Optimal Estimation of Gasoline LP-EGR via Unscented Kalman Filtering with Mixed Physics-based/Data-driven Components Modeling ⋆

Kwangmin Kim [*] **Jinsung Kim** [**] **Oheun Kwon** [**]
Se-Kyu Oh [**] **Yong-Wha Kim** [**] **Dongjun Lee** [*]

[*] *Department of Mechanical & Aerospace Engineering and IAMD,
Seoul National University, Seoul, Korea (e-mail:
{kkm0313,djlee}@snu.ac.kr).*
[**] *Hyundai-Kia Motors R&D Center, Gyeonggi-do, Korea (e-mail:
{jinsung.kim,oheuns,yongwhakim}@hyundai.com)*

Abstract: We propose a novel optimal estimation methodology for gasoline engine LP (low-pressure) EGR (exhaust gas recirculation) air-path system, which allows us to implement virtual sensors for oxygen mass fraction at the intake manifold and EGR mass flow rate at the LP-EGR valve, real sensors for them too expensive to deploy in production cars. We first decompose the LP-EGR air-path system into several sub-components; and opportunistically utilize physics-based modeling or data-driven modeling for each component depending on their model complexity. In particular, we apply the technique of MLP (multi-layer perceptron) as a means for data-driven modeling of LP-EGR/throttle valves and engine cylinder valve aspiration dynamics, all of which defy accurate physics-based modeling, that is also simple enough for real-time running. We further optimally combine these physics-based and data-driven modelings in the framework of UKF (unscented Kalman filtering), and also manifest via formal analysis that this mixed physics-based/data-driven modeling renders our estimator much faster to run as compared to the case of full data-driven MLP modeling. In doing so, we also extend the standard UKF theory to the more general case, where the system contains non-additive uncertainties both in the measurement and process models with cross-correlations and state-dependent variances, which stems from the inherent peculiar structure of our mixed physics-based/data-driven modeling approach, for the UKF formulation. Experiments are also performed to show the theory.

Keywords: Gasoline engine low-pressure exhaust gas recirculation (Gasoline LP-EGR), mixed physics-based/data-driven modeling, multi-layer perceptron (MLP), unscented Kalman filtering (UKF), non-additive cross-correlated measurement and process noises with state-dependent variacnes

1. INTRODUCTION

With the emission regulation for automotive vehicles becoming stringent worldwide, automotive engine technologies have been evolving to improve fuel economy. One of such technologies is the exhaust gas recirculation (EGR), which adds some inert and higher-specific-heat-ratio gas to the fresh air, thus, reducing in-cylinder combustion temperature. This EGR technology has first been researched and commercialized on diesel engines for NOx reduction purpose (Grondin et al., 2009), and later finds its way into gasoline engines to suppress knocking tendency for better fuel economy (Kiwan et al., 2017).

Two representative configurations of this EGR are: 1) high-pressure (HP) EGR, where the exhaust gas from the exhaust manifold is directly channeled to the intake mani-

fold (Kolmanovsky et al., 2000); and 2) low-pressure (LP) EGR, where the gas from the turbine downstream is fed into the compressor upstream through some intermediate components (Wiese et al., 2017). See Fig. 1 for the configuration of LP-EGR system we consider in this paper. Since the temperature of EGR to the intake manifold is much lower, the LP-EGR is typically considered substantially more effective than the HP-EGR.

It is however challenging to accurately control the amount of EGR to the intake manifold for this LP-EGR system. This is because the pressure difference at the EGR valve P_{egrv} (See Fig. 1) is typically fairly low in the LP-EGR system, making it impossible to accurately model the EGR flow rate with real-time runnable model complexity, e.g., orifice equation (Kiwan et al., 2017). This then means that, even if the aspirated fresh air flow rate W_a can be correctly measured with HFM (hot-film air-mass) sensor, since the EGR flow rate W_{egrv} can neither be
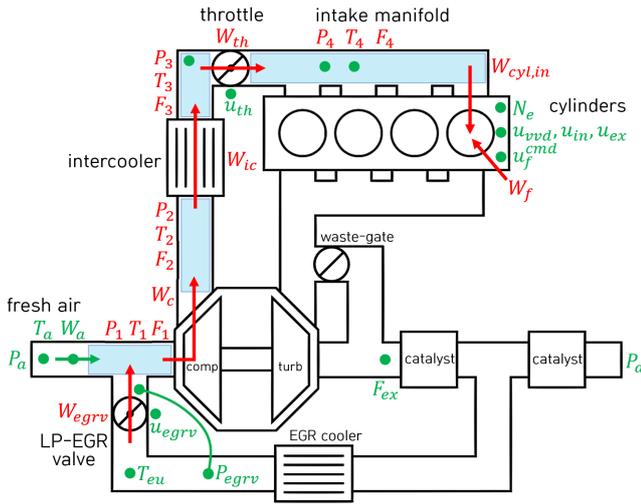
Fig. 1. Gasoline LP-EGR airpath system considered in this paper with sensors (green dots), intermediate volumes (blue boxes) and model variables (red letters).

modeled nor measured accurately as stated above, it can never be accurately known or controlled either. Feedback control can be adopted to enforce robustness against such model inaccuracy. This feedback control however requires the values of output variables (i.e. oxygen content at the intake manifold F_4 and EGR flow rate at the LP-EGR valve W_{egrv} in this paper) as well as other state variables. Adding sensors to directly real-time measure these variables in production cars is generally infeasible in terms of cost, robustness and form-factor (e.g. oxygen concentration sensor at intake manifold).

In this paper, we propose a novel optimal state estimation framework for this LP-EGR system. We focus on the system with the configuration as shown in Fig. 1, although the framework presented here can be applied to other configurations or other systems as well. We first decompose the LP-EGR system into sub-components as in Fig. 1, and opportunistically apply physics-based or data-driven modeling for each component. More precisely, we model all the "volumes" in Fig. 1 using their physics, since they possess rather simple dynamics, allowing for simple and accurate physics-based modeling. On the other hand, for the LP-EGR vale and the throttle valves, which do not subsume such simple/accurate physics-based models due to the low pressure drop as stated above, we apply data-driven modeling. For this, we in particular utilize MLP (multi-layer perceptron (Goodfellow et al., 2016; Shi et al., 2019; Moriyasu et al., 2018; Bottelli et al., 2013)), which is known to be able to approximate fairly rich class of functions and whose learning also known to converge well. We further apply this MLP modeling to the cylinder aspiration flow, whose accurate and real-time runnable physics-based modeling is virtually impossible with its dynamics depending on phenomena too difficult to accurately capture with simple and real-time runnable model (e.g. flow rate into cylinder near valve opening/closing times (Andersson and Eriksson, 2004; Kocher et al., 2012)).

