

A Novel Singular Perturbation Technique for Model-Based Control of Cold Start Hydrocarbon Emission

Mohammad Reza Amini and Mahdi Shahbakhti
Michigan Technological Univ.

Ali Ghaffari
K N Toosi Univ. of Technology

ABSTRACT

High hydrocarbon (HC) emission during a cold start still remains one of the major emission control challenges for spark ignition (SI) engines in spite of about three decades of research in this area. This paper proposes a cold start HC emission control strategy based on a reduced order modeling technique. A novel singular perturbation approximation (SPA) technique, based on the balanced realization principle, is developed for a nonlinear experimentally validated cold start emission model. The SPA reduced model is then utilized in the design of a model-based sliding mode controller (SMC). The controller targets to reduce cumulative tailpipe HC emission using a combination of fuel injection, spark timing, and air throttle / idle speed controls. The results from the designed multi-input multi-output (MIMO) reduced order SMC are compared with those from a full order SMC. The results show the reduced SMC outperforms the full order SMC by reducing both engine-out and tailpipe HC emission. This improved performance is caused by setting an optimum priority among the control inputs through the incorporation of SPA that evaluates the significance of the control variables/states on HC emission. The presented SPA approach can be applied for different automotive control applications that face the challenge of how to set the priority among different control actuators which often have conflicting impacts on desired control targets.

CITATION: Amini, M., Shahbakhti, M., and Ghaffari, A., "A Novel Singular Perturbation Technique for Model-Based Control of Cold Start Hydrocarbon Emission," *SAE Int. J. Engines* 7(3):2014, doi:10.4271/2014-01-1547.

INTRODUCTION

The automotive industry is facing the challenging and often conflicting requirements of reducing harmful exhaust emissions and reducing fuel consumption while maintaining drivability performance. These challenges will dramatically increase with the upcoming Low Emission Vehicle (LEV) III standards and the California Air Resource Board (CARB) fuel economy standards. Vehicle cold start is one major challenge which directly affects carbon monoxide (CO) and HC emission, fuel economy, and vehicle drivability. It is known that over 80% of un-burnt HC in SI engines are formed during the early period of cold engine operation below an ambient temperature of 20°C [1]. During the cold start phase, if the ambient temperature drops to -7°C, emission levels increase by 2.5 - 4 times more than in the case of 20°C [2,3]. High emission production rates during this time and the fact that this is an unavoidable part of the standard driving cycle dictate the importance of dealing with the HC emission problem during the cold start phase. The

main reasons for high HC production are poor fuel-air mixing, cold cylinder walls, and poor evaporation of cold fuel. Cold cylinder walls cause the injected fuel to condense. This results in poor fuel-air mixing conditions which lead to unstable combustion. Moreover, because of the low conversion efficiency of catalytic converters at low temperatures, un-burnt HC cannot be mitigated by the exhaust aftertreatment system until catalyst "light-off" (i.e. operating temperature) is reached.

Figure 1 provides an overview of some of the proposed solutions to tackle high HC emission during the cold start phase. The proposed solutions are divided into two major groups: active and passive measures [4]. Active measures are those in which extra equipment is added either to the engine or to the exhaust aftertreatment system to meet the desired requirements [5, 6, and 7]. Two major drawbacks in these measures include extra cost and often extra required energy supply.

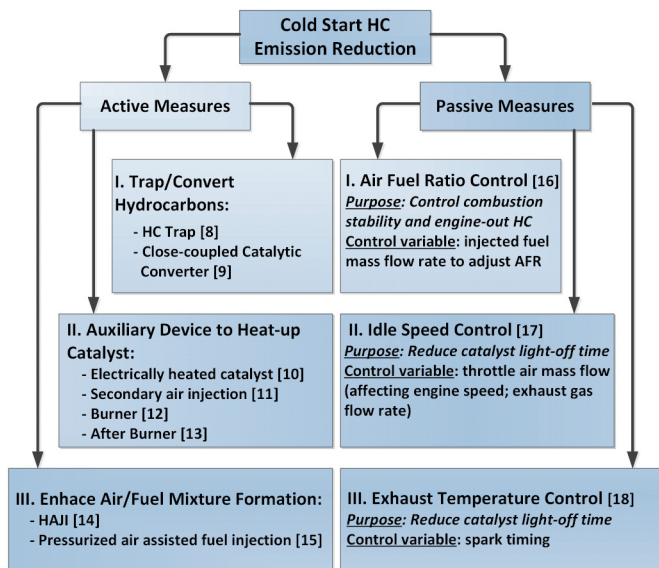


Figure 1. Some of major methods for reducing cold start HC emission in SI engines.

