$ and $P_{EGR}$ is simple and given by the following equations.

$$\frac{\partial f_{P_{Im}}}{\partial P_{Im}} = \frac{V_d RPM}{-120V_{Im}} \left( P_{Im} \frac{\partial \eta VE}{\partial P_{Im}} + \eta VE \right)$$  \hspace{1cm} (47)

$$\frac{\partial f_{P_{Im}}}{\partial P_{EGR}} = \frac{V_d RPM}{-120V_{Im}} \left( P_{Im} \frac{\partial \eta VE}{\partial P_{EGR}} \right)$$  \hspace{1cm} (48)

$$\frac{\partial f_{P_{Im}}}{\partial m_{ITV}} = \frac{T_{Im} R}{V_{Im} 6 \times 10^6}$$  \hspace{1cm} (49)

$$\frac{\partial f_{P_{Im}}}{\partial ICL} = \frac{V_d RPM P_{Im}}{-120V_{Im}} \left( \frac{\partial \eta VE}{\partial ICL} \right)$$  \hspace{1cm} (50)

$$\frac{\partial f_{P_{Im}}}{\partial ECL} = \frac{V_d RPM P_{Im}}{-120V_{Im}} \left( \frac{\partial \eta VE}{\partial ECL} \right)$$  \hspace{1cm} (51)

$$\frac{\partial f_{P_{EGR}}}{\partial P_{Im}} = \frac{V_d RPM}{-120V_{Im}} \left( P_{EGR} \frac{\partial \eta VE}{\partial P_{Im}} \right)$$  \hspace{1cm} (52)

$$\frac{\partial f_{P_{EGR}}}{\partial P_{EGR}} = \frac{V_d RPM P_{EGR}}{-120V_{Im}} \left( \frac{\partial \eta VE}{\partial P_{EGR}} + \eta VE \right)$$  \hspace{1cm} (53)

$$\frac{\partial f_{P_{EGR}}}{\partial m_{EGR}} = \frac{\partial f_{P_{Im}}}{\partial m_{ITV}}$$  \hspace{1cm} (54)

$$\frac{\partial f_{P_{EGR}}}{\partial ICL} = \frac{V_d RPM P_{EGR}}{-120V_{Im}} \left( \frac{\partial \eta VE}{\partial ICL} \right)$$  \hspace{1cm} (55)

$$\frac{\partial f_{P_{EGR}}}{\partial ECL} = \frac{V_d RPM P_{EGR}}{-120V_{Im}} \left( \frac{\partial \eta VE}{\partial ECL} \right)$$  \hspace{1cm} (56)
It is visible in this analytical framework that there are common computations shared between the models. E.g. $\frac{\partial \eta_{VE}}{\partial ICL}$ is common in (50) and (55) and is computed only once, further reducing the computational load. For $P_{bst}$ a hybrid approach is utilized. The numerical differentiation as shown in (42) is used for $\frac{\partial f_{P_{bst}}}{\partial P_{bst}}$ and $\frac{\partial f_{P_{bst}}}{\partial \omega_T}$ because of the saturation functions and singularities in the $\psi$ functions of the compressor model.

However, the linearization w.r.t. the inputs is still done analytically as follows.

$$\frac{\partial f_{P_{bst}}}{\partial \dot{m}_{TV}} = \frac{T_{bst}R}{-V_{bst}6\times10^6}$$  

$$\frac{\partial f_{P_{bst}}}{\partial \beta_{BOV}} = (P_{bst} - P_{cin}) \left( \frac{T_{bst}R}{-V_{bst}6\times10^6} \right)$$

The analytical differentiation of the $\omega_T$ dynamics results in long, complicated solutions which are marginally faster than the numerical solutions obtained from. Hence, the numerical approach was utilized in obtaining these partial derivatives but after some simplification as follows.

$$\frac{\partial f_{\omega_T}}{\partial P_{im}} = \frac{1}{1000/\tau_{T_{urb}}} \left( \frac{\partial \tau_{T_{urb}}}{\partial P_{im}} \right)$$

$$\frac{\partial f_{\omega_T}}{\partial P_{bst}} = -\frac{1}{1000/\tau} \left( \frac{\partial \tau_{Comp}}{\partial P_{bst}} \right)$$
Where, the numerical computation associated with \( \frac{\partial \tau_{\text{Comp}}}{\partial P_{\text{Im}}} \) and \( \frac{\partial \tau_{\text{Turb}}}{\partial P_{\text{BST}}} \) is eliminated due to the absence of the corresponding states in these equations, similar to the fixed zeros.

### 5.2.2.3 Analytical differentiation of artificial neural network sub-models

Majority of the functions used to generate the tracking and constraint output are the ANN models described in the previous chapter. Hence, the linearization of the outputs requires linearization of the ANNs w.r.t. their inputs. A typical ANN model has a graphical structure shown in Figure V-3.
This model can be mathematically represented as follows.

\[ y = s(W^n s(W^{n-1} s(... s(W^1 x + b^1) + b^{n-1}) + b^n) \]

Where, \(W^i\) and \(b^i\) are the weight and bias vectors for the \(i^{th}\) layer of the ANN model respectively. \(x\) is the input vector and \(n\) is the total number of layers in the function. \(s\) is the sigmoid activation function of the following form. Hence as shown in Figure V-3 the \(\alpha^i\) terms are computed as follows.

\[ \alpha^1 = W^1 x + b_1 \quad V-62 \]

\[ \alpha^2 = W^2 s(\alpha^1) + b_2 \quad V-63 \]

\[ \vdots \]

\[ \alpha^n = W^{n-1} s(\alpha^{n-1}) + b^{n-1} \quad V-64 \]

The function \(s\) is the sigmoid activation function of the following form.

\[ s(\alpha) = \frac{1}{1+e^{-\alpha}} \quad 65 \]

Similar to the ‘Back Propagation’ algorithm used to train the ANN models, the differential of this function w.r.t. its argument is computed as follows.

\[ s'(\alpha) = s(\alpha)(1 - s(\alpha)) \quad 66 \]
It is noteworthy that $s(\alpha)$ is in fact computed during the forward execution of the ANN model and is already available. Hence the partial derivative of the model with respect to its inputs is easily computed using the chain rule as follows.

$$\frac{\partial y}{\partial x} = s'(W^ns(W^{n-1}s(...s(W^1x+b^1)+b^{n-1})+b^n)(W^n)^T s'(W^{n-1}s(...s(W^1x+b^1)+b^{n-1}))((W^{n-1})^T ... s'(W^1x+b^1)(W^1)^T$$

Where, (66) is utilized to compute $s'$ using results of the computation of $s$ from the forward execution of the model. Figure V-4 shows the difference in execution time between the analytical differentiation method shown above and the numerical differentiation for different combinations of inputs. The execution of both methods was performed on a desktop computer. The analytical method is over ten times faster than the numerical approach.
Figure V-4. Comparison between execution times of $IMEP_n$ ANN model linearization using numerical and analytical approach.

Figure V-4 shows the difference in execution time between the numerical linearization and the hybrid model linearization for the entire system model using the methods described in this section. Similar to the difference in execution time for the ANN models shown in Figure V-5. The execution time for the system model linearization using the hybrid approach is also over ten times faster than the numerical approach.
5.3 Model discretization in time

Implementation of the system model in MPC requires discretization of the system model in time. The non-linear model is open-loop stable in continuous time domain, hence it is necessary to ensure that the discrete-time, linear model used for MPC is open-loop stable as well to preserve accuracy. The manifold filling and emptying dynamics have time constants which vary with engine load and speed. At high engine flow rates, these time constants becomes significantly shorter than $\delta t$. Such fast dynamics could be
eliminated from the model at these conditions, but that would require a variable-order model which is not desirable for MPC formulation. Euler’s finite difference approximation is a commonly used method for conversion of a model from continuous time to discrete time domain. However using this method for the linear system leads to numerical instability at some high load/speed conditions. Utilizing the State Transition Matrix (STM) method for time domain discretization generally preserves numerical stability to a higher degree compared to the finite difference approach [83]. The discrete time model is obtained using the following equation.

\[
\delta x_{k+1} = e^{A\delta t} \delta x_k + \left( \int_0^{\delta t} e^{A(\delta t-\tau)} d\tau B \right) \delta u_k
\]

Where, \(e^{A\delta t}\) and \(e^{A(\delta t-\tau)}\) are matrix exponentials computed using the Pade approximation. Hence, the discrete time model of the following form is derived from the STM method.

\[
\delta x_{k+1} = A_d \delta x_k + B_d \delta u_k
\]

\[
\begin{bmatrix} y_k \\ z_k \end{bmatrix} = \begin{bmatrix} y_0 \\ z_0 \end{bmatrix} + C_d \delta x_k + D_d \delta u_k
\]

Due to the nature of discrete-time process execution on an electronic micro-controller, the commands generated by the MPC algorithm are available only at the beginning of the next time step. This results in a unit time-step delay in the control actions. However, the prediction model () has a non-zero D matrix which does not
account for this delay and also isn’t favorable for MPC-QP formulation. In order to address this issue the discrete time model is modified as follows.

\[
\begin{bmatrix}
\delta x_{k+1} \\
\delta \bar{x}_{k+1}
\end{bmatrix} =
\begin{bmatrix}
A_d & B_d \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
\delta x_k \\
\delta \bar{x}_k
\end{bmatrix} +
\begin{bmatrix}
0 \\
I
\end{bmatrix} \delta u_k
\]

\[
\begin{bmatrix}
y_k \\
z_k
\end{bmatrix} =
\begin{bmatrix}
y_0 \\
z_0
\end{bmatrix} +
\begin{bmatrix}
C_d & D_d
\end{bmatrix}
\begin{bmatrix}
\delta x_k \\
\delta \bar{x}_k
\end{bmatrix}
\]

Augmentation of the original state vector with \( \bar{x}_k \) automatically induces a unit step delay in the model and eliminates the \( D \) matrix. The prediction model (71) to (72) is time-varying. Hence, it is derived at every stage within the prediction horizon using the method shown above and the set of these derived models is utilized to formulate the MPC-QP.

