Accuracy Versus Simplicity in Online Battery Model Identification

Abbas Fotouhi, Member, IEEE, Daniel J. Auger, Senior Member, IEEE,
Karsten Propp, and Stefano Longo, Senior Member, IEEE

Abstract—This paper presents a framework for battery modeling in online, real-time applications where accuracy is important but speed is the key. The framework allows users to select model structures with the smallest number of parameters that is consistent with the accuracy requirements of the target application. The tradeoff between accuracy and speed in a battery model identification process is explored using different model structures and parameter-fitting algorithms. Pareto optimal sets are obtained, allowing a designer to select an appropriate compromise between accuracy and speed. In order to get a clearer understanding of the battery model identification problem, “identification surfaces” are presented. As an outcome of the battery identification surfaces, a new analytical solution is derived for battery model identification using a closed-form formula to obtain a battery’s ohmic resistance and open circuit voltage from measurement data. This analytical solution is used as a benchmark for comparison of other fitting algorithms and it is also used in its own right in a practical scenario for state-of-charge estimation. A simulation study is performed to demonstrate the effectiveness of the proposed framework and the simulation results are verified by conducting experimental tests on a small NiMH battery pack.

Index Terms—Battery, electric vehicle (EV), equivalent circuit modeling, identification, optimization.

I. INTRODUCTION

OWADAYS, batteries are used in a wide range of applications varying from small smartphones and laptops to large high-power electric road vehicles, aircraft, and marine craft. In response to this growing demand, much research is focused on the development of new battery technologies. Integral to this is battery modeling, which is vital for many applications. In an electric vehicle (EV), it is important to understand state-of-charge (or remaining capacity), which is vital for any kind of range prediction. In addition, good models are essential for safe charging/discharging, optimal usage of batteries, and fast charging. A variety of approaches to battery modeling have been discussed in [1]. Electrochemical battery model is the most accurate form, and can explain key behaviors of a battery at the microscopic scale based on the chemical reactions occurring inside the cells. When focusing on the accuracy, electrochemical battery models are the first choice in this area [2]. However, it is important to strike a balance between model’s simplicity and accuracy when it is going to be embedded in microprocessors and provide accurate results in real time [3]. In other words, it is important to have models that are accurate enough, but not unnecessarily complicated. (Exactly what this means will depend on a particular application scenario.) For this reason, researchers have investigated other modeling approaches. One of these approaches is electrical circuit modeling or “equivalent circuit network” (ECN) modeling. Having less complexity than full electrochemical models, ECN models have been used in a wide range of applications and various battery types [4]–[7].

ECN battery models are often parameterized using experimental data. Depending on the application, battery parameterization can be performed offline or online. Because a battery model’s parameters change due to variations of battery state of charge (SoC), temperature, aging, etc., adaptive modeling approaches must often be applied to ensure a model is an accurate representation at any point in time. One of the existing approaches for online modeling of a time-varying system is the application of formal “system identification” techniques with origins in control theory and signal processing. There are many different system identification problems, and a good introduction can be obtained from Lennart Ljung’s seminal work on the subject [8]. The parameterization of an ECN model can be classified as an identification problem in which the model’s structure is fixed while the model’s unknown parameters are determined based on measured experimental data. The goal is to find a model that its output has the least deviation from the measured test data. Consequently, in this particular case, the battery identification problem is an “optimization problem” in which the parameters are optimized to have the least error in comparison to the test data. A mathematical algorithm which is used to fit the model to the measured data is called “identification algorithm” or “optimization algorithm” in this paper.

There are surprisingly few studies in the literature in which battery parameterization is carried out using formal “system identification” techniques. Genetic algorithm (GA) is used for battery model identification in [4] by considering complex models containing the ECN parameters, SoC and temperature at the same time and an offline optimization procedure is used.
to fit the model to experimental data. Similarly in [5] and [6], more than 30 parameters are obtained for two battery equivalent circuit models using multiobjective GA. The main reason that GA method had been used in those studies was said to be its greater benefits compared to other methods when an analytic solution does not exist and when the number of unknown parameters are large [5]. In [7], [9], and [10], an alternative optimization algorithm, particle swarm optimization (PSO) is employed to identify battery parameters from measured test data in the time-domain. In [7], 12 different electrical circuit model structures have been investigated and compared to each other with respect to their accuracy. In [9], a reduced-order electrochemical lithium-ion battery model is parameterized and in [10], a fractional-order model (FOM) is studied and its coefficients and differentiation orders are identified using PSO. The FOM was derived from a modified Randles model and its accuracy was compared to commonly used ECN models. In almost all of the previous studies in this area, the “accuracy” of battery model has been the main or only concern. However, the “speed” of the model identification process can be crucial as well, particularly in online applications. This topic has been touched in [11] where a control-oriented approach was used for battery model identification using subspace method. Because of the relatively small number of relevant studies in the literature, this paper has also focused on the identification time as discussed below.

