

OPEN ACCESS

Effective Parameterization of PEM Fuel Cell Models—Part II: Robust Parameter Subset Selection, Robust Optimal Experimental Design, and Multi-Step Parameter Identification Algorithm

To cite this article: Alireza Goshtasbi *et al* 2020 *J. Electrochem. Soc.* **167** 044505

View the [article online](#) for updates and enhancements.



Effective Parameterization of PEM Fuel Cell Models—Part II: Robust Parameter Subset Selection, Robust Optimal Experimental Design, and Multi-Step Parameter Identification Algorithm

Alireza Goshtasbi,^{1,*} Jixin Chen,² James R. Waldecker,² Shinichi Hirano,² and Tulga Ersal^{1,z}

¹Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan 48109, United States of America

²Ford Motor Company, Dearborn, Michigan 48121, United States of America

The second part of this two-part study develops a systematic framework for parameter identification in polymer electrolyte membrane (PEM) fuel cell models. The framework utilizes the extended local sensitivity results of the first part to find an optimal subset of parameters for identification. This is achieved through an optimization algorithm that maximizes the well-known D-optimality criterion. The sensitivity data are then used for optimal experimental design (OED) to ensure that the resulting experiments are maximally informative for the purpose of parameter identification. To make the experimental design problem computationally tractable, the optimal experiments are chosen from a predefined library of operating conditions. Finally, a multi-step identification algorithm is proposed to formulate a regularized and well-conditioned optimization problem. The identification algorithm utilizes the unique structure of output predictions, wherein sensitivities to parameter perturbations typically vary with the load. To verify each component of the framework, synthetic experimental data generated with the model using nominal parameter values are used in an identification case study. The results confirm that each of these components plays a critical role in successful parameter identification.

© 2020 The Author(s). Published on behalf of The Electrochemical Society by IOP Publishing Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 License (CC BY, <http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse of the work in any medium, provided the original work is properly cited. [DOI: 10.1149/1945-7111/ab7092]



Manuscript submitted September 30, 2019; revised manuscript received January 9, 2020. Published February 14, 2020.

Physics-based models of polymer electrolyte membrane (PEM) fuel cells have been extensively used to shed light on the underlying physical phenomena and optimize the cell design.¹ However, in most cases, predictions by such models are only qualitatively valid^{2–5} and quantitative agreement with experimental data has remained mostly elusive. While the qualitative descriptions have proven useful in many cases, certain application areas, such as model-based control,^{6,7} require models with more predictive power. An important aspect of a model's prediction capability is the parameters used to run the model. Therefore, effective parameter identification is a critical step in developing quantitatively predictive models. Additionally, parameter identification is often used as a tool to infer a fuel cell's physical characteristics (e.g., catalyst activity or mass transport resistance) from experimental data^{8,9} and diagnose degradation,¹⁰ which further highlights the need for effective identification methods. Despite such a need, the parameter identification problem is relatively understudied in the fuel cell literature.¹¹ This is an important gap that we aim to address in this work.

One of the main challenges associated with parameter identification in a large-scale model is that such models usually have a large number of parameters, many of which may not be practically identifiable given the available measurements. Under such circumstances, it is imperative to select a subset of parameters for identification.¹² Particularly, the selected subset must only include parameters that can be identified using the available data. Moreover, the parameters must be selected in a way that ensures the predictive capability of the model is maintained.¹³

These are conflicting requirements that necessitate close inspection of output sensitivities to different parameters to ensure that the selected subset has the required characteristics. Accordingly, several methods have been proposed in the literature to ensure the selected subset is optimal.^{14–18} Among these, methods that optimize a scalar metric of the Fisher information matrix (FIM) to maximize the precision of parameter estimates are commonly utilized.^{18–20} Other approaches that directly focus on minimizing the model's prediction error have also been proposed.^{13,21} However, to date, these methods have not been applied to fuel cell models, which leaves significant room for improvements in model parameterization.

Another important aspect of parameter identification is the data used for this purpose. In particular, the data must contain enough information that allows the effects of different model parameters to be distinguished. The degree to which a data set is informative is controlled by the experimental conditions used to collect the data and the corresponding model sensitivities. Therefore, model-based experimental design approaches have attracted considerable attention in the literature.^{22–27} These techniques seek to find optimal experiments that are informative for parameter identification. This is done, for example, by using the model to find conditions that maximize the sensitivity of the predicted output to model parameters.²⁸ Other proposed schemes maximize a scalar measure of the FIM,^{26,27} similar to parameter subset selection algorithms.

The idea of optimally designing experiments to reveal a parameter of interest is not new to the fuel cell community. Using differential conditions to achieve uniform distributions in a cell that allows characterization of membrane electrode assemblies (MEAs)²⁹ and using low reactant concentrations to characterize the transport resistances^{9,30} are examples of this idea. However, these approaches are based on expert knowledge and are typically limited to specialized studies that focus on a particular behavior of the cell. Moreover, they are not guaranteed to be optimal for parameter identification. Therefore, there is a need to adopt the above-mentioned methods to fuel cell applications and rigorously optimize the experiments for effective model parameterization. We also note that recent battery literature has extensively utilized model-based optimal experimental design with considerable success, as evidenced by the improved parameter identification accuracy.^{31–36} This further highlights the potential benefits of using these methods for enhanced modeling of fuel cells, as many of the underlying equations are structurally similar to those used in battery models.

Once the subset of parameters to be identified is selected and the optimal conditions are used to collect the experimental data, the parameters can be identified. This last step is formulated as an optimization problem, whose solution typically minimizes the output prediction error of the model. Importantly, the structure of the available data can be utilized to improve identification results.²⁸ Therefore, formulation of this step also requires consideration of model sensitivities to ensure the optimal solution is achieved. Here again, the fuel cell modeling literature is yet to capitalize on such potentials, leaving an important gap that should be addressed.

*Electrochemical Society Student Member.

^zE-mail: tersal@umich.edu

The above discussion points to several possibilities to improve parameter identification in fuel cell models. Accordingly, here we develop a systematic framework for effective model parameterization that addresses some of the current shortcomings of the fuel cell literature. While the framework is generally applicable to fuel cell models described by differential-algebraic equations, we use a pseudo-2D two-phase and non-isothermal model^{37–39} for this study as an example. Having investigated the sensitivities of various model predictions to different parameters in the first part of this work,⁴⁰ here we focus on the problems of parameter subset selection and optimal experimental design. Particularly, we use our extended local sensitivity analysis results to select a subset of parameters for identification that is robust to initial assumptions about the nominal parameter values. We use similar procedures for robust optimal experimental design to obtain a set of experiments that are maximally informative for identification of the selected parameters. Finally, a multi-step parameter identification approach is proposed and tested with synthetic experimental data generated by the model.

The rest of the paper is organized as follows. First, a brief overview of the model and the considered parameters is given, where the first part of this two-part study is also reviewed. Next, the problem of parameter subset selection is addressed. The robust optimal experimental design procedure is then discussed followed by a description of the proposed identification algorithm. Lastly, the results for an identification case study are presented before closing with some concluding remarks.

Overview of the Model and Part I of the Study

In this work, we use a physics-based model that has been developed recently for automotive applications.^{37–39} Particularly, the model is pseudo-2D and captures non-isothermal and two-phase phenomena and has been extensively validated with performance and resistance data obtained experimentally from state-of-the-art fuel cell stacks.³⁹ Further details about the model and its validation can be found in the original publications.^{37–39}

The sensitivities of several of the model predictions to a variety of parameters were examined in the first part of this work using an extended local analysis.⁴⁰ Particularly, cell voltage, high frequency resistance (HFR), and normalized membrane water crossover were studied for their sensitivity to 50 of the model parameters. The local analysis was conducted about several points in the parameter space to obtain a more global picture of the model sensitivities. Additionally, the sensitivities were generated using a large library of 729 different operating conditions with variations in temperature, pressure, relative humidity, and stoichiometric ratio. The parameters considered included geometrical and structural features of the cell, along with a number of fitting coefficients. The results were analyzed for sensitivity magnitudes and their relative independence, as they have implications for parameter identifiability. It was also

demonstrated that many of the model parameters are not identifiable from cell-level measurements. This unidentifiability is often due to insensitivity of a parameter or multiple parameters having the same impact on the predicted outputs, which points to opportunities for model refinement, e.g., by re-parameterizing the model.^{41–43} It was further shown how additional measurements can enhance identifiability of model parameters by increasing their corresponding impact on the predicted outputs and reducing the correlation between the effect of different parameters (see the first part of the study for further discussion⁴⁰).