We then optimally combine these physics-based and data-driven models along with the sensor measurements by fusing their means and uncertainties in the framework of

UKF (unscented Kalman filtering (Julier et al., 2000)). For this, we extend the standard theory of UKF to the case of mixed physics-based/data-driven models in general, which inherently has a peculiar uncertainty structure for UKF design: 1) the data-driven modeling uncertainties enter into the nonlinear measurement function as well as into the nonlinear state propagation function at the same epoch (i.e. non-addivite cross-correlated measurement and process noises), and 2) the variances of these data-driven modeling uncertainties further depend on the states. Both of these properties are not addressed in the previous works of UKF theories (Julier and Uhlmann, 2004; Chang, 2014; Yu et al., 2016; Tronarp et al., 2016), yet, our proposed UKF theory applies to the more general type of models with such properties, including the mixed physics-based/data-driven modeling approach in this paper. We further perform a formal analysis to manifest that the use of our mixed physics-based/data-driven modeling with our UKF algorithm designed for its peculiar uncertainty structure renders the estimation run much faster than the use of the purely data-driven modeling with the standard UKF, which is a good property for real-world implementations typically constrained by ECU (electronic control unit) limitations (e.g. industrial control applications such as production cars).

For related works, Chen and Wang (2013); Castillo et al. (2013); Grondin et al. (2009) developed state estimation algorithms for oxygen concentration or EGR mass fraction in diesel LP-EGR systems using Luenberger observer or Kalman filter based on physics-based models, which are limited on diesel engines because gasoline LP-EGR typically has much lower pressure difference (thus inaccurate) than diesel LP-EGR. Moriyasu et al. (2018) modeled the diesel HP-EGR system by using a pure single large MLP and applied standard UKF, yet, cannot work for real-time estimation purpose due to the high computational load of their large size MLP that is neither affordable to their ECU nor to their PC. Moriyasu et al. (2018) also conducted their work based on the ideal data from a simulation model and did not take account the modeling uncertainty of MLP into the estimation algorithm. Kaufmann et al. (2019) takes account the pure MLP model and its modeling error as a measurement and its variance in the EKF, whose estimation structure is standard and much simpler than for our proposed mixed physics/MLP-based model. Note that, the theoretical backgrounds and the frameworks to combine the MLP-based modeling with the traditional estimation/control methods are also unmature until nowadays, e.g., the stability analysis and stable control design are only recently proposed in Shi et al. (2019).

The rest of the paper is organized as follows. We present the mixed physics-based/data-driven modeling in Sec. 2. We then extend the standard UKF theory to address the peculiar uncertainty structure stemming from our problem along with a formal analysis on its computation load in Sec. 3. Experimental results to validate our proposed estimation algorithm are presented in Sec. 4. We summarize the paper and provide the concluding remarks in Sec. 5.

2. MIXED PHYSICS-BASED/DATA-DRIVEN MODELING

For the modeling of the gasoline LP-EGR system in Fig. 1, we divide the system into the four volumes (i.e. the compressor upstream, the compressor downstream, the intercooler downstream, and the intake manifold), whose physics-based modelings are easy and accurate, and the components in between or at the end of the volumes (i.e. the LP-EGR valve, the compressor, the intercooler, the throttle valve, the cylinder, and the fuel injector), whose physics-based modelings are often relatively difficult or inaccurate. The random variables that we want to estimate are the oxygen concentrations F_i, the temperatures T_i and the pressures P_i of the four volumes, and the mass flow rate at the LP-EGR valve W_{egrv}, where the subscripts $i = 1, 2, 3, 4$ denote the compressor upstream, the compressor downstream, the intercooler downstream and the intake manifold, respectively.

2.1 Physics-based Modeling of Volume Dynamics

With the standard physics laws (i.e. the ideal gas law, the mass conservation law, the energy conservation law, and the perfect mixing) assumed on each volumes, we have the standard form of volume dynamics model that are widely used in the automotive control fields (Grondin et al., 2009):

$$\dot{P}_1 = \frac{R\gamma}{V_1}(W_a T_a + W_{egrv} T_{eu} - W_c T_1) \tag{1}$$

$$\dot{T}_1 = \frac{RT_1}{P_1 V_1}[W_a(\gamma T_a - T_1) + W_{egrv}(\gamma T_{eu} - T_1)$$
$$+ W_c(\gamma - 1)T_1] \tag{2}$$

$$\dot{F}_1 = \frac{RT_1}{P_1 V_1}[W_a(F_a - F_1) + W_{egrv}(F_{eu} - F_1)] \tag{3}$$

$$\dot{P}_2 = \frac{R\gamma}{V_2}(W_c T_{co} - W_{ic} T_2) \tag{4}$$

$$\dot{T}_2 = \frac{RT_2}{P_2 V_2}[W_c(\gamma T_{co} - T_2) - W_{ic}(\gamma - 1)T_2] \tag{5}$$

$$\dot{F}_2 = \frac{RT_2}{P_2 V_2}W_c(F_1 - F_2) \tag{6}$$

$$\dot{P}_3 = \frac{R\gamma}{V_3}(T_{io}W_{ic} - T_3 W_{th}) \tag{7}$$

$$\dot{T}_3 = \frac{RT_3}{P_3 V_3}[W_{ic}(\gamma T_{io} - T_3) - W_{th}(\gamma - 1)T_3] \tag{8}$$

$$\dot{F}_3 = \frac{RT_3}{P_3 V_3}W_{ic}(F_2 - F_3) \tag{9}$$

$$\dot{P}_4 = \frac{R\gamma}{V_4}(W_{th}T_3 - W_{cyl,in}T_4) \tag{10}$$

$$\dot{T}_4 = \frac{RT_4}{P_4 V_4}[W_{th}(\gamma T_3 - T_4) - W_{cyl,in}(\gamma - 1)T_4] \tag{11}$$

$$\dot{F}_4 = \frac{RT_4}{P_4 V_4}W_{th}(F_3 - F_4) \tag{12}$$

where V_i $(i = 1, 2, 3, 4)$ are the volumes; R is the gas constant of the air; γ is the specific heat ratio of the air; W_a is the fresh air mass flow rate measured from the HFM sensor; T_a and T_{eu} are the temperatures of the fresh air and the LP-EGR gas, respectively, measured from each sensors; F_a and F_{eu} are the oxygen concentrations of the fresh air and the LP-EGR gas that are assumed to be constant and zero, respectively, in this paper; W_{egrv}, W_c, W_{ic}, W_{th} and $W_{cyl,in}$ are the mass flow rates through the LP-EGR vavle, the compressor, the intercooler, the throttle valve, and the cylinder inlets, respectively, which are random variables; T_{co} and T_{io} are the gas temperatures at the outlets of the compressor and the intercooler, respectively, which are random variables.

