

A Nonlinear Model Predictive Control Strategy with a Disturbance Observer for Spark Ignition Engines with External EGR

Qilun Zhu and Robert Prucka
 Clemson University

Michael Prucka and Hussein Dourra
 FCA US LLC

ABSTRACT

This research proposes a control system for Spark Ignition (SI) engines with external Exhaust Gas Recirculation (EGR) based on model predictive control and a disturbance observer. The proposed Economic Nonlinear Model Predictive Controller (E-NMPC) tries to minimize fuel consumption for a number of engine cycles into the future given an Indicated Mean Effective Pressure (IMEP) tracking reference and abnormal combustion constraints like knock and combustion variability. A nonlinear optimization problem is formulated and solved in real time using Sequential Quadratic Programming (SQP) to obtain the desired control actuator set-points. An Extended Kalman Filter (EKF) based observer is applied to estimate engine states, combining both air path and cylinder dynamics. The EKF engine state(s) observer is augmented with disturbance estimation to account for modeling errors and/or sensor/actuator offset. The complete control system demonstrates strong disturbance rejection and closed loop stability. RMS error of IMEP reference tracking is 1.1% for engine cycles without active combustion constraints. Mean computation time for the proposed control system is 1.07 ms in the rapid-prototype engine controller, suggesting it is feasible for future ECUs. The proposed model predictive engine control strategy actively compensates for air path delay, resulting in higher EGR percentage and reduced combustion phasing retard as compared to conventional engine control strategies. FTP drive cycle simulation results of both conventional map-based engine control and the proposed E-NMPC on the same engine model suggest approximately 6.5% of fuel economy improvement.

CITATION: Zhu, Q., Prucka, R., Prucka, M., and Dourra, H., "A Nonlinear Model Predictive Control Strategy with a Disturbance Observer for Spark Ignition Engines with External EGR," *SAE Int. J. Commer. Veh.* 10(1):2017, doi:10.4271/2017-01-0608.

INTRODUCTION

Engine manufacturers need to optimize their control strategies to achieve balance between performance and fuel economy. For traditional engine control strategies (e.g. map-based), that are designed to provide engine torque demanded by the driver, achieving one of the two objectives is often accompanied by the sacrifice of the other. Balancing between fuel economy and torque response can be handled by Model Predictive Control (MPC) strategies if future engine torque demand previews are available [1], [2], [3], [4]. Within each sampling time, the MPC engine control is able to find the optimal actuator set points that minimize fuel consumption and torque (or Indicated Mean Effective Pressure, IMEP) tracking error for the preview horizon, using a control oriented engine model. During this real time optimization process, constraints of actuator saturation, engine knock and combustion variability can be imposed to guarantee the feasibility of computed control actions. While the development of driver modelling and autonomous driving technology enables the preview of future torque demands, this topic is not investigated by this research. Considering the future IMEP demand is available, an Economic Nonlinear Model Predictive Control (E-NMPC) strategy for SI engines with external EGR is discussed in this research. Engine states

are estimated by an Extended Kalman Filter (EKF) enhanced with disturbance observation. The proposed control system is evaluated in terms of closed loop stability, IMEP reference tracking performance, fuel economy gain and computation time.

The engine models required to formulate the proposed E-NMPC include air path, torque generation, knock and combustion variation. Control oriented engine air path and torque generation models are well established [5]. The SI engine system is modeled and controlled in the engine cycle domain in this paper. While this approach agrees with the discrete nature of both MPC and IC engines, it also benefits from the fact that most control oriented knock and combustion variation models were constructed in the engine cycle domain. Previous researchers have proposed that knock can be predicted by integrating the Arrhenius function output for the end gases [6,7]. Coefficient of variation (COV) of IMEP is commonly used as an indicator of combustion variability for IC engines. Previous literature has illustrated that the stability of turbulent combustion can be correlated to the time scale ratio between turbulence motion and flame propagation [8,9]. Lee *et al.* [10] suggested that the COV of IMEP has strong correlation with combustion phasing. If the phasing

the modeling of COV of IMEP from traditional combustion stability research. Finally, regression models of COV of IMEP have been proposed by [11] and [12].

The proposed E-NMPC penalizes fuel consumption for the preview horizon. This penalty differentiates the current work from a tracking MPC, which is designed to minimize least square error from a reference only. The non-convexity of this problem results in multiple local optimal solutions. Global Nonlinear Programming (NLP) solvers, like dynamic programming and particle swarm, can be employed on MPC applications [13,14] and guarantee the closed loop stability [15,16]. However, these global NLP solvers require numerous evaluations of the system model, making them infeasible for MPC applications requiring a fast sampling time. Most model predictive engine control researchers have selected sub-optimal strategies to reduce the computational demand [17,18,19]. Linear parameter variant (LPV) MPC is a widely adopted sub-optimal predictive controller for nonlinear systems [20,21]. However, the LPV MPC is not suitable for the investigated engine control problem due to the high nonlinearity of the system. In this case, the optimization cannot converge to a local minimum or guarantee the feasibility of computed control actions. The later situation is more undesirable since it could lead to misfire or knock.

