Abstract—A rapid parameter identification procedure for a prototype automotive electromechanical brake is proposed, which is intended to support post-production quality control checks and quick model-based controller tunings. Instead of using a few different experimental manoeuvres that isolate the identification of certain sets of model parameters, an experiment optimisation is performed to single out the best trajectory that facilitates the identification of the whole sets of parameters while minimising covariance in the estimates. Utilising the measurements obtained from the optimal experimental design, both output error method and prediction error method are employed to estimate the parameters, and their estimation accuracy and evaluation speed are compared. Experimental results show that the proposed method has reduced the time required for experiments significantly with improved estimation accuracy.

I. INTRODUCTION

With throttle-by-wire now a standard feature in many production vehicles, brake-by-wire (BBW) is envisaged to be the next lineup by-wire technology to be implemented in future production vehicles. Compared to the incumbent hydraulic brakes, BBW technology offers reductions in components, space requirements, environmental impact, and increased design flexibility through software upgrades. BBW actuation is supported by the electromechanical brake (EMB), where the clamping force is generated from an electric motor connected through a reduction mechanism to the brake pads.

Mathematical models that sufficiently represent the EMB and capture the relevant dynamics are essential for simulation studies and the construction of model-based controllers. Kwak et al. [1] derived a ten degree-of-freedom physical model of an EMB with planetary gear transmission similar to the EMB used in this work. This model includes the dynamics of the sun gear, planetary gears, nut carrier, spindle and brake calliper, whereby nonlinear effects such as disc gap clearance, Coulomb frictions and gear backlashes are considered. However, a lower-order model is preferred to reduce the complexity in parameter identification and controller development. Using modal analysis, Kwak et al. [1] further shown that only the first mode is critical for the brake operation, therefore analytically justified the use of one degree-of-freedom lumped parameter model previously proposed by Maron et al. [2] and Lüdemann et al. [3]. The lumped parameter model can be derived using the torque equation around the motor rotational axis, whereby the overall inertia, friction and stiffness are individually lumped into three terms.

The conventional approach to parameter identification of EMB involves running several distinctive experimental manoeuvres specifically designed for the identification of a certain set of parameters [4]. The stiffness characteristics is identified by varying the motor position stepwise while measuring the clamp force. The friction model parameters are determined from three sets of experiments, namely the break-away tests without clamping for static friction, break-away tests with clamping for load-dependent static coefficient, and constant velocity tests for the Coulomb friction, viscous friction and Stribeck characteristics. While these tests are effective in identifying and validating the structure of the model when it is first constructed, the time required for running these experiments can be lengthy. In our case, nine hours of data collection was needed even with scripted automatic testings, in addition to extra hours for post-processing of data. During the development phase, repeated parameter identification is often required to evaluate the changes in parameters. Furthermore, parameter identification is also required for post-production quality control checks and model-based controller tunings. It is clear that the time required for executing the conventional EMB parameter identification routine is overwhelming for these purposes.

In order to support quick quality control checks post-production and rapid model-based controller tunings, a fast and accurate parameter identification procedure is desirable. Furthermore, measurement noises that are inevitable in practice have to be taken into account for improved estimation accuracy. These considerations motivate the development of a rapid model parameter identification procedure for an EMB that incorporates noise models.

The design of the rapid parameter identification procedure is commenced in two stages. First, the optimal experimental design is conducted in Section III. This involves planning an optimal output trajectory to minimise the time required for experiments needed to infer model parameters from a given number of input-output measurements. This trajectory has to be achievable by complying to the input and output constraints. In the second stage, parameter estimation algorithms are devised in Section IV to identify model parameters from the input-output measurements. Output error method (OEM) and the prediction error method (PEM) are formulated, followed by discussions on their accuracy and ease of implementation. Finally, the parameterised EMB model is validated against independently obtained experimental measurements.