An example of some recent advanced active methods is the development of a novel catalyst for cold start emission control [19]. The developed catalyst stores HC and NO_x at low temperatures and then converts a considerable portion of the stored HC/NO_x during the warm up period. An HC trap is another option to reduce HC emission; however, due to various issues like limited durability and high platinum group metal (PGM) concentrations, HC traps are not widely used [20]. In this regard, a new trapping material and system structure are proposed and verified [20].

In contrast to active measures, passive measures try to manage engine control variables to optimize combustion and exhaust aftertreatment efficiency. Example of SI engine control variables include throttle air flow rate, injected fuel flow rate, and spark timing. The major benefits of passive measures are preventing extra equipment cost and increasing the whole system flexibility to meet multi-objective targets. This study centers on developing a passive measure for reducing HC emission in SI engines.

One important requirement for developing a cold start emission control strategy is having an accurate and computationally efficient low-order physical model for real-time control. This model should have the capability to be easily implemented and utilized in the design of model-based controllers to evaluate system responses under real environmental conditions and uncertainties. Several engine cold start models and controllers have been reported [4, 21]. Some recent cold start studies include development of control-oriented models [21, 22], controller design [23, 24], and increasing robustness to implementation imprecision [25, 26]. Study in [27] presents a methodology that integrates appropriate physics-based models of the engine and catalyst into a numerical optimization scheme. Andrianov et al. presented a general physics-based integrated model of a SI engine, the exhaust system and a

three-way catalyst [28]. The model calculates fuel consumption, as well as CO, NO, and HC tailpipe emissions under transient driving conditions including the cold start and the warm up period.

Reliable data from the oxygen sensor is not generally available during the first cycles of the cold start period until the sensor reaches its operating temperature. Given the importance of oxygen sensor data for AFR control, Cao and Li developed a novel closed loop control based on ionization current in a combustion cycle [29]. Since the engine's calibration is a quasi-static process with multidimensional searches across engine control variables, Keynejad and Manzie developed suboptimal control techniques on an engine model to decrease the search effort associated with calibration [30].

One approach for simplifying the controller structure, reducing the calculation load, and reducing the controller calibration effort is by utilizing methods of model order reduction shown in Figure 2. Three groups are identified in Figure 2. The first group includes *parameter space reduction* methods. These methods reduce the controller calibration load by identifying and removing insensitive parameters. The second group is *grey-box modeling* approach which combines physical (clear-box) models and empirical (black-box) models and tries to develop high fidelity models [31]. The third group includes *plant order/state reduction* methods. These methods attempt to evaluate the system's dynamics and eliminate fast and low energetic states of the system's dynamics. In order to keep the nonlinearity of the plant, nonlinear state/order reduction methods should be applied on the engine model. This study is focused on developing a nonlinear model order reduction method for cold start controller design.

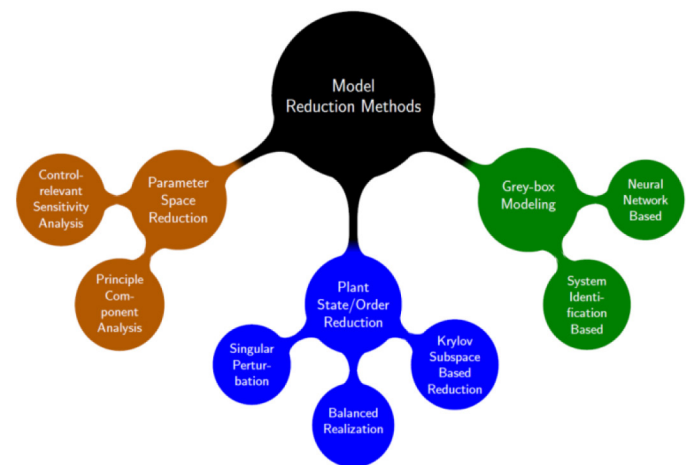


Figure 2. Overview of model reduction techniques for automotive control.