This research considers Sequential Quadratic Programming (SQP) to solve the real time NLP problem and obtain desired control actions. SQP is a continuous NLP algorithm based on Newton's method [22]. It is proven to converge to a local optimal solution near the starting point [23,24,25]. Previous research has discussed the possibility of applying SQP to NMPC [19]. The most important advantage of SQP is that it transforms complex NLP into a sequence of sub-level quadratic programming (QP) problems. Computing the Hessian matrix for each sub-QP is challenging, especially if the system model is complex and implicit. Numerical differentiation methods like algorithmic differentiation and finite difference are often necessary to linearize the model and obtain the Hessian matrix. This research exploits the Gauss-Newton-like structure of the investigated ENMPC to simplify the Hessian matrix computation. Compared to the Broyden - Fletcher - Goldfarb - Shanno (BFGS) rank-two update method [26,27] and disaggregated Hessian approximation [28], the Hessian of the proposed SQP strategy is inherently positive definite so well-developed convex QP algorithms can be applied.

The sub-QP problem can be solved using existing algorithms. Since the development of active set QP algorithms in the 1980s, this solver has been the fastest option for online operation [30], [31], [32], [33]. This style of active set algorithm is based on the fact that QP problems have closed form solutions if the active constraint set of the optimal solution is known. More recently developed parametric active set QP algorithms have been shown to exploit the benefits of warm start techniques [34,35]. They have better degeneracy handling capabilities and faster convergence rates if the initial guess of the active constraint set is not far from the optimal solution. For each SQP iteration of the proposed MPC strategy, the sequence of QPs

share similar structure and active constraint sets, making the parametric active set algorithm favorable for this application. In addition to active set methods, the primal-dual (or interior point) method is another option to solve QP for MPCs [36,37]. It is not widely considered for fast MPCs since it requires more computational effort to complete each iteration. Furthermore, the difficulty in finding a "warm" start point also makes it not suitable for fast online operation.

Engine state(s) estimation is required for the E-NMPC to compute the control actions. Most previous research discusses engine state(s) estimation separately in terms of different subsystems. Many publications can be found in air path system estimation [38], [39], [40], [41]. Cylinder composition estimation is extensively discussed in terms of air-to-fuel ratio control [42,43]. This research proposes a model based observer that utilizes sensors of different engine subsystems to generate a more informed estimation of multiple states simultaneously. This 'fusing' of sensors from the air path and in-cylinder reduces the impact of inaccuracy from individual sensors. It also provides possibilities to estimate states that are originally un-observable using decentralized estimators (e.g. residual gas mass).

Many estimation algorithms can be applied to this nonlinear observer application. The Extended Kalman Filter is considered in this research for its ability to optimally choose estimation gain based on the power of sensor noise and actuator disturbances. Furthermore, the engine model is already linearized within the execution of the sequential quadratic programming based MPC controller. The EKF can leverage this advantage to reduce the computation load. However, the performance of the EKF degrades if non-white-noise disturbances are present in the system. These situations occur due to modelling error, sensor/actuator fault and failure of a low level controller to track properly, so a disturbance observer is proposed to address these issues [44,45,46]. Similar to the concept of integral control that eliminates steady state tracking error, the disturbance observer approach augments the original state(s) estimation with integration states of potential error sources. With inputs from sensors the EKF can identify which error source(s) cause the disturbances instead of lowering confidence in the sensors.

This paper is organized as follows. First, an overview of the investigated engine system and proposed control structure is provided, and the control oriented engine model is formulated. The proposed SQP MPC strategy is then described in detail. The EKF engine state estimation routine with disturbance observer is then derived, followed by a discussion of simulation and experimental results. Experimental results describing how the E-NMPC tracks a step changing IMEP reference without violating constraints is then provided. The computational efficiency of the proposed algorithm is then quantified. Simulation results are then used to compare the fuel economy of a map-based engine control strategy and the proposed E-NMPC using the same engine model. Finally, conclusions are drawn related to the contribution of this research.