This paper is aimed at addressing the compromise between simplicity (i.e., speed) and complexity (i.e., accuracy) of a battery model by investigating different model structures and data fitting algorithms. Like some other studies, ECN models are used here. Based on the results available from [7], this paper is not aimed at investigating all the model structures again and just a subset (the Thevenin-based models) is assessed here. As a contribution to knowledge, this paper focuses on different fitting algorithms and their speed (execution time) in the face of different levels of model complexity. Three identification algorithms are investigated: 1) gradient descent (GD); 2) GA; and 3) prediction error minimization (PEM). The identification algorithms are evaluated using battery experimental data and then they are analyzed with regard to their speed and accuracy. The battery experimental data used in this paper was obtained by testing a six-cell NiMH battery pack using a low-cost test bench that was proposed in [12]. In addition, in order to get a clearer understanding of the battery model identification problem, “identification surface” showing how parameter choice relates to accuracy, is presented and discussed—this has not elsewhere appeared in the literature and is another contribution of this paper. In an identification surface, fitting error is plotted as a function of the unknown battery parameters. This visual analysis of the battery parameterization problem led us to derive a new analytical formula which is able to determine exact values of battery’s open circuit voltage (OCV) and internal ohmic resistance (IOR). The proposed analytical solution is used as a benchmark to validate the other identification algorithms. Finally, the proposed framework is evaluated in a realistic case study where EV’s battery SoC is estimated based on battery model identification.

II. CONCEPTUAL STRUCTURE OF THIS PAPER

As mentioned before, this paper is focused on simplicity (speed) of the battery model identification procedure particularly for online applications. In the proposed framework in this paper, a complex optimization problem is divided into two separate simpler problems which are: 1) parameter identification and 2) state estimation. Avoiding unnecessary complexities in an “application-oriented” approach is a contribution of this paper. Fig. 1 demonstrates the general conceptual structure of this paper in which the battery parameterization and state estimation are performed at separate parts. In this concept, the battery model would be as complex as really needed and not more than that. The idea is to simplify the battery model identification procedure by reducing the number of unknown parameters in an application-oriented way. In order to clarify this more, two examples are presented in the following.

Case 1: Assume that the target is to estimate battery SoC. In the proposed concept, just the needed battery parameters are identified and then they are served by an estimator which is able to predict SoC. In this scenario, if OCV is enough for SoC estimation of a particular battery type for example, the estimator is trained (offline) to be able to predict SoC based on OCV. Consequently, the identification part should just provide OCV for the estimator in online application. This case study has been investigated in more detail in Section V.

Case 2: In another case, assume that the target is to estimate battery state-of-health (SOH). First, we should decide which parameters are suitable to be used by the estimator for this purpose. Knowing that the answer depends on the battery chemistry, let us assume that OCV and IOR are enough for a particular battery type. So, the estimator should be trained offline to be able to detect SOH based on these two parameters. It would be possible by repeating a number of tests at different ages and the results are stored as lookup tables or polynomials at the estimator’s database. Then, the identification part should just provide the needed parameters (i.e., OCV and IOR in this case) for the estimator in online application.

Other case studies can be considered as well with regard to different targets, i.e., SoC, SOH, state-of-power [13], etc. or different battery chemistry, i.e., lead-acid, NiMH, lithium-ion, etc. However, in all possible cases, the minimum number of estimator’s inputs which are needed for that particular purpose, should be found. This subject is discussed in more detail in Section V. As illustrated in Fig. 1, the inputs of the estimator are the outputs of the identification part. Artificial intelligence (AI)-based algorithms such as neural networks [14] and fuzzy logic [15] are good candidates to be the estimator in the proposed scenario due to their powerful

Fig. 1. Online battery SoC estimation based on parameter identification.
ability to approximate nonlinear functions. The advantage of this new structure is to separate online and offline parts by integrating system identification techniques together with AI techniques.

It is necessary to explain the role of the temperature in this framework as well. The temperature effect is considered in the estimator and not during the identification procedure. An idea could be repeating the estimator’s training at different temperatures, however, the temperature effect is usually a challenge in battery modeling. In such a scenario, different trained estimators are used at different ranges of temperature using the input data from a temperature sensor. In online application, the temperature would be measured and served as a separate input to the estimator demonstrated in Fig. 1.