Plant state/order reductions methods for linear dynamical systems are described in detail in [32, 33, 34]. One well known linear reduction method is the singular perturbation approximation (SPA) method. This method divides system dynamics into fast and slow parts [34] and chooses the slow part as the important part of the system's dynamics.

In this paper, a new SPA method for the nonlinear cold start emission control problem is developed. The new SPA method helps to simplify the model by excluding unimportant system dynamics. In addition, as will be explained in the paper, the integration of the developed SPA method for the controller design provides a means to set priorities among different control inputs for reducing cold start emission. This is critical for cold start control since there are conflicting control inputs and it is challenging to set priorities among those different control inputs. For instance, delaying spark timing helps to increase the exhaust gas temperature for heating-up the catalyst, but it leads to higher engine-out HC emission. AFR control and idle speed control provide other effective control inputs for reducing cold start emission. The integration of SPA with controller design will help to set priorities among the control inputs for an optimum trade-off between reducing engine-out emission and reducing catalytic converter light-off time.

The contribution of this study is twofold. Firstly, we developed a novel nonlinear model order reduction method based on a singular perturbation approximation. Secondly, a reduced order nonlinear controller is designed using the developed SPA order reduction technique to control HC emissions from an SI engine during the cold start phase. Compared to designing a full order controller, our new approach reduces the calibration load in the controller design and shows better performance in the emission reduction by a more efficient adjustment of control inputs.

ENGINE COLD START MODEL

A control-oriented cold start emission model from [21] is used in this study for modeling and model-based controller design. The model captures major dynamics affecting engine-out HC emission and catalytic converter efficiency. The model is nonlinear with five states and three inputs, including air mass flow rate into the intake manifold ($\dot{m}_{ai}$), commanded mass flow rate of injected fuel into the cylinders ($\dot{m}_{fc}$), and spark timing (Δ) after top dead center (ATDC). The model is parameterized for a 2.4-Liter, 4-cylinder, DOHC 16 valve Toyota 2AZ-FE engine. Further modeling procedure are detailed in [35] and [36]. The states and equations describing the model are as follows:

$$x = \begin{bmatrix} m_a \\ \omega_e \\ \dot{m}_f \\ T_{cat} \\ T_{exh} \end{bmatrix}, \dot{x} = \begin{bmatrix} \dot{m}_a \\ \dot{\omega}_e \\ \dot{\dot{m}}_f \\ \dot{T}_{cat} \\ \dot{T}_{exh} \end{bmatrix} = \begin{bmatrix} \dot{m}_{ai} - \dot{m}_{ao} \\ \frac{1}{J}(\tau_{net} - \tau_L) \\ \frac{1}{\alpha_f}(\dot{m}_{fc} - \dot{m}_f) \\ \frac{1}{m_{cp}}(\dot{Q}_{gen} - \dot{Q}_{in} - \dot{Q}_{out}) \\ \frac{1}{\alpha_e}(ST \times AFI - T_{exh}) \end{bmatrix} \quad (1)$$