Having analyzed model sensitivities, our exclusive focus in this part of the study is on parameter identifiability. Accordingly, we note that while we had considered 50 parameters for sensitivity analysis, many of these parameters, such as those describing the channel geometry, are known for a given stack. Although these parameters were included in our sensitivity analysis to ensure that our results are not constrained to a particular cell design, they are not suitable candidates for identification. Therefore, we narrow down the list of considered parameters in this part to focus on those that may be considered for model parameterization. Specifically, we limit our analysis here to a set of 31 parameters that are given in the Appendix Table A.I, which also provides the bounds for each parameter. Note that the specified bounds, which are determined based on the uncertainty of the corresponding parameters reported in the literature, ensure the parameters maintain their physical significance. Regarding the outputs of interest, here we assume that voltage and HFR are measured in the experiments as they often are in real-world testings. Therefore, our discussion here is based on the combined sensitivity of the voltage and HFR predictions. A summary of the sensitivity results is shown in Fig. 1 for the 31 parameters that are subject to this part of the study.

Robust Parameter Subset Selection

Attempting to identify all of the model parameters using only voltage and HFR data results in an ill-conditioned inverse problem.⁴⁴ This necessitates the need for parameter subset selection so that the identification algorithm can focus exclusively on the parameters that are practically identifiable with the available measurements. We recall that a parameter is practically identifiable when it has an observable impact on the output predictions that is uniquely distinguishable from the impact of other parameters.¹⁷ Therefore, a large sensitivity magnitude alone does not guarantee identifiability of a parameter; the corresponding sensitivity vector must also be independent from that of other parameters. Accordingly, this section provides an algorithm that selects an identifiable subset of parameters. This subset selection procedure may also be viewed as a regularization technique to alleviate the ill-conditioning of the inverse problem.^{45,46} Briefly, the proposed method is as follows:

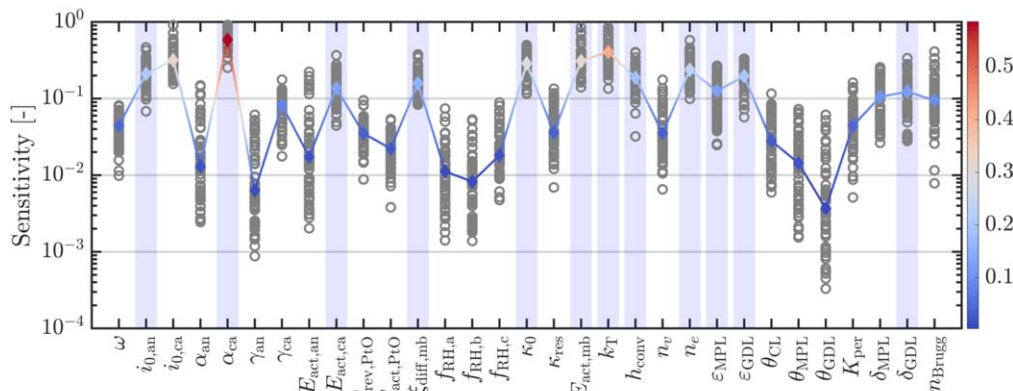


Figure 1. Summary of the sensitivity results based on voltage and HFR predictions and the results of the parameter subset selection method. The sensitivities at all sample points in the parameter space are shown as gray circles and the corresponding median is shown with the colored line, where the color changes based on the median sensitivity. The parameters selected by the robust parameter subset selection method are highlighted with a light blue shade.

1. Inspect the parameter set for highly collinear pairs and remove the parameter with smaller sensitivity magnitude from such pairs.
2. Choose the number of parameters, n_θ , to be included in the subset. This is done using singular value decomposition.
3. Find the locally optimal subset of parameters at each point in the parameter space, for which sensitivity information has been calculated.
4. Find the globally optimal subset of parameters for the entire parameter space. The local solutions from the previous step are used to guide the optimization.

Below, each step is described in detail.

We start by inspecting the parameter sensitivities for highly collinear pairs. Given two sensitivity vectors s_i^k and s_j^k for parameters i and j obtained at point k in the parameter space, their collinearity index is defined as:

$$\psi_{i,j}^k = \cos(\phi_{i,j}^k) = \frac{|s_i^{kT} s_j^k|}{\|s_i^k\|_2 \|s_j^k\|_2}, \quad [1]$$

where $\phi_{i,j}^k$ is the angle between the sensitivity vectors and the collinearity index $\psi_{i,j}^k$ varies between 0 and 1, indicating orthogonal and collinear sensitivity vectors, respectively. This collinearity index is calculated for all parameter pairs at all points in the parameter space, for which we had generated sensitivity data in the first part of the study.⁴⁰ The median of the calculated indices is then used as the representative collinearity index for each parameter pair over the entire parameter space:

$$\bar{\psi}_{i,j} = \text{med}_k \psi_{i,j}^k, \quad [2]$$

where k is the index for the sample point in the parameter space. Only one parameter from a pair whose median collinearity index is high can be practically identified.²⁰ Therefore, for pairs with $\bar{\psi}_{i,j}$ exceeding a threshold value of ψ_{th} , the parameter with lower sensitivity magnitude is discarded. Here we use $\psi_{th} = 0.9$, which results in removal of $i_{0,ca}$ and $E_{rev,PtO}$ from the list of parameters due to their significant collinearity with α_{ca} and ω , respectively. We also note that a large threshold is chosen to ensure a certain degree of flexibility is maintained in the resulting set of parameters. More specifically, smaller thresholds can lead to removal of parameters that can have a noticeably different impact than their collinear counterparts only at a few selective operating conditions.

In the second step we determine the number of parameters, n_θ , that are needed to capture the variations in the data due to perturbations to the remaining parameters. Singular value decomposition⁴⁷ is a powerful tool for this purpose. In particular, at each sample point in the parameter space we calculate the singular values of the sensitivity matrix. Let these singular values be ordered as follows:

$$\sigma_1^k > \sigma_2^k > \dots > \sigma_{n_\theta}^k, \quad [3]$$

where σ_i^k denotes the i -th largest singular value of the sensitivity matrix obtained at the k -th sample point in the parameter space and n_θ is the number of parameters under consideration after the lower sensitivity parameters in collinear pairs have been removed ($n_\theta = 29$). We then seek to find the smallest r such that:

$$\bar{\sigma}_r^k = \frac{\sum_{j=1}^r \sigma_j^k}{\sum_{i=1}^{n_\theta} \sigma_i^k} \geq \bar{\sigma}_{th}, \quad [4]$$

where $\bar{\sigma}_r^k$ monotonically increases with r and provides a qualitative measure as to how much of the total variations in the sensitivity directions are captured by the first r singular vectors. Moreover, $\bar{\sigma}_{th}$ is a threshold value that determines the number of selected singular values. Choosing a threshold of $\bar{\sigma}_{th} = 0.9$, we have found that, on average, 12 singular values are required to satisfy the above

inequality at the sample points in the parameter space. Accordingly, we focus on selecting a subset of size $n_I = 12$ for identification.

The third step of the algorithm is finding the locally optimal subset of parameters for identification. To this end, we must determine which of the model parameters have the largest *unique* impact on the predicted outcomes. This is done using the sensitivity matrix given by:

$$S^k = \left. \frac{\partial \tilde{y}}{\partial \theta} \right|_{\theta=\theta^k}, \quad [5]$$

where $\tilde{y}$ is the vector of output predictions and θ is the vector of parameters and the sensitivity matrix is evaluated at the k -th sample point in the parameter space. While our sensitivity calculations are slightly different,⁴⁰ the structure of the resulting matrix is the same. To select a parameter subset, we want to choose columns of the sensitivity matrix that ensure the corresponding parameters can be identified with the required precision. The parameter estimation precision is determined by the corresponding covariance (Σ_θ), which is constrained by the inverse of the FIM through the Cramer-Rao bound⁴⁸:

$$\Sigma_\theta \geq (F^k)^{-1}, \quad [6]$$

where F^k is the FIM at the k -th sample point in the parameter space, which is given by:

$$F^k = S^{kT} S^k. \quad [7]$$

Note that the definition of the FIM also includes the covariance of the measurement noise, which, for the subset selection procedure, can be assumed to be identity without loss of generality.²⁰ With the Cramer-Rao bound and the FIM defined, we recall that the objective of this step is to select a subset of parameters, whose corresponding covariance is minimal. This can be achieved by minimizing the lower bound given by the FIM. To this end, a scalar metric of the FIM is optimized. A number of scalar metrics have been suggested based on different criteria.⁴⁹ Among them, the D-optimality criterion is most commonly used,²⁰ which maximizes the determinant of the FIM obtained from selected columns of the sensitivity matrix. Specifically, such an FIM is given by:

$$F_I^k = S_I^{kT} S_I^k = (S^k L)^T (S^k L), \quad [8]$$

where L is the selection matrix defined as:

$$L = [e_{i_1}, e_{i_2}, \dots, e_{i_{n_I}}]. \quad [9]$$

In the above equation, e_j denotes the j th standard unit vector and $\{i_1, i_2, \dots, i_{n_I}\}$ is the set of selected parameter indices. We can now define the D-optimal selection criterion explicitly:

$$\varphi_D^k = \max \varphi_D(F_I^k) = \max \log \det(F_I^k), \quad [10]$$

where φ_D is the D-optimality criterion. The D-optimal solution, φ_D^k , maximizes the determinant of the FIM, which has the effect of minimizing the volume of the confidence ellipsoid for the parameter estimates. To find φ_D^k , an integer program must be solved²⁰:

$$\begin{aligned} & \text{maximize} && \varphi_D(F_I^k) \\ & \text{subject to} && F_I^k = (S^k L)^T (S^k L) \\ & && L = [e_{i_1}, e_{i_2}, \dots, e_{i_{n_I}}] \\ & && i_j \in \{1, \dots, n_\theta\}, j = 1, \dots, n_I. \end{aligned} \quad [11]$$