### 5.4 MPC Cost function formulation

The formulation of the MPC-QP largely depends on the chosen objective function that is to be minimized. A tracking objective function is chosen to formulate the MPC-QP as follows.

\[
J(x(k), U(k)) = \bar{y}(k + N)^T Q_N \bar{y}(k + N) + \sum_{i=k}^{k+N-1} [\bar{y}(i)^T Q \bar{y}(i) + \bar{u}(i)^T R \bar{u}(i)]
\]

\[
\bar{y}(k) = y(k) - y_{Ref}(k)
\]

\[
\bar{u}(k) = u(k) - u_{Ref}(k)
\]

\[
U(k) = [u(k), u(k + 1), ... u(k + N - 1)]^T
\]
Where,

\[ Q \in \mathbb{R}_{>0}, Q_N \in \mathbb{R}_{>0}, R \in \mathbb{R}^{7 \times 7} \]

\( Q \) and \( Q_N \) are the tuning matrices for the stage and terminal penalties on IMEP\( n \) tracking error. \( Q_N \) is chosen large enough for closed loop stability \([55][56]\). \( R \) is the tuning matrix to penalize tracking error of the inputs and is always positive definite.

5.5 Reference input generation

The reference input \( u_{Ref} \) is the input vector which minimizes fuel consumption of the engine at steady state conditions. It is also utilized to generate the reference trajectory of the prediction model around which the MPC finds the optimal solution. \( \dot{m}_{ITV_{\text{ref}}}, \dot{m}_{WG_{\text{ref}}} \) and \( \dot{m}_{EGR_{\text{ref}}} \) which are part of \( u_{Ref} \) are particularly important to derive the reference trajectory of the model in the prediction horizon. The steady state values of \( \dot{m}_{ITV_{\text{ref}}} \) under boosted conditions can only be achieved by increasing the turbine power by sufficiently reducing \( \dot{m}_{WG} \). Hence using the steady state \( \dot{m}_{ITV_{\text{ref}}} \) can lead to a reference trajectory which is physically impossible and undesirable. E.g. \( P_{Bst} \) may drop well below \( P_{Im} \) in the reference trajectory, even low enough to drop the compressor pressure ratio below 1, resulting in choked operation. At this condition, the numerically derived \( \frac{\partial(m_{\text{Comp}})}{\partial P_{Bst}} \) becomes 0, decoupling \( P_{Bst} \) from the compressor mass flow rate model. This results in only \( \omega_T \) influence on the compressor mass flow rate. Since \( \omega_T \) has the slowest dynamics, the compressor mass flow rate practically is ‘locked’ at the choked position corresponding to \( \omega_T \) at that instant in the reference trajectory. For large tip-in
maneuvers, this problem results in an infeasible MPC-QP problem due to the imposed \( \xi_{ITV} \) constraint. Similarly, utilizing the steady state \( \dot{m}_{WG_{ref}} \) and \( \dot{m}_{EGR_{ref}} \) to obtain the reference trajectory results in the \( \xi_{WG} \) and \( COV_{IMEP_n} \) constraints being violated in the reference trajectory. To address these problems in deriving \( u_{Ref} \), the following steps have been taken. \( \dot{m}_{ITV_{ref}} \) is derived using the orifice flow model and the corresponding steady state throttle position. \( \dot{m}_{WG_{ref}} \) and \( \dot{m}_{EGR_{ref}} \) are obtained as follows

\[
\dot{m}_{WG_{ref}} = W_{G \text{ratio}}(\dot{m}_{ITV_{ref}})
\]

\[
\dot{m}_{EGR_{ref}} = EGR_{\text{ratio}}(\dot{m}_{ITV_{ref}})
\]

5.6 MPC Quadratic program problem formulation and optimization

The non-linear constrained minimization problem is defined as follows.

\[
\arg \min_{U(k)} J(x(k), U(k))
\]

Subject to:

\[
z(i) \leq [0.15 \ 1 \ 0 \ 0 \ 0]^T
\]

Based on the cost function Error! Reference source not found. and the prediction model, an MPC - Quadratic Program is formulated as follows.

\[
\min_U \frac{1}{2} U^T H U + L^T U
\]

\[
MU \leq F
\]

Where,

\[
H \in \mathbb{R}^{70 \times 70}_{>0}, L \in \mathbb{R}^{70}, M \in \mathbb{R}^{70 \times 70}_{>0}, F \in \mathbb{R}^{70}, U \in \mathbb{R}^{70}
\]
Solving the MPC-QP described above results in the optimal solution $U^*$ which is a sequence of sub-optimal inputs in the vicinity of the reference inputs $U_{ref}$ out of which the input vector which corresponds to the current time step $u^*(k)$ is chosen and applied to the engine system.

5.7 Simulation with virtual engine in the loop

A 1D-dynamical engine model designed in GT-Suite was coupled with the control system in order to perform co-simulation as shown in Figure V-6. Co-simulation setup between GT-Power and Simulink. The simulation was performed at 3000 RPM.

Figure V-6. Co-simulation setup between GT-Power and Simulink
Figure V-7 IMEPn response of MPC co-simulated with GT-Power

5.7.1 IMEPn tracking response

The closed-loop IMEPn response of the engine is shown in Figure V-7. It is immediately visible that the MPC tracks the commanded IMEPn with some steady state offset. This occurs due to the $u_{ref}$ calibration being derived from the experimental engine data. This reference input corresponds to a steady state IMEPn which is higher than the desired value. Due to the tuning penalty $R$ on the reference inputs, the MPC generates
inputs which are a ‘trade-off’ between the steady-state reference inputs $u_{ref}$ and the inputs required to improve the tracking of the desired IMEPn.

Figure V-8 Control actuator trajectories for MPC cosimulation with GT-Power, air-path actuators are on the left hand side and cylinder actuators are on the right hand side. Black dashed lines are the steady-state reference inputs and blue lines are the inputs generated by MPC.
5.7.2 Air-path actuator trajectories

The control actuator trajectories for the IMEPn transient are shown in Figure V-8. The MPC shows differences in air-path actuation for the two tip-in maneuvers shown. For the first tip-in at 2 seconds, the MPC commands minimal change in $\dot{m}_{Egr}$ whereas, for the second tip-in at 12 seconds the MPC commands a complete stoppage of $\dot{m}_{Egr}$ for a short period of time. $\dot{m}_{WG}$ is also stopped temporarily at the time of tip-in followed by a return to it’s reference value. $\beta_{BOV}$ actuation on the second tip-in mitigates the drop in $P_{Bst}$ caused by the rapid increase of $\dot{m}_{ITV}$. For the tip-out at 7 seconds MPC mainly commands $\beta_{BOV}$ to address the surge constraint.

5.7.3 Engine-cylinder actuators

The $SPK$ (Spark timing) actuator is significantly retarded from the reference value during both the tip-in maneuvers. MPC also advances the intake cam position $IVO$ during the tip-in maneuvers for volumetric efficiency and air-flow improvement. During the tip-out the MPC advances $SPK$ to handle the $COV_{1IMEPn}$ constraint.

5.7.4 Constraint handling

The last three constraint outputs in the constraint vector are purely air-path constraints. Out of these constraints the constraints $P_{Im} - P_{Bst} < 0$ and $\dot{m}_{WG} - \dot{m}_{Eng} < 0$ are always satisfied on the GT-Power model independent of the MPC commands. These constraints are in place merely to ensure physically feasible behavior in the control oriented model (30) to (37). The surge constraint handling is demonstrated
in Figure V-9 where, the compressor operation momentarily crosses the surge-line but returns back to the normal operating region of the compressor.