II. ELECTROMECHANICAL BRAKE MODEL

An experimentally validated lumped parameter EMB model is considered:

\[ \dot{x} = f(x, T_m), \quad y = h(x), \]  

(1)

where \( x = [x_1, x_2]^T \) is the state with motor position \( x_1 \) and motor velocity \( x_2 \), and \( T_m \) is the motor torque, which is also the
input of the system. The nonlinear function \( f(\cdot) \) is given by
\[
f(x, T_m) = -\frac{1}{2} \left( N F_{cl}(x_1) - T_f(x_1, x_2) - T_m \right) \left( x_2 - 1 \right) J \left( NF_{cl}(x_1) - \overline{T}_f(x_1, x_2) - T_m \right)
\]
where \( F_{cl}, J, N, T_f \) are the clamp force, lumped inertia, gear ratio and friction torque respectively. The stiffness characteristics is experimentally found to be hyperelastic, and can be modelled using a third-order polynomial function:
\[
F_{cl}(x_1) = \begin{cases} k_1 x_1^2 + k_2 x_1^2 & \text{if } x_1 \geq 0, \\ 0 & \text{if } x_1 < 0. \end{cases}
\]

The friction model adopted here is based on the friction model presented in Hess and Soom [5], but modified to also include the load-dependent friction coefficient and smooth transition at zero velocity, given by:
\[
T_f = T_c x_2 + \left( T_c + \frac{T_m - T_c}{1 + (x_2/\omega_s)^2} \right) \tanh(x_2),
\]
where \( T_c, T_v \) and \( \omega_s \) are the Coulomb friction torque, viscous friction coefficient and the characteristic velocity of the Stribeck friction respectively. The load-dependent static friction torque \( T_s \) is given by
\[
T_s(F_{cl}) = T_{s0} + \mu_s F_{cl},
\]
where
\[
h(x) = [ F_{cl}(x_1), x_2 ]^T.
\]

Since the experimental measurements are conducted in discrete sampling period and are influenced by noises, the continuous-time EMB model (1) is discretised and augmented with noise terms, given by:
\[
x_{d}(k+1) = f_d(x_d(k), T_m(k), \theta) + B_v v_i(k),
\]
\[
y_{d}(k) = h(x_d(k), \theta) + v_o(k),
\]
where \( x_{d}(k) = [x_1(k), x_2(k)]^T \) is the state vector at time step \( k \) and \( f_d(\cdot) \) is calculated using Euler forward method:
\[
f_d = x_d + t_s f(x_d(k), T_m(k), \theta),
\]
where \( t_s \) is the step size. The plant dynamics are affected by the process noise, \( v_i(k) \) and the associated matrix \( B_v = [0, 1]^T \). The model output is given by \( y_{d}(k) \), and the measured output \( y_{m}(k) \) is corrupted by the output measurement noise, \( v_o(k) \). Note that the EMB parameters, \( \theta = [\theta_1, \theta_2, \ldots, \theta_7]^T = [k_1, k_2, T_c, T_{s0}, T_v, \mu_s, \omega_s]^T \) are now included as function arguments in the model (7), (8).

III. OPTIMAL EXPERIMENTAL DESIGN

A. Problem Formulation

Optimal experimental design is performed in order to plan an output trajectory utilised for inferring the model parameters of stiffness (3) and friction (4). To this end, the following standing assumption is made.

Assumption 1. There exists an EMB plant model (7), (8) with known bounds of the model parameters.

Remark 1. An initial calibration is typically performed, and a priori estimation of the model parameters is presumed to be available. Additionally, the changes of gear ratio \( N \) and the lump inertia \( J \) are expected to be insignificant over time. On the other hand, the stiffness and friction characteristics are prone to large changes due to wear and operating conditions. Hence, during the rapid parameter estimation process, \( N \) and \( J \) are assumed to be fixed to reduce dimensionality and only \( \theta \) is to be estimated.

The values listed in Table I are obtained using the conventional identification approach alluded in the introduction, and will be used as the initial estimates in the model-based approach to experimental design.