This is a combinatorial problem, whose solution may be obtained by exhaustive search for small problems.⁵⁰ However, the search space grows combinatorially and this approach becomes infeasible even for moderately-sized problems. Therefore, clustering methods have been utilized to reduce the size of the problem.^{20,51} Optimization routines such as genetic algorithm (GA) may also be used, albeit

with no guarantee of convergence to the true optimal solution.¹⁹ Alternatively, the popular orthogonal method^{14,52,53} can be used to obtain a sub-optimal solution. Particularly, Chu et al.¹⁸ showed that the orthogonal method is a forward selection algorithm that yields a sub-optimal solution for φ_D^* . Forward selection in this context refers to the fact that at each iteration, the algorithm only picks the best parameter to be added to the already selected list of parameters. In that sense, the orthogonal method leads to a greedy algorithm.⁵⁴ Variations of this algorithm have been suggested to consider multiple parameters at each stage, achieving more optimal results.¹⁸ The advantage of the original orthogonal method lies in its simple implementation. For example, the QR algorithm in MATLAB can be readily used for this purpose.¹⁴ Accordingly, we use the orthogonal method to find an initial sub-optimal solution that is then refined through GA.

Since only local sensitivity data are used in the above process, the resulting subset of parameters is only locally optimal and not applicable throughout the entire parameter space. A more robust approach is therefore needed. Two such approaches have been commonly used in the literature based on the R-optimality and ED-optimality criteria.²⁶ In particular, the R-optimal selection criterion is based on a worst-case scenario and yields a selection that maximizes the minimum of the D-optimality criterion over the parameter space^{25,26}:

$$\varphi_R^* = \max_k \min_k \varphi_D(\mathbf{F}_I^k) = \max_k \min_k (\log \det(\mathbf{F}_I^k)), \quad [12]$$

On the other hand, the ED-optimality criterion maximizes the expected value of the D-optimality criterion over the parameter space^{24,26}:

$$\varphi_{ED}^* = \max \mathbb{E}[\varphi_D(\mathbf{F}_I)] = \max \mathbb{E}[\log \det(\mathbf{F}_I)], \quad [13]$$

Either criterion may be used to select a robust subset of parameters that is less dependent on the nominal parameter values. We have found the R-optimality criterion to be sensitive to outlier points in the parameter space. Therefore, we use the ED-optimality criterion in this work. The corresponding optimization problem is an integer program similar to that presented in Eq. 11, which is solved using

GA. In this case, the local solutions obtained during the previous step are used in the initial GA population to guide the optimization. The GA algorithm is run multiple times and the best solution is selected.

Using the procedure outlined in this section, we have selected a subset of 12 parameters for identification, which are highlighted in Fig. 1. The selected parameters cover a range of physical characteristics that impact the voltage and HFR predictions under a variety of operating conditions. For example, the selected subset includes parameters that impact the reaction kinetics, ohmic resistance, and mass transport losses. These are all parameters that may be inferred from voltage and HFR measurements alone. There are other parameters that could affect the internal conditions, but have no discernible impact on the measured signals, and therefore are not selected by the subset selection algorithm. Moreover, the selected subset may be affected by a range of modeler's decisions including the original parameter set considered (out of which a subset is to be selected), the bounds used for each parameter, and the operating conditions included in the predefined library. We also emphasize that this selection is optimal in an average sense, meaning that a better solution might exist in some regions of the parameter space. However, using the ED-optimality criterion helps ensure some degree of robustness to initial assumptions about the parameter values.

Robust Optimal Experimental Design for Parameter Identification

Having selected an identifiable subset of parameters, we now consider the experimental data used for identification. Importantly, we note that the data used for this purpose must be informative to ensure successful identification. This requirement motivates model-based optimal experimental design (OED), wherein the model is used to find the operating conditions that render the model predictions most sensitive to the parameters of interest. This idea is schematically shown in Fig. 2, where a numerical example is also illustrated.

The OED procedure requires solving an optimal control problem to obtain the most informative inputs. However, this approach is not

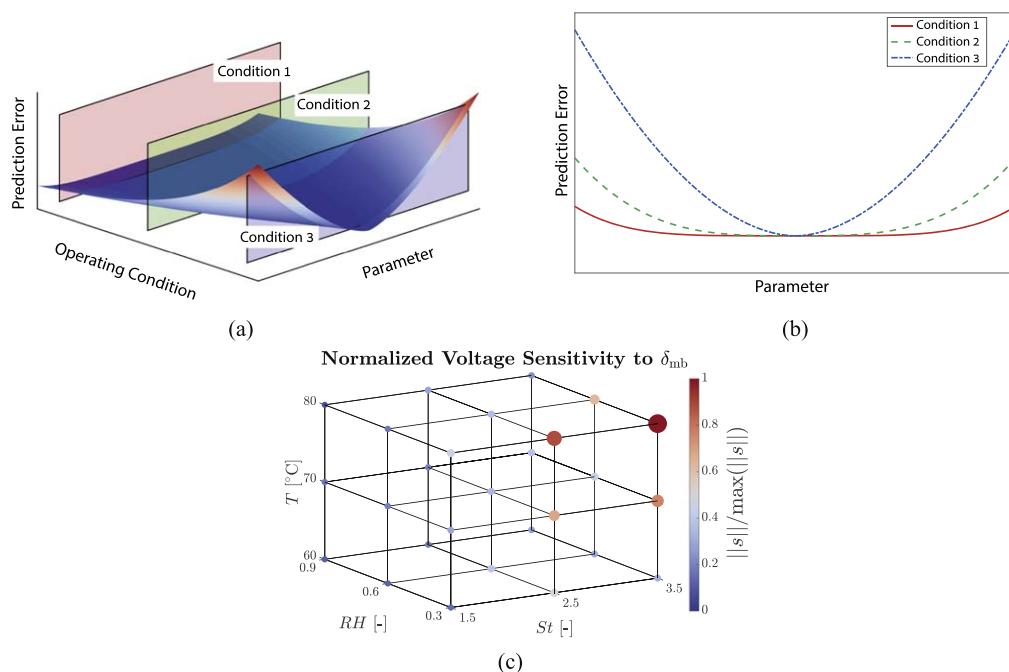


Figure 2. Optimal experimental design concept: (a) Schematic of model prediction errors as a function of parameter and operating condition, (b) schematic error variations with parameter for three example conditions, where the prediction error curve has a flat region for condition 1 rendering the parameter unidentifiable, while it is bowl-shaped for condition 3, which is ideal for parameter identifiability, and (c) example voltage sensitivities to membrane thickness, where hot and dry conditions are found to lead to increased sensitivity and improved identifiability of the membrane thickness.

computationally tractable for a large-scale model. Accordingly, methods such as input vector parameterization have been suggested,²⁶ which searches over a discrete number of inputs to find the optimal solution. Alternatively, selecting an optimal set of conditions from a predefined library has been proposed as a computationally efficient method to solve the OED problem.³² Since we have generated sensitivity data for a library of conditions in part I,⁴⁰ we use the latter approach to design maximally informative experiments for parameter identification.

Note that when working with a library of operating conditions, the OED problem is very similar to the parameter subset selection discussed earlier and the same procedures apply. Specifically, the sensitivity matrix obtained at the k -th sample point in the parameter space using the entire library of operating conditions consists of blocks that correspond to the individual experiments:

$$\mathbf{S}_I^k = \begin{bmatrix} (\mathbf{S}_I^k)_1 \\ (\mathbf{S}_I^k)_2 \\ \vdots \\ (\mathbf{S}_I^k)_{n_I} \end{bmatrix} \quad [14]$$

Here $(\mathbf{S}_I^k)_i$ denotes the sensitivity of the output predictions to the selected subset of parameters (set I) under the i -th set of operating conditions from the library of n_I conditions. We recall that in this work we have $n_I = 729$ conditions.