Figure V-9 Compressor operation during co-simulation of GT-Power with MPC

Since the knock model is not tuned and the $COV_{IMEPn}$ cannot be evaluated using the GT-Power model, they will be evaluated in the experimental testing section.
VI. LOWER-LEVEL CONTROL AND ESTIMATION

The commands generated by MPC require conversion into actuator set-points for real-time application. This lower-level actuation is categorized into two sections, air-path control and engine-cylinders control. Real time estimation of the values of the states is also required in order to formulate the MPC-QP problem.

6.1 Valve mass flow rate control

The air-path control consists of resolving the valve mass flow rate commands $\dot{m}_{ITV}$, $\dot{m}_{WG}$, and $\dot{m}_{EGR}$ into valve position set-points using the orifice flow equation as follows.

$$Ae = \frac{\dot{m} \sqrt{RT}}{\psi \left( \frac{P_{in}}{P_{out}} \right)}$$  \hfill (76)

$$\theta = lut(Ae)$$  \hfill (77)

Where, $lut$ is the lookup table function to convert commanded valve area $Ae$ to a valve position $\theta$. $\psi$ is the flow function which has an infinite gradient when the pressure difference $\Delta P$ across the valve approaches as shown in Figure VI-1. This results in ‘chattering’ in the valve position due to small changes in $\Delta P$. 


Figure VI-1. The gradient of the flow function becomes very steep as the pressure difference across the valve diminishes.

In order to address this issue a simple filter with the following structure is applied.

\[
\theta_{\text{Final}}(k) = w(k)\theta_{\text{Init}}(k) + (1 - w(k)) \\
w(k) = |g\Delta P(k)|
\]

Where, \( g \) is a tuning parameter which can be tuned heuristically to mitigate chattering at \( \Delta P(k) \approx 0 \) while also having sufficiently fast transient response. The engine-cylinders control is simply the realization of the commands SPK, IVO, and EVC.
6.2 Low-pressure EGR $\Delta P$ control

Recent findings show the difficulty in control of mass flow rate across the LPE system [37][38]. The standard orifice model accuracy deteriorates rapidly in a highly pulsating $\Delta P_{LPE} \approx 0$. Increasing $\Delta P_{LPE}$ to atleast 5 kPa was observed to maintain the orifice flow model’s accuracy within 5%. Hence, the PCT valve is actuated using

$$ Ae = \frac{\dot{m}\sqrt{RT}}{\psi(F_{In})P_{In}} $$

(76). The mass air-flow sensor unit which is shortly upstream of the PCT valve is used for the $\dot{m}$ and $P_{In}$ input. $P_{Out}$ is simply set to $P_{In} - 5 \text{ kPa}$. Since, the turbine outlet pressure is always higher than $P_{In}$, $\Delta P_{LPE} \geq 5 \text{ kPa}$ is guaranteed. However, this method leads to much higher $\Delta P_{LPE}$ at higher power conditions where the ‘natural’ $\Delta P_{LPE}$ without PCT throttling is close to 5 kPa. This method results in unnecessary compressor inlet throttling, potentially limiting peak torque of the engine. Although this is a caveat in this method, it is simple to implement and ensures acceptable accuracy in controlling the mass flow rate through the LPE system.

6.3 Turbo speed estimation and state measurement

Since the prediction model is physics based, the states have physical meaning as well. $P_{Im}$ and $P_{Bst}$ measurements are part of the production engine sensor set. An intake oxygen sensor is installed at the inlet of the intake manifold for real time measurement of $P_{Egr}$. These measurements are also low pass filtered to minimize noise and oscillations induced from the discrete-event gas exchange processes. The cutoff frequency high
enough to maintain transient accuracy within 50ms. \( \omega_T \) is estimated using the Fixed Point Iteration method described in [89]. The compressor model is utilized for estimation.

\[
\hat{\omega}_T(k) = \hat{\omega}_T(k-1) + K \left( \frac{\alpha \left( \Pi_{\text{comp}} - \hat{\Pi}_{\text{comp}} \right)}{\partial \Pi_{\text{comp}}/\partial \hat{\omega}_T} + \frac{(1-\alpha)(\dot{m}_{\text{comp}} - \hat{\dot{m}}_{\text{comp}})}{\partial \dot{m}_{\text{comp}}/\partial \hat{\omega}_T} \right)
\]

80

\[
\Pi_{\text{Comp}} = \frac{P_{\text{Bst}}}{P_{\text{Cin}}}
\]

81

\[
\alpha = \frac{1}{\frac{\partial \Pi_{\text{Comp}}}{\partial \dot{m}_{\text{Comp}}}}
\]

82

Where, \( \hat{\Pi}_{\text{Comp}} \) and \( \hat{\dot{m}}_{\text{Comp}} \) is the predicted pressure ratio and compressor mass flow rate using the estimated turbo speed \( \hat{\omega}_T(k-1) \) from the previous iteration and the compressor model. \( \alpha \) dictates how much emphasis is placed on utilizing \( \Pi_{\text{Comp}} \) over \( \dot{m}_{\text{Comp}} \) for estimation of \( \hat{\omega}_T(k) \). Figure VI-2 shows the estimated vs measured turbocharger speed for a fully transient engine test. Some of the scatter in the Figure VI-2 is caused by the noise in measurement of turbocharger speed. The accuracy of the turbo speed estimation using this method was observed to be within the 15% error band even for low speed extrapolated region of the compressor map. The RMS error and max error for this test is 0.28 kRad/s and 1.5 kRad/s respectively.
Figure VI-2. Real-time estimated vs measured turbocharger speed using fixed point iteration method.
VII. EXPERIMENTAL VALIDATION OF MPC

7.1 Experimental control system hardware configuration

The architecture of the Rapid-control prototyping system interfaced with the engine and it’s ECU is shown in Figure VII-1. The MPC derived in the previous section requires higher computational ability compared to standard engine controllers. In order to maintain real-time execution ability using this formulation, a Dspace DS1006 has been utilized to execute the MPC program similar to [90][91]. The DS1006 executes in two timer-task mode. The MPC program operates at the sample time $\delta t$. At the end of each step of it’s execution $u^*$ is sent to a communication block which executes at a sample time of 2ms. This block consists of the CAN transmit subsystems which send $u^*$ to the ES910 over a 1Mbit/s J1939 CAN bus. The same channel is used to receive states $x$ and parameters $k$ from the ES910. These are utilized to execute the next step of the MPC program. During testing, the message transmission and receive time was observed to be less than 2ms which is negligible compared to the unit-step delay induced from discrete time execution.

The lower-level control program is deployed onto an ETAS ES910 real-time simulation controller. Because of the additional actuators and sensors associated with LP-EGR and WG control an ES930 I/O extension unit has also been added to the ES910 controller. The $P_{Egr}$ measurement is also performed using an Oxygen sensor controller box which outputs analog voltage as a function of Oxygen concentration. The lower level control layer also has the position controllers for the LP-EGR, PCT and BOV
valves. These are conventional PID based controllers. Because of the faster dynamics of the lower-level actuators, the sample time of the low-level controller and estimator is set to 1ms.

Figure VII-1. Rapid control prototyping setup for experimental validation of MPC
7.2.1 Experimental test results 2900 RPM: Aggressive tip-in and tip-out maneuvers

Analysis of the MPC controller performance is done by comparing against the open-loop fuel-economical control input $u_{Ref}$. The analysis is done for step commands of $IMEP_n$ at 3000 RPM to show the tip-in and tip-out response of both the control approaches. Figure VII-2 shows the tip-in response of the controllers. It is evident that the tip-in response of the MPC is much quicker than the open-loop control. The MPC response also has a slight overshoot followed by stable steady state tracking. The air-path actuator trajectories have been shown in Figure VII-3, Figure VII-4 and Figure VII-5.

![Graph showing tip-in IMEPn response for MPC and Open-loop control strategy](image)

Figure VII-2. Tip-in IMEPn response for MPC and Open-loop control strategy

At the instant of the tip-in the commanded reference $\dot{m}_{ITV}$ is momentarily increased and dropped based on the reference steady state intake throttle position which corresponded to a pressure difference of approximately 15kPa across the intake throttle. Because of this
the MPC commanded $\dot{m}_{ITV}$ impulse is slightly higher as the forward throttle flow constraint is inactive in the reference trajectory. MPC completely stops $\dot{m}_{WG}$ and $\dot{m}_{EGR}$ to accelerate turbocharger spool-up. $\dot{m}_{EGR}$ converges to the reference value in the latter part of the transient. The cylinder inputs for both the control strategies is shown in Figure VII-6 and Figure VII-7. MPC retards the spark timing and advance the intake cam timing heavily compared to the reference inputs during the tip-in maneuver.

![Intake Throttle Mass Flow Rate Trajectories](image1)

Figure VII-3. Intake throttle mass flow rate trajectories for MPC and Open-loop control

![Wastegate Mass Flow Rate Trajectories](image2)

Figure VII-4. Wastegate mass flow rate trajectories for MPC and Open-loop control
Figure VII-5. External cooled EGR mass flow rate trajectories for MPC and Open-loop control shows that MPC deliberately stops EGR flow on the aggressive tip-in.

