Effective Parameterization of PEM Fuel Cell Models—Part I: Sensitivity Analysis and Parameter Identifiability

To cite this article: Alireza Goshtasbi et al 2020 J. Electrochem. Soc. 167 044504

View the article online for updates and enhancements.
Mathematical models have become indispensable tools in fuel cell research and development efforts. These models, similar to those of other electrochemical energy systems, such as batteries, usually have tens of parameters that may affect their predictions. This is especially the case for physics-based models that are derived from first principles, as such models often incorporate various physical attributes of a cell along with detailed kinetics of electrochemical reactions. Given that the prediction capabilities of these models depend not only on their structure (i.e., the first principles and assumptions used in deriving the model), but also on the availability of accurate parameter values, the problem of parameter identification calls for critical investigation.

Despite its centrality to model accuracy and reliability, parameter identification has not received much attention in the fuel cell literature, where arbitrarily assigning parameter values,1–5 or doing so based on expert knowledge,6–9 and citing qualitative agreement with experimental data as evidence for model validity is common.10–12 In other instances, parameter identification is treated as material characterization,9 wherein the model parameters are obtained through component characterization techniques. While these methods are crucial to enhance understanding of fuel cell behavior and improve model performance, they are often costly and require specialized equipment. Moreover, even extensive characterization efforts could only unveil a handful of parameters, typically with considerable uncertainty.10 For instance, significant efforts have been aimed at determining thermal conductivity of various layers in a fuel cell assembly, often resulting in values that vary by orders of magnitude for the same material set.11,12 Lastly, even obtaining all model parameters through such ex situ experiments would not guarantee accurate model predictions. This lack of prediction accuracy is mainly due to the inherent and inevitable simplifications of mathematical models. However, this does not indicate that the models do not have predictive power. Rather, it motivates effective parameter identification as a necessary step toward unleashing the full predictive capability of physics-based models. Against this backdrop, we seek a parameter identification approach that can be effective in accurately determining model parameters even with limited and non-invasive measurements.

The fact that parameter identification remains an understudied subject in the fuel cell community is widely attributed to the computational cost of available physics-based models that do not lend themselves to systematic optimization-based model parameterization, which typically require a large number of model evaluations. However, some in the fuel cell community have already utilized these techniques. For example, nonlinear least squares have been used to fit models of oxygen reduction reaction (ORR) to kinetic data.13,14 Others have used polarization data to identify cell-level model parameters with these methods.15,17 While these works stand out in terms of their efforts for systematic parameter identification, they are still limited in their application. Specifically, they only identify a few of the model parameters using optimization techniques and rely on literature values for the other parameters. More importantly, the choice of parameters to be identified through optimization and the operating conditions used to collect experimental data is made arbitrarily or based on the researchers’ physical intuition about the problems. This amounts to an important gap in the fuel cell modeling literature, which the present work aims to address.

Fortunately, the problem of parameter identification is common among many fields, some of which offer a vast body of literature on the topic. Most importantly, the recent battery modeling literature has focused extensively on parameter identification and closely related problems. In particular, sensitivities of the well-known models have been studied,18–21 parameter identifiability of the models have been investigated,22–26 various methods for simultaneous identification of multiple model parameters have been developed,27 and Bayesian statistics have been used to infer parameter distributions.28 More recently, there has been a move toward optimally designing experiments for parameter identification.29–32 In addition to this recent battery literature, other studies in fields ranging from system biology33,34 to process industry35 offer further basis for systematic parameter identification efforts.

Drawing motivation from this literature, our objective is to fill the above-mentioned gap by developing a systematic approach for parameter identification tailored specifically to fuel cell applications. In particular, we aim to answer the following questions: How are different model outputs affected by model parameters? What are the major difficulties associated with large-scale parameter identification...
problems? How many and which of the model parameters can be accurately identified given a particular set of measurements? How can we optimally design the experiments for the purpose of parameter identification? And finally, how do we ensure that our results are robust to initial assumptions about the values of model parameters?

To answer these questions, in this two-part study we develop a systematic procedure for parameter identification with limited and non-invasive measurements using as an example a recently developed physics-based model that accounts for two-phase phenomena and non-isothermal behavior of an operating fuel cell. Importantly, the computational cost of the model is low enough for a large number of model evaluations. In this part of the study we use the aforementioned model to investigate and analyze identifiability of its parameters using a sensitivity-based approach. Particularly, we start with an extensive sensitivity analysis that considers a large number of model parameters. This analysis is conducted considering a variety of operating conditions that can be used to collect experimental data. Moreover, in order to make our approach robust to initial assumptions about the parameter values, we carry out the sensitivity analysis at several points in the parameter space to obtain a more global picture of the sensitivity results. Using these sensitivity results, we study the identifiability of the model parameters and select a subset of parameters for identification. We also demonstrate how different subset selection algorithms can lead to different identifiability properties. These results are further expanded and utilized in the second part of the study, wherein we develop a framework for effective model parameterization and propose a robust optimal experimental design to generate more informative data for parameter identification.

The rest of the paper is organized as follows. First, the parameter identification problem is outlined. This is followed by a description of the methods used for sensitivity analysis and a discussion of the main results, where sensitivities of various model outputs to the considered parameters are investigated. Lastly, identifiability analysis and relevant issues are discussed, before concluding with a brief summary.

Parameter Identification Problem

The parameter identification problem consists of finding the model parameters given some measured data, whose behavior is to be captured by the model. To make this more concrete, first we note that the model can be described using a set of differential-algebraic equations (DAEs):

\[
\begin{align*}
\dot{x} &= f(x, z, u, \theta), & x(0) &= x_0 \\
0 &= g(x, z, u, \theta), & z(0) &= z_0 \\
\hat{y} &= h(x, z, u, \theta),
\end{align*}
\]

where \(x\) is the vector containing dynamic states of the model (e.g., membrane temperature and hydration), \(\hat{y}\) is the vector of predicted outputs (e.g., cell voltage), \(z\) is the algebraic variables (e.g., solid phase potential when double layer capacitance is ignored), \(u\) is the vector of inputs (e.g., channel pressure and coolant temperature), and \(\theta\) denotes the vector of model parameters (e.g., exchange current density for oxygen reduction reaction). The vector valued functions \(f, g, h\) represent the system dynamics, algebraic constrains, and output equations, respectively. Note that while fuel cell models are typically described with partial differential equations (PDEs), all such models may be written in DAE format after proper spatial discretization. We further note that while a model can make predictions about internal states of the cell (i.e., \(\hat{y}\) can contain elements of \(x\)), we use \(\hat{y}\) to denote only the model predictions for which experimental data can be measured. Henceforth, we refer to \(\hat{y}\) as predicted outputs, with the understanding that the corresponding measurements can be taken in an actual experimental setting. Given the above generic model description, parameter identification is to find a suitable vector \(\theta\) such that the model predictions given by \(\hat{y} = h(x, z, u, \theta)\) match the available measurements, \(y\), as closely as possible.

The parameter identification problem can be solved using a variety of methods, among which least squares is the most common approach. However, an identifiability analysis must precede parameter identification. The goal of such analysis is to determine whether the parameters of interest can be uniquely identified given a particular set of measurements: e.g., can the thermal conductivity of cell layers be determined using only voltage measurements? Identifiability can be analyzed in both a local and a global sense. A parameter is locally identifiable if it can be determined uniquely in a neighborhood around a nominal parameter value, given a set of measurements; no two parameter values that are close enough would result in the same model predictions. On the other hand, global identifiability indicates that no two parameter values, no matter how far apart, can give the same predictions. This means that among all possible parameter values, there is only one that can best represent the given measurements.

Given these definitions, an identifiability analysis of model parameters prior to parameter identification is warranted. To carry out such analysis, one must understand what the root causes of unidentifiability are. Particularly, what can render a model parameter unidentifiable? An obvious answer is the case where a model parameter has no impact on the predicted outputs. A more subtle case is when a combination of model parameters impacts the predicted outputs together. For instance, if a model has two parameters, \(a\) and \(b\), but they always appear as a product \((ab)\) throughout the model, then neither of these parameters will be identifiable. In fuel cell models, this happens, for example, with exchange current density and volumetric surface area of Pt in the catalyst, as they always appear as a product in the Butler-Volmer model of reaction kinetics. These examples amount to what is known as structural unidentifiability, i.e., the model structure does not allow for unique identification of these parameters even if every variable of interest could be measured with perfect precision.

This notion of identifiability, while useful, falls short in practical settings. In particular, a model parameter might be structurally identifiable but require a lot of measurements with high precision to be identified. In such cases, practical identifiability yields better insight, as it accounts for the limited quantity and precision of the measurements. Readers interested in a more formal treatment of the subject are encouraged to consult the review paper by Miao et al. and references therein.