With the additional standard assumption that the compressor upstream is almost completely opened to the atmosphere through the fresh air inlet (i.e. $P_1 \simeq P_a$, thus, $\dot{P}_1 \simeq 0$), which then applied to (1), we have the relations between the compressor mass flow rate and the other variables:

$$W_c = \frac{W_a T_a + W_{egrv} T_{eu}}{T_1} \tag{13}$$

Similarly, with the assumption of almost no pressure loss at the intercooler (i.e. $P_2 \simeq P_3$, thus, $\dot{P}_2 \simeq \dot{P}_3$), which then applied to (4) and (7), we have the relatioins between the intercooler mass flow rate and the other variables:

$$W_{ic} = \frac{W_c \frac{T_{co}}{V_2} + W_{th}\frac{T_3}{V_3}}{\frac{T_2}{V_2} + \frac{T_{io}}{V_3}} \tag{14}$$

We have the standard model of the intercooler outlet gas temperature:

$$T_{io} = T_2 - \epsilon_{ic}(T_2 - T_{cool}) \tag{15}$$

where T_{cool} is the coolant temperature (e.g. $T_{cool} = T_a$ in case of air cooling); ϵ_{ic} is the cooling efficiency that is assumed to be constant in this paper for the simplicity of the writing, while can also be easily incorporated as a bias state for estimation as in the experiment in Sec. 4.

In addition, with the standard assumption of stoichiometric reaction between the oxygen and the fuel in the cylinder, we have the standard relation

$$F_{ex} = \frac{W_{cyl,in}F_4 - \text{OFR}_s \cdot W_f}{W_{cyl,in} + W_f} \tag{16}$$

where F_{ex} is the oxygen concentration of the exhaust gas measured from the lambda sensor; W_f is the the mass flow rate of the fuel into the cylinder; OFR_s is the stoichiometric reaction mass ratio between the oxygen and the gasoline fuel.

The propagation equations of the random variables that we want to estimate are then summarized by (2)-(6) and (8)-(12) with P_1 and P_3 replaced by P_a and P_2, respectively, and with W_c, W_{ic} and T_{io} in the equations replaced by (13), (14) and (15), respectively. This propagation model can then be presented by:

$$x[k] = f(x[t], y_p[t] + \tilde{y}_p[t], m[t]) \tag{17}$$

where k is the time step number; $t = k - 1$ is the prior step of k;

$$x := [T_1, T_2, T_3, T_4, P_2, P_4, F_1, F_2, F_3, F_4]^T$$

is the state random variable;

$$y_p := [W_a, T_a, T_{eu}]^T$$

is the measurement of the boundary conditions for this physics-based model; $\tilde{y}_p[t]$ is the uncertainty in this measurement of boundary conditions;

$$m := [W_{egrv}, W_{th}, W_{cyl,in}, T_{co}, W_f]^T \tag{18}$$

Forward propagation
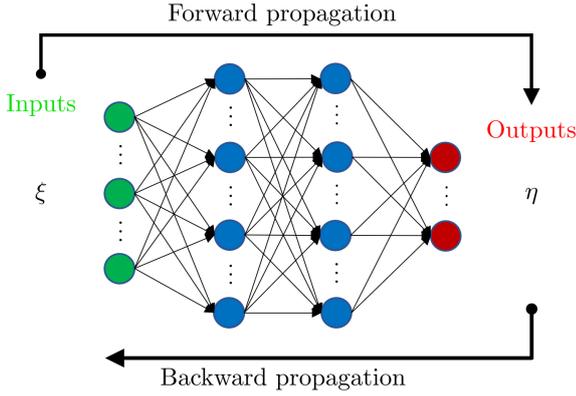
Fig. 2. Typical structure of the MLP (multi-layer perceptron).

is the random variable that are not yet modeled in this Sec. 2.1 because their physics-based models are typically inaccurate (e.g. W_{egrv}) or difficult (e.g. $W_{cyl,in}$).

2.2 Data-Driven Component Modeling via MLP

For the data-driven modeling of each random variables in m (18), we apply the technique of MLP, which is a typical forward neural network (–see Fig. 2) that calculates its output from its input by forward propagation based on weights and nonlinear activation functions s.t.

$$\eta_0 = \xi \tag{19}$$
$$\eta_l = \phi(\Omega_l \eta_{l-1}) \quad \text{for} \quad l = 1, \cdots, L \tag{20}$$
$$\eta_{L+1} = \Omega_{L+1} \eta_L \tag{21}$$

where ξ is the input vector to the MLP; $\eta_l \in \Re^{n_l}$ is the output vector of the l-th layer; n_l is the number of neurons in the l-th layer; $\Omega_l \in \Re^{n_l \times n_{l-1}}$ is the weight matrix multiplied to η_{l-1} to calculate the input vector to the l-th layer; $\phi(\cdot)$ is the activation function in each hidden neuron; L is the total number of the hidden layers between the input and the output layers of the MLP; η_{L+1} is the output vector of the MLP. From (19)-(21), we have the explicit function between the input/output of the MLP:

$$\eta_{L+1} = \Omega_{L+1} \phi(\Omega_L \phi(\cdots \phi(\Omega_1 \xi) \cdots)) =: \Phi(\theta, \xi) \tag{22}$$

that is parameterized by the weights $\theta := \{\Omega_1, \cdots, \Omega_{L+1}\}$, and this parameter θ is trained offline to fit the data set of the input/output pair $\{\xi^{\text{data}}, \eta^{\text{data}}\}$, typically by minimizing the entire loss between the function and the data set s.t.

$$\theta^* = \operatorname{argmin} \mathcal{L}\left(\Phi\left(\theta, \{\xi^{\text{data}}\}\right), \{\eta^{\text{data}}\}\right) \tag{23}$$

This minimization typically requires the gradients of the cost function w.r.t. the weights (e.g. gradient decent) that are calculated offline by backward propagation s.t.

$$\frac{\partial \mathcal{L}}{\partial \eta_l} = \frac{\partial \mathcal{L}}{\partial \eta_{l+1}} \frac{\partial \eta_{l+1}}{\partial \eta_l} = \frac{\partial \mathcal{L}}{\partial \eta_{l+1}} \begin{bmatrix} \phi'\left(\Omega_{l+1}^1 \eta_l\right) \Omega_{l+1}^1 \\ \vdots \\ \phi'\left(\Omega_{l+1}^{n_{l+1}} \eta_l\right) \Omega_{l+1}^{n_{l+1}} \end{bmatrix} \tag{24}$$

$$\frac{\partial \mathcal{L}}{\partial \Omega_{l+1}^{ij}} = \frac{\partial \mathcal{L}}{\partial \eta_{l+1}} \frac{\partial \eta_{l+1}}{\partial \Omega_{l+1}^{ij}} = \frac{\partial \mathcal{L}}{\partial \eta_{l+1}^i} \phi'\left(\Omega_{l+1}^i \eta_l\right) \eta_l^j \tag{25}$$

for $l = L, \cdots, 0$, where Ω_l^i is the i-th row of Ω_l; Ω_{l+1}^{ij} is the j-th element of Ω_{l+1}^i; η_l^j is the j-th element of η_l. Note that, the bias weights are also included while omitted in (19)-(25) for simple explanation.