Similar to the air-path actuators, the deviation of the MPC inputs from the reference input is mainly to improve the turbocharger spool-up. Figure VII-8, Figure VII-9 and Figure VII-10 shows the modeled volumetric efficiency, measured exhaust temperature and measured turbocharger speed for both of the control strategies. The advanced intake cam timing is clearly to improve the transient volumetric efficiency of the cylinders whereas the retarded spark timing is to increase the exhaust temperature.
Figure VII-6. Spark timing trajectories for MPC and Open-loop control

Figure VII-7. Camshaft position timing trajectories for MPC and Open-loop control
Figure VII-8 ANN Modeled volumetric efficiency using MPC inputs and open-loop inputs for tip-in maneuver

Figure VII-9 Measured turbine inlet temperature for MPC and Open-loop control shows that MPC commands late spark timing to increase specific exhaust gas enthalpy for turbocharger response
Turbocharger speed for MPC increases much more rapidly than open-loop control. These multiple coordinated inputs result in the turbocharger speed increasing significantly faster than the open-loop control. It is noteworthy, that once the IMEPn tracking has reached steady state, the deviation between the MPC and the open-loop inputs is minimal.

The most relevant constraint for the tip-in maneuver is the $K I^2$ constraint. The MPC advances the spark timing well beyond the reference value between 3 and 4 seconds. However, MPC also commands $\dot{m}_{EGR}$ beyond the reference value as well during this time period. The measured $KI^2$ for cylinder 1 is shown in Figure VII-11. Due to the spark timing advancement, the $KI^2$ value increases well beyond the open-loop control trajectory. However, due to the addition of extra $\dot{m}_{EGR}$ into the engine, the constraint is not violated.
Figure VII-11 Measured $K_I^2$ for MPC does not exceed the limit of 1 despite advanced spark timing during tip-in.

7.2.2 Experimental test results 2900 RPM: Aggressive tip-out maneuver

Because the open-loop control strategy does not have active compressor surge handling ability, only the MPC trajectories are shown. Figure VII-12 shows that there is a slight instability in IMEPn after the tip-out.
Figure VII-12. Tip-out IMEPn response for MPC and Open-loop control strategy

The air path trajectories for the tip-out maneuver are shown in Figure VII-13. The MPC commands the intake throttle mass flow rate to zero to reduce the intake manifold pressure. MPC commands stoppage of EGR mass flow rate for combustion stability concerns. Simultaneously the MPC commands an increase in waste-gate mass flow rate and blow off valve mass flow rate to address the tip-out surge problem. Figure VII-15 shows that even though the MPC commands the blow off valve to open instantaneously, the compressor operation actually crosses over into the surge region. This is mainly because, the boosted engine operation before the tip-out is very close to the surge line. At the instant of the tip-out the blow off valve command issued by the MPC cannot be instantly fulfilled by the lower level controller due to bandwidth limitations. The true compressor mass flow rate dynamics during the surge is also governed by very fast limit-cycle like behavior which can be modeled using the classical Moore-Greitzer model.
mentioned in [53]. However, addition of this model introduces dynamics which are significantly faster than the MPC sample time to and should effectively be used in the lower level control layer.

Apart from compressor surge avoidance, tip-out dilution from trapped external EGR in the air-path is a concern as well. Since this EGR cannot be evacuated by any other means other than by passage through the engine itself, the MPC can only control the engine actuators to handle this problem. It is visible from Figure VII-14 that MPC commands over-advanced spark timing at the instant of the tip-out to maintain combustion stability. Although, internal residual gas fraction is not included in the control model, the COV model indirectly captures the effect of excessive valve overlap on COV which is the most likely reason for valve timing manipulation at the instant of tip-out.
Figure VII-13 Air-path trajectories for aggressive tip-out maneuver

Figure VII-14 Cylinder actuator trajectories for aggressive tip-out maneuver
Figure VII-15 Compressor operation during the aggressive tip-out
VIII. CONCLUSIONS AND RESEARCH CONTRIBUTIONS

Summary of research work

An evaluation of state-of-the-art control methods for turbocharged engine control was performed in order to identify the lapses in research and potential for improvements was identified. The synergies between various actuators on turbocharged SI engines is an open research topic. The recent research work published for MPC based subsystem control of turbocharged engines was examined in detail and it was quickly identified that majority of research articles utilize linear system models which are either physics based or of black-box nature and derived offline. Although this approach works well for air-path control, is questionable when engine cylinder actuators are also included in the MPC framework. Hence, in order to assess the error associated with the linearity approximation a control oriented physics based model of the turbocharged engine air-path was developed with ANN based sub-models for outputs and constraints. A study was performed to evaluate the deviation of the non-linear model from the linearized model over the prediction horizon which revealed that the linear approximation error is not trivial for the highly non-linear and the most important outputs of the engine, IMEPn. Extending this study further, a simple analysis was performed to evaluate the number of re-linearizations required within the prediction horizon for a given required accuracy level and operating point. The results of this analysis show that the model has varying levels of linear approximation error depending on the operating condition at which the analysis is done and the frequency of the reference inputs applied to the model.
Under the knowledge that the engine model is non-linear and MPC strategies require linearization of the model, avenues to speed up the linearization process have been explored. Analytical differentiation has been widely utilized to reduce the number of computations associated with model linearization. The resulting approach adopted has been demonstrated to be at least 10 times faster in obtaining a linearized system model. This is a significant advantage with cascaded benefits especially when the linearization is done more frequently within the prediction horizon. Based on this approach, an NMPC where linearization is done at every stage within the prediction horizon is formulated. A preliminary co-simulation with 1D-Engine simulation has been performed and the results have demonstrated the ability of the control system to coordinate air-path and engine cylinder actuators simultaneously while tracking the desired IMEPn within 1 bar.

In order to facilitate experimental validation of the NMPC a lower level control and turbocharger speed estimation layer was developed. Simple inversion of the orifice flow model was performed with an additional adaptive filter to mitigate chattering of the intake throttle control at very low pressure differential values. A fixed-point iteration based turbocharger speed estimation algorithm was utilized to estimate the turbocharger rotational speed using the production sensor set. A two micro-controller based rapid-control-prototyping system was configured to facilitate real-time execution of the NMPC algorithm on a computationally powerful micro-controller and the actuator control and turbo speed estimation on a simpler machine. A sufficiently fast communication system was established between the CAN controllers. Using this layout, real-time experimental validation of the NMPC algorithm on a 2.0 L Turbocharged gasoline engine with cooled
external EGR was performed. In the absence of a production type engine control algorithm for this engine configuration, the NMPC performance was evaluated against the reference open loop control which is optimized for fuel economy. The NMPC performs control actuator maneuvers similar to those observed using the GT-Power co-simulation. The NMPC demonstrates the ability to manipulate Spark and cam timing to improve engine breathing and exhaust gas enthalpy to maximize turbocharger response while simultaneously controlling the air-path actuators. However, the surge control has pitfalls due to the proximity of the compressor operation to the surge line and the bandwidth limitations of the lower level control which were unaccounted for in the GT-Power co-simulation.

**Broader impacts and practical aspects**

The algorithm outlined in this work is derived mainly from a physics based air-path model coupled with ANN based engine model derived from steady state data. These models were derived using 40 hours of steady state engine dynamometer test data. This time period is negligible compared to the time required to acquire transient data to tune black-box models of the engine. Furthermore, the physics based nature of the air-path implies that changes in air-path geometry and turbocharger hardware can be quickly implemented in the model without the need to re-identify parameters associated with air path dynamics. These attributes of this approach imply that the algorithm can be ported to a different engine system with minimal engine testing.
Due to demonstrated ability of the algorithm to co-ordinate multiple actuators simultaneously, calibration efforts required are minimal with only 15 constants which require tuning, which is negligible compared to the large number of lookup tables and maps associated with transient control of turbocharged engines on production ECUs.

**Future work and potential for improvement**

The NMPC proposed in this research is based on the solution to an MPC-QP formulated based on the current state of the system, user defined objective function and the reference inputs applied to the system. Although an MPC-QP is convex due to linearized models used in the formulation, the original Non-linear optimization problem can be non-convex. As a result, applying the sub-optimal control inputs derived from the MPC-QP optimization may violate one or more constraints. This approach was used in this research because 2 out of the 5 constraints on the mainly to impose physical feasibility in the control model and can never be violated in reality. The remaining 3 operational constraints could possibly be violated with this approach. In order to find sub-optimal solutions which satisfy the non-linear constraints, non-linear optimization techniques like Sequential Quadratic Programming and Interior-point optimization methods could be explored. However, these methods would require additional computational resources.