With the above block structure in mind, we note that while parameter subset selection is concerned with optimally choosing the columns of the sensitivity matrix to contain the most information, the OED focuses on choosing the best $(\mathbf{S}_I^k)_i$ blocks of the matrix for that same purpose. Therefore, D-optimality criterion is also commonly used for experimental design.^{22,32,35,36} The ED-optimality criterion may likewise be used for robust design.^{24,26} We note, however, that the orthogonal method is no longer applicable, since we are operating on the rows of the sensitivity matrix rather than its columns. Instead, the local D-optimal solution can be obtained through convex relaxation of the underlying integer program³²:

$$\begin{aligned} & \text{maximize} && \log \det(\tilde{\mathbf{F}}_I^k) \\ & \eta_1, \eta_2, \dots, \eta_{n_I} \\ & \text{subject to} && \tilde{\mathbf{F}}_I^k = \sum_{j=1}^{n_I} \eta_j (\mathbf{S}_I^k)_j^T (\mathbf{S}_I^k)_j \\ & && 0 \leq \eta_j \leq \frac{1}{n_u}, \quad j = 1, \dots, n_I \\ & && \sum_{j=1}^{n_I} \eta_j = 1, \end{aligned} \quad [15]$$

where n_u is the experimental budget, i.e., the number of experiments to be selected from the library. Determining the experimental budget requires consideration of the actual experimental cost, as well as the computational cost required to evaluate the model for the selected experiments. Additionally, the number of selected experiments must provide enough information for parameter identification. Taking these into consideration, here we set the number of experiments to $n_u = 6$.

The solution to the above convex program is a set of n_u experiments that are D-optimal at sample point k in the parameter space. In other words, this solution is only locally optimal. Therefore, to find an experimental design that is less dependent on the initial parameter values, the ED-optimality criterion of Eq. 13 is used.

Briefly, the robust OED procedure is as follows:

1. Choose the number of experiments, n_u , to be selected from the library.
2. Solve the convex program 15 to obtain the D-optimal design at the sampled points in the parameter space.
3. Use GA to obtain the ED-optimal design considering all of the sampled points. The D-optimal designs from the previous step are used in the initial population to guide the GA.

Table I. Robust optimal experimental design.

Exp #	p_{ch} [bar]	T_{cool} [°C]	RH_{an} [-]	RH_{ca} [-]	St_{an} [-]	St_{ca} [-]
1	2.5	80	30	30	2.5	2.0
2	1.5	80	30	90	1.5	2.0
3	1.5	40	90	90	2.0	2.5
4	1.5	80	90	30	2.0	2.5
5	1.5	80	30	30	2.5	2.0
6	1.5	80	90	30	1.5	2.5

Table II. Experimental design using LHS.

Exp #	p_{ch} [bar]	T_{cool} [°C]	RH_{an} [-]	RH_{ca} [-]	St_{an} [-]	St_{ca} [-]
1	2.68	59	80	72	2.3	2.1
2	2.05	48	54	51	2.3	1.5
3	1.79	62	65	43	1.8	2.2
4	2.85	70	85	83	1.7	1.9
5	3.33	75	48	32	2.1	2.5
6	2.45	43	33	62	1.5	1.7

The experimental design resulting from this procedure is shown in Table I. The identification results based on this robust OED procedure are compared with another experimental design using Latin Hypercube Sampling (LHS). Importantly, we use the same number of experiments ($n_u = 6$) for the LHS design. The resulting LHS design is given in Table II and the identified parameters using both experimental design procedures are compared below.

Systematic Parameter Identification

So far, we have selected a subset of parameters for identification and designed informative experiments for that purpose. The last critical piece of the model parameterization framework is the identification algorithm. A good identification algorithm should be sufficiently constrained. An algorithm with too many degrees of freedom can lead to over-fitting and poor parameter identification.⁴⁴ Regularization methods such as ridge regression^{45,46} and LASSO⁵⁵ are commonly used to constrain the identification process and formulate a well-conditioned problem. However, in the context of parameter identification, application of these methods typically requires prior information about the parameter values. Since we assume no prior information about the parameters, here we focus on using the structure of the data to constrain the identification procedure. We also recall that the parameter subset selection approach, described above, indeed regularizes the problem to a certain extent and the identification algorithm of this section is developed for further regularization.

The underlying principle of the algorithm proposed here is that the cell response shows varying sensitivity to parameter perturbations depending on the load as seen in Fig. 3. For instance, a parameter that affects reactant transport resistance might only become identifiable at higher loads. Therefore, we use this particular sensitivity structure to identify the selected parameters in multiple steps. Particularly, we divide the polarization curve and the corresponding HFR data into 3 sections based on the load (see Fig. 3): (1) low current density or kinetic region, (2) medium current density or ohmic region, and (3) high current density or mass transport region. Alternatively, the regions may be defined based on the sensitivity data, for example, by classifying the regions based on average voltage sensitivity (low, medium, and high) under different conditions. However, here we utilize the existing knowledge in the fuel cell literature, where the three aforementioned regions are often well-defined. Encoding this knowledge into the framework allows for simpler formulation.

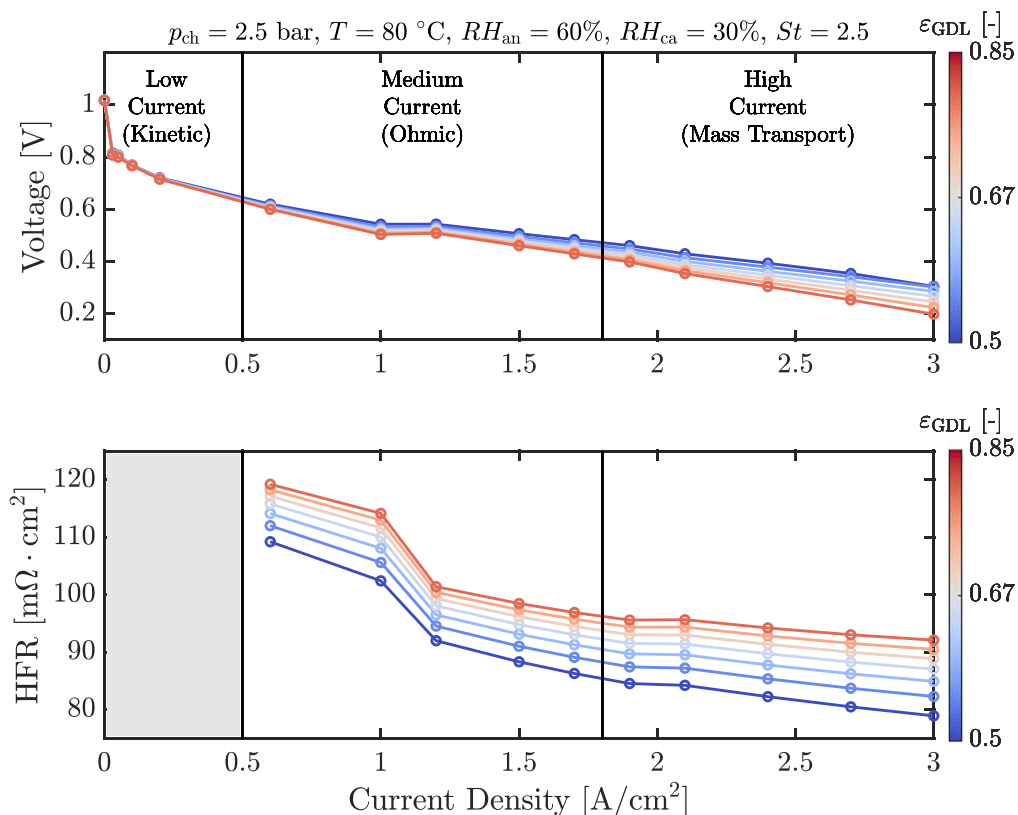


Figure 3. Sample output variations to changes in GDL porosity, where different load regions are identified. The sensitivity of the predicted outputs varies with the load.

Following the above division, the selected parameters are grouped based on the corresponding sensitivities of the predicted outputs in each region. Specifically, the first parameter group consists of those parameters that are identifiable in the low current density region of the polarization curve, while the second and third groups include the parameters that become identifiable at medium and high current densities, respectively. Identifiability in this context is measured by the average output sensitivity to the parameter of interest; a parameter is deemed identifiable in a particular region if the corresponding sensitivity is greater than a threshold value S_{th} , which is chosen to be 0.1 in this work. This grouping procedure based on output sensitivities is shown in Fig. 4.

Following the above grouping of parameters, the proposed multi-step identification algorithm is as follows:

1. Identify the parameters in the first group using data obtained at low current densities. The parameters in the other two groups are fixed at their nominal values.
2. Identify the parameters in the second group using data obtained at medium current densities. Parameters in the third group are kept at their nominal values, while those that were identified in the first step are further refined by allowing them to vary within a contracted search space.
3. Identify the parameters in the third group using data obtained at high current densities. Previously identified parameters are allowed to vary, but their respective search spaces are contracted. Particularly, the search spaces for the parameters in the first group are shrunk further compared to their ranges in the second step of the algorithm.
4. Refine all identified parameters using the entire data set. The search space for all parameters are further contracted in this step to constrain the problem. This step is only used to refine the identified values.