SYSTEM OVERVIEW

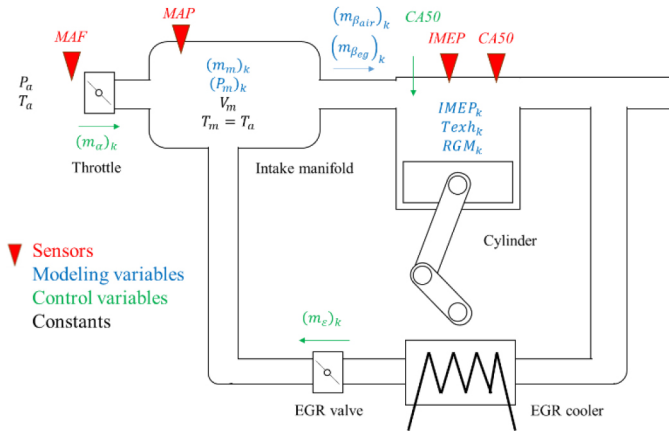


Figure 1. Block diagram of the SI engine system with high pressure external EGR.

This research focuses on IMEP control of SI engines with external EGR. Fuel injection control is assumed to maintain a stoichiometric air-to-fuel ratio (AFR), maximizing catalyst efficiency. Manifold temperature is assumed to be constant since the EGR is cooled with a heat exchanger. Finally, the gases in the air-path system are considered incompressible. Figure 1 shows the engine configuration with labels of sensors, control and modeling variables.

The proposed MPC is designed in the engine cycle domain. The MPC manipulates throttle air mass flow per cycle, m_a , EGR valve mass flow per cycle, m_e , and combustion phasing CA50 (crank angle at which 50% of total heat release occurs). These variables are sent to lower level controllers, with faster update frequencies, as references. These lower level controllers can be map-based or low-dimensional optimal control strategies that have low computation load (e.g. an optimal CA50 control strategy [47]). The two-layer supervisory control structure (shown in Figure 2) exploits the frequency separation of different system dynamics and removes nonlinearities from upper level controllers, making it favorable for fast MPC applications (e.g. [48]).

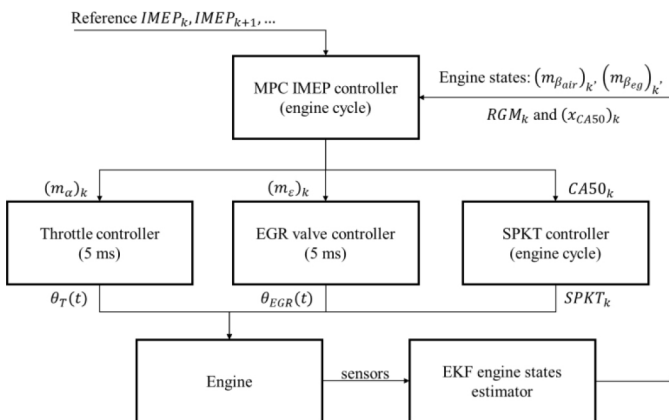


Figure 2. Two-layer supervision engine control structure.

The 5 ms throttle and EGR valve controllers are based on the inversion of the orifice flow model. The spark timing (SPKT) is controlled by directly inverting an empirically calibrated CA50

map. The proposed engine IMEP control hierarchy transfers the nonlinear orifice valve flow and combustion phasing models to the faster sub-level controllers. This process reduces model complexity for the MPC and lowers computation load. Another advantage of this control structure is that feedback control and lookup tables can be applied to the sub-level controllers since they only deal with low order nonlinear dynamics. These controllers can have faster update frequencies than the MPC, better exploiting the available bandwidth of the control actuators. This structure also allows for fine tuning of individual actuator responses to compensate for dynamics ignored by the MPC loop.

CONTROL ORIENTED ENGINE MODEL

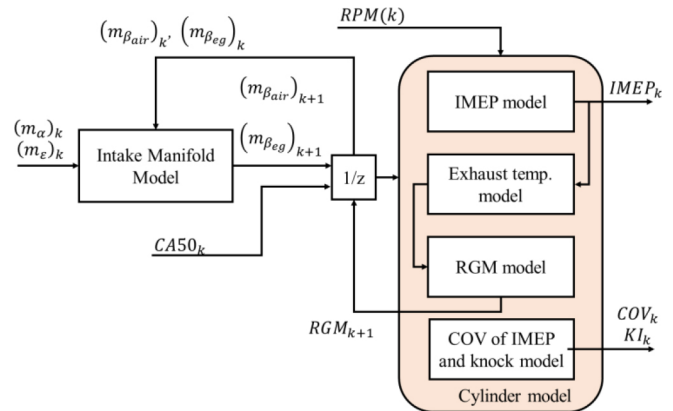


Figure 3. Block diagram of the control oriented engine model

The control oriented engine model consists of intake manifold and cylinder models. The intake manifold model computes the amount of air and EGR flow into the cylinder, using a lumped volume method and speed-density for cylinder charge. The cylinder model computes IMEP and exhaust temperature using an energy balance method. The Residual Gas Mass (RGM) is also computed by the cylinder model using the camshaft phasing, engine speed, MAP and exhaust temperature result from the energy balancing model. Figure 3 shows the block diagram of the control oriented engine model.