III. BATTERY MODEL IDENTIFICATION

Based on a concept presented in [8], three main parts of a system identification process are: 1) experiment design; 2) model structure selection; and 3) fitness criterion selection. These three parts are discussed for battery model identification in this section. Then, a graphical approach is applied which provides a deeper understanding about the battery model identification problem. Consequently, an analytical solution is derived for the battery identification problem in a simplified case when the battery model has just two unknown parameters that are OCV and ohmic resistance.

A. Battery Experiment

As mentioned before, experimentation is a part of the system identification procedure. The experiments should be designed well to cover the system’s different operating conditions. Particularly for batteries, charge-discharge test at different current and temperature levels is a common approach. In these tests, battery is subjected to a charge or discharge current as the input and battery’s terminal voltage is measured as the output. Such a test is performed in this paper as well. In a case study, a six-cell pack of NiMH batteries was tested using the test bench presented in [12]. The battery tester includes a current sink for physical discharge and a microcontroller used as data acquisition device to communicate with MATLAB. The demanded currents and the voltage measurements are processed by an MATLAB function sending, receiving data and storing them in the form of column vectors for time, current, and voltages. The current is controlled through the output voltage of an operational amplifier in the current sink. Communication between host computer and current sink is realized via an Arduino Uno microcontroller [16] with MATLAB/Simulink connectivity. A pulse-width modulation signal from the microcontroller generates the control voltage to set the current to its desired value. To connect the microcontroller board to MATLAB, the Arduino-IO Package is used which transfers the signals via a serial USB connection to MATLAB workspace and vice versa. More details of the proposed test bench can be found in [12].

The NiMH battery pack was selected due to its simple and save handling as well as its convenient output voltage. Specifications of the battery pack are listed in Table I. The experiment was conducted by applying consecutive discharge current pulses to the battery and measuring the battery terminal voltage. Data is saved in time domain with a sampling rate of 1 Hz. Fig. 2 illustrates the battery measurements during an experiment. The test started from fully charged state (8.5 V) and continued until the terminal voltage dropped below the cut-off voltage (6 V) which means depleted charge state. The discharge rate is 1C that is 2.4 A and length of each pulse is 40 s with a relaxation time of 60 s in between.

B. Battery Model Structure

ECN battery models are used in this paper. The ECN battery models are constructed by putting resistors, capacitors and voltage sources in a circuit. The simplest form of an ECN battery model is internal resistance model \((R_{\text{model}})\) [17]. \(R_{\text{model}}\) includes an ideal voltage source \((V_{\text{OC}})\) and a resistance \((R_0)\) as depicted in Fig. 3 in which \(V_t\) is the battery terminal voltage and \(I_L\) is the load current. Adding one RC network to \(R_{\text{model}}\) increases its accuracy by considering the battery polarization characteristics. This model, called Thevenin model [18] or IRC model, is illustrated in Fig. 3 in
Fig. 3. ECN battery models: (a) R model, (b) 1RC model, and (c) 2RC model.

TABLE II
ECN BATTERY MODELS’ SPECIFICATIONS

<table>
<thead>
<tr>
<th>Name</th>
<th>Equations</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resistance Model (R Model)</td>
<td>$V_t = V_{OC} - R_O \cdot I_L$</td>
<td>$V_{OC}, R_O$</td>
</tr>
<tr>
<td>Thevenin Model (1RC model)</td>
<td>$\frac{dV_t}{dt} = \frac{I_L}{C_p} - \frac{V_t}{R_p C_p}$</td>
<td>$V_{OC}, R_O, R_P, C_P$</td>
</tr>
<tr>
<td>2RC model</td>
<td>$\frac{dV_t}{dt} = \frac{I_L}{C_2} - \frac{V_t}{R_2 C_2}$</td>
<td>$V_{OC}, R_O, R_1, C_1, R_2, C_2$</td>
</tr>
</tbody>
</table>

which $V_t$ is cell’s terminal voltage, $V_{OC}$ is OCV, $R_O$ is internal resistance, $R_P$ and $C_P$ are equivalent polarization resistance and capacitance, respectively. Adding more RC networks to the battery model (2RC model in Fig. 3) may improve its accuracy but it increases the complexity as well. So, a compromise is necessary when computational effort and time are vital. The ECN battery models’ specifications are presented in Table II.