The general idea is that the output sensitivities to parameter perturbations typically increase as the load is increased. Therefore, it is helpful to start the identification process by only using the data at low current densities and as our confidence in the parameter estimates increases, move to higher loads and identify the remaining parameters. The increased sensitivity at higher loads also means that it is critical to allow previously identified parameters to be refined as the load is increased. Such cumulative fitting enables more accurate parameter estimation.^{31,32}

Alternative implementations.—Modified implementations of the general idea behind the presented multi-step identification algorithm are possible. A few potentially useful modifications are highlighted below:

- As we move to higher loads, our implementation disregards lower current density data points. Alternatively, one may use low current data at later stages of the identification algorithm along with larger weights for the higher current measurements (i.e., weighted least squares).
- In our implementation the problem is regularized by contracting the search space for previously identified parameters. Alternatively, the same effect may be achieved by penalizing deviations from previously identified parameters. For example, ridge regression or LASSO formulations can be utilized in later stages of the algorithm.
- We only group the parameters based on their sensitivity variations with the current density. Alternatively, this grouping can be carried out by considering all of the operating conditions, such as the humidity and temperature. However, even though this approach can lead to better regularized formulation, it would require more experimental conditions to be effective.

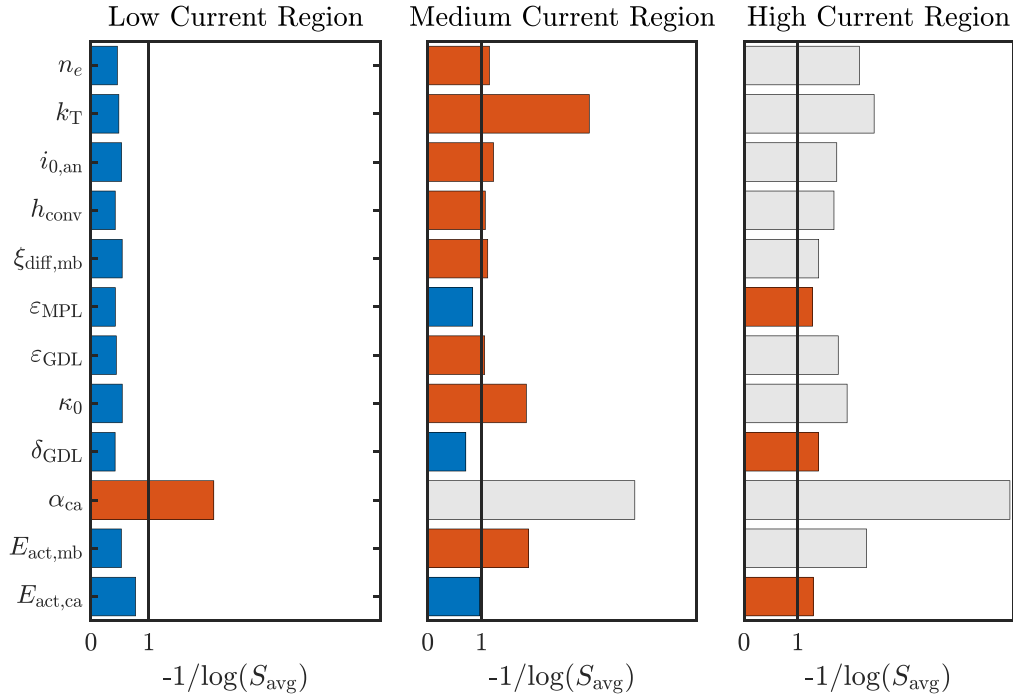


Figure 4. Combined voltage and HFR sensitivity for the entire library of operating conditions averaged among the samples in the parameter space. The blue, red, and gray bars denote, respectively, the parameters that are yet to be identified, those that are being identified in the current step, and those that were identified in a previous step and are being refined in the current step. The identifiability criterion is based on a sensitivity threshold that is chosen to be 0.1.

While we do not pursue any of the above modifications in this work, they offer some directions for further improvement of the proposed identification algorithm.

Verification of the Proposed Framework

As described above, the proposed model parameterization framework has three main components, namely, parameter subset selection, optimal experimental design, and the multi-step identification procedure. In this section, we demonstrate the utility of each of these components. To this end, we use the model to generate synthetic experimental data with a set of nominal parameter values (θ^*). The synthetic data are then used to identify the model parameters starting from perturbed values. Particularly, the optimization cost function is the weighted norm of the error:

$$J(\theta) = w_V \frac{\|e_V(\theta)\|_2}{\sqrt{n_V}} + w_R \frac{\|e_R(\theta)\|_2}{\sqrt{n_R}}, \quad [16]$$

where e_V and e_R denote the prediction error vectors for voltage (measured in mV) and HFR (measured in $\text{m}\Omega \cdot \text{cm}^2$), respectively, n_V and n_R are the number of elements in e_V and e_R , and w_V and w_R are the corresponding weights. Since the HFR varies in a smaller range compared to voltage, here we use $w_V = 1$ and $w_R = 10$. Given that we apply large perturbations to the nominal values and the above cost has many local minima, a gradient-based optimizer would fail to find the optimal solution. Use of a global optimizer is thus warranted. We have experimented with both GA and particle swarm (PS) algorithms and found that with the same level of effort to tune the optimizer hyper-parameters, the latter can find a more optimal solution. Therefore, we use PS to identify the selected parameters.

To highlight the importance of each part of the framework, we consider 4 different cases for identification:

1. Only the selected parameters (see Fig. 1) are perturbed from their nominal values. The OED conditions (Table I) are used along with the multi-step identification procedure to identify the parameters.
2. Only the selected parameters (see Fig. 1) are perturbed from their nominal values. The LHS conditions (Table II) are used along with the multi-step identification procedure to identify the parameters.
3. Only the selected parameters (see Fig. 1) are perturbed from their nominal values. The OED conditions (Table I) are used along with a single-step identification procedure, where all of the selected parameters are identified using the entire data set.
4. All parameters, except those that were removed during the collinearity analysis, are perturbed from their nominal values. The OED conditions (Table I) are used along with the systematic identification procedure to identify the parameters.

In all of the above cases, the applied parameter perturbations are random. In particular, the perturbations to the selected parameters are drawn from $\mathcal{U}(-0.3, 0.3)$ (i.e., uniform distribution between -0.3 and 0.3), while the perturbations to the parameters that are not included in the identification (case 4) are drawn from $\mathcal{U}(-0.15, 0.15)$. We also note that these perturbations are applied to the scaled parameter values and would amount to significant deviations from the nominal values on the original parameter axes. Additionally, it should be pointed out that the nominal values do not coincide with any of the sample points in the parameter space that are used to generate the sensitivity data, select the parameter subsets, and design the optimal experiments. These measures are taken to properly examine the robustness of the proposed framework.

The identified parameter values for the first two cases are provided in Table III along with the nominal and perturbed values. Moreover, the identification results for all of the above scenarios are shown in Fig. 5 in terms of the relative error of the identified parameters. Particularly, Fig. 5a compares the results from the first two scenarios and highlights the utility of OED. As seen in the figure, the maximum relative error in the parameter estimates is only 8.4% with the OED (for n_e), whereas the LHS design results in a maximum relative error of about 170% (for $\xi_{\text{diff,mb}}$). Moreover, the mean relative errors are 3.9% and 28.7% for the optimal and LHS experimental designs, respectively. In addition to these relative