During experimental design, the input is assumed to be free from noise, that is \( v_i = 0 \). However, the output is assumed to be contaminated by noise with zero-mean and Gaussian distribution, that is \( v_o \sim N(0, R) \). The sample covariance, \( R \) can be estimated using [6]:
\[
R = \frac{1}{n} \sum_{k=1}^{n} (y_{m}(k) - \bar{y}_{d}(k)) (y_{m}(k) - \bar{y}_{d}(k))^T.
\]

To obtain \( y_{m}(\cdot) \) and \( y_{d}(\cdot) \), an effort can be made to drive the plant using an arbitrary chosen input trajectory \( \{T_{m}(k)\}_{k=1}^{n} \), where the output \( y_{m}(\cdot) \) are measured. Then, the model output, \( y_{d}(\cdot) \) can be computed from (7) using the input sequence. Note in practise, process (or input) noise exists, although it is difficult to estimate and incorporated in the experimental design. In Section IV-B, the process noise will be estimated during parameter estimation using PEM.

Furthermore, the experimental design can be approached by optimising the input sequence, \( U := \{T_m(k), k = 1, \ldots, n\} \) such that the resultant output trajectory \( Y := \{y_m(k), k = 1, \ldots, n\} \) yields a small covariance in the estimated parameters. As established by the Cramér-Rao lower bound, the lower bound of the maximum likelihood estimator covariance is given by the inverse of the Fisher information matrix, \( M^{-1} \), where \( M \) is defined by [7]:
\[
M = E \left( \frac{\partial \log p(Y|U, \theta)}{\partial \theta} \left( \frac{\partial \log p(Y|U, \theta)}{\partial \theta} \right)^T \right).
\]

Assumption 2. The Fisher information matrix for the input output pairings in the trajectory, \( M \) is nonsingular.

Remark 2. A nonsingular \( M \) is necessary for the existence of the Cramér-Rao lower bound matrix and can be provided through appropriate design of the experimental trajectory.

Several scalar measures of performance based on \( M \) is presented in [7], for example the A-optimality that minimises the average variance of the parameters, where \( \text{min}_{U} \text{tr}(M^{-1}) \) is evaluated; and the D-optimality that minimises the volume of the uncertainty ellipsoids, where \( \text{min}_{U} \det (M^{-1}) \) is evaluated. The D-optimality is adopted in this work, attracted by its
invariant property under scale changes in the parameters and linear transformations of the output. This property is useful as the parameters in the EMB model have magnitudes differing by factors of up to 10^6.

To facilitate meaningful comparison between variances and to avoid numerical sensitivity during optimisation, the parameters and output are scaled. The normalised parameter vector, \( \theta = \,[\bar{\theta}_1, \ldots, \bar{\theta}_7]^T \) is introduced as:
\[
\bar{\theta}_i = \theta_i / \theta_{nom,i}, \quad i = 1, 2, \ldots, 7
\]
where the original parameters in SI units \( \theta \) are normalised to their respective nominal values \( \theta_{nom,i} \) taken from Table I.

Note that the output ranges in SI units are given by
\[
\{ h = [F_{cl}, x_2]^T : F_{cl} \in [0, 30000], x_2 \in [-300, 300] \},
\]
where they differ by two orders of magnitude. To avoid heavy weighting on one of the measured outputs, \( h \) is scaled using the following expression:
\[
\bar{h} = \text{diag} \left([10000^{-1}, 100^{-1}]\right) h.
\]

Note that the scaled outputs \( \bar{h} \) have reduced ranges with the same order of magnitude, given by
\[
\{ \bar{h} = [\bar{h}_1, \bar{h}_2]^T : \bar{h}_1 \in [0, 3], \bar{h}_2 \in [-3, 3] \}.
\]
Additionally, the scalar performance measure of the \( D \)-optimal design can be converted into a convex function using \( \log \det (M^{-1}) \) [8]. It can be shown that \( M \) is a positive semidefinite symmetric matrix. If \( M \) is nonsingular, then \(-\log \det (M) = \log \det (M^{-1})\) can be employed in order to avoid numerical issues during the matrix inversion when \( M \) has a bad condition number.