The air mass at intake manifold (m_a), the engine speed (ω_e), the mass flow rate of injected fuel into cylinders ($\dot{m}_f$), the catalyst temperature (T_{cat}) and the exhaust gas temperature (T_{exh}) are five system states. More details on the functions and the parameters in Eq. (1) are found in the Appendix. The HC production rate from the engine ($\dot{H}C_{eng}$) is calculated along with the catalytic converter efficiency (η_{cat}) to give the HC emission rate out of the tailpipe ($\dot{H}C_{tp}$) [4]:

$$\dot{H}C_{eng} = \frac{\dot{m}_f(r_c - 1)}{r_c} e^{-a\left(\frac{\theta_{EVO} - \theta_0}{\delta\theta}\right)^n} \quad (2)$$

$$\eta_{cat} = 0.98 \left(1 - \exp \left[-5 \left(\frac{AFR - 0.7}{AFR_{st} - 0.3} \right)^{15} \right] \right) \left(1 - \exp \left[-0.2 \left(\frac{T_{cat} - 30}{150} \right)^5 \right] \right) \quad (3)$$

$$\dot{H}C_{tp} = \dot{H}C_{eng}(1 - \eta_{cat}) \quad (4)$$

where, AFR_{st} indicates stoichiometric air fuel ratio, r_c the compression ratio. $\delta\theta$, θ_0 , θ_{EVO} are parameters to calculate the fuel burn rate as detailed in the Appendix. Eq.(1), Eq.(2), Eq. (3), Eq.(4) altogether shape the dynamical system equations in which the output of the model is the engine's HC production rate. The final output is the tailpipe HC flow rate. The difference between engine-out and tailpipe HC emission depends on the catalytic converter performance in the exhaust system.

It can be easily concluded from the model equations that among all five states, the intake manifold air mass and engine speed are affected only by air mass flow as the control input. Commanded fuel mass flow rate is affected by injected fuel mass rate and all three inputs impact the temperatures of the catalyst and exhaust gases. Although all three inputs and m_a , ω_e and $\dot{m}_f$ affect the $\dot{H}C_{eng}$ and η_{cat} , only the last two states (T_{cat} and T_{exh}) have influence on the catalytic converter efficiency. Consequently, the final outcome of this model, HC tailpipe flow rate, is a function of all states and inputs. Thus, in order to control and reduce the amount of tailpipe HC emission, all aspects of the engine cold start model should be considered.

The cold start model has been validated with experimental data from the same engine in the previous work [37]. A comparison between measured and simulated cumulative tailpipe HC emission is illustrated in Figure 3. It is observed that at the end of the simulation period (200 seconds), there is less than 2% error between measured and simulated results.

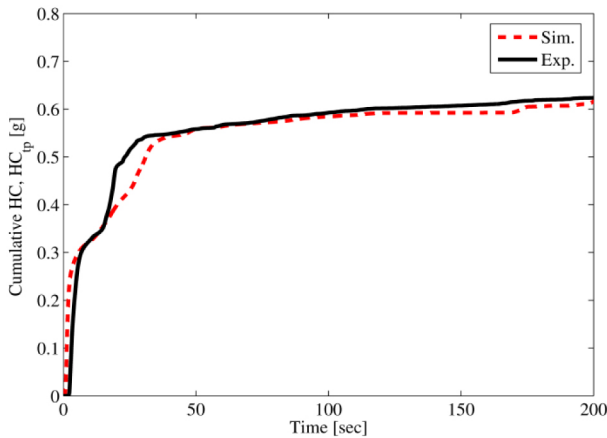


Figure 3. Comparison between measured and simulated cumulative tailpipe HC emission [37].

LINEAR SPA MODEL ORDER REDUCTION

The linear balanced direct truncation (DT) technique simplifies matrices of system dynamics with respect to differences among Henkel singular values of balanced realization form of a system [32]. Linear balanced realization transfers the space of physical states of the model into another linear space in which new non-physical states form a linear combination of physical states. Furthermore, in this new space, controllability and observability gramians are equal and diagonal. Since the magnitudes of singular values indicate the level of each state's energy, significant differences among these singular values can be considered as a proof of the state's dominance [32]. Consider a linear asymptotically stable with minimal realization system with state space equations as follows:

$$S = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbb{R}^{(n+p) \times (n+m)}, \begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \quad (5)$$

This system has a balanced realization where controllability gramian (P), observability gramian (Q) and the diagonal matrix consisting of Hankel singular values (HSV, Σ) are equal:

$$P = Q = \Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n), \quad \text{where: } \sigma_i \geq \sigma_{i+1}, i = 1, 2, \dots, n-1 \quad (6)$$

The balanced system in Eq.(1) and (2) can be divided into Eq. (7) subsystems ($n > r$):

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}_{n \times n}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = [C_1 \quad C_2], \quad \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \quad (7)$$

The subsystem $(A_{11}, B_1, C_1, \Sigma_1)_r$ could be a good approximation of the original model if there is a large difference between the HSVs ($\sigma_r \gg \sigma_{r+1}$) [32]. Singular Perturbation Approximation (SPA) method divides a system into fast and slow dynamics.