(a) Mass flow rate at the LP-EGR valve   (b) Mass flow rate into the cylinders
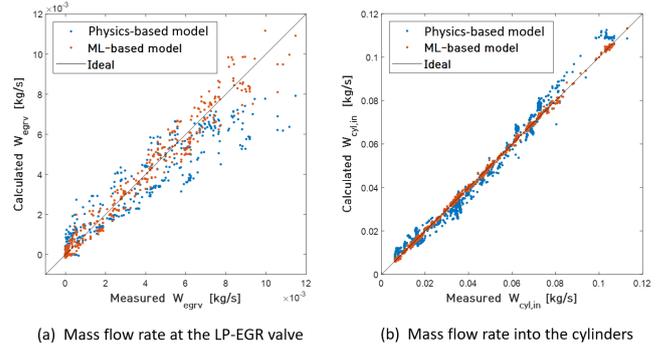
Fig. 3. The modeling errors of the MLP-based models (red) and the physics-based models (blue).

For the MLP-based modeling of W_{egrv} in m (18), we first choose its input vector as $\xi_{W_{egrv}} = [u_{egrv}, P_{egrv}, T_{eu}]^T$, which is physically reasonable to describe W_{egrv} (–see Fig. 1), where u_{egrv} is the EGR valve opening angle measured from its encoder; P_{egrv} is the differential pressure across the LP-EGR valve measured from the differential pressure sensor. We do not incorporate F_{eu} because its value is almost constant (i.e. 0) and its incorporation does not clearly improve the model in our experience. We then have the data set of the input/output pair $\{\xi_{W_{egrv}}^{\text{data}}, W_{egrv}^{\text{data}}\}$ for modeling. We choose the size of the MLP and train the weights by standard procedures, e.g., in our experiment in Sec. 4 we simply use the neural fitting toolbox in MATLAB, where the size of the MLP is one-hidden layer (i.e. $L = 1$), and the number of the hidden neurons n_1 is, by our experience, chosen to minimize the test set error to avoid overfitting. We then have the MLP-based model of W_{egrv}:

$$W_{egrv}^{MLP}(\xi_{W_{egrv}}) := \Phi(\theta_{W_{egrv}}^*, \xi_{W_{egrv}}) \tag{26}$$

Note that, we have the fitting errors (–see Fig. 3(a)), i.e., the difference between the data and the trained MLP-based model:

$$\tilde{W}_{egrv}^{MLP,\text{fit}} := W_{egrv}^{\text{data}} - W_{egrv}^{MLP}(\xi_{W_{egrv}}^{\text{data}}) \tag{27}$$

and similarly we can define the modeling error, i.e., the difference between the real and the trained MLP-based model:

$$\tilde{W}_{egrv}^{MLP} := W_{egrv} - W_{egrv}^{MLP}(\xi_{W_{egrv}}) \tag{28}$$

Because our data set (i.e. steady-state engine calibration data that is time-averaged on each operation points) does not contain noise, we assume each data represents the real value, i.e., $(\xi_{W_{egrv}}^{\text{data}}, W_{egrv}^{\text{data}}) = (\xi_{W_{egrv}}, W_{egrv})$, thus, the fitting error represents the modeling error from (27)-(28), i.e., $\tilde{W}_{egrv}^{MLP} \simeq \tilde{W}_{egrv}^{MLP,\text{fit}}$. By noting this, we can see in Fig. 3(a) that the modeling errors are clearly smaller and their distribution is more unbiased in the data-driven MLP-based model than in the typical physics-based model (i.e. orifice equation (Grondin et al., 2009)), which is yet, often large relative to the real value.

Such remaining modeling error in the data-driven model can be probabilistically considered in the optimal estimation stage in Sec. 3, by seeing the error as a random uncertainty with some modeled variance $Q_{W_{egrv}}$ s.t.

$$\tilde{W}_{egrv}^{MLP} \sim (0, Q_{W_{egrv}})$$

For this, we extract the variance $Q_{W_{egrv}}$ from the fitting error (i.e. modeling error) data $\{\tilde{W}_{egrv}^{MLP,\text{fit}}\}$, and this again can be modeled as a function of the other variable, e.g., $Q_{W_{egrv}}(\xi_{W_{egrv}})$ (Khosravi et al., 2011). Here, we rather simply model this variance by the function of a scalar $W_{egrv}^{MLP}(\xi_{W_{egrv}}) \in \Re^1$ to avoid a lack of data in a large input dimension, thus, have:

$$\tilde{W}_{egrv}^{MLP} \sim \left(0, Q_{W_{egrv}}\left(W_{egrv}^{MLP}(\xi_{W_{egrv}})\right)\right) \qquad (29)$$

We then have the propagation equation between random variables (i.e. the input vector $\xi_{W_{egrv}}$, the output W_{egrv}, and the model uncertainty $\tilde{W}_{egrv}^{MLP}$):

$$W_{egrv} = W_{egrv}^{MLP}(\xi_{W_{egrv}}) + \tilde{W}_{egrv}^{MLP} \qquad (30)$$

Similar to the procedure of (26), (29) and (30) for W_{egrv}, we obtain the propagation equations for the other random variables in m (18), respectively:

$$W_{th} = W_{th}^{MLP}(\xi_{W_{th}}) + \tilde{W}_{th}^{MLP} \qquad (31)$$

$$W_{cyl,in} = W_{cyl,in}^{MLP}(\xi_{W_{cyl,in}}) + \tilde{W}_{cyl,in}^{MLP} \qquad (32)$$

$$T_{co} = T_{co}^{MLP}(\xi_{T_{co}}) + \tilde{T}_{co}^{MLP} \qquad (33)$$

$$W_f = W_f^{MLP}(\xi_{W_f}) + \tilde{W}_f^{MLP} \qquad (34)$$

with

$$\tilde{W}_{th}^{MLP} \sim \left(0, Q_{W_{th}}\left(W_{th}^{MLP}(\xi_{W_{th}})\right)\right) \qquad (35)$$

$$\tilde{W}_{cyl,in}^{MLP} \sim \left(0, Q_{W_{cyl,in}}\left(W_{cyl,in}^{MLP}(\xi_{W_{cyl,in}})\right)\right) \qquad (36)$$

$$\tilde{T}_{co}^{MLP} \sim \left(0, Q_{T_{co}}\left(T_{co}^{MLP}(\xi_{T_{co}})\right)\right) \qquad (37)$$

$$\tilde{W}_f^{MLP} \sim \left(0, Q_{W_f}\left(W_f^{MLP}(\xi_{W_f})\right)\right) \qquad (38)$$

where $W_{th}^{MLP}(\cdot)$, $W_{cyl,in}^{MLP}(\cdot)$, $T_{co}^{MLP}(\cdot)$ and $W_f^{MLP}(\cdot)$ are the trained MLP-based models of each; $Q_{W_{th}}(\cdot)$, $Q_{W_{cyl,in}}(\cdot)$, $Q_{T_{co}}(\cdot)$ and $Q_{W_f}(\cdot)$ are the variances of the model uncertainties of each; the input vector of each are chosen to be physically reasonable to describe each output:

$$\xi_{W_{th}} := [u_{th}, P_2, P_4, T_3, F_3]^T \qquad (39)$$

$$\xi_{W_{cyl,in}} := [P_4, N_e, u_{vvd}, u_{ex}, u_{in}]^T \qquad (40)$$

$$\xi_{T_{co}} := [W_c, P_2, T_1]^T \qquad (41)$$

$$\xi_{W_f} := u_f^{cmd} \qquad (42)$$

where u_{th} is the throttle valve opening angle measured from its encoder; N_e is the rpm of the engine measured from the crankshaft angle hall sensor; u_{vvd} is the variable valve duration input of the cylinder; u_{ex} and u_{in} are the cylinder exhaust and intake ports cam angles measured from their encoders, respectively; u_f^{cmd} is the fuel volume flow rate command.