**Procedure 1 (Optimal Experimental Design).** Given \( n \) possible input \( U := [T_m(1), \ldots, T_m(n)]^T, \, f_d(\cdot) \) subject to (8) and Assumption 2, the optimal experimental trajectory is found by the solution to the following:
\[
\begin{align*}
\min_{\theta, U} & \quad -\log \det (M), \\
\text{s.t.} & \quad x_d(1) = x_0,\\
& \quad x_d(k + 1) = f_d(x_d(k), T_m(k), \theta), \quad k = 1, \ldots, n\\
& \quad |T_m(k)| \leq T_{m \text{ max}},\\
& \quad F_{cl}(k) \leq F_{cl \text{ max}},\\
& \quad |x_2(k)| \leq x_2 \text{ max}.
\end{align*}
\]

Note that the \( M \) in (13) can be numerically approximated using the following expressions [6], [7]:
\[
\begin{align*}
\frac{\partial f_d}{\partial \theta} &= \left[ \frac{\partial f_d}{\partial \theta_1}, \ldots, \frac{\partial f_d}{\partial \theta_7} \right]^T, \\
\frac{\partial f_d}{\partial x_d} &= \left[ \frac{\partial f_d}{\partial x_{d1}}, \ldots, \frac{\partial f_d}{\partial x_{d7}} \right]^T, \\
\chi_{\theta_i,k+1} &= f_{d,k} \chi_{\theta_i,k} + \frac{\partial f_{d,k}}{\partial \theta_i}\chi_{\theta_i,k+1},
\end{align*}
\]
where \( \chi_{\theta_i,1} = 0 \) and \( k = 1, \ldots, n \). The arguments of \( f_d(\cdot) \) and \( \bar{h}(\cdot) \) are omitted due to space constraint and the time steps are written in the subscripts.

**B. Results and Discussion**

The optimisation problem (13) is solved using the derivative-free optimisation function, \texttt{ga} in MATLAB, which is based on genetic algorithm [9]. The sampling period is chosen to be \( t_s = 0.001 \) seconds. However, instead of updating the input sequence \( U \) at every sampling period, it is updated once every 20th sample to reduce the number of optimised variables. For the ease of repeatability, the initial conditions of the experiment are selected to be zero clamp force and zero velocity, where \( x_d(1) = [0, 0]^T \). Additionally, the length of experiment can also be freely chosen, which can be reflected by varying the \( n \) in (13).

![Fig. 1. Optimal trajectories for parameter identification versus various length of experiments.](image)

The input and output trajectories of four optimal experimental designs with varying length from 0.1 seconds to 0.4 seconds are plotted in Figure 1. Regardless of the length of experiments, the clamp force trajectories resemble the step responses, while the velocity transients consist of initial acceleration, operation at maximum speed and deceleration to zero. The performance measures of the optimal trajectories with varying lengths are illustrated in Figure 2, where the criteria for \( A \)-optimal design and \( D \)-optimal design are presented. The values of both criteria decrease with the experimental length, implying a longer experiment provides better parameter estimation accuracy. This is because a longer experiment allows more time for full system dynamics to be excited and more data to be collected. However, a decreased rate of improvement is observed, where the reduction in the performance criteria is more significant for experiments shorter than 0.18 seconds. This is because for shorter experiments, the maximum values of clamp force transients increase significantly with a slight increase in exponential length. However, for experimental length longer than 0.18 seconds, the maximum values of clamp force trajectories remain at approximately 23 kN, as shown in Figure 1. To strive a balance between the estimation accuracy and the time required for experiments, 0.3 seconds is chosen as the length of experiment, and the corresponding optimal trajectory will be applied in the next subsection for parameter estimation. While the experimental length of 0.2 seconds is adequate, a slightly longer experiment is chosen to account for situations where softer brake pads are used and longer rise-times are expected compared to the nominal case.
IV. PARAMETER ESTIMATION

The optimal experimental clamp force trajectory determined in Section III is utilised as the reference trajectory.