A variety of methods have been reported in the literature to investigate both practical and structural identifiability issues. Structural identifiability analysis requires the structure of the model to be known and, having its roots in control theory, usually utilizes tools such as differential algebra and implicit function theorem, to name a few. One of the more recent developments in this area is the use of extended observability analysis for this purpose. Despite such rich history, these methods are usually limited to small problems and do not lend themselves to large-scale models. Fortunately, practical identifiability analysis does not suffer from such a shortcoming and may be readily applied to large-scale problems. In this respect, Monte Carlo methods, correlation analysis, and sensitivity-based methods have been developed in the literature. Among these, the methods based on sensitivity analysis are most commonly utilized, as they scale easily and also provide information about the structure of the model without the need for a full-blown structural identifiability analysis. Accordingly, we use a sensitivity-based identifiability analysis method to closely investigate the structure of our model and study the identifiability of its parameters.

Sensitivity analysis is a rich topic in itself with a vast body of literature. A good sensitivity analysis requires consideration of different aspects of the problem. Therefore, in the next sections we discuss this in detail and present our sensitivity analysis results before moving on to identifiability analysis.
**Sensitivity Analysis: Methods**

Sensitivity analysis is commonly used as a tool to investigate the impact of different input factors (e.g., operating conditions or model parameters) on outputs of interest. Examples of sensitivity analysis in the fuel cell literature usually include perturbing input factors of interest individually to quantify each factor’s significance.\(^5.47-49\)

More recent studies have conducted rigorous sensitivity analyses on fuel cell models.\(^{50-52}\) In this regard, the recent work by Vetter et al.\(^{51}\) stands out, as they consider a relatively large set of parameters and also conduct a global sensitivity analysis, where they find that the parameters impacting membrane water uptake and transport to be of profound importance. It is based on this important conclusion that they suggest efforts should be aimed at accurate estimation of such parameters. This statement offers a different viewpoint for sensitivity analysis: while sensitivity results have been typically used to guide research efforts in optimizing the design and material properties from a performance standpoint, they can be interpreted in a different light to enhance the predictive power of fuel cell models. In other words, in most of the previous cases sensitivity analysis can be considered as a use case for the model to improve fuel cell designs, whereas Vetter et al.’s suggestion amounts to using sensitivity analysis in order to refine the model itself by proper parameter identification. In this sense, the present study is similar to that of Vetter et al.\(^{51}\) However, we extend their results in several important ways: (1) the model used in this study is more comprehensive than Vetter et al.’s and it incorporates transport effects along the flow channels, (2) we consider a larger set of model parameters, and (3) we extensively analyze the sensitivity results in terms of the dependencies between the parameters as suggested by sensitivity directions and not only sensitivity magnitudes that were used in Vetter et al.’s study. For completeness, a detailed account of our methods is presented after a brief overview of the model and the parameters considered for this study.

**Fuel cell model and parameters.**—The model used in this study is a pseudo-2D transient model of PEM fuel cells that was developed for automotive applications and explained in detail in earlier publications.\(^{56-58}\) Briefly, the model is non-isothermal, accounts for two-phase flow, and incorporates transport in the through-the-membrane and along-the-channel directions. When appropriate, reduced-order sub-models are utilized to achieve computational efficiency. While the model formulation allows for counter-flow simulations and also captures the transport under the lands, the version of the model adopted for this study is co-simulations and also capturing the transport under the lands, making the computational cost for the current study manageable. Further details of the model can be found in the original publications\(^{56-58}\) and are omitted for brevity.

The model has more than a hundred parameters. This includes physical constants and parameters that describe the stack geometry, reaction kinetics, physical properties of the material set, as well as several fitting coefficients used to match the model predictions with experimental data. Many of the model parameters are known (e.g., the ionomer water uptake isotherms) and are not subject to sensitivity analysis. For the purpose of this study, we have chosen 50 of the model parameters from the parameter classes mentioned above. Most of the selected parameters can be potential candidates for identification using experimental data. Although geometrical parameters are usually known for a given fuel cell stack, they are nonetheless included in our sensitivity analysis, because the design geometry can impact the sensitivity of other model parameters. Including the geometrical parameters is thus important to ensure that our results are not constrained by a specific stack design. The chosen parameters are shown in the Appendix Table A.1 along with their description, units, and corresponding ranges. A complete list of model parameters containing their default values is provided in the original modeling work and is not repeated here for space considerations.\(^{58}\)

**Sensitivity calculations.**—A variety of methods have been suggested in the literature to calculate model sensitivities. Among them, the derivative-based approach is the most commonly used due to its easy and possibly efficient implementation. This approach estimates the sensitivities as follows:

$$ S = \frac{\partial \bar{y}}{\partial \bar{\theta}} $$

where \( S \) is the sensitivity matrix, whose \((i, j)\)-th entry denotes the sensitivity of the \(i\)-th predicted output to the \(j\)-th parameter. When multiple outputs and parameters are considered, a scaled version is typically used to ensure the results are not affected by the choice of units. Naturally, the sensitivity results obtained with this method are only valid in the immediate neighborhood of the nominal parameter values that are used to evaluate the model.

To better explore the entire range of individual parameters and obtain sensitivities that are more representative of the sensitivity values throughout this range, here we adopt a sample-based approach. In particular, to calculate the sensitivity of parameter \(i\) at any point in the parameter space, we start by mapping the parameter range onto a unit interval using linear or logarithmic min-max scaling:

$$ \hat{\theta}_i = \frac{\theta_i - \theta_{i,\text{min}}}{\theta_{i,\text{max}} - \theta_{i,\text{min}}} $$

$$ \log \hat{\theta}_i = \log \frac{\theta_i - \log \theta_{i,\text{min}}}{\log \theta_{i,\text{max}} - \log \theta_{i,\text{min}}} $$

where \( \hat{\theta}_i \) is the scaled parameter value that lies between 0 and 1, and \( \theta_{i,\text{min}} \leq \theta_i \leq \theta_{i,\text{max}} \). Logarithmic scaling is used when the parameter range spans more than two orders of magnitude, while linear scaling is used otherwise (see Appendix Table A.1). We then take \( n_s \) equidistant samples from the unit interval and evaluate the model outputs using the corresponding parameter values. The \((j, i)\)-th element of the sensitivity matrix is then given by:

$$ s_{j,i} = \left( \frac{n_s}{\sum_{k=1}^{n_s} y_{j,k}} \right) \frac{1}{n_s} \sum_{k=1}^{n_s} \left( \frac{y_{j,k+1} - y_{j,k}}{\hat{\theta}_{i,k+1} - \hat{\theta}_{i,k}} \right) $$

where \( y_{j,k} \) denotes the \(j\)-th output evaluated using the \(k\)-th sample and \( \hat{\theta}_{i,k} \) is the \(k\)-th sample of parameter \(i\). Note that the sensitivity is normalized by the mean of the output and is therefore dimensionless. We use \( n_s = 7 \) samples in this work.

The above approach to calculating parameter sensitivities is motivated by the fact that the parameter perturbations during the analysis should be on par with those applied during parameter identification.\(^{59}\) However, the successful application of such a method depends on the particular problem structure. Specifically, if perturbing a parameter can result in non-monotonic variations in the model output, then this analysis may break down. Nonetheless, the outputs of our model vary monotonically in most cases. Non-monotonic behavior is typically restricted to certain regions in a parameter interval, which can be effectively captured by increasing the number of samples. This method of sensitivity calculation allows us to better investigate the entire range of any given parameter rather than restricting the analysis to the immediate neighborhood of the nominal value.

While exploring the entire range of the parameter under consideration, the above method keeps the other parameters fixed at their nominal values. Therefore, it still results in a local analysis. A global analysis is needed to inspect the parameter interaction, wherein all parameters are varied simultaneously to calculate the sensitivity results.\(^{54}\) However, global analysis typically utilizes Monte-Carlo methods, which require evaluating the model at a large number of sampled points in the parameter space for convergence of the multi-dimensional integrals. Therefore, such analysis often becomes computationally infeasible for large-scale models.
Given the shortcomings of local and global analyses, we seek an approach that strikes a balance between its validity region and the computational burden associated with it. Accordingly, we conduct an extended local analysis, wherein sensitivities are analyzed locally about several sampled points in the parameter space. This extended analysis yields a more global picture of the model sensitivities. Analogous methods have been used as efficient screening tools to narrow down the number of parameters. Based on the result of this extended analysis, the following statistics can be calculated:

\[ s_i = \text{med} \left\| s_i^k \right\|_2 \]
\[ s_i^k = \frac{\text{var} \left\| s_i^k \right\|}{\sqrt{k}} \]

where \( s_i^k \) denotes the sensitivity vector for the \( i \)-th parameter evaluated at the \( k \)-th sample point in the parameter space, \( s_i \) is the median of sensitivity magnitudes for parameter \( i \), and \( s_i^k \) denotes the corresponding standard deviation. We use the median, which is chosen due to its robustness to outliers, as an indicator of the overall parameter sensitivity. On the other hand, the standard deviation indicates how significantly a given parameter’s sensitivity depends on other model parameters.