C. Battery Identification Algorithm

Generally, a system identification problem can be considered an optimization problem where we are trying to find the optimum values of a model’s parameters to minimize the identification error. It should be noted that here we are talking about predefined model structures with unknown parameters. So, the parameter vector ($\theta$) is determined so that the prediction error ($\varepsilon$) is minimized, defined as follows:

$$\varepsilon(t_k, \theta) = y(t_k) - \hat{y}(t_k|t_{k-1}; \theta)$$  \hspace{1cm} (1)

where $y(t_k)$ is the measurement data at time $k$ and $\hat{y}(t_k|t_{k-1}; \theta)$ is the model’s prediction at time $k$ using the parameters $\theta$. In this case, the battery model’s parameters are optimized so that the least difference between the measured terminal voltage and the model’s output is achieved. Consequently, a fitness function like the root mean square error (RMSE) can be used as follows:

$$\text{RMSE} = \left[ \frac{1}{N} \sum_{k=1}^{N} \varepsilon(t_k, \theta)^2 \right]^\frac{1}{2}.$$  \hspace{1cm} (2)

As discussed in the introduction section, a mathematical algorithm is needed to fit the model to the measured data by minimizing RMSE and is called identification algorithm or optimization algorithm here. Three identification algorithms are investigated which are GD algorithm [19], GA [20], and PEM algorithm [8]. There are lots of references in the literature about these methods and explaining details of these algorithms is out of the framework of this paper. Here, the important point is to assess their performance and efficiency in battery model identification specifically in online applications. The contribution of this paper is that both speed and accuracy of the algorithms are investigated and different versions of one algorithm are also considered. For example, GD’s speed and accuracy depend on its step size. Another factor that can affect GD algorithm’s performance is its initial condition. In order to reduce this effect, hundreds of uniformly distributed points (250 for R model and 420 for 1RC model) are first tried and the best would be the real initial point of the GD algorithm. Similarly for GA, the performance changes by varying GA parameters [20]. Table III contains detailed information about GA options in MATLAB which is set in this paper. Here, the population size and the generation number are changed to produce faster or more accurate GA versions. Different versions of GD algorithm and GA are developed and evaluated with respect to accuracy and speed. Consequently, Pareto diagrams [21] are obtained and analyzed to demonstrate the tradeoff between accuracy and speed which is discussed later. For PEM algorithm, just one version is used which is fast and accurate comparing to the other methods. In addition to the above mentioned algorithms, an analytical solution is derived in order to obtain ohmic resistance and OCV which is discussed later.

D. Battery Identification Surface

It would be useful to know that what type of optimization problem is the battery parameterization problem. To address
this question and in order to get a more clear understanding about the battery model identification problem, “identification surfaces” are presented and discussed as a novelty of this paper. In an identification surface, identification error value (RMSE) is plotted as a function of the unknown battery parameters. As an example in 2-D space, Fig. 4 demonstrates identification surfaces for the $R_{\text{model}}$ which has two variable parameters: 1) $V_{\text{OC}}$ and 2) $R_{\text{O}}$. The range of the parameters, [6.5 8.5] (V) for $V_{\text{OC}}$ and [0.04 0.2] (ohm) for $R_{\text{O}}$, are selected based on primary identification results which are discussed in Section IV. Three surfaces are illustrated in the figure at (a) 10%, (b) 50%, and (c) 90% SoC levels. It can be seen that in all cases, there is a minimum point at which RMSE is minimum, which is the optimization solution point. It is also clear from the figure that the minimum point moves by change of the battery SoC.

The surfaces shown in Fig. 4 are obtained by using the measurement history over the past 1% SoC change. There is a point on each surface at which RMSE has the least value which is the optimization solution. It should be noted that the optimum point of the identification surface moves with regard to SoC, however, it keeps its general shape as demonstrated in Fig. 4.

E. Analytical Solution of the Battery Model Identification Problem

In this section, a new analytical solution is derived for battery model identification from measurements data. The proposed analytical solution is calculated in a special case that the battery model has just two unknown parameters ($R_{\text{model}}$). The idea of this solution is obtained as an outcome of the battery identification surfaces. It was demonstrated for $R_{\text{model}}$ that the battery identification problem has a smooth surface with a unique minimum point. However, lots of local minima would appear by adding RC networks to the model. Limiting ourselves to the case of $R_{\text{model}}$, the proposed exact solution can be used individually in practice and as a benchmark to evaluate other methods as well.