Assumption 3. The EMB clamp force tracking controller is assumed to be sufficiently fast relative to the optimal experimental trajectory.

Remark 3. The dynamics of the transfer function from clamp force reference to the actual clamp force can be neglected and the clamp force profile identified in Section III can be explicitly used in the parameter identification. A near-time-optimal clamp force controller has been suggested in [10].

The experimentally obtained input and output measurements will be used to infer the EMB model parameters in the following using both OEM and PEM.

A. Output Error Method (OEM)

The OEM assumes that the noise only contaminates the output measurements, i.e. there is no process noise. The output measurement noise is assumed to be zero-mean Gaussian $v_i \sim N(0, R)$. Note exact knowledge of initial condition $x_0$ is not available due to noise.

Procedure 2 (Parameter estimation using OEM). Given the plant model (7) and a sequence of measured input $U := [T_m(1), \ldots, T_m(n)]^T$ and measured output $Y := [y_m(1), \ldots, y_m(n)]^T$, the OEM finds the model parameters $\theta$ and the initial condition $x_0$ in order to minimise the error between the measured and the model outputs, which can be formulated as the following [6]:

$$\min_{x_0, \theta} \left\{ \frac{1}{2} \sum_{k=1}^{n} \left( (y_m(k) - \hat{y}_d(k))^T R^{-1} (y_m(k) - \hat{y}_d(k)) \right) + \frac{n}{2} \log \det (R) \right\}$$

s.t. $x_d(1) = x_0$, (7) and (9).

B. Prediction Error Method (PEM)

The PEM takes account of both the output measurement noise and the process noise, given by $v_o \sim N(0, R)$ and $v_i \sim N(0, Q)$ respectively. Note that the effect of input measurement noise can be included in the process noise.

Instead of relying on the model output from (7), PEM utilises a one-step ahead prediction model constructed using the extended Kalman filter (EKF). The EKF is a nonlinear version of the Kalman filter that involves repeated linearisation of the model during evaluation [11]. It is evaluated at two stages, namely the predict phase and the update phase. During the predict phase, the states and the covariance of the states are predicted one-step ahead using measurements at $k-1$, evaluated using the following:

$$\hat{x}(k|k-1) = f_d(\hat{x}(k-1|k-1), T_m(k-1), \theta),$$

$$P(k|k-1) = A(k-1) P(k-1|k-1) A(k-1)^T + B_o Q B_o^T, \quad (15)$$

where $A(k-1)$ is the state transition matrix. These predictions are subsequently corrected in the update phase using measurements at $k$. This involves the calculation of the Kalman gain, given by

$$K(k) = P(k|k-1) C(k)^T S(k)^{-1}, \quad (16)$$

where $S(k)$ denotes the residual covariance, obtained from

$$S(k) = C(k) P(k|k-1) C(k)^T + R. \quad (17)$$

The states and covariance are then updated using the following expressions with the Kalman gain:

$$\hat{x}(k|k) = \hat{x}(k|k-1) + K(k) [y_m(k) - h(\hat{x}(k|k-1), \theta)],$$

$$P(k|k) = [I - K(k) C(k)] P(k|k-1). \quad (18)$$

The state transition matrix $A(k-1)$ and the observation matrix $C(k)$ are defined to be the following Jacobians:

$$A(k-1) = \frac{\partial f_d}{\partial x_d} \big|_{x(k-1|k-1), T_m(k-1)}, \quad C(k) = \frac{\partial h}{\partial x_d} \big|_{x(k|k-1)} \quad (19)$$