So far, we have focused on the magnitude of the sensitivity vectors. While this is an informative measure to quantify the impact of each parameter, it does not offer any insight into parameter interactions. For instance, if two parameters have large sensitivity magnitudes but have similar effects on the outputs, they cannot be uniquely identified and the resulting parameter identification problem is ill-conditioned. Therefore, to quantify the similarity between two sensitivity vectors, \( s_i^k \) and \( s_i^l \) evaluated at sample point \( k \) in the parameter space, we use the following collinearity index:

\[ \psi_{ij}^k = \cos(\phi_{ij}^k) = \frac{\left\| s_i^k \right\| \left\| s_j^k \right\|}{\left\| s_i^k \right\|_2 \left\| s_j^k \right\|_2} \]

where \( \phi_{ij}^k \) denotes the angle between the sensitivity vectors and the collinearity index, \( \psi_{ij}^k \), lies between 0 and 1, indicating orthogonal and collinear sensitivity directions, respectively.

It should be noted that if the pair of sensitivity vectors have small magnitudes, this collinearity index is of small utility. In such a case, the parameters are non-influential and even if they are completely independent of one another, they cannot be effectively identified. Therefore, in our analysis in the next section, we only highlight the collinearity indices for the influential parameters as characterized by their high sensitivity magnitudes. Furthermore, note that the above definition of collinearity only allows for analyzing pairs of parameters. Therefore, we only use this definition to inspect parameter pairs and visualize the similarity between the effects of different parameter pairs without the need for coordinate transformations. However, when selecting parameter subsets for optimal identification results, we utilize more elaborate methods that allow for analyzing the linear dependence between a set of selected parameters (see the section on Identifiability Analysis and Parameter Subset Selection).

### Sampling of the parameter space for extended local analysis

As mentioned previously, we conduct local sensitivity analysis about several points in the parameter space to obtain a more globally valid picture of the model sensitivities. Choosing the appropriate method to sample the parameter space is thus critical. Particularly, our emphasis is on efficiently exploring the entire parameter space with the fewest number of points. This is especially important due the fact that the space under consideration is a 50-dimensional space. To this end, we map this high dimensional parameter space into a unit hypercube using min-max scaling as described before. We then use the quasi-random Sobol sequence to efficiently sample the entire space. The main advantage of this sequence is that it has low discrepancy, meaning that the gaps between the generated points are relatively uniform. Using MATLAB built-in function for the Sobol sequence, we generate 350 samples in the 50-dimensional unit hypercube.

One caveat of generating the points in the parameter space using the Sobol sequence is that it is physics-agnostic. Therefore, many of the parameter combinations might yield unreasonable results even if the ranges for each of the parameters were reasonably constrained. To remedy this, we use an initial screening step. This step entails simulating the model for the parameter values generated by the Sobol sequence using automotive operating conditions. In particular, we use 4 different conditions that represent a variety of driving scenarios. Most importantly, we also have experimental data under these conditions with state-of-the-art membrane electrode assembly (MEA) materials (see Stack A data in the original modeling work for further information). For each sample in the Sobol sequence, the model’s voltage predictions are compared with the corresponding experimental measurements and the sample is accepted if:

\[ ||e_{\text{exp}}||_{\infty} = ||V_{\text{model}} - V_{\text{data}}||_{\infty} \leq \epsilon_{\text{max}}. \]

where \( V_{\text{model}} \) and \( V_{\text{data}} \) are the model predictions and the available measurements, respectively, and \( \epsilon_{\text{max}} \) denotes the error threshold that is chosen to be 300 mV in this work. With this choice of threshold, 55 samples are accepted. Given the high dimensionality of the parameter space under consideration, this is a very small number of samples, which is constrained by our computational capabilities. Nonetheless, as is shown in the next section, the particular structure of the model leads to reasonable convergence of the sensitivity results even with this small number of samples.

### Library of operating conditions

For a complete sensitivity and identifiability analysis, a variety of operating conditions must be considered. This is especially important, since some operating conditions may be more informative for parameter identification. Accordingly, we consider the entire space of potential operating conditions and use a full-factorial design with 3 levels to explore this space. In particular, we consider variations in the channel pressure (\( p_{ch} \)), the coolant inlet temperature (\( T_{\text{cool}} \)), the anode and cathode relative humidity (\( RH_{an}, RH_{ca} \)), and the anode and cathode stoichiometric ratio (\( St_{an}, St_{ca} \)) as follows:

\[ p_{ch} \in [1.5, 2.5, 3.5] \text{ [bar]}, \]
\[ T_{\text{cool}} \in [40, 60, 80] \text{ [°C]}, \]
\[ RH_{an} \in [30, 60, 90] \text{ [-]}, \]
\[ RH_{ca} \in [30, 60, 90] \text{ [-]}, \]
\[ St_{an} \in [1.5, 2.0, 2.5] \text{ [-]}, \]
\[ St_{ca,\text{led}} \in [1.5, 3.0, 4.5] \text{ [-]}, \]
\[ St_{ca,\text{bed}} \in [1.5, 2.0, 2.5] \text{ [-]}, \]

where the cathode stoichiometric ratio changes based on the current density: \( St_{\text{bed,calc}} \) values are used at current densities below 1 A cm\(^{-2}\) and \( St_{\text{bed,calc}} \) values are used at higher current densities. Moreover, a cross-pressure of 0.1 bar is maintained between the anode and cathode, i.e., the anode pressure is 0.1 bar higher than the cathode. The current density values (\( i_{\text{dens}} \)) and the variations in the corresponding coolant temperature (\( \Delta T_{\text{cool}} \)) are the same for all operating conditions:

\[ i_{\text{dens}} = [0.0, 0.03, 0.05, 0.1, 0.2, 0.6, 1.0, 1.2, 1.5, 1.7, 1.9] \]
\[ 2.1, 2.4, 2.7, 3.0] \text{ [A cm}^{-2}\text{]}, \]
\[ \Delta T_{\text{cool}} = [0.5, 0.5, 0.5, 1.0, 2.0, 5.0, 10, 10, 10, 10, 12, 12, 12] \text{ [°C]} \]

The 15 points chosen on the polarization curve can roughly be divided into the three commonly known regions, namely, the kinetic, ohmic, and mass transport regions. Given the 6 variable factors that divide into the three commonly known regions, namely, the kinetic, ohmic, and mass transport regions. Given the 6 variable factors that determine the operating conditions and the 3 levels chosen for each one, a full factorial design yields a total of 729 operating conditions. For each parameter value, the model is evaluated using all of these operating conditions to generate the required sensitivity data.
Data collection method.—For any given set of parameter values and choice of operating conditions, the model is simulated starting from open-circuit voltage and the load is increased in a step manner to reach the highest current density at 3 A·cm\(^{-2}\). Each load is maintained for 80 s and the data for the last 10 s are averaged for the quasi-steady state values that are used in this study. Here, we analyze sensitivity of 3 model outputs to the different parameters. These outputs include cell voltage, high frequency resistance (HFR), and membrane water crossover. The cell voltage data are collected for all 15 points on the polarization curve. However, the HFR and water crossover data are only recorded for the last 10 points on the polarization curve corresponding to the ohmic and mass transport regions. The data at lower current densities are left out of the analysis due to considerable uncertainty in the corresponding measurements during actual experiments.

Sensitivity Analysis: Results and Discussion

Convergence of sensitivity results.—An important consideration is the convergence of the sensitivity results given the small number of points chosen to sample a high-dimensional parameter space. In particular, the question is whether 55 samples are enough to claim the results are representative of the sensitivities throughout the parameter space. Since median and standard deviation are the statistics of choice for our analysis, Fig. 1 shows the corresponding results for increasing number of samples in the parameter space. As seen in the figure, while the statistics change as we increase the number of samples from 5 to 45, they remain relatively unchanged when the samples are increased to 55. We therefore conclude that given the particular structure of the problem, the small number of samples used here appears to be enough to obtain a valid representation of model sensitivities. Nonetheless, we acknowledge that further rigorous analysis is required to concretely determine the number of samples needed for convergence.