Assume that the identification error criterion in below has to be minimized

$$\text{RMSE} = \frac{1}{\sqrt{n}} \left( \sum_{i=1}^{n} (V_{t,i} - \hat{V}_{t,i})^2 \right)^{0.5}$$  \hspace{1cm} (3)

where $V_{t,i}$ and $\hat{V}_{t,i}$ are measured and estimated values of the battery terminal voltage, respectively. Using $R_{\text{model}}$, the estimated value can be obtained as follows:

$$\hat{V}_{t} = V_{\text{OC}} - R_{O}.I.$$  \hspace{1cm} (4)

The model has just two variable parameters ($V_{\text{OC}}$ and $R_{O}$), so RMSE can be considered as a function of these parameters as follows:

$$\text{RMSE} = f(V_{\text{OC}}, R_{O}).$$  \hspace{1cm} (5)

In a rectangular coordinate system, the gradient of a function like $f$, is the vector field whose components are the partial derivatives of $f$ as follows:

$$\nabla f = \frac{\partial f}{\partial V_{\text{OC}}} e_1 + \frac{\partial f}{\partial R_{O}} e_2$$  \hspace{1cm} (6)
where \( e_1 \) and \( e_2 \) are the orthogonal unit vectors pointing in the coordinate directions.

Because the minimum value of RMSE is desired, the corresponding optimum values of the parameters are obtained by putting the gradient equal to zero

\[
\nabla f = 0 \Rightarrow \frac{\partial f}{\partial V_{OC}} = 0 \quad \text{and} \quad \frac{\partial f}{\partial R_O} = 0.
\]

(7)

The inputs of this optimization problem are current and voltage measurements whereas the outputs are \( V_{OC} \) and \( R_O \). Assume that we have a set of measurement data containing current \( (I) \) and terminal voltage \( (V_t) \) as follows:

\[
\begin{align*}
V_t &= [V_{t,1}, V_{t,2}, V_{t,3}, \ldots, V_{t,n}] \\
I &= [I_1, I_2, I_3, \ldots, I_n]
\end{align*}
\]

(8)

(9)

where \( n \) is the length of identification horizon or the number of data points at each identification iteration. It should be noted that the measurement vectors are constant during the optimization process. On the other hand, the estimated values of the battery terminal voltage are obtained from the model as follows:

\[
\hat{V}_t = \left[ \hat{V}_{t,1}, \hat{V}_{t,2}, \hat{V}_{t,3}, \ldots, \hat{V}_{t,n} \right]
\]

\[
= [V_{OC} - R_O I_1, V_{OC} - R_O I_2, \ldots, V_{OC} - R_O I_n].
\]

(10)

Having \( V_t \) and \( \hat{V}_t \), RMSE can be rewritten as follows:

\[
\text{RMSE} = \left( \frac{(V_{t,1} - V_{OC} + R_O I_1)^2 + \cdots + (V_{t,n} - V_{OC} + R_O I_n)^2}{n} \right)^{1/2}.
\]

(11)

Now, the gradient terms are obtained as follows:

\[
\begin{align*}
\frac{\partial \text{RMSE}}{\partial V_{OC}} &= \frac{1}{2\sqrt{n}} -2(V_{t,1} - V_{OC} + R_O I_1) - \cdots - 2(V_{t,n} - V_{OC} + R_O I_n) \\
&= \frac{1}{2\sqrt{n}} \left( \frac{(V_{t,1} - V_{OC} + R_O I_1)^2 + \cdots + (V_{t,n} - V_{OC} + R_O I_n)^2}{n} \right)^{1/2}
\end{align*}
\]

(12)

and putting the gradient equal to zero, we have

\[
\frac{\partial \text{RMSE}}{\partial V_{OC}} = 0 \Rightarrow V_{OC} - \frac{R_O}{n} \cdot \sum_{i=1}^{n} I_i - \frac{1}{n} \cdot \sum_{i=1}^{n} V_{t,i} = 0.
\]

(13)

Similarly, for \( R_O \), we have

\[
\begin{align*}
\frac{\partial \text{RMSE}}{\partial R_O} &= \frac{1}{2\sqrt{n}} 2(V_{t,1} - V_{OC} + R_O I_1)(I_1) + \cdots + 2(V_{t,n} - V_{OC} + R_O I_n)(I_n) \\
&= \frac{1}{2\sqrt{n}} \left( \frac{(V_{t,1} - V_{OC} + R_O I_1)^2 + \cdots + (V_{t,n} - V_{OC} + R_O I_n)^2}{n} \right)^{1/2}
\end{align*}
\]

(14)

and

\[
\frac{\partial \text{RMSE}}{\partial R_O} = 0 \Rightarrow V_{OC} - \frac{R_O}{n} \cdot \sum_{i=1}^{n} I_i^2 - \frac{\sum_{i=1}^{n} V_{t,i} I_i}{\sum_{i=1}^{n} I_i} = 0.
\]

(15)