The tuning of the NMPC is also an open question, wherein the duality between Q and R implies that IMEPn tracking affects the cost associated with tracking of the reference inputs and vice versa. Drive cycle or extended transient simulation could be
performed with different tuning values to study impact on fuel-economy. Improved transient turbocharger response may not necessarily imply reduced fuel economy, as consideration of interaction with transmission and shifting strategies is also a significant factor in fuel economy.
A. APPENDIX

A Control Algorithm for Low Pressure – EGR Systems using a Smith Predictor with Intake Oxygen Sensor Feedback

Due to ever increasing stringency in emissions and fuel economy regulations development of cost-effective and production viable concepts for engine efficiency improvement is necessary. Utilization of cooled EGR in downsized turbocharged gasoline engines has been shown to improve efficiency at many engine operating conditions [1][2][3]. Previous research work in the domain of LP-cEGR control focuses mainly on utilization of mean value models of air and EGR paths for determination of EGR and air fraction [12]. An open-loop EGR valve mass flow model is utilized in these methods [4][5]. The primary challenges associated with control of LP-cEGR systems are; (1) mass flow modeling across EGR valve due to low-pressure differential across the system [4][5], and (2) EGR control is difficult during transients due to long length of EGR path [4][5][12]. Additionally, the methods described in [4] and [12] also utilize an intake throttle valve upstream of the EGR mixing location, which increases pressure differential across the EGR system. Additional throttle valves for EGR control are not used for this research.

Universal Exhaust Gas Oxygen (UEGO) sensors are widely used in engine exhaust systems for control and diagnostics. These sensors measure the concentration of oxygen in the sampled gas. Previous research has demonstrated the feasibility of using UEGO
sensors to accurately measure the EGR fraction in the intake system [7][8]. The potential of these sensors to provide real-time EGR fraction feedback control is explored in this research. Water condensation from exhaust [1] can severely damage the high temperature sensing element in these sensors. Also, in order to ensure proper EGR and Air mixing at the sensor location, the sensor location has to be adequately downstream of the mixing location. However, moving the sensor farther downstream of the EGR valve increases the EGR valve to sensor time delay. This severely restricts the calibrated ‘aggressiveness’ of the feedback control system. In systems with dead time delays, Smith Predictor based feedback control is a widely used method for process control [13][6]. Smith Predictor based control is a predictive method of process control where the measured output of the process is delayed by a known time duration. Fluid flow through a pipe and engine cylinders are common examples of processes with dead time delay.

The paper begins with description of experimental setup including the engine and control system hardware. As required by the Smith Predictor method, models are defined for the EGR and air path. These consist of open-loop modeling of EGR valve mass flow and the transport delay across the EGR and air path. These models are validated in isolation to evaluate their accuracy. Finally, the complete control structure is tested at steady state engine operating conditions to evaluate performance and stability.

A downsized turbocharged inline 4 gasoline spark-ignition (SI) engine with specifications described in Table 1 is used for testing. An AC engine dynamometer controls engine speed and/or loading for all experiments. A set of production-intent sensors is installed on the engine for basic engine functionality (air, fueling, ignition,
valve timing and turbo control). A low-pressure and cooled EGR (LP-cEGR) configuration is retrofitted on the engine. Exhaust gases are extracted downstream of the turbine and upstream of the catalyst. EGR passes through a cooler and is delivered to the intake air-path system upstream of the compressor. The EGR cooler is a tube-core type chosen for low pressure loss characteristics. Twin liquid-to-air intercoolers have been used to allow high boost/load capability in the dynamometer cell, where air/air heat exchanger effectiveness is low.

<table>
<thead>
<tr>
<th>Engine type</th>
<th>Turbocharged Inline 4 cylinder SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Displaced volume</td>
<td>1998 cc</td>
</tr>
<tr>
<td>Bore x Stroke</td>
<td>86 x 86 mm</td>
</tr>
<tr>
<td>Compression ratio</td>
<td>9.5:1</td>
</tr>
<tr>
<td>Valve train</td>
<td>DOHC Continuously variable valve timing with 4 valves per</td>
</tr>
<tr>
<td>Fuel system</td>
<td>Direct Injection</td>
</tr>
<tr>
<td>EGR system</td>
<td>Low-pressure cooled EGR</td>
</tr>
<tr>
<td>Base engine controller</td>
<td>BOSCH SI Engine controller</td>
</tr>
<tr>
<td>Rapid prototype</td>
<td>ETAS ES910</td>
</tr>
<tr>
<td>controller</td>
<td></td>
</tr>
</tbody>
</table>
The overall EGR control system architecture is outlined in Figure 2, and a description of the signals is provided in Table 2. A production-intent engine controller is modified to include software hooks on specific engine control parameters. An ETAS ES910 Rapid Prototype Controller (RPC) is used in conjunction with the base engine controller using CAN Communication Protocol (CCP) bypass communication. This is done to establish data interface between the base engine controller and the rapid prototype controller. A variant of an oxygen sensor is also installed post compressor (Figure 1.). This sensor is similar in operation to UEGO sensors, however it has been developed and optimized to be operated specifically in the intake environment, where temperature range and gas composition is different from the exhaust. EGR fraction is calculated based on change in intake oxygen concentration caused by EGR dilution [7][8]. Hereafter, the intake oxygen sensor is referenced as the ‘EGR sensor’ used for feedback control.

Figure A-1 Control system setup
EGR Valve Modeling for Control

Mass flow rate across the EGR valve is modeled using an isothermal orifice flow model in equation (76). This model is specifically used for exhaust gases at lower temperatures.

\[
\dot{m}_{EGR} = \frac{C_D A_{Value} P_{In}}{\sqrt{R T_{EGR} }} \left\{ \frac{2 P_{Out}}{P_{In}} \left[ 1 - \frac{P_{Out}}{P_{In}} \right] \right\}^{1/2}
\]

\[83\]

\(C_D\) is the discharge coefficient and \(A_{Value}\) is the valve open area. \(P_{In}\) is the valve inlet pressure. To determine pressure at the inlet of the EGR valve, the pressure at the inlet of the EGR cooler and the pressure drop across the EGR cooler needs to be known. Calculation of pressure drop across the EGR cooler requires prior knowledge of mass flow across the EGR system. For the sake of simplicity, this research considers EGR valve inlet pressure to be the same as EGR cooler inlet pressure. A physics based model is utilized to predict this pressure as a function of engine operating conditions [68]. \(P_{Out}\)
is the valve outlet pressure which is considered to be the same as $P_{Comp-In}$. $R$ is the specific gas constant for stoichiometric exhaust gas. LP-cEGR systems are required to operate at a much lower pressure differentials as compared to HP-EGR systems. The pressure ratio correction is very sensitive to minor inaccuracies in the pressure ratio at such low pressure differences. In addition, the pressure at the inlet of the EGR valve pulsates[4][5], which influences the flow through the EGR system. The model shown in (83) holds strictly true only for steady flow conditions. Pressure pulsations reduce the accuracy of the model as the flow is unsteady at the valve and averaged pressures are used as inputs to the model.

In this research the discharge coefficient used in the model has been characterized as a function of engine speed and EGR valve angle. This is done to partially capture the effect of the primary frequency of pressure pulsations on the discharge coefficient since the primary frequency of exhaust pressure pulsations is directly proportional to the engine speed [11]. The discharge coefficients were obtained by running an optimization routine which minimizes the error between the modeled mass flow and the measured mass flow. The measured mass flow is derived from experimental sweeps of EGR valve position between fully closed and fully open at different engine operating conditions. The engine load was swept from 3 to 11 Bar BMEP and the engine speed was swept from 1200 to 2700 RPM. The final result of this optimization routine is the two-dimensional discharge coefficient map shown in Figure A-2. Typical maximum values of discharge coefficient for butterfly type valves are close to 0.6 [11]. The values obtained for this research were significantly lower (0.15) mainly due to consideration of EGR cooler inlet pressure as
valve inlet pressure. A minor inaccuracy in the mapping between EGR valve angle and open area has also influenced the discharge coefficient.

Modeled and measured mass flow rates along with the +/-10% error lines are shown in Figure A-3. The majority of operating points lie within the +/-10% range. Engine speed has been used to capture the effect of exhaust pressure pulsations. The primary frequency of exhaust pressure pulsations is a linear function of engine speed and number of cylinders/exhaust blowdown events per cycle. This has been done to partially characterize the effect of exhaust pressure pulsation frequency on the EGR mass flow. Effects of peak amplitudes and pulsation shape may also have an impact, but are outside of the scope of this research.

![Dimensional Discharge coefficient map obtained from optimization routine](image-url)

Figure A-2 Dimensional Discharge coefficient map obtained from optimization routine
Figure A-3 Measured vs modeled EGR valve mass flow rate

**Transport delay model**

The dead-time delay between EGR valve and the EGR sensor consists of two components. The first is transport delay from the EGR mixing location to EGR sensor. The second component is sensor response time, consisting of transport from exterior of sensor to sensing element and electrical signal delay. The sensor response time is approximated as a linear function of flow velocity in the control volume in which the sensor is mounted. A simplified equation based on the assumption of plug flow is derived using the steady state continuity equation for mass flow rate through a control volume.
This method is utilized to estimate the transport delay from the entry to the exit of the control volume. The equation used to estimate the transport delay $\tau_i$ across a generic control volume ‘i’ in the EGR system is given by equation (84).

$$\tau_i = \frac{L_i P_i A_i}{\dot{m}_i R T_i}$$

In this equation $L_i$ and $A_i$ are the length and flow area respectively. $\dot{m}_i$, $P_i$ and $T_i$ are the mass flow rate, pressure and temperature in the control volume respectively. Although the equation is derived for strictly steady state conditions, it is still used for the quasi steady state operating conditions tested in this research where only EGR valve transients were performed.