Single measurement.—We begin our analysis of the sensitivity results with individual model predictions, i.e., we consider the voltage, HFR, and water crossover sensitivities separately. The results are shown in Fig. 2. The figure shows the results for all points in the parameter space and highlights the median values. Starting with the voltage sensitivity results (Fig. 2a), we observe that the ORR kinetic parameters have the highest sensitivities. Particularly, the ORR charge transfer coefficient (\(\alpha_{ca}\)) has the highest sensitivity among all parameters followed by the ORR...
reference exchange current density ($i_{0,ca}$). Other important parameters, as characterized by their high sensitivity magnitudes, are those that control the membrane transport properties. These include parameters impacting water transport through the membrane ($\xi_{\text{diff,mb}}$), as well as those affecting electron and proton transport efficiency ($R_{\text{elec}}, \kappa_0, E_{\text{act,mb}}$), which influence the resistance of the cell. Finally, the parameters controlling the thermal ($k_T, h_{\text{conv}}$) and mass transport behavior of the cell ($n_a, \varepsilon_{MPL}, \varepsilon_{GDL}$) also demonstrate relatively high voltage sensitivities. On the other hand, the parameters related to hydrophobicity of the porous layers ($\theta_{\text{MPL}}, \theta_{\text{GDL}}$) and some geometrical features ($w_{\text{rib}}$) show negligible sensitivities.

At this point, it is critical to note that all these sensitivity results depend on the underlying model and are only useful when interpreted with the model structure in mind. This makes interpretation of the results a crucial step in the overall analysis. For example, a high sensitivity value for a given parameter does not necessarily mean that the parameter is also physically important. It only
indicates that the particular structure of the model with all the underlying assumptions and the ranges considered for the parameter values renders that particular parameter sensitive. The same can be said of insensitive parameters, where their insensitivity is not always indicative of their lack of importance in a physical setting. This point cannot be overemphasized, as sensitivity results are commonly misinterpreted in the literature. For instance, Vetter et al.51 used their sensitivity results as a corroborating evidence that contact resistances are important in fuel cell modeling. While we do not contest this particular claim, we note that these sensitivity results were generated with a model that promotes the importance of contact resistances. Therefore, making broad conclusions about the physical importance of a parameter founded on model-based sensitivities is not always warranted. Accordingly, we believe that these results are best used to inspect and enhance the model in conjunction with experiments and expert knowledge. For instance, if the results indicate that the measurements should be insensitive to a particular parameter, but the actual experimental data or expert opinion indicate otherwise, this can be used as a basis to revisit the model assumptions and make the necessary changes to ensure the model sensitivities match those of the physical system.

Following the above arguments, we note that some of the insensitivities captured in our results are artifacts of the modeling assumptions. Examples of these include the contact angle of the porous layers and the landing width. The literature has established the importance of porous layer water retention capabilities, which are affected by the layer’s hydrophobicity.57-60 The lack of voltage sensitivity in our results mostly stems from the fact that these properties directly affect liquid water distribution in the porous layers, which does not have an appreciable impact on the voltage predictions under dry conditions. However, our results do not show significant voltage sensitivity to the contact angles even under wetter conditions. This insensitivity indicates that the continuum Leverett approach used is not sufficient to capture the experimental variations with only one contact angle as the variable.51 Therefore, if the application at hand requires more accurate description of the two-phase flow phenomena, model enhancement is required. Such improvements may also take the form of reparameterizing the sub-models. The same can be said of the land width, where experimental evidence and more detailed models clearly show the importance of this factor.8,62,63 The insensitivity observed in our results is due to the fact that the version of the model used in the current study does not capture the in-plane transport phenomena, thereby leaving the landing width as a non-influential parameter.

The HFR sensitivities are shown in Fig. 2b, where fewer parameters have high sensitivity compared to the voltage. In particular, mostly those parameters that directly affect the cell resistance are deemed to be sensitive ($R_{\text{elec}}, \kappa_0, E_{\text{act,mb}}, \delta_{\text{mb}}$). Other notably sensitive parameters are those that control membrane water crossover ($\zeta_{\text{diff,mb}}$) and heat transport across the cell ($k_T, h_{\text{conv}}$). Particularly, the thermal conductivity of the cell layers appears to be one of the most sensitive parameters ($k_T$). This can be attributed to the fact that heat transport can profoundly impact the membrane water content, thereby affecting the cell resistance.

Lastly, the membrane water crossover sensitivities are shown in Fig. 2c, where the scaling factor for membrane water diffusivity ($\zeta_{\text{diff,mb}}$) and the thermal conductivity of the cell layers ($k_T$) are found to be the most influential parameters. The high sensitivity of water crossover to $\zeta_{\text{diff,mb}}$ is expected, as this parameter directly controls water diffusivity in the membrane. On the other hand, the high sensitivity of $k_T$ may seem counter-intuitive. However, the thermal balance of the cell is known to contribute significantly to the overall water balance across the MEA64-66 and this result indicates that the model indeed captures such behavior.

Considering all of the 3 output sensitivities at the same time, we note that the voltage seems to be the richest signal that is affected by

Figure 3. Collinearity of the voltage sensitivity vectors. The off-diagonal values that are larger than 0.707 are highlighted, as they indicate an acute angle of less than 45° between the corresponding sensitivity vectors, which leads to considerable similarity between the corresponding parameter effects.
many of the model parameters. This represents an opportunity for parameter identification, as variations in the model parameters can be observed in voltage predictions. On the other hand, this also amounts to an important challenge, as many parameters impact voltage simultaneously and their effects are typically difficult to deconvolute. Fortunately, the other output predictions demonstrate a rather selective sensitivity to parameter variations, where only a few parameters affect the HFR and water crossover predictions. Therefore, these additional output predictions can be used in conjunction with voltage to better deconvolute the impact of different parameters for successful identification. This can be viewed as common knowledge, as additional measurements are typically known to help with parameter identification. Nonetheless, the methods presented in this work can be used to quantify how much utility there is in adding any particular measurement and determine which of the model parameters can be identified with acceptable precision given the available measurements.

The relative independence of sensitivity vectors is another important measure that must be taken into account. Accordingly, the collinearity indices based on the 3 output predictions and defined by Eq. 10 are shown in Figs. 3–5. Note that the figures are generated based on local analysis data and only highlight the pairs where the magnitude of the sensitivity vector for either parameter is non-negligible and the angle between the corresponding sensitivity vectors is less than 45°. Starting with the collinearity between voltage sensitivity vectors in Fig. 3, we observe that many kinetic parameters have similar impacts on the voltage predictions. Specifically, the ORR reference exchange current density (i_0,ca) and charge transfer coefficient (α,ca) are seen to have a large collinearity index. This is an expected result, as these two parameters have very similar effects in the voltage predictions determined by the Butler-Volmer equation. Other highly collinear pairs include (κ_0, E_act,mb) and (θ,mb, E_act,mb). Again, these results may be expected, as all these three parameters impact the membrane resistance, thereby influencing the ohmic voltage loss.

In addition to these expected results, we also observe some counter-intuitive behavior, where, for example, some of the ohmic parameters are found to be highly correlated with kinetic parameters as measured by their impact on voltage predictions. Particularly, the collinearity index for the (κ_0, i_0,ca) pair is greater than 0.9, while it is about 0.77 for the (κ_0, ECSA,ca) pair. This does not indicate that these parameter pairs are physically related to each other; rather, this means that the particular model structure may not allow for distinguishing their impact from one another using only voltage data. In other words, the same polarization curve may be obtained by tweaking either of the parameters in such correlated pairs. If the experimental results show a clearly distinguishable impact of such parameter pairs on the voltage measurements, then this result would point to the deficiency of the model in capturing this behavior and calls for model refinement.

Moving on to the collinearity indices based on HFR sensitivities shown in Fig. 4, we observe that almost all of the parameters that were determined to impact HFR in Fig. 2b appear to be largely collinear when considered in pairs. Therefore, HFR measurements alone may not be sufficient to identify all of these parameters. As for membrane water crossover sensitivities, Fig. 5 illustrates that many of the catalyst layer structural properties impact water crossover predictions in the same manner. For example, the anode catalyst layer Pt loading (L_Pt,an) and its ionomer to carbon ratio (IC,an) have a collinearity index of 0.96. This is due to the fact that in our model construction, they both affect the porosity of the catalyst layer; increasing either the Pt loading or the ionomer to carbon ratio reduces the porosity. Since porosity affects the water retention capabilities of the catalyst layer, which is adjacent to the membrane, variation in either parameter lead to similar impacts on the membrane water crossover predictions.
Finally, we point out that the results presented in Figs. 3–5 are generated using all operating conditions from the library discussed earlier. Therefore, even if the results indicate high similarity between two parameters, there might still be an opportunity of decorrelating their sensitivities by judiciously selecting the operating conditions. This is an important motivation for the optimal experimental design that is presented in the second part of this work.