In the engine cycle domain, cylinder air charge per cycle can be computed according to intake manifold air density and volumetric efficiency as:

$$m_{\beta_{air}} = \frac{m_{m_{air}}}{V_m} V_d \eta_v \quad (1)$$

where:

$m_{m_{air}}$ is air mass in the intake manifold;

η_v is volumetric efficiency;

V_d is engine displacement;

V_m is intake manifold volume.

The air mass balance of the intake manifold can be expressed as:

$$(m_{m_{air}})_{k+1} = (m_{m_{air}})_k + (m_{\alpha})_k - (m_{\beta_{air}})_k \quad (2)$$

Re-arranging [equation \(1\)](#), and substituting it into [equation \(2\)](#):

$$(m_{\beta_{air}})_{k+1} = \frac{1}{K+1}(m_{\beta_{air}})_k + \frac{K}{K+1}(m_{\alpha})_k \quad (3)$$

where:

$$K = (\eta_V(k)V_d)/V_m$$

The cylinder EGR flow can be modeled similar to the air mass flow, as the follows:

$$(m_{\beta_{eg}})_{k+1} = \frac{1}{K+1}(m_{\beta_{eg}})_k + \frac{K}{K+1}(m_{\epsilon})_k \quad (4)$$

The manifold pressure is then computed by rearranging the speed density equation, as:

$$(P_m)_k = \frac{((m_{\beta_{air}})_k + (m_{\beta_{eg}})_k)RT_m}{(\eta_V)_k V_d} \quad (5)$$

It can be observed from [equations \(3\), \(4\), and \(5\)](#) that manifold dynamics are independent of engine speed (regardless of the slowly varying volumetric efficiency η_V), unlike most time domain models. In-cylinder gas composition includes air, exhaust gas, fuel and other minor species that are neglected in this research. The amount of air and fuel can be determined by $m_{\beta_{air}}$, assuming a stoichiometric AFR. The fuel mean effective pressure (P_f) can be computed as:

$$(P_f)_k = LHV \cdot \frac{(m_{\beta_{air}})_k}{\sigma_0 V_d} \quad (6)$$

where:

σ_0 is stoichiometric AFR

LHV is low heating value of the fuel

The amount of in-cylinder exhaust gas is the summation of $m_{\beta_{eg}}$ and residual gas mass (RGM). This research utilizes a modified semi-empirical model originally proposed by *Fox et al.* [49] that separates RGM into two parts; (1) trapped residual at exhaust valve closing (EVC) due to un-swept cylinder volume, and (2) exhaust gas backflow into the cylinder and intake runner during the valve overlap period. After adding terms ΔP_{exh} and ΔP_m to account for wave tuning dynamics to the original Fox model [50], the residual gas mass for each engine cycle can be calculated according to:

$$RGM_k = C_1 \frac{P_{exh}}{R(T_{exh})_{k-1}} V_c + C_2 \sqrt{\frac{P_{exh}}{R(T_{exh})_{k-1}} ((P_{exh} + \Delta P_{exh}) - (P_m + \Delta P_m)_{k-1}) A_{flow} \frac{OLV}{\omega_e}} \quad (7)$$

where:

$P_{exh} \approx 110 \text{ kPa}$ is the exhaust pressure.

T_{exh} is the exhaust temperature.

R is gas constant.

V_c is the cylinder clearance volume.

A_{flow} is effective flow area during valve overlap period.

OLV is overlap volume which is the cylinder volume difference between EVC and IVO.

C_1, C_2 are calibration factors.

ω_e is engine speed.

The total fraction of in-cylinder exhaust gas c_{eg} can be generated as:

$$(c_{eg})_k = \frac{(m_{\beta_{eg}})_k + RGM_k}{(m_{\beta_{air}})_k + (m_{\beta_{eg}})_k + RGM_k} \quad (8)$$

With the information of cylinder composition, IMEP can be modeled using the Willans approximation method [5]:

$$IMEP_k = e_k (P_f)_k - P_0 \quad (9)$$

where:

e is the 'slope' factor.

$P_0 = P_{exh} - P_m$ is pumping effective pressure (PMEP).

The slope factor, e , is related to engine speed (ω_e), $CA50$ and c_{eg} . The final value of e is the product of multiple slope factors with dependence on fewer inputs:

$$e = e_{\omega}(\omega_e) e_{\zeta}(CA50) e_{eg}(c_{eg}, \omega_e) \quad (10)$$