Combining (13) and (15) and eliminating \( V_{OC} \), we have

\[
\frac{R_O}{n} \cdot \sum_{i=1}^{n} I_i + \frac{1}{n} \cdot \sum_{i=1}^{n} V_{t,i} - R_O \cdot \sum_{i=1}^{n} I_i^2 - \frac{\sum_{i=1}^{n} V_{t,i} I_i}{\sum_{i=1}^{n} I_i} = 0
\]

\[
\Rightarrow R_O = \left( \frac{\sum_{i=1}^{n} V_{t,i} I_i}{\sum_{i=1}^{n} I_i} - \frac{\sum_{i=1}^{n} V_{t,i} I_i}{\sum_{i=1}^{n} I_i} \right) \left( \frac{\sum_{i=1}^{n} I_i^2}{n} - \frac{\sum_{i=1}^{n} I_i^2}{\sum_{i=1}^{n} I_i} \right).
\]

(16)

(17)

Having \( R_O \), we can obtain \( V_{OC} \) by using any of (13) or (15). These \( V_{OC} \) and \( R_O \) are the optimum battery model’s parameters that give us the least RMSE. It should be noted that the proposed analytical solution is obtained for \( R \) model and not just for NiMH battery type. In other words, the above formula can be used for any battery chemistry. By changing the battery type, the identification surfaces will change a bit in shape, however, the general trend remains fixed. The proposed formula gives the average values of IOR and OCV by using charge and discharge data at the same time.

In the next sections, the analytical solution is used to validate the results of GD, GA, and PEM algorithms in case of the \( R \) model. In addition, the proposed analytical solution is used independently in real-time applications when IOR and OCV are enough like NiMH battery as depicted in Fig. 1.

IV. IDENTIFICATION RESULTS

A. Battery Parameterization

As an example of battery parameterization results, the four parameters of IRC model are obtained using PEM algorithm as presented in Fig. 5. The identification process is repeated over the whole range of SoC at 20 °C temperature at regular intervals; called “identification window” or “identification horizon.” The battery identification window can be a time window or SoC window. However, a combination of both is designed and used in this paper because of our interest in EV application. Since the power demand from an EV’s battery pack can change in a wide range, identifying the battery model at regular time intervals is not effective. On the other hand, EV battery’s SoC can change in few seconds when the power demand is very high. It may cause numerical problems for the identification process when the number of data points is not enough to identify the parameters. In Fig. 5, identification process is repeated every 1% change in SoC. So, the battery model is identified using the measurement’s history in the past 1% SoC. However, if the identification window’s length is less than 120 s, the window’s length is extended to the past 120 s.

B. Battery Identification’s Time and Accuracy Using Different Algorithms

In this paper, battery parameterization is performed using different identification algorithms which are GD, GA, and PEM. The battery model structure is fixed during the identification process and the parameters are obtained to minimize the fitting error between the model’s output and the experimental data. Indeed, the identification problem is an optimization problem in this case. The fitting quality is measured by RMSE criterion, introduced in (2). So, the minimum value of RMSE is desired. On the other hand, this paper is focused on the
identification time as well. In order to demonstrate the trade-off between accuracy and speed of the identification process, a set of feasible solutions, called Pareto optimal set [21], is obtained by repeating the identification process at different levels of accuracy as illustrated in Fig. 6(a) and (b) for GD and GA, respectively.

For PEM algorithm just one version is used because of its different features. Comparison between the algorithms is discussed in more detail in the next section. Average identification time and RMSE by applying different identification algorithms are presented in Fig. 7(a) and (b), respectively, and in Table IV as well. The identification time is the time that is needed to complete the battery parametrization process. Because the identification process needs to be repeated at different SoC levels, average value of the execution time of the identification code is used as an index here.

In addition, the effect of battery model structure on the tradeoff results is investigated by using different models introduced in Table II. As an example, two versions of each GD and GA algorithms are selected based on Pareto analysis to be compared with PEM method as demonstrated in Fig. 7. The analytical solution is used as a benchmark in the case of $R$ model. The compromise between speed and accuracy for different algorithms can be analyzed based on these results which is discussed in the results analysis section.
C. Battery Model Structures

The results demonstrate that the identification error has decreased by adding one RC network to the battery model while the identification time is increased by adding more complexity. So, it is important to investigate if it is beneficial to use the more accurate model or keeping the faster one. For this purpose, different model structures are examined and compared using a fixed algorithm (i.e., PEM). The identification results of PEM method for R, 1RC, and 2RC models are stated in Table V. RMSE values are quite similar for 1RC and 2RC models. In other words, there is no remarkable improvement in voltage prediction by adding the second RC network. This outcome is also in accordance with the results published in previous studies in [7]. Further discussions about the results are presented in the following sections.