Multiple measurements.—Having presented sensitivity results for different outputs individually, we now turn our attention to the case where multiple output predictions are considered at the same time. Here we consider three cases. In the first two cases, voltage prediction is considered along with either the HFR or membrane water crossover. For the last case, all three outputs are examined simultaneously. For these combined prediction cases, the sensitivity vectors are obtained by simply concatenating the corresponding sensitivity vectors for each individual output prediction. Therefore, the sensitivities to different parameters inherit the main characteristics of the individual output predictions as seen in Fig. 6. Particularly, we observe that adding HFR prediction to voltage results in increased sensitivity of the resistance parameters, such as $\kappa_0$ and $R_{\text{elec}}$, while the sensitivity of the important kinetic parameters, such as $\alpha_{\text{ca}}$, remain high. Similarly, including membrane water crossover prediction alongside voltage promotes the sensitivity of parameters that impact each output individually.

The above discussion shows that adding measurements can make more parameters sensitive enough for the purpose of identification. However, in addition to having an observable impact on the output predictions, the effect of parameters should be uniquely distinguishable from one another for practical identifiability. Therefore, the role of additional measurements in deconvoluting parameter effects should be investigated. Accordingly, Fig. 7 shows how the collinearity indices between pairs of sensitivity vectors change as additional output predictions are considered. We observe that adding measurements generally tends to decorrelate parameter effects. For instance, using only voltage data (Fig. 7a) can lead to correlations between some of the ohmic and kinetic parameters (e.g., $\kappa_0$ and $i_{0,\text{ca}}$ have a collinearity index of $\psi = 0.9$). When HFR is used along with voltage (Fig. 7b), these correlations are significantly reduced (e.g., $\psi = 0.55$ for $\kappa_0$ and $i_{0,\text{ca}}$). Adding the membrane water crossover to the measurements further decorrelates the effects of certain parameters. For example, the collinearity index is $\psi = 0.71$ for $L_{\text{Pt,an}}$ and $\alpha_{\text{ca}}$ with voltage and HFR measurements (Fig. 7b), but reduces to $\psi = 0.41$ when membrane water crossover is considered (Fig. 7d).

The analysis shows the magnitude of each parameter’s impact on the model predictions and its degree of independence from other parameters as additional measurements are considered. Notably, we have seen how adding measurements can promote identifiability by increasing the sensitivity while decorrelating the effect of different parameters. While this seems rather intuitive, the significance of these results lies in their quantification of the parameter sensitivities and their interactions. The methods presented herein can thus be used to focus the resources on identification of most crucial parameters and help with effective model parameterization. This is further demonstrated in the second part of the study.

Impact of operating conditions.—So far, our analysis has been based on data generated using all of the operating conditions in the library. However, it is conceivable that certain operating conditions can better reveal the impact of a particular parameter and improve its identifiability. To further demonstrate this, Fig. 8 shows how the sensitivity of model parameters can change with the operating condition. Particularly, voltage sensitivity to the gas diffusion layer thickness is shown in Fig. 8a for different operating conditions in the
library, where considerable variations in voltage sensitivity are observed as the operating conditions are varied. One main trend that is highlighted in the figure is the decreasing sensitivity with increased cathode flow rate. This is an expected result as the layer’s thickness mostly impacts mass transport losses, which are heightened at lower stoichiometric ratios.

To further demonstrate the impact of operating conditions, Fig. 8b shows the ordered voltage sensitivities for 5 different parameters. Particularly, the sensitivity at each operating condition is calculated. The results are then ordered in a decreasing fashion such that the first point on each line indicates the best operating condition at which the highest sensitivity is obtained. This ordering is done to better illustrate the effect of operating conditions on parameter sensitivities. The figure clearly shows how operating conditions can increase or diminish the sensitivity of a parameter. We can also observe that the degree to which the operating

Figure 6. Parameter sensitivities for multiple output predictions: (a) voltage and HFR, (b) voltage and membrane water crossover, and (c) voltage, HFR, and membrane water crossover. Results for all points in the parameter space are shown with gray circles. The sensitivity medians are also shown, where the line color changes with the sensitivity value.
conditions affect the sensitivity varies with the parameter. For example, the sensitivities of $f_{\text{RH}}, a, \delta_{\text{mb}},$ and $\delta_{\text{GDL}}$ show considerable dependence on the operating conditions as they can vary by almost two orders of magnitude based on the condition. The sensitivity of other parameters on the other hand, may demonstrate moderate to negligible dependence on the operating conditions, as is the case for $\omega$ and $i_{0,\text{ca}}$.

The observable impact of operating conditions on parameter sensitivities motivates a closer look at the sensitivity results. Specifically, the question is how much would a parameter sensitivity increase if only the best operating condition is used to calculate it rather than using all of the operating conditions. An interesting aspect of this is whether a parameter that is deemed insignificant as measured by its average sensitivity at all operating conditions, can have a discernible impact under particular conditions. This is illustrated in Fig. 9, where the voltage and HFR sensitivities based on three calculations are shown: (1) using all operating conditions, (2) using only the operating condition that results in highest parameter sensitivity, and (3) using the most sensitive loading condition of the best operating condition. For the third case, the voltage and HFR data at each condition are characterized based on the load region (kinetic, ohmic, and mass transport regions). Only data from the most sensitive region are then used to calculate the sensitivity. The figure shows that for most parameters, considering only one particular operating condition rather than all conditions in the library significantly enhances the sensitivity. Most notably, if we set a sensitivity threshold of 0.1 for identifiability, as shown in the figure, we observe that a number of parameters that are unidentifiable considering all operating conditions, are indeed identifiable at specifically chosen conditions. The channel width ($w_{\text{ch}}$) and the residual ionomer conductivity ($\kappa_{\text{res}}$) are examples of such parameters (see Fig. 9a). Such discernible impact of operating conditions on parameter identifiability motivates the need for optimal experimental design that is addressed in the second part of the study.

### Identifiability Analysis and Parameter Subset Selection

Using the sensitivity results obtained in the previous section, we now turn our attention to identifiability analysis. Recall that a parameter is identifiable if (1) it has an observable impact on the model predictions for which measurements are available, and (2) its impact is unique in the sense that the same behavior cannot be reproduced by variations in other parameters. To further analyze identifiability, we consider a least-squares setting. Particularly, assume that $y$ is the vector of available measurements. We further assume, for the purpose of discussion, that there exists at least one set of parameter values for which our model predictions, $\tilde{y}$, exactly match the measurements. We denote such a parameter set by $\theta^*$. Therefore, we have:

![Collinearity of sensitivity vectors based on single and multiple predicted outputs: using (a) voltage, (b) voltage and HFR, (c) voltage and water crossover, and (d) voltage, HFR, and water crossover as the predicted outputs.](image-url)
The dependence of \( \tilde{y} \) on the parameter set is made explicit. The question is whether \( *q \) is unique and how it can be found. To answer this, we formulate a nonlinear least-squares problem, which minimizes the following cost:

\[
\chi(\theta) = \sum_{i=1}^{N} (y_i - \tilde{y}_i)^2,
\]

where \( N \) is the number of available data points. Note that by restricting our analysis to the immediate neighborhood of \( *q \), we can use Taylor expansion to evaluate \( \tilde{y}_q \):

\[
\tilde{y}_q \approx \tilde{y}_q^* + \frac{\partial y}{\partial \theta}(\theta - \theta^*),
\]

where the higher order terms are neglected. Plugging Eq. 13 into 12 and using 11, we have:

\[
\chi(\theta) \approx \sum_{i=1}^{N} \left[ \frac{\partial y_i}{\partial \theta}(\theta - \theta^*) \right]^2 = (S\Delta\theta)^T(S\Delta\theta),
\]

where \( \Delta\theta = \theta - \theta^* \) and the definition of the sensitivity matrix (Eq. 4) is used to write the last equality. This last equation indicates that the optimal solution is obtained when \( (S^TS)\Delta\theta = 0 \). If the Hessian \( (H = S^TS) \) has full rank, then \( \Delta\theta = 0 \) is the unique solution, which indicates that \( \theta = \theta^* \), meaning that the parameters are locally identifiable at \( \theta^* \). In case the Hessian is rank deficient, there is an infinite number of solutions that yield a zero least-squares cost (belonging to the null space of \( S^TS \)). This means that the parameter identification problem does not have a unique solution. Note that rank-deficiency of the Hessian captures both requirements of identifiability; in theory, a zero sensitivity vector or one that is linearly dependent on other sensitivity vectors yields a rank deficient Hessian. In practice, this derivation indicates that a poorly conditioned Hessian matrix, which can be caused by small sensitivity magnitudes or highly dependent sensitivity vectors, can hinder parameter identifiability. Specifically, the condition number of the sensitivity matrix is defined as:

\[
s_H = \frac{\lambda_{\text{max}}(H)}{\lambda_{\text{min}}(H)},
\]

where \( \lambda_{\text{max}} \) and \( \lambda_{\text{min}} \) denote the largest and smallest eigenvalues of the Hessian, respectively. The significance of the condition number and, more generally, the eigenvalues of the Hessian, lies in the fact that they control the rate of reduction in the least-squares cost in the direction of each eigenvector: a large (small) eigenvalue means a rapidly (slowly) decaying cost along the direction of the corresponding eigenvector.