Procedure 3 (Parameter estimation using PEM). Given the EKF (15)–(19) and a sequence of measured input $U$ and measured output $Y$, the PEM finds the model parameters $\theta$, the initial condition $x_0$ and the process noise covariance $Q$ in order to minimise the error between the measured and the predicted outputs, which can be formulated as the following [6]:

$$\min_{x_0, \theta, Q} \left\{ \frac{1}{2} \sum_{k=1}^{n} \left( (y_m(k) - \tilde{y}_d(k))^T S^{-1} (y_m(k) - \tilde{y}_d(k)) \right) + \frac{n}{2} \log \det (S) \right\}$$

s.t. $\tilde{x}_d(1|1) = x_0$, $P(1|1) = P_0$, (15)–(19),

$$\tilde{y}_d(k) = h(\tilde{x}(k|k), \theta), \quad k = 1, \ldots, n.$$
a) OEM vs. PEM: The normalised estimated parameters are listed in Table II. Due to the presence of noise, the estimated parameters obtained from each set of experiment are not exactly identical, as demonstrated in Figure 5. To assist with comparison, the means of the estimations given by both methods are calculated, and are found to match closely.

However, discrepancies are observed between the estimations and the nominal values. To check the validity of these values, simulations results are compared against independently obtained experimental data. From the step responses illustrated in Figure 6, the simulated responses using parameters obtained from both OEM and PEM visually overlap with the experimentally obtained responses, indicating a satisfactory estimation.

Simulated responses obtained using the nominal parameter values are also plotted in Figure 6, and are compared to responses obtained using the newly acquired optimised parameters. It is evident that the simulated responses obtained using the optimised parameters match the experimental results better. This result indicates that the optimised parameters are more accurate than the nominal values.

The performance of estimation can be compared using the variance of the estimated parameters. The sum of variances given by OEM and PEM are 0.00480 and 0.00388 respectively. Since the sum of variances given by the PEM is smaller, this suggests that it renders better estimations. This can also be examined by comparing the 50% confidence region depicted in Figure 5, where the estimations provided by PEM have smaller confidence regions. This improvement is resulted from the consideration of process noise in the PEM, where the covariance Q is estimated to be 0.00020154, suggesting that the motor torque measurements have errors with standard deviation of 0.0142 Nm. Note that the estimated standard deviation in torque measurement error is relatively substantial compared to the friction torque at very slow speed, as it has the same order of magnitude compared to the break-away friction torque.

Further model validation is performed using experiment with sinusoidal trajectory, as shown in Figure 7. There are some disagreements in the simulated and the experimental responses. A closer look at the velocity responses indicates that the model shows lockups when the velocity crosses zero while the clamp force transient is at its peak. However, the velocity lockups are not observed in the experimental measurements. This suggests that the response is very sensitive at very low speed, as a small error in motor torque measurement or minor estimation error in friction characteristics at very low speed may result in velocity lockups, which then leading to significant deviation in the clamp force trajectory. Nevertheless, in the period where velocity lockups do not happen, for example in between 1.5 seconds to 1.7 seconds, the simulated responses are qualitatively matching the measurements well.
clamp force and velocity deviate from the measurements after 0.2 seconds, hinting an underestimation of friction torque. The sum of variances of the estimated parameters is 0.00729, which is significantly higher compared to the optimal design.

The major attractive feature offered by the optimal experimental design and the parameter estimation proposed in this paper is the speed and accuracy in obtaining a useful set of model parameters for an EMB. Compared to the nine hours of data collection needed for the running four distinctive sets of steady state experiments [4], the rapid procedure proposed in this paper only requires one minute of data collection. This is because the full system dynamics is exploited in the rapid procedure and transient responses are used to infer the model parameters, as opposed to the steady state responses used in the slower routine. Moreover, the parameter estimations provided by the rapid procedure are also more accurate, as demonstrated by the validation results in Figure 6.

VI. CONCLUSIONS

The need for a rapid model parameter identification for an electromechanical brake (EMB) is addressed in this paper. First, the optimal experimental trajectory that minimises the data collection time while ensuring the measurements provide sufficiently rich information about the plant dynamics is determined — needed for accurate estimation of parameters.

As opposed to the nine hours of data collection involved in the parameter identification routine proposed in [4], a total of one minute of experimentation is conducted utilising the optimal experimental trajectory. It was observed that the PEM provides a marginally better estimation accuracy compared to the OEM, but required significantly longer evaluation time. This may be an important consideration for practical implementation.

REFERENCES