Note that, the effects of the cylinder valves $[u_{vvd}, u_{ex}, u_{in}]^T$ in (40) to $W_{cyl,in}$ are too complicated to model based on some physics laws. Similar to the LP-EGR valve, the accuracy of this MLP-based cylinder flow model is better than that of the typical physics-based model (i.e. speed-density model (Grondin et al., 2009)) as compared in Fig. 3(b).

The propagation models (30)-(34) can then be presented by:

$$m[k] = g(x[k], y_m[k] + \tilde{y}_m[k]) + \tilde{m}[k] \qquad (43)$$

where

$$y_m := [u_{egrv}, P_{egrv}, T_{eu}, u_{th}, N_e, u_{vvd}, u_{ex}, u_{in}, u_f^{cmd}]^T$$

is the measurement of the boundary conditions for the MLP-based models; $\tilde{y}_m[k]$ is the uncertainty in this measurement of the boundary condition;

$$g(x[k], y_m[k] + \tilde{y}_m[k]) := \begin{bmatrix} W_{egrv}^{MLP}(\xi_{W_{egrv}}) \\ W_{th}^{MLP}(\xi_{W_{th}}) \\ W_{cyl,in}^{MLP}(\xi_{W_{cyl,in}}) \\ T_{co}^{MLP}(\xi_{T_{co}}) \\ W_f^{MLP}(\xi_{W_f}) \end{bmatrix} \qquad (44)$$

is simply a bunch of the MLP-based models;

$$\tilde{m}[k] := \left[\tilde{W}_{egrv}[k], \tilde{W}_{th}[k], \tilde{W}_{cyl,in}[k], \tilde{T}_{co}[k], \tilde{W}_f[k]\right]^T \qquad (45)$$

is the uncertainty vector specifying the modeling error. From (29) and (35)-(38), the model uncertainty vector can be characterized by:

$$\tilde{m}[k] \sim (0, Q_{\tilde{m}}(g(x[k], y_m[k] + \tilde{y}_m[k]))) \qquad (46)$$

where $Q_{\tilde{m}} := diag\left(Q_{W_{egrv}}, Q_{W_{th}}, Q_{W_{cyl,in}}, Q_{T_{co}}, Q_{W_f}\right)$.

2.3 Mixed Physics/MLP-Based Model

We obtain the state propagation model of the entire system, by substituting the MLP-based model (43) obtained in the Sec. 2.2 into the physics-based state dynamics model (17) obtained in the Sec. 2.1:

$$x[k] = F\left(x[t], y[t] + \tilde{y}[t], \tilde{m}[t]\right) \qquad (47)$$

where

$$y := [W_a, T_a, T_{eu}, u_{egrv}, P_{egrv}, u_{th}, N_e, u_{vvd}, u_{ex}, u_{in}, u_f^{cmd}]^T$$

is the measurement of the boundary conditions of the entire system (i.e. the union of y_p and y_m); $\tilde{y}[t]$ is the uncertainty in this measurement of the boundary condition.

With the measurements that contain the information of the state random variables:

$$z := [P_2, P_4, T_4, F_{ex}]^T$$

we obtain the measurement model, by substituting the MLP-based model (43), thus (32) and (34), into the physics-based model (16):

$$z[k] = H(x[k], y[k] + \tilde{y}[k], \tilde{m}[k]) + \tilde{z}[k] \qquad (48)$$

where $\tilde{z}[k]$ is the measurement noise of $z[k]$.

We checked that the state random variable x is observable from this state propagation/measurement model (47)-(48) when they are linearized w.r.t. x.

Remark 1 The forward propagation of MLP (19)-(21) implies that the network output calculation is typically fast as long as the number of neurons is small. Furthermore, for the weight training (23) based on the backward propagation of the gradients (24)-(25), $\phi(\cdot)$ is typically required to be gradient-smooth, i.e., $\phi'(\cdot) \in \mathcal{C}^1$, implying that the gradient of the network outputs w.r.t. the network inputs is smooth, i.e., $\frac{\partial \Phi}{\partial \xi}(\cdot, \xi) \in \mathcal{C}^1$. These structural characteristics of typical MLP means that the state model including such MLPs can be used: 1) for real-time estimation/control purpose as long as the size of each MLP is sufficiently small, and 2) for the linearization-based estimation/control methods (e.g. with $\frac{\partial F}{\partial x} \in \mathcal{C}^1$ in EKF/LQR) as long as $\phi'(\cdot) \in \mathcal{C}^1$.

3. UKF DESIGN FOR MIXED PHYSICS-BASED/DATA-DRIVEN MODEL

For the optimal estimation of the state $x[k]$ from many different information (i.e. the sensors $y[k]$ and $z[k]$ and the nonlinear models (47)-(48)) that contain uncertainties as well (i.e. the measurement noises $\tilde{y}[k]$ and $\tilde{z}[k]$ and the modeling error $\tilde{m}[k]$), we adopt UKF, which utilizes σ-points (set of finite number of points encoding the state mean and covariance) to capture the nonlinear distortion of the state probability distribution through nonlinear functions that even exacerbates in the experiment, and then similar to KF (Kalman Filtering) obtain the *maximum a posteriori* optimal state estimate given the current state measurement and based on the variances of the uncertainties.