Since the identifiability of the model parameters depend on the eigenvalues of the Hessian, the spectra of these eigenvalues are shown in Fig. 10 for sensitivity matrices obtained at several points in the parameter space. The figure shows interesting spectra, wherein, the eigenvalues are evenly spread over many orders of magnitude. The literature refers to this phenomenon as ‘sloppiness’ of the model. It has been suggested that accurate parameter identification is
impossible under such circumstances and the focus should instead be on predictive behavior of the model.68 This assertion has since been challenged in the literature and it has become clear that sloppiness and identifiability are two very distinct concepts and neither implies the other.70–72 Most recently, Chis et al. have even argued that sloppiness is not a useful concept and the focus should be on parameter identifiability.70 Nonetheless, sloppiness can have implications for identifiability: it is more difficult to distinguish the identifiable

Figure 9. Variations in (a) voltage and (b) HFR sensitivities as the best operating conditions are used.

Figure 10. Sloppiness of the eigenvalue spectra of the least squares cost Hessian.
parameters from the unidentifiable ones in a sloppy model. Particularly, note that the eigenvalue spectra in Fig. 10 show no discernible gap between the large and small eigenvalues. Therefore, the influential eigenvectors cannot be easily selected.

How can one address this difficulty due to model sloppiness? We believe that tools from the literature on identifiability analysis remain relevant. In particular, we focus on selecting subsets of parameters that can be identified using the available measurements. To determine whether the selected subset is identifiable, we use the metric proposed by Brun et al., which amounts to a threshold on the smallest eigenvalue of the resulting Hessian:

$$\lambda_{\text{min}}(H) \geq \lambda_0,$$

where $H$ is the Hessian matrix obtained from the sensitivities of the selected parameters and $\lambda_0$ is the chosen threshold. Recall that the smallest eigenvalue controls the slowest rate of decay of the least-squares cost. By ensuring that the minimum eigenvalue is greater than some threshold, we are controlling the decay rate of the cost.

Taking the eigenvalue threshold into account, we rank parameters with two approaches: (1) based on the sensitivity magnitude and (2) using the orthogonal method. The orthogonal method uses the Gram-Schmidt orthogonalization procedure and selects the parameter with the largest sensitivity first. At each following step of the algorithm, the component of each parameter sensitivity that is orthogonal to the sensitivity vectors of the previously selected parameters is calculated. The parameter with the largest orthogonal component is then selected as the next most influential parameter. This is implemented using the MATLAB QR factorization.

The advantage of the orthogonal method over the magnitude based approach is in the fact that it accounts for linear dependencies between parameter sensitivities. Therefore, between two parameter subsets of equal size selected with these methods, the one based on the orthogonal method has less linear dependency and is therefore expected to have a larger minimum eigenvalue. This is shown in Fig. 11a, which, using sensitivity data at one of the points in the parameter space, shows the minimum Hessian eigenvalue for different sizes of selected parameter subsets. The figure clearly shows that parameter selection based on the orthogonal method yields a better conditioned Hessian. For instance, if we choose an eigenvalue threshold of $\lambda_0 = 1$, and choose a subset of 36

![Diagram](image-url)
parameters using the magnitude-based approach, the selected subset would be deemed unidentifiable as the minimum eigenvalue is 0.92. On the other hand, a subset of the same size obtained with the orthogonal method yields a minimum eigenvalue of 3.93 and is therefore identifiable according to the chosen threshold.

To analyze the effect of additional measurements on improvements in parameter identifiability, Fig. 11b shows the minimum Hessian eigenvalue using different measurements. It is observed that adding measurements can indeed improve the condition of the Hessian, thereby enhancing parameter identifiability. We emphasize again, that while the main results may seem intuitive, it is the quantitative and rigorous treatment of such behavior that constitutes the main contribution of this work.

To make these results more tangible, Table I shows the top 10 parameters ranked based on the voltage sensitivities using both the sensitivity magnitude and the orthogonal method. We note that when using the voltage sensitivity magnitude, the reference ORR exchanger current density ($i_{0,ca}$) is the second most important parameter. However, the orthogonal method yields a much lower rank for this parameter, since its sensitivity is highly collinear with that of the transfer coefficient ($\alpha_{ca}$). Similarly, the orthogonal method promotes ranking of parameters that have the largest impact, which is independent of previously ranked parameters. The table also shows parameter rankings based on multiple measurements using the orthogonal method. We observe that, for instance, using HFR alongside voltage as the outputs of interest, improves the ranking of parameters that impact the cell resistance ($R_{dec}$ is ranked second when HFR is used as an additional output). Analogously, considering membrane water crossover as an additional output of interest promotes the ranking of parameters that have a profound impact on this output, such as $\Gamma_{diff,mb}$.

Finally, we emphasize that the analysis in this section is conducted using sensitivity data obtained locally about one of the sample points in the parameter space. Moreover, all of the operating conditions in the library are used in the analysis. Therefore, even though these results might indicate that a large number of model parameters may be identifiable, such an identification would require a prohibitively large set of experiments. We discuss these points in more detail in the second part of the study, where we refine our analysis to select a subset of parameters that is expected to be optimal in the entire parameter space and use experimental design as a tool to improve parameter identification.

Overall, these results underline the importance of rigorous identifiability analysis as an important tool to complement expert intuition when selecting parameters for identification. Particularly, we note that while many of our results might appear intuitive, their significance lies in the quantitative procedure used to obtain them and such conformation with expert knowledge further verifies the proposed methodology. The rigorous and systematic nature of the presented procedure makes it applicable even in the absence of expert knowledge and could lead to new insights. Due to the very same characteristics of the procedure, even though the example results are specific to the particular fuel cell model considered, the procedure itself can be applied to other models, as well.

**Summary and Conclusions**

A systematic procedure is presented on a pseudo-2D, non-isothermal, and two-phase model of a PEM fuel cell to study the output sensitivities to a variety of the model parameters. In particular, three common measurements are considered for the sensitivity analysis, namely, the cell voltage, HFR, and membrane water crossover. A large set of model parameters are investigated using an extended local sensitivity analysis, which is carried out by performing local analyses about several points in the parameter space. The parameters that have the highest impact on individual output predictions are determined. Moreover, the collinearity between pairs of parameters is investigated, where some physically unrelated pairs are found to have similar impacts on model predictions. The effect of operating conditions on parameter sensitivities is shown, where sensitivity of some parameters is found to vary considerably with the conditions. It is also shown that a number of parameters that generally have low sensitivity, can become more sensitive under certain operating conditions. Finally, the sensitivity results are used to analyze identifiability of model parameters, where it is found that the parameter selection method and additional measurements can clearly influence identifiability of the selected subset of parameters.

The sensitivity and identifiability analysis methodologies presented in this paper can be used to optimize the parameter identification process. In particular, with these sensitivity and identifiability results in hand, the parameter subset selection for identification can be executed in a robust and optimal manner, and the experiments that will generate the data for parameterization can be made maximally informative to identify the chosen subset of parameters with the highest overall accuracy and confidence. These opportunities are pursued in the second part of this paper.

**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, their units, corresponding ranges, as well as the method used to scale them are given. Importantly, note that all parameters considered here are bounded to specific ranges, which are determined either based on the corresponding uncertainty in the parameter values reported in the literature (as is the case with candidate fitting parameters) or the potential design envelope (as is the case with geometrical parameters).