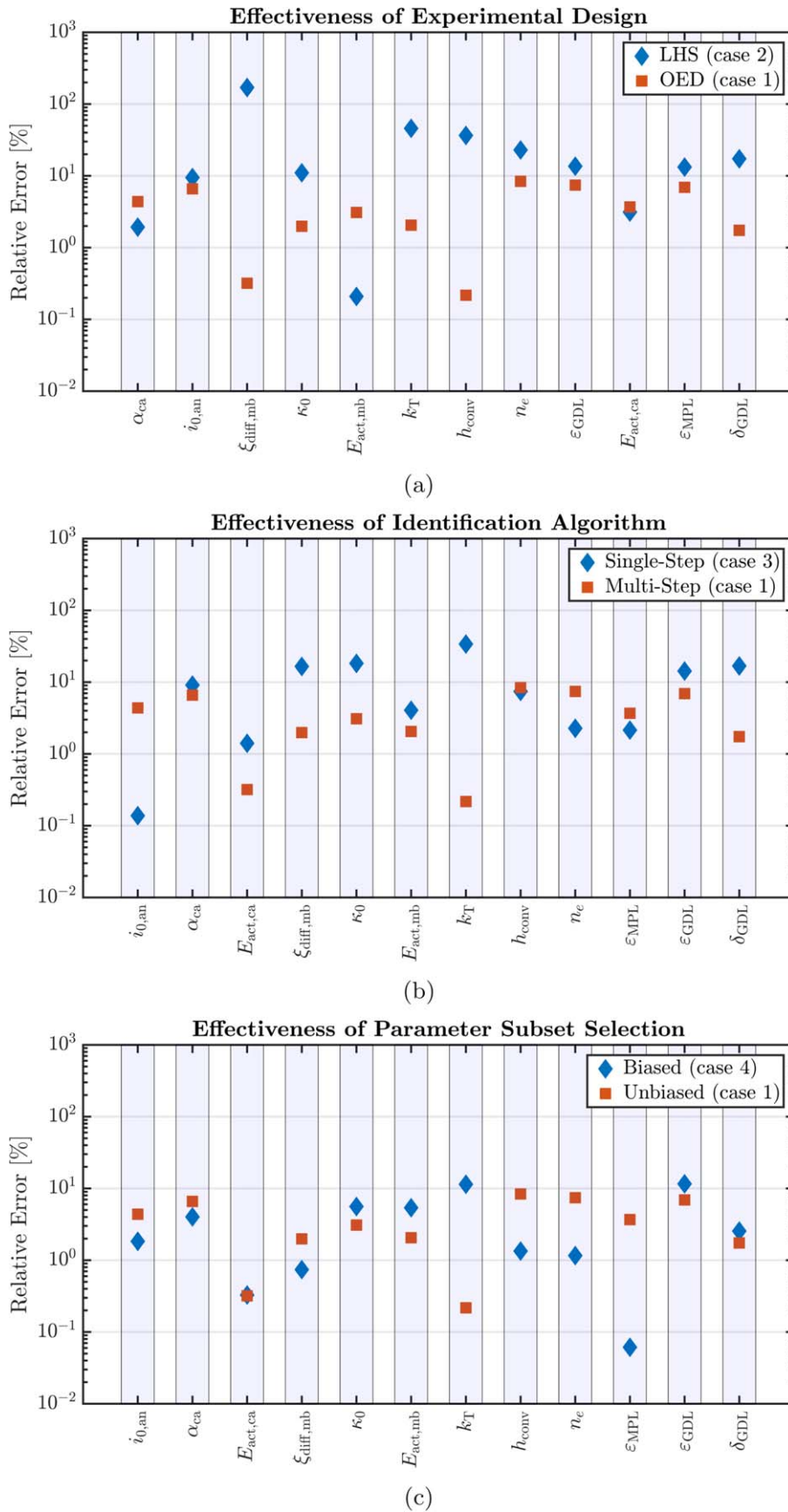


Figure 5. Relative error in parameter estimates obtained using synthetic data. Case 1 is compared to case: (a) 2, (b) 3, and (c) 4 to demonstrate the effectiveness of optimal experimental design, the multi-step parameter identification algorithm, and the subset selection method, respectively.

errors, the model voltage and HFR predictions are shown in Fig. 6. Particularly, the predictions with the nominal and identified parameter values are shown for both OED and LHS cases. Two main

observations follow: first we note that the OED conditions lead to more pronounced variations in the predictions; therefore, they are more informative. Second, we observe that for both sets of

Table III. Parameter identification results.

Parameter	True	Perturbed	LHS	Robust OED
α_{ca}	0.7	0.731	0.713	0.730
$i_{0,an}$	0.01	0.002 9	0.010 9	0.009 3
$E_{act,ca}$	67.0	58.09	64.91	64.53
$\xi_{diff,mb}$	1.0	0.317	2.698	0.997
κ_0	0.4	0.286	0.356	0.408
$E_{act,mb}$	15	13.54	14.97	14.54
k_T	1.0	0.715	0.542	1.021
h_{conv}	1.0	0.241 3	0.635	1.002
n_e	1.8	2.166	1.387	1.649
ε_{MPL}	0.60	0.581 9	0.520	0.642
ε_{GDL}	0.78	0.754	0.674	0.722
δ_{GDL}	135	156	112	137

See Table A-I for parameter descriptions and units.

conditions, the identified parameter values reproduce the synthetic experimental data very well. This result emphasizes that a good fit to experimental data does not readily imply an accurate parameter identification. These observations, combined with the relative parameter estimation errors mentioned above, demonstrate the importance of having maximally informative experimental data for parameter identification, which can be obtained through model-based OED.

To show the effectiveness of the proposed multi-step identification algorithm, Fig. 5b compares the relative errors obtained using the proposed algorithm with those obtained from a single-step algorithm. The single-step algorithm uses the entire data set to identify all of the selected parameters. Note that in both cases we use the data based on OED conditions of Table I. The figure shows that

the maximum and mean relative errors for the single-step case are 33.8% and 10.5%, respectively. This showcases the improved identification capability that is achieved by constraining the problem using the data structure.

Lastly, the subset selection approach is verified by comparing the identification results of case 1 with those of case 4, wherein the parameters that are not included in the selected subset are also perturbed from their nominal values. These results are shown in Fig. 5c. We note that if the subset selection algorithm is successful, then the parameters that are not included in the subset are not expected to have a significant impact on the output predictions. Therefore, perturbations to these parameters should not lead to considerable bias in the identified parameters. We further note that the parameters that were left out of the selected subset due to their collinearity with another parameter are not perturbed in case 4, as their perturbations can result in bias in the estimates of their collinear counterparts. As can be seen in the figure, perturbing the parameters outside of the selected subset does not have a significant impact on the relative errors in the parameter estimates. In particular, the maximum and mean relative errors are 11.6% and 3.8%, respectively, which are on par with those of case 1. Therefore, the subset selection algorithm has been successful.

These results highlight the importance of three key elements in successful parameter identification: selecting a suitable subset of parameters for identification, optimally designing experiments to generate maximally informative data for identification, and constructing a sufficiently regularized identification algorithm. Moreover, by extending the analysis beyond the vicinity of a point in the parameter space, these procedures can be made robust to initial assumptions about the nominal parameter values. We also emphasize that while a particular fuel cell model is used as an example to demonstrate the various steps involved in the parameter identification framework, the proposed procedures are generally

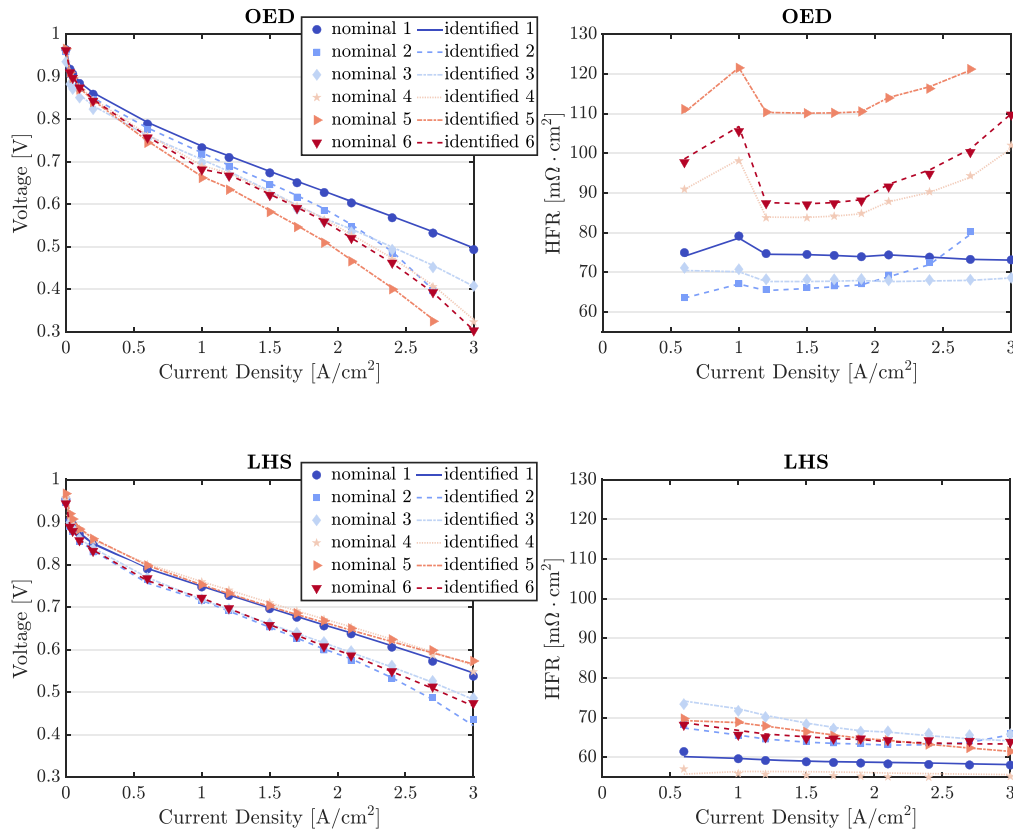


Figure 6. Polarization and HFR data obtained using OED and LHS conditions shown in Tables I and II, respectively. Results for both the nominal parameter values (markers) and the identified values (lines) are shown. The identified values are based on case 1 (for OED) and case 2 (for LHS) and are shown in Table III. Same axes limits are used to allow for direct comparison of the two experimental design results.

applicable to different types of models, including other fuel cell models.