3.1 Analysis & Consideration of Uncertainties into UKF

Recall that the uncertainties $\tilde{m}[t]$ and $\tilde{y}[t]$ enter into the nonlinear state propagation function at time t (47) as well as $\tilde{m}[k]$ and $\tilde{y}[k]$ into the nonlinear measurement function at time k (48) (i.e. non-additive process and measurement noises). In order to consider the nonlinear distortions of the probability distributions of these uncertainties as well as the state, we define the augmented state vector containing all of them similar to (Julier and Uhlmann, 2004):

$$x_a[t] := [x[t]^T, \tilde{y}[t]^T, \tilde{m}[t]^T, \tilde{y}[k]^T, \tilde{m}[k]^T]^T \quad (49)$$

In our mixed physics-based/data-driven modeling, state $x[t]$ and uncertainties $\tilde{y}[t]$ and $\tilde{m}[t]$ are cross-correlated given the measurement $z[t]$ because $z[t]$ contains their information as in (48) at time t, and these same $x[t]$, $\tilde{y}[t]$ and $\tilde{m}[t]$ enter into the nonlinear state propagation at time t (47), which implies the σ-points that are to be propagated through (47) to predict the next step state $x[k]$ should reflect their cross-correlations given $z[t]$. For this purpose, we define:

$$x_b[t] = [x[t]^T, \tilde{y}[t]^T, \tilde{m}[t]^T]^T \quad (50)$$

whose probability distributions should be updated together by $z[t]$, having:

$$\hat{x}_a[t|t] := \mathrm{E}\left[\begin{bmatrix} x_b[t] \\ \tilde{y}[k] \\ \tilde{m}[k] \end{bmatrix} \middle| z[t], \cdots, z[1] \right] = \begin{bmatrix} \hat{x}_b[t|t] \\ 0 \\ 0 \end{bmatrix}$$

$$P_a[t|t] := \mathrm{E}\left[\begin{bmatrix} x_b[t] - \hat{x}_b[t|t] \\ \tilde{y}[k] \\ \tilde{m}[k] \end{bmatrix} [\cdot]^T \middle| z[t], \cdots, z[1] \right] \quad (51)$$

$$= \begin{bmatrix} P_b[t|t] & 0 & 0 \\ 0 & Q_{\tilde{y}} & 0 \\ 0 & 0 & \hat{Q}_{\tilde{m}}[k|t] \end{bmatrix} \quad (52)$$

where

$$\hat{Q}_{\tilde{m}}[k|t] := \mathrm{E}\left[Q_{\tilde{m}}(g(x[k], y_m[k] + \tilde{y}_m[k])) | z[t], \cdots, z[1] \right] \quad (53)$$

The standard generation of $2N + 1$ σ-points (Julier et al., 2000) for the mean $\hat{x}_a[t|t]$ and the covariance $P_a[t|t]$ has the form of:

$$\mathcal{X}_a^{0,i,i+N}[t|t] = \hat{x}_a[t|t] \ , \ \hat{x}_a[t|t] \pm \gamma \left(\sqrt{P_a[t|t]} \right)_i$$

$$= \begin{bmatrix} \hat{x}_b[t|t] \\ 0 \\ 0 \end{bmatrix} ,$$

$$\begin{bmatrix} \hat{x}_b[t|t] \\ 0 \\ 0 \end{bmatrix} \pm \gamma \begin{bmatrix} \sqrt{P_b[t|t]} & 0 & 0 \\ 0 & \sqrt{Q_{\tilde{y}}} & 0 \\ 0 & 0 & \sqrt{\hat{Q}_{\tilde{m}}[k|t]} \end{bmatrix}_i$$

$$=: \begin{bmatrix} \mathcal{X}_b^i[t|t] \\ \mathcal{X}_{\tilde{y}}^i[k] \\ \mathcal{X}_{\tilde{m}}^i[k] \end{bmatrix} \quad (54)$$

where $\gamma = \alpha\sqrt{N + \kappa}$ and $i = 1, \cdots, N$; $N := dim(x_a)$; α is the scaling parameter for the higher-order (≥ 3) error terms in UKF; $\kappa = 3 - N$ is the tuning parameter for partially matching the 4th order moment for, e.g., Gaussian distribution. The corresponding weights for the σ-points are:

$$\mathcal{W}_m^0 = \frac{\lambda}{N + \lambda} \ , \quad \mathcal{W}_c^0 = \frac{\lambda}{N + \lambda} + (1 - \alpha^2 + \beta)$$

$$\mathcal{W}_m^i = \mathcal{W}_c^i = \frac{1}{2(N + \lambda)} \quad \text{for} \ i = 1, 2, \cdots, 2N$$

where $\lambda = \alpha^2(N + \kappa) - N$.

Note that, if the variances of the uncertainties are not state-dependent, e.g., $Q_{\tilde{m}}(\cdot)$ is modeled as a constant, the $2N + 1$ σ-points are straightforwardly computed by (54) given $\hat{x}_b[t|t]$, $P_b[t|t]$, $Q_{\tilde{y}}$ and $Q_{\tilde{m}}$, yet, our mixed physics-based/data-driven modeling has the state-dependent variances of modeling uncertainties, which implies we cannot compute the entire equation of (54) at once before we obtain $\hat{Q}_{\tilde{m}}[k|t]$. However, thanks to the causality of (47) and (46) that $x[k]$ is not a function of the uncertainties (e.g. $\tilde{m}[k]$) whose variances are functions of $x[k]$, we only need to compute $\mathcal{X}_b^i[t|t]$, i.e., the first $dim(x_b)$ rows of (54), which is computable without $\hat{Q}_{\tilde{m}}[k|t]$, to predict the next step state $x[k]$ by (47).

3.2 UKF Algorithm Procedure to Estimate the State

Given $\hat{x}_b[t|t]$ and $P_b[t|t]$, we first calculate $\mathcal{X}_b^i[t|t]$ for $i = 0, \cdots, 2N$ in (54), and then compute:

$$\mathcal{X}_x^i[k|t] = F\left(\mathcal{X}_x^i[t|t], y[t] + \mathcal{X}_{\tilde{y}}^i[t|t], \mathcal{X}_{\tilde{m}}^i[t|t] \right) \quad (55)$$

where $\mathcal{X}_x^i[t|t]$, $\mathcal{X}_{\tilde{y}}^i[t|t]$, $\mathcal{X}_{\tilde{m}}^i[t|t]$ are the elements of $\mathcal{X}_b^i[t|t]$ corresponding to $x[t]$, $\tilde{y}[t]$, $\tilde{m}[t]$, respectively. We then predict $x[k]$:

$$\hat{x}[k|t] = \sum_{i=0}^{2N} \mathcal{W}_m^i \mathcal{X}_x^i[k|t]$$

After $y[k]$ (thus $y_m[k]$) is given at time k, we calculate the variance of modeling uncertainty $\tilde{m}[k]$ in (53):

$$\hat{Q}_{\tilde{m}}[k|t] = \mathrm{E}\,[\, Q_{\tilde{m}}(g(\hat{x}[k|t], y_m[k]))$$
$$+ \nabla_{(x, y_m)} Q_{\tilde{m}}(g(\hat{x}[k|t], y_m[k])) \cdot (\tilde{x}[k|t], \tilde{y}[k])$$
$$+ \mathcal{O}\left((\tilde{x}[k|t], \tilde{y}[k])^2 \right) \mid z[t], \cdots, z[1]\,]$$
$$= Q_{\tilde{m}}(g(\hat{x}[k|t], y_m[k]))$$
$$+ \mathrm{E}\left[\mathcal{O}\left((\tilde{x}[k|t], \tilde{y}[k])^2 \right) \mid z[t], \cdots, z[1]\,\right]$$
$$\simeq Q_{\tilde{m}}(g(\hat{x}[k|t], y_m[k])) \quad (56)$$

which in turn is a simple substitution of $\hat{x}[k|t]$ and $y_m[k]$ into the variance function (46). Note that, the error is 2nd order high relative to the variance $\hat{Q}_{\tilde{m}}[k|t]$, thus 4th order moment, which implies that the accuracy of the $\hat{Q}_{\tilde{m}}[k|t]$ computation by simple substitution (56), though not using σ-points (similar to the mean propagation of EKF), is compatible to the standard UKF.