---

**Table I. Parameter rankings with different methods and multiple measurements.**

| Ranking | Voltage (Magnitude) | Voltage (Orthogonal) | Voltage, HFR (Orthogonal) | Voltage, HFR, Water Crossover (Orthogonal) |
|---------|---------------------|----------------------|---------------------------|------------------------------------------|
| 1       | $\alpha_{ca}$       | $\alpha_{ca}$        | $\alpha_{ca}$             | $\alpha_{ca}$                            |
| 2       | $h_{ch}$            | $h_{ch}$             | $R_{dec}$                 | $R_{dec}$                                |
| 3       | $n_e$               | $n_e$                | $n_f$                     | $n_f$                                    |
| 4       | $\delta_{GDL}$     | $K_T$                | $\delta_{GDL}$            | $\delta_{GDL}$                           |
| 5       | $\delta_{GDL}$     | $\delta_{GDL}$       | $\delta_{GDL}$            | $\delta_{GDL}$                           |
| 6       | $h_{an}$            | $h_{conv}$           | $h_{conv}$                | $h_{conv}$                              |
| 7       | $\delta_{MPL}$     | $\delta_{MPL}$       | $\delta_{MPL}$            | $\delta_{MPL}$                           |
| 8       | $E_{act,ca}$        | $i_{0,ca}$           | $E_{act,mb}$              | $E_{act,mb}$                            |
| 9       | $\delta_{mb}$      | $\delta_{mb}$        | $\delta_{mb}$             | $\delta_{mb}$                           |
| 10      | $E_{act,ca}$        | $i_{0,ca}$           | $E_{act,mb}$              | $E_{act,mb}$                            |

**References**

[1] Jones, D. A. et al. Parameter identifiability analysis of a PEM fuel cell model. *Journal of The Electrochemical Society*, 2020 167 044504.
Table A·I. Nomenclature and parameter list.

| Parameter | Description [Units] | Range |
|-----------|----------------------|-------|
| $\omega$ | Energy parameter for Temkin isotherm [kJ·mol$^{-1}$] | [1, 20]$^a$ |
| $\alpha_{\text{an}}$ | HOR$^b$ transfer coefficient [-] | [1.2, 2.0]$^a$ |
| $\alpha_{\text{ca}}$ | ORR$^b$ transfer coefficient [-] | [0.4, 1.0]$^a$ |
| $i_{\text{an}}$ | HOR exchange current density [A·cm$^{-2}$] | $10^{-3}$, $10^{-1}$$^b$ |
| $i_{\text{ca}}$ | ORR exchange current density [A·cm$^{-2}$] | $10^{-3}$, $10^{-1}$$^b$ |
| $\gamma_{\text{an}}$ | HOR reaction order [-] | [0.5, 1]$^a$ |
| $\gamma_{\text{ca}}$ | ORR reaction order [-] | [0.5, 1]$^a$ |
| $E_{\text{act,an}}$ | HOR activation energy [kJ·mol$^{-1}$] | [5, 20]$^a$ |
| $E_{\text{act,ca}}$ | ORR activation energy [kJ·mol$^{-1}$] | [30, 80]$^a$ |
| $E_{\text{Pto,PO}}$ | Reversible potential for PtO formation [V] | [0.75, 0.85]$^a$ |
| $k_{\text{cross}}$ | Scaling factor for membrane permeability [-] | [0.01, 10]$^a$ |
| $\xi_{\text{eff,mem}}$ | Scaling factor for membrane diffusivity [-] | [0.1, 10]$^a$ |
| $f_{\text{RH,a}}$ | Fitting parameter for ORR RH dependency [-] | [0.4, 0.9]$^a$ |
| $f_{\text{RH,b}}$ | Fitting parameter for ORR RH dependency [-] | [0.2, 4]$^a$ |
| $f_{\text{RH,c}}$ | Fitting parameter for ORR RH dependency [-] | [0.3, 1]$^a$ |
| $R_{\text{elec}}$ | Total electronic resistance [μΩ·cm$^{-2}$] | [30, 80]$^a$ |
| $\eta_{\text{res}}$ | Residual ionomer conductivity [S·cm$^{-1}$] | [0.2, 0.6]$^a$ |
| $E_{\text{act,mb}}$ | Membrane activation energy [kJ·mol$^{-1}$] | [5, 30]$^a$ |
| $EW_{\text{an}}$ | Membrane ionomer equivalent weight [g·mol$^{-1}$] | [600, 1200]$^a$ |
| $EW_{\text{ca}}$ | CL$^e$ ionomer equivalent weight [g·mol$^{-1}$] | [600, 1200]$^a$ |
| $k_T$ | Scaling factor for thermal conductivity [-] | [0.1, 10]$^a$ |
| $h_{\text{conv}}$ | Heat convection coefficient [W·cm$^{-2}$·K$^{-1}$] | [0.05, 10]$^a$ |
| $n_e$ | Saturation correction exponent for diffusivity [-] | [1, 3]$^a$ |
| $n_{\text{t}}$ | Porosity correction exponent for diffusivity [-] | [1, 3]$^a$ |
| $\delta_{\text{MPL}}$ | MPL porosity [-] | [0.4, 0.65]$^a$ |
| $\varepsilon_{\text{GDL}}$ | GDL$^b$ porosity [-] | [0.5, 0.85]$^a$ |
| $\theta_{\text{CL}}$ | CL contact angle [°] | [90.5, 100]$^a$ |
| $\theta_{\text{MPL}}$ | MPL contact angle [°] | [110, 150]$^a$ |
| $\theta_{\text{GDL}}$ | GDL contact angle [°] | [100, 130]$^a$ |
| $\kappa_{\text{per}}$ | Scaling factor for permeability [-] | [0.02, 0.5]$^a$ |
| $w_{\text{ch}}$ | Channel width [cm] | [0.1, 0.4]$^a$ |
| $w_{\text{rib}}$ | Rib (land) width [cm] | [0.02, 0.3]$^a$ |
| $h_{\text{ch}}$ | Channel height [cm] | [0.01, 0.1]$^a$ |
| $L_{\text{ch}}$ | Channel length [cm] | [15, 40]$^a$ |
| $\delta_{\text{m}}$ | Membrane thickness [μm] | [8, 30]$^a$ |
| $\delta_{\text{MPL}}$ | MPL thickness [μm] | [20, 70]$^a$ |
| $\delta_{\text{GDL}}$ | GDL thickness [μm] | [100, 240]$^a$ |
| $ECSA_{\text{an}}$ | Anode CL ECSA [m$^2$·g$^{-1}$] | [40, 90]$^a$ |
| $ECSA_{\text{ca}}$ | Cathode CL ECSA [m$^2$·g$^{-1}$] | [40, 90]$^a$ |
| $\varepsilon_{\text{CL,an}}$ | Anode CL porosity [-] | [0.3, 0.7]$^a$ |
| $\varepsilon_{\text{CL,ca}}$ | Cathode CL porosity [-] | [0.3, 0.7]$^a$ |
| $L_{\text{Pt,an}}$ | Anode CL Pt loading [mg·cm$^{-2}$] | [0.05, 0.3]$^a$ |
| $L_{\text{Pt,ca}}$ | Cathode CL Pt loading [mg·cm$^{-2}$] | [0.05, 0.4]$^a$ |
| $\epsilon_{\text{IC,an}}$ | Anode CL I/C ratio [-] | [0.6, 1.3]$^a$ |
| $\epsilon_{\text{IC,ca}}$ | Cathode CL I/C ratio [-] | [0.6, 1.3]$^a$ |
| $w_{\%_{\text{Pt,an}}}$ | Anode CL Pt weight percentage [-] | [0.2, 0.7]$^a$ |
| $w_{\%_{\text{Pt,ca}}}$ | Cathode CL Pt weight percentage [-] | [0.2, 0.7]$^a$ |
| $m_{\text{Brogue}}$ | Bruggeman coefficient for CL [-] | [1, 3]$^a$ |

a) Linear scaling. b) Logarithmic scaling. c) Hydrogen Oxidation Reaction. d) Oxygen Reduction Reaction. e) Catalyst Layer. f) Microporous Layer. g) Gas Diffusion Layer.

References

1. D. Gertieisen, T. Heilmann, and C. Ziegler, “Modeling the phenomena of dehydration and flooding of a polymer electrolyte membrane fuel cell.” J. Power Sources, 187, 165 (2009).
2. G. A. Futter, P. Gazdzicki, K. A. Friedrich, A. Latz, and T. Jahnke, “Physical modeling of polymer-electrolyte membrane fuel cells: understanding water management and impedance spectra.” J. Power Sources, 391, 148 (2018).
3. P. Schneider, C. Sadeler, A.-C. Scherzer, N. Zamel, and D. Gertieisen, “Fast and reliable state-of-health model of a PEM cathode catalyst layer.” J. Electrochem. Soc., 166, F322 (2019).
4. B. Jayasankar and K. Karan, “O2 electrochemistry on Pt: a unified multi-step model for oxygen reduction and oxide growth.” Electrochim. Acta, 273, 367 (2018).
5. 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, F901 (2016).
6. B. Barnes, D. Spernjak, G. Luo, L. Hao, K. S. Chen, C.-Y. Wang, R. Mukundan, and R. L. Borup, “Validation of a two-phase multidimensional polymer electrolyte membrane fuel cell computational model using current distribution measurements.” J. Power Sources, 236, 126 (2013).
7. 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).
8. A. Goshatschi, P. Garcia-Salabert, J. Chen, K. Talukdar, D. G. Sanchez, and T. J. informative. "Through the membrane transition, phenomena in PEM fuel cells: A modeling study." *J. Electrochem. Soc.*, **166**, F3154 (2019).
9. S. K. Babu, H. T. Chung, P. Zelenay, and S. Listler, "Modeling electrochemical performance of the hierarchical morphology of precious group metal-free cathode for polymer electrolyte fuel cells." *J. Electrochem. Soc.*, **164**, F1037 (2017).