The criterion in SPA for separating the fast and the slow part of response is the difference between the HSVs [33]. Assume that the system of Eq.(5) has separate slow and fast dynamics, x_s and z respectively; This equation can be rearranged to:

$$\begin{cases} \dot{x}_s = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}_{n \times n} \begin{bmatrix} x_s \\ z \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u, \quad \begin{bmatrix} x_s(0) \\ z(0) \end{bmatrix} = \begin{bmatrix} x_{s0} \\ z_0 \end{bmatrix} \\ y = [C_1 \quad C_2] \begin{bmatrix} x_s \\ z \end{bmatrix} + Du, \quad x_s \in \mathbb{R}^r, \quad z \in \mathbb{R}^{n-r}, \quad r < n \end{cases} \quad (8)$$

If z indicates the fast part of the dynamics, it means that the effects of these states diminish soon and what would dictate the system response, is its slow part [34]. Since fast part loses its influence on the system response soon, its variation can be set to zero. If $\dot{z} = 0$, Eq.(8) can be simplified as follows:

$$\begin{cases} \dot{x}_s = \bar{A}x_s + \bar{B}u_s, \quad x_s(0) = x_{s0} \\ y_s = \bar{C}x_s + \bar{D}u_s \end{cases} \quad (9)$$

Where x_s means slow states and:

$$\begin{cases} \bar{A} = A_{11} - A_{12}A_{22}^{-1}A_{21} \\ \bar{B} = B_1 - A_{12}A_{22}^{-1}B_2 \\ \bar{C} = C_1 - C_2A_{22}^{-1}A_{21} \\ \bar{D} = D - C_2A_{22}^{-1}B_2 \end{cases} \quad (10)$$

The subsystem $(\bar{A}, \bar{B}, \bar{C}, \bar{D})$ could be a good approximation of original model if there is a large difference between the HSVs ($\sigma_r \gg \sigma_{r+1}$) [34].

The performance of SPA linear reduction methods is determined by testing the techniques on a linearized cold start model. At the first step, the nonlinear cold start emission model is linearized around an operation point. Linearization of the nonlinear model is done using the first order approximation of Taylor expansion with respect to a certain operation point before the catalyst light-off point. The resulting state space realization includes:

$$A = \begin{bmatrix} -16.6292 & -3.9948 & 1.4478 & 0.0120 & -0.0032 \\ -0.1929 & -4.0673 & 4.3134 & 0.0245 & -0.0028 \\ 0.1232 & -4.2955 & -1.6335 & -0.0275 & 0.0133 \\ -0.0125 & -0.0052 & 0.0001 & -0.0128 & -0.8406 \\ 0.0028 & 0.0007 & -0.0005 & 0.0058 & -23.8002 \end{bmatrix},$$

$$B = \begin{bmatrix} -3.1609 & 128.5742 & -0.0000 \\ 16.0822 & 2.1161 & 0.0000 \\ 5.8482 & -0.4477 & 0.0000 \\ 0.0039 & 0.0478 & -0.0000 \\ 0.0001 & -0.0106 & 0.0000 \end{bmatrix},$$

$$C = [135.4482 \quad 16.2628 \quad -5.9097 \quad -0.0503 \quad 0.0081],$$

$$D = [0 \quad 0 \quad 0.0051] \quad (11)$$

It becomes possible to separate more energetic and slower states by calculating the HSVs. Eq.(12) indicates the HSVs of the engine model. It can be seen that there is a large difference of several magnitude orders among the first three values and