We then calculate $\mathcal{X}_{\tilde{y}}^i[k]$ and $\mathcal{X}_{\tilde{m}}^i[k]$ for $i = 0, \cdots, 2N$ in (54) based on $\hat{Q}_{\tilde{m}}[k|t]$, and then compute:

$$\mathcal{Z}^i[k] = H\left(\mathcal{X}_x^i[k|t], y[k] + \mathcal{X}_{\tilde{y}}^i[k], \mathcal{X}_{\tilde{m}}^i[k]\right) \quad (57)$$

We compute:

$$\hat{z}[k] = \sum_{i=0}^{2N} \mathcal{W}_m^i \mathcal{Z}^i[k]$$

$$P_z[k] = \sum_{i=0}^{2N} \mathcal{W}_c^i \left(\mathcal{Z}^i[k] - \hat{z}[k]\right)(\cdot)^T + R[k]$$

where $R[k] := \mathrm{E}[\tilde{z}[k]\tilde{z}[k]^T]$ is the variance of the measurement uncertainty.

We calculate $\mathcal{X}_b^i[k|t]$ using $\mathcal{X}_{\tilde{y}}^i[k|t] = \mathcal{X}_{\tilde{y}}^i[k]$ and $\mathcal{X}_{\tilde{m}}^i[k|t] = \mathcal{X}_{\tilde{m}}^i[k]$:

$$\mathcal{X}_b^i[k|t] := \begin{bmatrix} \mathcal{X}_x^i[k|t] \\ \mathcal{X}_{\tilde{y}}^i[k|t] \\ \mathcal{X}_{\tilde{m}}^i[k|t] \end{bmatrix} = \begin{bmatrix} \mathcal{X}_x^i[k|t] \\ \mathcal{X}_{\tilde{y}}^i[k] \\ \mathcal{X}_{\tilde{m}}^i[k] \end{bmatrix} \quad (58)$$

which means that we reuse the σ-points of uncertainties that were already used in the calculation of $\mathcal{Z}^i[k]$ in (57) (i.e. $\mathcal{X}_{\tilde{y}}^i[k]$ and $\mathcal{X}_{\tilde{m}}^i[k]$) for the right computation of the cross-correlation between $\mathcal{X}_b^i[k|t]$ and $\mathcal{Z}^i[k]$ in (59).

We then compute by the standard procedure:

$$\hat{x}_b[k|t] = \sum_{i=0}^{2N} \mathcal{W}_m^i \mathcal{X}_b^i[k|t] = [\hat{x}[k|t]^T, 0, 0]^T$$

$$P_b[k|t] = \sum_{i=0}^{2N} \mathcal{W}_c^i \left(\mathcal{X}_b^i[k|t] - \hat{x}_b[k|t]\right)(\cdot)^T$$

(whch, by calculation, further becomes:

$$= diag\left(\, P[k|t]\,,\, Q_{\tilde{y}}\,,\, Q_m(g(\hat{x}[k|t], y_m[k]))\,\right)\,\,)$$

$$P_{x_b z}[k|t] = \sum_{i=0}^{2N} \mathcal{W}_c^i \left(\mathcal{X}_b^i[k|t] - \hat{x}_b[k|t]\right)\left(\mathcal{Z}^i[k] - \hat{z}[k]\right)^T$$

$$K_b[k] = P_{x_b z}[k|t] P_z[k]^{-1} \quad (59)$$

Given the measurement $z[k]$, we compute the Kalman correction:

$$\hat{x}_b[k|k] = \hat{x}_b[k|t] + K_b[k](z[k] - \hat{z}[k]) \quad (60)$$

$$P_b[k|k] = P_b[k|t] - K_b[k]P_z[k]K_b[k]^T$$

which is the *maximum a posteriori* optimal fusion of the predicted state information $\hat{x}_b[k|t]$ and the current information of $x_b[k]$ contained in $z[k]$. The algorithm then returns the state estimate $\hat{x}[k|k]$ and its covariance $P[k|k]$ from $\hat{x}_b[k|k]$ and $P_b[k|k]$, respectively, while proceed to the begginning of this Sec. 3.2 with $t \leftarrow k(= t+1)$ for recursive estimation.

Remark 2 We analyze the computational load of UKF for the cases of: 1) our proposed mixed physics/MLP-based model (47) containing the small splitted MLPs (44), and 2) the pure single large MLP model:
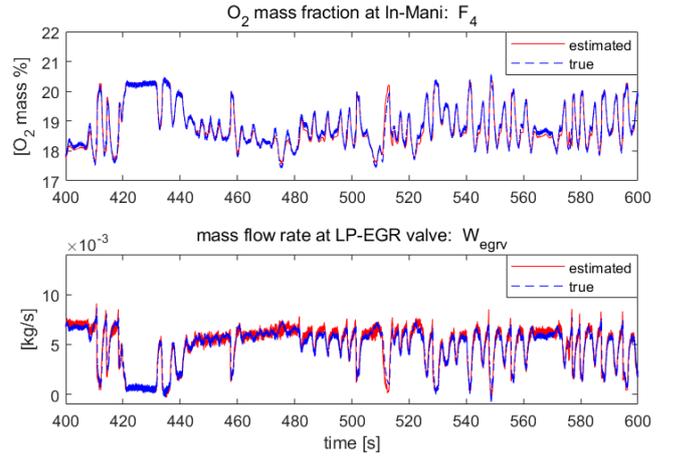
Fig. 4. Experimental results: oxygen mass fraction at intake manifold F_4 (top) and mass flow rate at the LP-EGR valve W_{egrv} (bottom) in real engine experiment estimated by our state estimation algorithm (red) and their ground truths (blue-dotted).

$$x[k] = g_{\mathrm{pure}}(x[t], y[t] + \tilde{y}[t]) + \tilde{m}_{\mathrm{pure}}[t] \quad (61)$$

For simple comparison, we suppose that the number of input, output and hidden neurons in each splitted MLPs in (44) are $\frac{1}{dim(m)}$ times smaller than that in the pure single large MLP g_{pure} of (61), respectively, which means that the number of the multiplications of network weights in the forward calculations in each splitted MLPs is $\frac{1}{dim(m)^2}$ times smaller than in g_{pure}. This means that the number of total calculations in g (44) is $\frac{1}{dim(m)}$ times smaller than in g_{pure} (61) because the number of MLPs in g is $dim(m)$. This implies that the computational cost of our proposed mixed physics/MLP-based model would be lower than the pure single large MLP model. In UKF algorithm, the $2N+1$ ($N = dim(x)+2dim(y)+2dim(m)$) σ-points are propagated through our proposed mixed physics/MLP-based model, while $2N_{\mathrm{pure}}+1$ ($N_{\mathrm{pure}} = dim(x)+2dim(y)$) σ-points through the pure single large MLP model. By combining the cost of the propagation model itself and the number of σ-points in the UKF, the final computational cost of the UKF with our mixed physics/MLP-based model would be:

$$\frac{1}{dim(m)} \times \frac{2N+1}{2N_{\mathrm{pure}}+1}$$
$$= \frac{1}{dim(m)} + \frac{4}{2dim(x) + 4dim(y) + 1}$$

times smaller than the UKF with the pure single large MLP model, e.g., $\frac{1}{4}$ times in our LP-EGR system with $dim(x) = 10$, $dim(y) = 11$ and $dim(m) = 5$.