10. N. Zamel and X. Li, "Effective transport properties for polymer electrolyte membrane fuel cells—A focus on the gas diffusion layer." *Prog. Energy Combust. Sci.*, **39**, 111 (2013)
11. O. Abou El Naga, G. G. W. Bock, N. Hussain, S. Pasquali, a da Plessis, S. le Roux, F. Seland, H. Su, and B. G. Pollet, "Thermal conductivity in the thickest regions of micro porous layer coated porous transport layers for the PEM fuel cell." *Int. J. Hydrogen Energy*, **40**, 16775 (2015).
12. H. Sadeghifar, N. Djilali, and M. Bahrami, "Effect of polytetrafluoroethylene (PTFE) and micro porous layer (MPL) on thermal conductivity of fuel cell gas diffusion layers: modeling and experiments." *J. Power Sources*, **248**, 632 (2014).
13. M. Moore, A. Puiz, and M. Secanell, "Investigation of the ORR using the double-trap intrinsic kinetic model." *J. Electrochem. Soc.*, **166**, F670 (2013).
14. M. Markiewicz, C. Zaltis, and A. Kucernak, "Performance measurements and modelling of the ORR on fuel cell electrocatalysts—the modified double trap model." *Electrochim. Acta*, **179**, 126 (2015).
15. B. Carne and N. Djilali, "Systematic parameter estimation for pem fuel cell stacks." *J. Power Sources*, **144**, 83 (2005).
16. P. Jain, L. T. Biegler, and M. S. Jhon, ""Identiﬁcation of the Doyle-Fuller-Newman model from experimental cycling of a lithium-ion cell: parameters identiﬁcation and ranking." *J. Electrochem. Soc.*, **164**, F1037 (2017).

17. A. Pozzi, G. Ciaramella, S. Volkwein, and D. M. Raimondo, "Global sensitivity analysis of cylindrical LiFePO4 battery performance using multi-physics experiments for a lithium-ion cell: parameters identiﬁcation and ranking." *J. Electrochem. Soc.*, **164**, B514 (2012).
18. T. C. Yau, P. Sauriol, X. T. Bi, and J. Stumper, "Parameter sensitivity analysis and development of a two-dimensional real-time model of proton-exchange-membrane fuel cells." *Energy Conv. Manage.*, **162**, 276 (2018).
19. D. Zhao, M. Dou, D. Zhou, and F. Gao, "Study of the modeling parameter effects on the polarization characteristics of the PEM fuel cell." *Int. J. Hydrogen Energy*, **41**, 22316 (2016).
20. C. Min, Y. He, X. Liu, B. Yin, W. Jiang, and W. Tao, "Parameter sensitivity examination and discussion of PEM fuel cell simulation model validation Part II: Results of sensitivity analysis and validation of the model." *J. Power Sources*, **160**, 374 (2006).
21. X.-D. Wang, J.-L. Xu, and D.-J. Lee, "Parameter sensitivity examination for a composite three-dimensional, two-phase, non-isothermal model of polymer electrolyte membrane fuel cell." *Int. J. Hydrogen Energy*, **37**, 15706 (2012).
22. D. Zhou, T. T. Nguyen, E. Breaz, D. Zhao, S. Clénét, and F. Gao, "Global parameters sensitivity analysis and development of a two-dimensional real-time model of proton-exchange-membrane fuel cells." *Energy Conv. Manage.*, **162**, 276 (2018).
23. R. Vetter and J. O. Schumacher, "Experimental parameter uncertainty in proton exchange membrane fuel cell modeling. Part II: Sensitivity analysis and importance parameter estimation." *J. Power Sources*, **9521** (2016).
24. B. Laoun, M. W. Naceur, A. Khelifa, and A. M. Kannan, "Global sensitivity analysis of proton exchange membrane fuel cell model." *Int. J. Hydrogen Energy*, **41**, 9521 (2016).
25. D. V. Raman, J. Anderson, and A. Papachristodoulou, "Delineating parameter unidentifiability in complex models." *Phys. Rev. E*, **95**, 032314 (2017).
26. A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, in *Global Sensitivity Analysis: The Primer* (John Wiley & Sons, New Jersey) (2008).
27. Y. Chu and J. Hahn, "Parameter set selection via clustering of parameters into pairwise indistinguishable groups of parameters." *Indust. Eng. Chem. Res.*, **48**, 6000 (2009).
28. A. Thomas, G. Maranzana, S. Didjerjean, J. Dillet, and O. Lottin, "Transport characteristics of saturated gas diffusion layers treated with hydrophobic coatings." *Chem. Eng. Sci.*, **176**, 503 (2018).
29. Q. Meyer et al., "Effect of gas-diffusion layer properties on water distribution across air-cooled, open-cathode polymer electrolyte fuel cells: a combined ex situ X-ray tomography and in-operando neutron imaging study." *Electrochim. Acta*, **211**, 478 (2016).
30. A. García-Salabert, J. T. Gostick, J. V. Zenzuyk, G. Hwang, M. Vera, and A. Z. Kabrhel, "On the limitations of volume-averaged descriptions of gas diffusion layers in the modeling of polymer electrolyte fuel cells." *ECS Trans.*, **50**, 133 (2017).
31. S. Lee, H. Jeong, B. Ahn, T. Lim, and Y. Son, "Parameter study of the channel design in the bipolar plate of PEM fuel cells." *Int. J. Hydrogen Energy*, **33**, 5691 (2008).
32. S. G. Goebel, "Impact of land width and channel span on fuel cell performance." *J. Power Sources*, **196**, 7350 (2011).
33. S. Kim and M. Meneses, "Investigation of temperature-driven water transport in polymer electrolyte fuel cell: phase-change-induced flow." *J. Electrochem. Soc.*, **166**, B535 (2019).
34. R. S. Fu, J. S. Preston, U. Pasupuggalli, T. Shiomi, S. Miyazaki, Y. Tabuchi, D. S. Wang, and D. L. Jacobson, "Water transport across a polymer electrolyte membrane under thermal gradients." *J. Electrochem. Soc.*, **158**, B301 (2020).
35. A. Thomas, G. Maranzana, S. Didjerjean, J. Dillet, and O. Lottin, "Measurements of electrode temperatures, heat and water flows in PEMFCs: conclusions about transfer mechanisms." *J. Electrochem. Soc.*, **166**, F191 (2013).
67. K. S. Brown and J. P. Sethna, “Statistical mechanical approaches to models with many poorly known parameters.” Phys. Rev. E, 68, 021904 (2003).
68. R. N. Gutenkunst, J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, and J. P. Sethna, “Universally sloppy parameter sensitivities in systems biology models.” PLoS Comput. Biol., 3, e189 (2007).
69. C. Tönsing, J. Timmer, and C. Kreutz, “Cause and cure of sloppiness in ordinary differential equation models.” Phys. Rev. E, 90, 023303 (2014).
70. O.-T. Chis, A. F. Villaverde, J. R. Banga, and E. Balsa-Canto, “On the relationship between sloppiness and identifiability.” Math. Biosci., 282, 147 (2016).
71. A. White, M. Tolman, H. D. Thames, H. R. Withers, K. A. Mason, and M. K. Transtrum, “The limitations of model-based experimental design and parameter estimation in sloppy systems.” PLoS Comput. Biol., 12, e1005227 (2016).
72. M. Secrèr, T. Toni, and M. P. Stumpf, “The ABC of reverse engineering biological signalling systems.” Mol. BioSyst., 5, 1925 (2009).
73. R. Brun, P. Reichert, and H. R. Küunsch, “Practical identifiability analysis of large environmental simulation models.” Water Resour. Res., 37, 1015 (2001).
74. B. F. Lund and B. A. Foss, “Parameter ranking by orthogonalization—applied to nonlinear mechanistic models.” Automatica, 44, 278 (2008).