**EGR fraction models**

Implementation of the Smith Predictor methodology requires definition of the EGR fraction model with no transport delay and the EGR fraction model with modeled transport delay. These models are described as per the schematic in Figure A-4.

![Figure A-4 Modeled low pressure EGR path](image)
The EGR fraction model with no transport delay is used to predict the EGR fraction at mixing point A. Under the assumption that the mixing is instantaneous and perfect, the EGR fraction at A, $EGR_f_A$ is given by equation (85).

$$EGR_f_A = \frac{m_{EGR}}{m_{Air} + m_{EGR}}$$  \hspace{1cm} (85)

The EGR fraction at location B and time, $t$, $EGR_f_B(t)$ which is also the EGR sensor location is the transport delayed value of the EGR fraction at A. The modeled EGR fraction at B is given by equation (86).

$$EGR_f_B(t) = EGR_f_A(t - \tau(t))$$  \hspace{1cm} (86)

Where $\tau(t)$ is the modeled transport delay at time $t$ from location A to location B, and is determined using equation (87).

$$\tau(t) = \tau_{AtoC}(t) + \tau_C(t) + \tau_{CtoB}(t)$$  \hspace{1cm} (87)

Where $\tau_{AtoC}(t)$ and $\tau_{CtoB}(t)$ are the modeled transport delays from A to the compressor and compressor to B respectively. $\tau_C(t)$ is the transport delay within the compressor which has been assumed to be zero. These delays are calculated separately because the pressure and temperature within the control volumes from A to the compressor and from the compressor to B are different.
**EGR fraction measurement using Oxygen sensing**

The EGR mole fraction, $EGRf_{B Sens}$, is the measurement performed by the EGR sensor. It is calculated using the measured mole fraction of oxygen $x_{O_2 Sens}$ with equation (88).

$$EGRf_{B Sens} = \frac{x_{O_2 Amb} - x_{O_2 Sens}}{x_{O_2 Amb} - x_{O_2 Exh}}$$  

$x_{O_2 Amb}$ and $x_{O_2 Exh}$ are the mole fractions of oxygen in fresh air and the exhaust gases respectively. In order to derive the EGR mass fraction, the following assumptions are made:

- Effect of ambient humidity on $x_{O_2 Amb}$ is ignored
- $x_{O_2 Exh}$ is negligible due to mostly stoichiometric operation
- The difference between EGR mass fraction and EGR mole fraction is considered negligible for control simplicity

**Feedback control with Smith Predictor**

A classical Smith Predictor was implemented as a dead time delay compensator. Previous research shows that Smith Predictor based classical feedback control for EGR systems in Large Diesel Engines improves control stability [61]. The application of the Smith predictor is visible in the EGR fraction controller shown in Figure A-5. Linear, continuous approximations of subsystems have been shown. Closed loop feedback control of EGR fraction is performed using a PID controller.
\[ P_{Plant_B}(s) \approx P_{Plant_A}(s) \cdot e^{-\tau_{Plant}(t)s} \]  
89

\[ P_{Pred_A}(s) = P_{Model_A}(s) + (P_{Plant_A} \cdot e^{-\tau_{Plant}(t)s} - P_{Model_A}(s) \cdot e^{-\tau(t)s}) \]  
90

The primary objective of this architecture is to eliminate the delay dynamics from the output. The EGR Path + Sensor subsystem system has the output \( EGRf_{B Sens} \) which is inherently delayed by the physical transport delay \( \tau_{Plant}(t) \) as shown in equation (89) \( P_{Plant_A}(s) \) is the EGR Path + Sensor subsystem without the transport delay. The output of augmented system approaches that of the Mixing point EGR fraction model as the term in the parenthesis in equation (90) tends to 0 i.e. as the output of \( P_{Model_A}(s) \) tends to \( P_{Plant_A}(s) \) and \( \tau(t) \) to \( \tau_{Plant}(t) \). This emphasizes the importance of accuracy of the EGR flow model and transport delay model. Effectively, the output \( EGRf_{A Pred} \) of the
augmented system $P_{Pred,A}(s)$ has minimal delay dynamics in it. This reconstructed output along with the desired EGR fraction command $EGRf_{BComm}$ is used to generate the error upon which the PID controller operates. Therefore, instability and oscillation induced by dead time delays are mitigated.

**Experimental test results**

The benefits of LP-cEGR are seen at different engine operating conditions due to different mechanisms [36]. This research mainly focuses on LP-C-EGR control at low to mid load. This is done partially due to the generally longer transport delay $\tau(t)$ associated with these load conditions. Longer transport delays can exist for high load and very low engine speeds but for this research these conditions were not considered.

**Transport delay model**

The transport delay model was tested experimentally by performing step changes in EGR valve position at quasi steady engine operating conditions. These conditions were established by operating the engine at fixed actuator positions/commands; such as the intake throttle valve, valve timing, waste-gate and blow-off valve position, etc. The number of engine cycles elapsed between the EGR valve step and the effect of the step at the EGR sensor was compared with the delay calculated by the transport delay model. This is done at various engine operating conditions to account for varying engine cycle
time. The majority of test points lie within a +/- 1 engine cycle band, as shown in Error! Reference source not found..

Figure A-6 Modeled vs measured transport delay

**EGR fraction models**

Recall that $EGR_{f_B}$ is the transport delayed version of $EGR_{f_A}$ using the transport delay model. In order to evaluate accuracy of the EGR fraction models EGR valve steps were performed, and modeled and measured EGR fractions at the EGR sensor location are compared Figure A-7 to Figure A-9. show this comparison at various engine operating conditions for different EGR valve position steps. At all of the operating conditions, the fast change in $EGR_{f_B}$ and $EGR_{f_{B\text{Sens}}}$ caused by the EGR valve position step are within +/- 1 engine cycle. The steady state difference between $EGR_{f_B}$ and $EGR_{f_{B\text{Sens}}}$ is highest
at 2100 RPM and 2.5 Bar BMEP. This is attributed to the inaccuracy of the EGR mass flow model at near unity pressure ratio.

Figure A-7. Comparison between modeled and measured EGR fraction at EGR sensor location; engine operated at 1700 RPM 3 Bar BMEP
Figure A-8. Comparison between modeled and measured EGR fraction at EGR sensor location; engine operated at 2100 RPM 5 Bar BMEP
Figure A-9. Comparison between modeled and measured EGR fraction at EGR sensor location; engine operated at 2100 RPM 2.5 Bar BMEP

**Feedback control with Smith Predictor**

In order to test the feedback control architecture outlined in Figure A-5, step inputs are commanded for desired EGR fraction. These tests are performed using conventional PID control in Figure 11, Figure 13 and Figure 15. Step response of EGR fraction controller with conventional PID controller at 2000RPM 9 Bar BMEP as well as Smith Predictor based control as seen in Figure 12, Figure 14 and Figure 16. Step response of EGR fraction controller with Smith Predictor at 2000RPM 9 Bar BMEP The controller gains Kp, Ki and Kd were determined heuristically as 130, 6 and 0.01 respectively by step
response analysis. They were kept constant for all of the tests. In the case of the conventional PID control there are two main observations:

- The EGR valve position overshoots the position required to reach the desired EGR fraction
- The EGR valve position appears to be 90 degrees out of phase with the sensed EGR fraction, thereby inducing oscillation in the EGR fraction

Oscillations can be mitigated with appropriate controller tuning. However, the controller gains would also have to be scheduled as per engine operating condition as the flow within the EGR and air path is mainly dependent on engine operating condition.

Figure A-10. Step response of EGR fraction controller with conventional PID controller at 1500RPM 7 Bar BMEP
In case of the Smith Predictor based controller the following observations are made:

- Oscillations related to phase lag between EGR valve and sensor are greatly reduced (compared to PID)
- Overshoots in EGR fraction are greatly reduced (compared to PID)
- Rise time and fall time (time duration from 10% to 90% of the setpoint value) are increased over conventional PID
There are small overshoots in the EGR valve position, but they are small enough in magnitude and time to not affect EGR fraction significantly, as evident in Figure A-13 at 3 and 27 seconds. There are some steady state oscillations in EGR fraction as well. These were determined to be a result of a ‘sticky’ EGR valve and controller combination. The increase in rise and fall time are mainly due to the difference between modeled and measured EGR fractions. The comparison between the performance criteria of these two control strategies is summarized in Table A-1.