4. EXPERIMENT

Our proposed UKF algorithm in Sec. 3 with our mixed physics/MLP-based modeling in Sec. 2 is validated by a real engine experiment through the agressive transient cycle (Fig. 4), which is totally different to the engine operation data for the data-driven modeling of the MLPs in Sec. 2.2. We only used a typical production sensor setup (i.e. y and z defined in the Sec. 2.3) with its sampling time more than 20ms and nonuniform for running the

algorithm, while measure the ground-truth of the intake manifold oxygen concentration from the expensive non-production sensor, ETAS ES636. In Fig. 4, we can see the proposed state estimation algorithm is accurate and stable/robust in the aggressive transient operation, where typical EKF easily diverges due to the nonlinearities exacerbated by the long/non-uniform sampling time and the possibly large sensing/modeling uncertainties. Note that, in this estimation experiment, we also estimated W_{egrv} by including this variable into the state vector x and its MLP-based model (30) into the measurement model (48) with its modeling uncertainty.

5. CONCLUSION

We propose a novel optimal estimation methodology for gasoline engine LP-EGR system, which allows us to implement virtual sensors for oxygen mass fraction at the intake manifold and EGR mass flow rate at the LP-EGR valve, real sensors for them too expensive to deploy in production cars. We obtain the mixed physics/MLP-based model of the system, by substituting relatively small MLPs of the subcomponents which defy accurate physics-based modeling (i.e. LP-EGR/throttle/cylinder valves, etc.) into the physics-based model of the other parts, rather than obtaining single large MLP for the entire system. We optimally combine these physics-based and data-driven modelings in the framework of UKF, and also manifest via formal analysis that this mixed physics/MLP-based modeling renders our estimator much faster to run as compared to the case of full data-driven MLP modeling. In doing so, we also extend the standard UKF theory to the more general case, where the system contains non-additive uncertainties both in the measurement and process models with cross-correlations and state-dependent variances, which stems from the inherent peculiar structure of our mixed physics-based/data-driven modeling approach, for the UKF formulation. We validate the proposed framework, which combines the mixed physics/MLP-based modeling and the UKF designed for its peculiar uncertainty structure, by real engine experiment with a typical sensor setup of production cars.

REFERENCES

Andersson, P. and Eriksson, L. (2004). Cylinder air charge estimator in turbocharged si-engines. In *SAE Technical Paper*.

Bottelli, S., Waschl, H., Savaresi, S., del Re, L., and Formentin, S. (2013). Data driven estimation of exhaust manifold pressure by use of in-cylinder pressure information. *SAE Int. J. Engines*, 6(1), 659–668.

Castillo, F., Witrant, E., Talon, V., and Dugard, L. (2013). Simultaneous air fraction and low-pressure egr mass flow rate estimation for diesel engines. In *IFAC Symp. System Structure & Control*, 731–736.

Chang, G. (2014). Marginal unscented kalman filter for cross-correlated process and observation noise at the same epoch. *IET Radar Sonar Navig.*, 8(1), 54–64.

Chen, P. and Wang, J. (2013). Observer-based estimation of air-fractions for a diesel engine coupled with aftertreatment systems. *IEEE Trans. Control Systems Technology*, 21(6), 2239–2250.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep learning. –. MIT Press.

Grondin, O., Moulin, P., and Chauvin, J. (2009). Control of a turbocharged diesel engine fitted with high pressure and low pressure exhaust gas recirculation systems. In *Conf. Decision & Control*, 6582–6589.

Julier, S.J. and Uhlmann, J.K. (2004). Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3), 401–422.

Julier, S.J., Uhlmann, J.K., and Durrant-Whyte, H.F. (2000). A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Trans. Automatic Control*, 45(3), 477–482.

Kaufmann, E., Gehrig, M., Foehn, P., Ranftl, R., Dosovitskiy, A., Koltun, V., and Scaramuzza, D. (2019). Beauty and the beast: Optimal methods meet learning for drone racing. In *Proc. Int. Conf. Robotics & Automation*, 690–696.

Khosravi, A., Nahavandi, S., Creighton, D., and Atiya, A.F. (2011). Comprehensive review of neural network-based prediction intervals and new advances. *IEEE Trans. Neural Networks*, 22(9), 1341–1356.

Kiwan, R., Stefanopoulou, A., Martz, J., Surnilla, G., Ali, I., and Styles, D. (2017). Effects of differential pressure sensor gauge-lines and measurement accuracy on low pressure egr estimation error in si engines. In *SAE Technical Paper*, 731–736.

Kocher, L.E., Stricker, K., Alstine, D.G.V., Koeberlein, E., and Shaver, G.M. (2012). Oxygen fraction estimation for diesel engines utilizing variable intake valve actuation. In *American Control Conf.*, 4963–4968.

Kolmanovsky, I., Sun, J., and Druzhinina, M. (2000). Charge control for direct injection spark ignition engines with egr. In *American Control Conf.*, 34–38.

Moriyasu, R., Ueda, M., Ikeda, T., Nagaoka, M., Jimbo, T., Matsunaga, A., and Nakamura, T. (2018). Real-time mpc design based on machine learning for a diesel engine air path system. In *IFAC Conf. Engine & Powertrain Control, Simulation and Modeling*.

Shi, G., Shi, X., O'Connell, M., Yu, R., Azizzadenesheli, K., Anandkumar, A., Yue, Y., and Chung, S.J. (2019). Neural lander: Stable drone landing control using learned dynamics. In *Proc. Int. Conf. Robotics & Automation*, 9784–9790.

Tronarp, F., Hostettler, R., and Sarkka, S. (2016). Sigma-point filtering for nonlinear systems with non-additive heavy-tailed noise. In *Proc. Int. Conf. Information Fusion*.

Wiese, A., Stefanopoulou, A., Buckland, J., and Karnik, A.Y. (2017). Modelling and control of engine torque for short-circuit flow and egr evacuation. In *SAE Technical Paper*.

Yu, H., Zhang, X.J., Wang, S., and Song, S.M. (2016). Alternative framework of the gaussian filter for nonlinear systems with synchronously correlated noises. *IET Sci. Meas. Technol.*, 10(4), 306–315.