V. APPLICATION OF BATTERY MODEL IDENTIFICATION FOR SOC ESTIMATION

In a case study, the proposed identification formula, developed in Section III part E, has been used for battery SoC estimation. Referring back to Fig. 1, this section is about the 1st case which was presented in Section II. The goal is testing the proposed identification procedure in a realistic scenario that is battery SoC estimation in an EV. As mentioned in Section II, one of the contributions of this paper is achieving the target through the simplest way using the proposed framework shown in Fig. 1. Since the final target is SoC estimation in this case, it is necessary to identify only the needed parameters for SoC estimation, otherwise we are just increasing computational effort in real-time with no additional gain. Our results and the results in [22]–[24] demonstrate that the OCV curve of NiMH battery contains enough information for SoC estimation. Consequently, only OCV parameter is used for SoC estimation in this case.

In order to simulate the NiMH battery pack under real conditions in an EV, Nissan Leaf EV is simulated [25] over two real world driving cycles: 1) the worldwide harmonized light vehicles test procedure (WLTP) class 3 and 2) the urban dynamometer driving schedule (UDDS). Then the power demand is scaled down (500:1) to be applicable for the small battery pack [26]. These simulations give us battery’s current and terminal voltage which can be used as the inputs of the identification part. Consequently, battery’s OCV is identified during the driving cycle simulations. The driving cycles are repeated until the battery is completely depleted in order to cover the whole SoC range. Simulation results are illustrated in Fig. 8 in which the EV moves on WLTP and UDDS driving cycles.

Next step is the use of the identification results for SoC estimation. As demonstrated in Fig. 1, the SoC estimator should be able to make the connection between OCV and SoC. An adaptive neuro-fuzzy inference system (ANFIS) [15], [27] is utilized for this purpose here. So, ANFIS stands as a function like $f$ in bellow which gives us SoC value by having OCV

$$
\text{SoC} = f(V_{OC})
$$

In this scenario, ANFIS should be trained offline using experimental data containing OCV values in the whole range of SoC. This information is stored as “ANFIS rules.” In online application, OCV is obtained in the identification part and then ANFIS predicts SoC as demonstrated in Fig. 9.

The OCV-SoC characteristic curve is roughly stable over temperature variations and it has been used for other battery types as well [28]. However, the temperature effect can be considered by repeating the training process at different temperatures. By this way, different ANFIS estimators can be used at different temperature ranges. Consequently, the nonlinearity which is caused by the temperature effect is handled by dividing the whole range into smaller parts. The proposed structure can be extended for SOH estimation as well. In that case, the effect of aging on OCV should be considered at the estimator. The OCV-SoC characteristic curve is stable over aging of the battery when normalized by age and age dependent battery capacity [29]. In the present case, the tests are done at constant temperature ($20^\circ$C) for a new battery. Specifications of the proposed ANFIS battery SoC estimator, such as type and number of membership functions (MFs), are presented in Table VI. More details about the ANFIS structure and its specifications can be found in [15].
Coulomb-counting (CC) method is used here as a benchmark for validation of SoC estimation results. It should be noted that CC method cannot be utilized in practice because of its practical limitations discussed in a number of previous studies [3], [26], [29], [30]. However, it is an ideal reference to evaluate other battery SoC estimation techniques. In CC method, SoC is calculated by integrating the load current to know how much capacity is used and remained. Supposing SoC₀ as the initial SoC at time t₀, battery’s SoC at time t is defined as follows:

\[ \text{SoC} = \text{SoC₀} - \left( \int_{t_0}^{t} \frac{\gamma i(t)}{C_t} \, dt \right), \quad 0 < \text{SoC} < 1 \]  

(19)

where i(t) is the current in ampere (A) and is assumed positive for discharging and negative for charging. Parameter \( \gamma \) is the battery Coulombic efficiency (which is a number a bit less than 1, however, here is considered 1) and \( C_t \) is cell’s total capacity in ampere-second (A.s) when the time is in second. Therefore, SoC is a number between 0 and 1 representing depleted and fully-charged states, respectively. The NiMH battery pack’s SoC is calculated using both CC and ANFIS methods as demonstrated in Fig. 10 for WLTP and UDDS simulations. The results demonstrate that the proposed identification algorithm is able to identify OCV parameter very well. The identification results are validated against the results of experimental data as demonstrated in the figure as well. The average SoC estimation error is around 2% whereas the maximum error is less than 10% as stated in Table VII.