Figure A-12. Step response of EGR fraction controller with conventional PID controller at 1300RPM 4 Bar BMEP
Figure A-13. Step response of EGR fraction controller with Smith Predictor
Figure A-14. Step response of EGR fraction controller with conventional PID controller at 2000RPM 9 Bar BMEP

Figure A-15. Step response of EGR fraction controller with Smith Predictor at 2000RPM 9 Bar BMEP
Table A-1 Conventional PID control vs Smith Predictor Control performance comparison

<table>
<thead>
<tr>
<th>Performance criteria</th>
<th>Conventional PID Control</th>
<th>Smith Predictor Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average overshoot [%]</td>
<td>36.3</td>
<td>12</td>
</tr>
<tr>
<td>Average rise time [sec]</td>
<td>0.823</td>
<td>1.23</td>
</tr>
<tr>
<td>Average fall time [sec]</td>
<td>0.383</td>
<td>1.13</td>
</tr>
<tr>
<td>Steady state behavior</td>
<td>Oscillating</td>
<td>Stable</td>
</tr>
</tbody>
</table>

Summary/Conclusions

A closed loop LP-cEGR feedback controller with Smith predictor based dead time delay compensator was implemented and tested. Open loop models for EGR fraction at the mixing and EGR sensor locations along with a transport delay model were developed and
validated. EGR valve step changes were performed to evaluate step response of a conventional PID controller with that of Smith Predictor based control.

The Smith Predictor control showed significant reduction in average overshoot as well as improved steady state stability. However, due to mismatch between EGR sensor feedback and modeled EGR fraction, rise time and fall time increased as compared to PID control. Accurate open loop modeling of EGR mass flow in LP-cEGR systems is still an area where development is required. The conventional orifice flow model begins to lose accuracy at low pressure differentials and with pressure pulsations. These conditions require additional calibration and tuning parameters to be usable for control.
Nonlinear Model Predictive Control of Dual Loop - Exhaust Gas Recirculation in a Turbocharged Spark Ignited engine

The external Exhaust Gas Recirculation (EGR) technology is widely used on Compression Ignition (CI) engines for its ability to reduce NOx emission [69]. This technology was introduced to Spark Ignition (SI) engines recently due to its ability to mitigate knock and reduce pumping loss[11]. However, SI engines require much more precise EGR concentration control than the CI engines. The employment of High Pressure (HP) and Low Pressure (LP) dual loop (DL) EGR provides more degrees-of-freedom for EGR concentration control. The LP-EGR loop has a characteristic transport delay much longer than the HP-EGR loop, which introduces additional difficulty in air fraction control in the intake manifold. Conversely, exhaust enthalpy exchange through the turbine is reduced in the case of HP-EGR thereby reducing turbine power and hence adversely affecting compressor mass flow rate. It is desirable to coordinate the dual loop EGR systems to achieve given EGR concentration target and maximize turbocharger efficiency.

In this work, a Nonlinear MPC (NMPC) is proposed to control the intake manifold pressure (MAP) and EGR concentration of a turbocharged SI engine with DL-EGR, by manipulating the throttle, EGR valves of HP and LP loops. For a given preview horizon of tacking references, the NMPC minimizes tracking error and control effort subject to the actuator limit and states constraints.

The control of turbocharged gasoline engine air handling system with EGR is a Multi-Input Multi-Output (MIMO) problem with nonlinear and coupled dynamics. NMPC has been utilized in [93], to track torque for a turbocharged SI engine where the throttle
and wastegate are considered as the control inputs. External EGR control has not been included in the aforementioned research. While references for the DL-EGR control of SI engine are limited, the control of the diesel engine air system with dual loop EGR has been extensively discussed in literature [69] and [94], where the authors successfully implemented a switched linear MPC with feed-forward and feedback controllers for the multi-input multi-output nonlinear DL EGR and VGT system. In these papers, multiple linear models were developed over engine speed and fueling operating regions and a switched linear MPC approach with quadratic cost function was employed to calculate control actions.

The EGR control strategy developed for CI engines may not be suitable for SI engines. The LP EGR transport dynamics considered in previous CI engine research are highly simplified and usually similar to a first order delay. This approximation may be tolerable for CI engine operation as the ignition process is more robust against EGR dilution. Such an assumption for a SI engine can potentially lead to undesirable operation like misfire and partial combustion due to over dilution of EGR. Hence, it is necessary to improve the accuracy of transport delay dynamics in the control oriented model of the airpath as shown in Figure A-16, wherein a NMPC has been utilized with a segmented boost manifold based transport delay model integrated into the airpath model. This research employs a similar approach in which a high order transport delay model is utilized to improve the prediction of LP-EGR fraction from the location of its delivery to the engine cylinders.
Control oriented air path model

The individual components of the air path have been modeled separately and ultimately integrated to form a control oriented air path as depicted in Error! Reference source not found..

Figure A-16. Schematic of Engine and Air path model.

Here, \( P, T, F, \dot{m}, \omega, \) and \( V \) are pressure, temperature, mass fraction, mass flow rate, rotational speed and volume respectively. Subscripts \( Exh, Bst, Im, HPE, LPE, ITV, Comp, Turb, I \) and \( E \) are the exhaust manifold, boost manifold (between compressor and throttle), intake manifold, high-pressure EGR, low-pressure EGR, intake throttle, compressor, turbine, ambient compressor inlet and post-turbine exhaust respectively. The
following section is dedicated to description of the component level models used in the air path model.

**Gas transport dynamics in boost manifold**

The boost manifold is typically the largest control volume in a turbocharged engine air path and hence consideration of the gas transport dynamics in the boost manifold is critical for transient control of EGR dilution at the cylinders. This manifold has been traditionally considered as a lumped volume in the turbocharged diesel engine air path. However, in typical turbocharged automotive engines the boost manifold is characterized more by thin long pipes rather than plenums. Therefore, a lumped volume based assumption for the boost manifold is done at the expense of accuracy. Experimental testing has confirmed the above mentioned phenomenon as shown in Figure 4 where a step input in EGR valve position results in a dead time delayed and mixing influenced response from the EGR measurement sensor installed at the throttle valve. The mixing influence cannot be attributed purely to mixing in the manifold but also to the filling and emptying dynamics of the sensor itself.
The consequence of excessive EGR dilution in diesel engines typically results in increased smoke and soot emission. However, with excessive EGR dilution, gasoline spark ignited engines can suffer from very high cyclic variability in torque production. Furthermore, the three way catalyst can encounter premature damage due to misfires caused by excessive dilution. The gas transport model chosen here (91) is based on conservation of mass and the ideal gas law in a control volume.

\[ \dot{F}_i = \frac{RT_i}{P_i V_i} (F_i \dot{m}_{in} - F_i \dot{m}_{out}) \]

where, \( F_i \) is the EGR fraction in the \( i^{th} \) control volume \( V_i \). The number of control volumes utilized to discretize the boost manifold is the same as the number of fraction states \( F_i \) used in the transport model. Hence, the order of the transport model is equal to the number of sections used to discretize the control volume.
Figure A-18. Comparison of step response of transport delay models with true dead time delay response

Higher order models show increasing accuracy and similarity to the ‘true’ dead time delayed response (if pure pipe plug flow is considered) as shown in Figure A-18. A 5th order transport delay model was chosen for this research as it has significantly higher accuracy than the 1st order model which is a single lumped volume for the entire boost manifold. Moreover, higher order models would have diminishing improvement in accuracy over the computational burden of additional state dynamics.
State space model and control problem formulation

Using the component level models described in the previous section, iso-thermal pressure dynamics and turbocharger rotor dynamics described in [14] the discrete time state-space model is derived as given below. Since the goal of this research was to evaluate control of mass flow rates of the EGR valves, temperature dynamics were ignored to keep system dimensionality low and allow for more transport delay states.

\[ x(k + 1) = f(x(k), u(k)) \]
\[ y = Cx = [P_{Im} F_{Im}]^T \]
\[ x = [P_{Im} P_{Bst} \omega_{Turbo} F_1 F_2 F_3 F_4 F_5 F_{Im} \bar{m}_{LPE}]^T \]
\[ u = [\bar{m}_{ITV} \bar{m}_{HPE} \bar{m}_{LPE}]^T \]

The output vector \( y \) consists of the intake manifold pressure \( P_{Im} \) and intake manifold EGR\% \( F_{Im} \). For gasoline SI engines the intake manifold pressure and EGR\% strongly associated with the engine load and these references are assumed to be available from an external source (eg. Torque management system, supervisory control). The remaining states \( P_{Bst}, \omega_{Turbo}, F_i \) and \( \bar{m}_{LPE} \) are the compressor outlet pressure, turbocharger rotational speed, the transport delay states and the LP-EGR mass flow rate respectively. The input vector \( u \) consists of the intake throttle and high pressure mass flow rate \( \bar{m}_{ITV} \) and \( \bar{m}_{HPE} \). The LP-EGR mass flow rate is controlled using input \( \bar{m}_{LPE} \), as this
allows tuning of the rate of change of the LP-EGR mass flow rate. The model is also normalized to mitigate potential numerical issues during execution.