Since effective model parameterization is crucial to successful predictive modeling, these results can have profound implications for model-based design and control strategies. In addition, parameter identification plays a significant role in inferring physical characteristics of a fuel cell from experimental data. For example, parameter identification with cell-level measurements such as voltage has been used to diagnose fuel cell degradation issues,¹⁰ determine catalyst activity,^{8,56} and evaluate mass transport losses.^{9,30} Numerous electrochemical impedance spectroscopy (EIS) studies provide further examples of such efforts.⁵⁷ Against this backdrop, the methods presented in this work also gain critical importance in further helping improve the current understanding of various aspects of fuel cell performance and durability by enabling more accurate parameter identification.

Summary and Conclusions

A framework for systematic parameter identification for PEM fuel cell models is developed. The framework consists of three main components: (1) selecting a robust subset of parameters for identification, (2) using model-based design of experiments to obtain a robust set of maximally informative data for identification of the selected parameters, and (3) using the structure of the data to formulate a regularized identification problem. This framework is based on the extended local sensitivity analysis that is explained in the first part of the study. Particularly, sensitivities of model predictions to parameter perturbations are used to determine the parameters that have the largest unique impact on the predicted outputs. This is formalized through the D-optimality criterion.

Importantly, rather than maximizing the local D-optimality criterion, we maximize its expected value over a number of sample points in the parameter space to ensure the resulting subset is less dependent on the nominal parameter values. The same procedure is applied to select the most informative operating conditions from a predefined library of conditions. Lastly, a multi-step identification procedure is proposed to regularize the optimization problem. The effectiveness of each of the components in the framework is verified through a case study, wherein synthetic experimental data generated with nominal parameter values are used to identify the selected parameters. The results demonstrate the efficacy of the proposed subset selection, optimal experimental design, and the multi-step identification procedures in improving parameter identification results. While the methods of this work are applied to a full cell model, they can also be utilized for sub-models that focus on a specific phenomenon. Therefore, we believe that adoption of such methods would help ensure that the models are properly parameterized and their full predictive capabilities are realized.

Acknowledgments

Financial and technical support by Ford Motor Company is gratefully acknowledged.

Appendix

The complete list of model parameters considered in this work is provided below. Particularly, the parameter descriptions and their units and corresponding ranges are given. Importantly, note that all parameters considered here are bounded to specific ranges, which are determined based on the corresponding uncertainty in the parameter values reported in the literature.

Table A-I. Nomenclature and parameter list.

Parameter	Description [Units]	Range
ω	Energy parameter for Temkin isotherm [$\text{kJ} \cdot \text{mol}^{-1}$]	[1, 20]
α_{an}	HOR ^{a)} transfer coefficient [-]	[1.2, 2.0]
α_{ca}	ORR ^{b)} transfer coefficient [-]	[0.4, 1.0]
$i_{0,\text{an}}$	HOR exchange current density [$\text{A} \cdot \text{cm}^{-2}$]	$[10^{-3}, 10^{-1}]$
$i_{0,\text{ca}}$	ORR exchange current density [$\text{A} \cdot \text{cm}^{-2}$]	$[10^{-9}, 10^{-6}]$
γ_{an}	HOR reaction order [-]	[0.5, 1]
γ_{ca}	ORR reaction order [-]	[0.5, 1]
$E_{\text{act,an}}$	HOR activation energy [$\text{kJ} \cdot \text{mol}^{-1}$]	[5, 20]
$E_{\text{act,ca}}$	ORR activation energy [$\text{kJ} \cdot \text{mol}^{-1}$]	[30, 80]
$E_{\text{rev,PtO}}$	Reversible potential for PtO formation [V]	[0.75, 0.85]
$E_{\text{act,PtO}}$	Activation energy for PtO formation [$\text{kJ} \cdot \text{mol}^{-1}$]	[5, 20]
$\xi_{\text{diff,mb}}$	Scaling factor for membrane diffusivity [-]	[0.1, 10]
$f_{\text{RH,a}}$	Fitting parameter for ORR RH dependency [-]	[0.4, 0.9]
$f_{\text{RH,b}}$	Fitting parameter for ORR RH dependency [-]	[0.2, 4]
$f_{\text{RH,c}}$	Fitting parameter for ORR RH dependency [-]	[0.3, 1]
κ_0	Factor for ionomer conductivity [$\text{S} \cdot \text{cm}^{-1}$]	[0.2, 0.6]
κ_{res}	Residual ionomer conductivity [$\text{S} \cdot \text{cm}^{-1}$]	$[10^{-4}, 0.006]$
$E_{\text{act,mb}}$	Membrane conductivity activation energy [$\text{kJ} \cdot \text{mol}^{-1}$]	[5, 30]
k_T	Scaling factor for thermal conductivity [-]	[0.1, 10]
h_{conv}	Heat convection coefficient [$\text{W} \cdot \text{cm}^{-2} \cdot \text{K}^{-1}$]	[0.05, 10]
n_v	Saturation correction exponent for diffusivity [-]	[1, 3]
n_e	Porosity correction exponent for diffusivity [-]	[1, 3]
K_{per}	Scaling factor for permeability [-]	[0.02, 50]
θ_{CL}	CL ^{c)} contact angle [$^\circ$]	[90.5, 100]
θ_{MPL}	MPL ^{d)} contact angle [$^\circ$]	[110, 150]
θ_{GDL}	GDL ^{e)} contact angle [$^\circ$]	[100, 130]
ε_{MPL}	MPL porosity [-]	[0.4, 0.65]
ε_{GDL}	GDL porosity [-]	[0.5, 0.85]
δ_{MPL}	MPL thickness [μm]	[20, 70]
δ_{GDL}	GDL thickness [μm]	[100, 240]
n_{Brugg}	Bruggeman coefficient for CL [-]	[1, 3]

a) Hydrogen Oxidation Reaction. b) Oxygen Reduction Reaction. c) Catalyst Layer. d) Microporous Layer. e) Gas Diffusion Layer.

References

1. A. Z. Weber et al., "A critical review of modeling transport phenomena in polymer-electrolyte fuel cells." *J. Electrochem. Soc.*, **161**, F1254 (2014).
2. A. Goshtasbi, P. García-Salaberri, J. Chen, K. Talukdar, D. G. Sanchez, and T. Ersal, "Through-the-membrane transient phenomena in PEM fuel cells: A modeling study." *J. Electrochem. Soc.*, **166**, F3154 (2019).
3. J. Zhou, D. Stanier, A. Putz, and M. Secanell, "A mixed wettability pore size distribution based mathematical model for analyzing two-phase flow in porous electrodes II. Model validation and analysis of micro-structural parameters." *J. Electrochem. Soc.*, **164**, F540 (2017).
4. I. V. Zenyuk, P. K. Das, and A. Z. Weber, "Understanding impacts of catalyst-layer thickness on fuel-cell performance via mathematical modeling." *J. Electrochem. Soc.*, **163**, F691 (2016).
5. D. Gerteisen, T. Heilmann, and C. Ziegler, "Modeling the phenomena of dehydration and flooding of a polymer electrolyte membrane fuel cell." *Journal of Power Sources*, **187**, 165 (2009).
6. A. Goshtasbi and T. Ersal, "LQ-MPC design for degradation-conscious control of PEM fuel cells." *American Control Conf.* Philadelphia, PA (IEEE) 1555 (2019) [DOI: 10.23919/ACC.2019.8814417].
7. J. Luna, E. Usai, A. Husar, and M. Serra, "Enhancing the efficiency and lifetime of a proton exchange membrane fuel cell using nonlinear model-predictive control with nonlinear observation." *IEEE Trans. Ind. Electron.*, **64**, 6649 (2017).
8. S. Jomori, K. Komatsubara, N. Nonoyama, M. Kato, and T. Yoshida, "An experimental study of the effects of operational history on activity changes in a PEMFC." *J. Electrochem. Soc.*, **160**, F1067 (2013).
9. T. A. Greszler, D. Caulk, and P. Sinha, "The impact of platinum loading on oxygen transport resistance." *J. Electrochem. Soc.*, **159**, F831 (2012).
10. L. M. Pant, Z. Yang, M. L. Perry, and A. Z. Weber, "Development of a simple and rapid diagnostic method for polymer-electrolyte fuel cells." *J. Electrochem. Soc.*, **165**, F3007 (2018).
11. P. Dobson, C. Lei, T. Navessin, and M. Secanell, "Characterization of the PEM fuel cell catalyst layer microstructure by nonlinear least-squares parameter estimation." *J. Electrochem. Soc.*, **159**, B514 (2012).
12. A. Miller, in *Subset Selection in Regression* (Chapman and Hall/CRC, Boca Raton, FL) (2002).
13. Y. Chu, Z. Huang, and J. Hahn, "Improving prediction capabilities of complex dynamic models via parameter selection and estimation." *Chem. Eng. Sci.*, **64**, 4178 (2009).
14. B. F. Lund and B. A. Foss, "Parameter ranking by orthogonalization—applied to nonlinear mechanistic models." *Automatica*, **44**, 278 (2008).
15. R. Brun, P. Reichert, and H. R. Künsch, "Practical identifiability analysis of large environmental simulation models." *Water Resour. Res.*, **37**, 1015 (2001).
16. T. Quaiser and M. Mönnigmann, "Systematic identifiability testing for unambiguous mechanistic modeling-application to JAK-STAT, MAP kinase, and NF- κ B signaling pathway models." *BMC Syst. Biol.*, **3**, 50 (2009).
17. H. Miao, X. Xia, A. S. Perelson, and H. Wu, "On identifiability of nonlinear ODE models and applications in viral dynamics." *SIAM Review*, **53**, 3 (2011).
18. Y. Chu and J. Hahn, "Generalization of a parameter set selection procedure based on orthogonal projections and the d-optimality criterion." *AIChE J.*, **58**, 2085 (2012).
19. Y. Chu and J. Hahn, "Parameter set selection for estimation of nonlinear dynamic systems." *AIChE J.*, **53**, 2858 (2007).
20. Y. Chu and J. Hahn, "Parameter set selection via clustering of parameters into pairwise indistinguishable groups of parameters." *Indust. Eng. Chem. Res.*, **48**, 6000 (2008).
21. K. A. McLean, S. Wu, and K. B. McAuley, "Mean-squared-error methods for selecting optimal parameter subsets for estimation." *Indust. Eng. Chem. Res.*, **51**, 6105 (2012).
22. C. P. Kitsos, in *Optimal Experimental Design for Non-linear Models: Theory and Applications* (Springer Science & Business Media, New York) (2014) [DOI: 10.1007/978-3-642-45287-1].
23. L. Pronzato, "Optimal experimental design and some related control problems." *Automatica*, **44**, 303 (2008).
24. L. Pronzato and É. Walter, "Robust experiment design via stochastic approximation." *Math. Biosci.*, **75**, 103 (1985).
25. L. Pronzato and É. Walter, "Robust experiment design via maximin optimization." *Math. Biosci.*, **89**, 161 (1988).
26. S. Asprey and S. Macchietto, "Designing robust optimal dynamic experiments." *J. Proc. Control*, **12**, 545 (2002).
27. R. Schenkendorf, X. Xie, M. Rehbein, S. Scholl, and U. Krewer, "The impact of global sensitivities and design measures in model-based optimal experimental design." *Processes*, **6**, 27 (2018).
28. A. Goshtasbi, J. Chen, J. R. Waldecker, S. Hirano, and T. Ersal, "On parameterizing PEM fuel cell models." *American Control Conf.* Philadelphia, PA (IEEE) 903 (2019) [DOI: 10.23919/ACC.2019.8815294].
29. J. P. Owejan, J. E. Owejan, and W. Gu, "Impact of platinum loading and catalyst layer structure on PEMFC performance." *J. Electrochem. Soc.*, **160**, F824 (2013).
30. D. R. Baker, D. A. Caulk, K. C. Neyerlin, and M. W. Murphy, "Measurement of oxygen transport resistance in PEM fuel cells by limiting current methods." *J. Electrochem. Soc.*, **156**, B991 (2009).
31. L. Zhang, C. Lyu, G. Hinds, L. Wang, W. Luo, J. Zheng, and K. Ma, "Parameter sensitivity analysis of cylindrical LiFePO₄ battery performance using multi-physics modeling." *J. Electrochem. Soc.*, **161**, A762 (2014).
32. S. Park, D. Kato, Z. Gima, R. Klein, and S. Moura, "Optimal experimental design for parameterization of an electrochemical lithium-ion battery model." *J. Electrochem. Soc.*, **165**, A1309 (2018).
33. A. Pozzi, G. Ciarabella, S. Volkwein, and D. M. Raimondo, "Optimal design of experiments for a lithium-ion cell: parameters identification of an isothermal single particle model with electrolyte dynamics." *Indust. Eng. Chem. Res.*, **58**, 1286 (2018).
34. A. P. Schmidt, M. Bitzer, Á. W. Imre, and L. Guzzella, "Experiment-driven electrochemical modeling and systematic parameterization for a lithium-ion battery cell." *J. Power Sources*, **195**, 5071 (2010).
35. M. J. Rothenberger, D. J. Docimo, M. Ghanaatpishe, and H. K. Fathy, "Genetic optimization and experimental validation of a test cycle that maximizes parameter identifiability for a Li-ion equivalent-circuit battery model." *J. Energy Storage*, **4**, 156 (2015).
36. J. Forman, J. Stein, and H. Fathy, "Optimization of dynamic battery parameter characterization experiments via differential evolution." *American Control Conf.* Washington, DC (IEEE) 867 (2013) [DOI: 10.1109/ACC.2013.6579945].
37. A. Goshtasbi, B. L. Pence, and T. Ersal, "Computationally efficient pseudo-2D non-isothermal modeling of polymer electrolyte membrane fuel cells with two-phase phenomena." *J. Electrochem. Soc.*, **163**, F1412 (2016).
38. A. Goshtasbi, B. L. Pence, and T. Ersal, "A real-time pseudo-2D bi-domain model of PEM fuel cells for automotive applications." *Dynamic Systems and Control Conf.* Tysons Corner, VA V001T25A001 (2017) [DOI: 10.1115/DSCC2017-5053].
39. A. Goshtasbi, B. L. Pence, J. Chen, M. A. DeBolt, C. Wang, J. R. Waldecker, S. Hirano, and T. Ersal, "A mathematical model toward real-time monitoring of automotive PEM fuel cells." *J. Electrochem. Soc.*, **167**, 024518 (2020).
40. A. Goshtasbi, J. Chen, J. Waldecker, S. Hirano, and T. Ersal, "Effective Parameterization of PEM Fuel Cell Models—Part I: Sensitivity Analysis and Parameter Identifiability." *Journal of the Electrochemical Society*, **167**, 044504 (2020).
41. A. Holiday, M. Kooshkbaghi, J. M. Bello-Rivas, C. W. Gear, A. Zagaris, and I. G. Kevrekidis, "Manifold learning for parameter reduction." *J. Comput. Phys.*, **392**, 419 (2019).
42. M. C. Eisenberg and M. A. Hayashi, "Determining identifiable parameter combinations using subset profiling." *Math. Biosci.*, **256**, 116 (2014).
43. J. D. Stigter and J. Molenaar, "A fast algorithm to assess local structural identifiability." *Automatica*, **58**, 118 (2015).
44. R. C. Aster, B. Borchers, and C. H. Thurber, in *Parameter Estimation and Inverse Problems* (Elsevier, London) (2018) [DOI: 10.1016/C2009-0-61134-X].
45. C. Kravaris, J. Hahn, and Y. Chu, "Advances and selected recent developments in state and parameter estimation." *Comput. Chem. Eng.*, **51**, 111 (2013).
46. D. P. Howsmon and J. Hahn, "Regularization techniques to overcome over-parameterization of complex biochemical reaction networks." *IEEE Life Sci. Lett.*, **2**, 31 (2016).
47. L. N. Trefethen and D. Bau III, in *Numerical Linear Algebra* (SIAM, Philadelphia) (1997).
48. X. Lin and A. G. Stefanopoulou, "Analytic bound on accuracy of battery state and parameter estimation." *J. Electrochem. Soc.*, **162**, A1879 (2015).
49. E. Walter, in *Identifiability of Parametric Models* (Pergamon Press, Oxford) (1987).
50. S. R. Weijers and A. Vanrolleghem, "A procedure for selecting best identifiable parameters in calibrating activated sludge model no. 1 to full-scale plant data." *Water Sci. Technol.*, **36**, 69 (1997).
51. W. Dai, L. Bansal, J. Hahn, and D. Word, "Parameter set selection for dynamic systems under uncertainty via dynamic optimization and hierarchical clustering." *AIChE J.*, **60**, 181 (2014).
52. K. Z. Yao, B. M. Shaw, B. Kou, K. B. McAuley, and D. Bacon, "Modeling ethylene/butene copolymerization with multi-site catalysts: parameter estimability and experimental design." *Polym. React. Eng.*, **11**, 563 (2003).
53. M. Fink, A. Attarian, and H. Tran, "Subset selection for parameter estimation in an HIV model." *PAMM: Proc. Appl. Math. Mech.*, **7**, 1121501 (2007).
54. T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, in *Introduction to Algorithms* (MIT Press, Cambridge, MA) (2009).
55. T. Hastie, R. Tibshirani, J. Friedman, and J. Franklin, "The elements of statistical learning: data mining, inference and prediction." *Math. Intelligencer*, **27**, 83 (2005).
56. Á. Kriston, T. Xie, D. Gamliel, P. Ganesan, and B. N. Popov, "Effect of ultra-low Pt loading on mass activity of polymer electrolyte membrane fuel cells." *J. Power Sources*, **243**, 958 (2013).
57. M. Heinzmann, A. Weber, and E. Ivers-Tiffée, "Advanced impedance study of polymer electrolyte membrane single cells by means of distribution of relaxation times." *J. Power Sources*, **402**, 24 (2018).