The objective of the control algorithm is to minimize the squared error between a reference value $y_{Ref}(k)$ and the system output $y(k)$ for the preview horizon. Hence, the optimal control problem has been designed to be an output tracking controller with the cost function defined as follows.

$$J(x(k), U(k)) = \bar{y}(k+N)^TH\bar{y}(k+N) + \sum_{i=k}^{k+N-1} \bar{y}(i)^TQ\bar{y}(i) + u(i)^TRu(i)$$

where,

$$\bar{y}(k) = y(k) - y_{Ref}(k)$$

$Q$ and $H$ are the tuning matrices for the penalty on output tracking error within the prediction horizon and the terminal output error. The tuning of these matrices was done to place more emphasis on the tracking of $F_{Im}$ as the consequence of poor tracking of this parameter can result in misfires. $R$ is the tuning matrix for penalty on magnitude of input $u$. The non-linear optimization problem is hence formulated as follows.

$$\min_{U(k)} J(x(k), U(k))$$

$$\dot{m}_{ITV}(i) \geq 0,$$

$$\dot{m}_{HPE}(i) \geq 0,$$

$$P_{Bst}(i) - P_{Im}(i) \geq 0,$$
\[ \omega_{Turbo}(i) \geq 0. \]

where, \( i = k, k + 1 \ldots, k + N - 1 \)

All of the constraints are imposed to maintain physical feasibility in the system. The turbocharger rotational speed has the longest time constant in the system (1.1 seconds) Hence, for a sample time of 0.1 seconds the horizon length \( N \) is 11. The numerical values chosen in the tuning matrices are chosen heuristically with more emphasis on the EGR% control. Fig. 6 shows the overall schematic representation of Simulation results and discussion the NMPC control structure. The NMPC problem was simulated using ACADO toolkit [95].
The simulation was performed to study the ability of NMPC to handle the following aspects of the system:

- Reference tracking of $P_{lm}$ and $F_{lm}$
- LP-EGR transport delay in the boost manifold
- Interaction between EGR loops and turbocharger
- Control of split ratio between HP-EGR and LP-EGR

The intake manifold pressure reference trajectory was deliberately chosen to have boosted and un-boosted values so that the interaction between the EGR circuits and the turbocharger performance could be studied. Arbitrary time varying references were chosen for EGR%. The simulation was performed for three engine speeds, 1000, 2000 and 3000 RPM with the same reference trajectories. The speeds were chosen because external cooled EGR has maximum benefit in knock mitigation at low speed high load conditions. In the following section simulation trajectories are shown for 2000 RPM.

**High-pressure EGR control only**

It is evident from Figure A-20 and Figure A-21 that in this case the system is able to track the reference of the intake manifold mass better than the low pressure only and dual loop cases. However, the intake manifold pressure reference tracking under a reference value above atmospheric pressure is the poorest in this case as seen between 5 and 7 seconds. This is because of the reduction of mass flow rate through the turbine caused by the high-pressure EGR system. The HP-EGR circuit effectively behaves like a wastegate and hence results in reduced turbine power. As per the perception of the driver, this
would seem as poor engine torque response. The controller overshoots the intake throttle and HP-EGR mass flow rates whenever there is a sudden change in reference. Another thing to note is that the boost pressure always stays above or equal to the intake manifold pressure implying that $\dot{m}_{ITV}$ is physically feasible.

Figure A-20 Pressure states and intake manifold EGR dilution trajectories for HP-EGR control only.
Figure A-21 Input trajectories for HP-EGR control only.
Low-pressure EGR control only

The most significant difference between this case and the HP-EGR case is that the intake manifold pressure reference tracking is significantly better as seen in Figure A-22 and Figure A-23. as all of the exhaust mass flow passes through the turbine resulting in higher turbine power to drive the compressor to higher boost levels. However, the reference tracking for EGR mass in the intake manifold is worse with oscillation at the EGR% reference changes at 5 and 8 seconds. Due to the transport delay model being incorporated into the NMPC, there is also some lead action visible at 0.5 seconds on the LP-EGR mass flow rate to meet the EGR% reference change at 1 second. The LP-EGR mass flow rate has an immediate effect on only the EGR% in the first section of the boost manifold which is far upstream of the intake manifold. The dilution dynamics in every consequent air-path section diminishes the effect of this input on the intake manifold EGR%. Hence, the tuning weight in matrix R associated with the rate of change of LP-EGR mass flow rate had to be modified to reduce oscillations in the LP-EGR mass flow rate.
Figure A-22 Pressure states and intake manifold EGR dilution trajectories for LP-EGR control only.
Figure A-23 Input trajectories for LP-EGR control only.
Dual-loop EGR control

In this case the intake manifold pressure and EGR mass reference tracking appear to be reasonably better compared to the LP-EGR only case as shown in Figure 9a. and 9b. Since the engine is operated throttled up to 5 seconds the NMPC commands only HP-EGR up to 4 seconds. At the transition from throttled to boosted operation LP-EGR flow is initiated to maximize turbine power. As the intake manifold pressure level increases due to turbo spool-up, a sharp HP-EGR mass flow rate spike is commanded at 5 and 8 seconds to increase intake manifold pressure further while maintaining the desired EGR% reference. The penalty on rate of change of LP-EGR mass flow rate was also relaxed compared to the LP-EGR only case. It is noteworthy that the commanded LP-EGR mass flow rate shows lesser oscillations despite the relaxed penalty indicating that NMPC effectively uses HP-EGR to ‘supplement’ the LP-EGR during boosted operation.
Figure A-24 Pressure states and intake manifold EGR dilution trajectories for Dual-loop EGR control.
Figure A-25 Input trajectories for Dual-loop EGR control.
**Additional comparison between different cases**

The intake manifold EGR% for different cases at 2000 RPM is shown in Figure A-25. Even though the HP-EGR case settles to the reference at 5 seconds, there is an error associated with the first order filling dynamics of the intake manifold visible between 4.9 and 5 seconds. This error is diminished for the LP-EGR and DL-EGR case as the primary EGR actuator in these cases is the LP-EGR mass flow rate which is farther downstream from the intake manifold.

![Figure A-26 Transient intake manifold EGR dilution for multiple EGR architectures](image)

The performance of the NMPC for dual and single loop EGR architectures has been summarized in TABLE A-2. At 1000 and 2000 RPM the DL-EGR shows reduced RMSE
and Max errors compared to the single loop architectures. However, at 3000 RPM the DL-EGR showed poorer performance than the single loop systems. This can be attributed to non-ideal tuning for the output tracking and input penalties.

TABLE A-2. Performance summary of EGR % control using NMPC

<table>
<thead>
<tr>
<th>Engine speed [RPM]</th>
<th>EGR architecture</th>
<th>RMSE [EGR%]</th>
<th>Max Error [EGR%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>HP-EGR</td>
<td>0.105</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>LP-EGR</td>
<td>0.13</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>DL-EGR</td>
<td>0.0489</td>
<td>0.6</td>
</tr>
<tr>
<td>2000</td>
<td>HP-EGR</td>
<td>0.13</td>
<td>1.53</td>
</tr>
<tr>
<td></td>
<td>LP-EGR</td>
<td>0.17</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>DL-EGR</td>
<td>0.07</td>
<td>0.061</td>
</tr>
<tr>
<td>3000</td>
<td>HP-EGR</td>
<td>0.16</td>
<td>2.07</td>
</tr>
<tr>
<td></td>
<td>LP-EGR</td>
<td>0.1</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>DL-EGR</td>
<td>0.2</td>
<td>2.34</td>
</tr>
</tbody>
</table>

Conclusion and future work

The proposed dual-loop EGR NMPC strategy is capable of coordinating the HP and LP EGR mass flow rate to track the intake manifold pressure and intake manifold EGR% references. Inclusion of turbocharger dynamics in the model facilitates optimal balancing between the EGR loops whilst minimizing turbocharger lag. The NMPC
guarantees constraints and hence physically feasible values for the control inputs. It also demonstrates the ability to consider the transport delay dynamics in the LP-EGR loop due to the utilization of a multi-segment transport delay model. For the lower engine speeds, DL-EGR control shows superior EGR% tracking performance. Inclusion of compressor and turbine bypass valve, engine cylinder models in addition to further refinement of the tuning parameters in the NMPC will be considered as future work.
REFERENCES


[38] Kiwan, R., Stefanopoulou, A., Martz, J., “Effects of differential pressure measurement characteristics on low pressure-EGR estimation error in Si-engines”, 2016, IFAC (International Federation of Automatic Control)


[44] Leufven, O., and Eriksson, L., “Surge and Choke Capable Compressor”, Model 2011 IFAC World Congress, Milano, Milano


[61] “Exhaust Gas Recirculation Control for Large Diesel Engines - Achievable Performance with SISO Design,” IFAC 9th Conference on Control Applications in Marine Systems, ISSN – 14746670.


[64] Normey-Rico, J., “Control of Dead-time Processes,” Springer publications


[92] Source of engine picture:
https://www.gmperformancemotor.com/parts/19328837.html