<table>
<thead>
<tr>
<th>Driving Cycle</th>
<th>Mean Error</th>
<th>Maximum Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLTP</td>
<td>2.13%</td>
<td>9.90%</td>
</tr>
<tr>
<td>UDDS</td>
<td>1.12%</td>
<td>3.58%</td>
</tr>
</tbody>
</table>

VI. RESULTS ANALYSIS AND DISCUSSION

The results of this paper can be classified into different groups which are discussed separately in the following parts.

A. Visual Investigation of the Battery Model Identification Problem

At the first stage of this paper, the battery model identification problem was handled using a visual approach. The goal was to get a deeper understanding of the problem. For this purpose, battery model identification is analyzed as an optimization problem and RMSE surfaces (Fig. 4) are obtained as a contribution of this paper. The simplest model structure (R model) was used to visualize the identification problem in

---

**TABLE VI**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value/Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of fuzzy inference system</td>
<td>Sugeno-type</td>
</tr>
<tr>
<td>Number of inputs</td>
<td>1</td>
</tr>
<tr>
<td>Number of outputs</td>
<td>1</td>
</tr>
<tr>
<td>Number of inputs’ MFs</td>
<td>5</td>
</tr>
<tr>
<td>Type of input’s MFs</td>
<td>Generalized bell-shaped</td>
</tr>
<tr>
<td>Type of output’s MFs</td>
<td>Linear</td>
</tr>
<tr>
<td>Defuzzification method</td>
<td>Weighted average</td>
</tr>
<tr>
<td>Training epoch number</td>
<td>20</td>
</tr>
<tr>
<td>Number of points in training dataset</td>
<td>50</td>
</tr>
<tr>
<td>Number of points in checking dataset</td>
<td>50</td>
</tr>
</tbody>
</table>

**TABLE VII**

<table>
<thead>
<tr>
<th>Driving Cycle</th>
<th>Mean Error</th>
<th>Maximum Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLTP</td>
<td>2.13%</td>
<td>9.90%</td>
</tr>
<tr>
<td>UDDS</td>
<td>1.12%</td>
<td>3.58%</td>
</tr>
</tbody>
</table>
two dimensions. This presentation type of the battery model identification problem, which is done here for the first time, helps us to clarify the problem more. Consequently, more suitable methods can be selected with regard to the number and type of the optimum points and smoothness of the optimization surface. For example, GD method was selected as a result of this visual analysis because of the smoothness of the surfaces. The analytical solution was also originated from this visual analysis. The battery model identification surface is investigated at different SoC levels. It is demonstrated that how the optimum point of the identification surface moves with regard to the battery SoC. It is also demonstrated that the general shape of the surface remains consistent with regard to the local and global optimum points and smoothness of the surface.

B. Development of New Analytical Solution for Obtaining Battery’s IOR and OCV From Measurement Data

As an achievement of this paper, a new analytical solution was derived for battery parameterization. A closed-form formula was presented for obtaining battery’s IOR and OCV from measurement data including current and voltage. The advantages of the proposed formula are its speed and accuracy and its applicability for any battery type. A limitation of the proposed formula is that it gives the average values of IOR and OCV by using charge and discharge data at the same time. Using the average value of battery OCV during charge-discharge is a common way to reduce battery model’s complexity especially in online applications. More details on the calculation of the average battery OCV can be found in [31]. Another issue that should be considered before using the analytical formula is the measurement frequency. The number of measurement data points should be large enough in order to get consistent identification results. In this paper, the identification window was at least 60 s which includes 60 data points measured at 1 Hz.

C. Investigating Different Algorithms for Online Battery Parameterization

Another part of the results was obtained by investigating different identification algorithms. Five identification codes are developed based on three different algorithms to investigate the tradeoff between accuracy and speed. The results in Table IV demonstrate that the $R$ model’s parameters can be identified using any of the algorithms and there is no major difference between them because of the simplicity of the model. However, the results show something different for $1RC$ model. Consequently, the choice of the battery identification algorithm highly depends on the model’s complexity. This paper aims at demonstrating such an application-oriented point of view where different algorithms are selected based on the amount of complexity that is needed for each particular application.

It is concluded that GD and GA algorithms are easily adjustable by changing their parameters. This flexibility makes them suitable to do the compromise between precision and speed for different applications. On the other hand, the results demonstrate that PEM method is a suitable algorithm for battery model identification which has both speed and accuracy at the same time. However, there are other important criteria that can be considered for algorithm selection which are the simplicity of programming and being easily embeddable. With this aspect, GD algorithm is the simplest one that can be easily programmed and embedded for real-time battery model identification purposes. Again it should be noted that the discussion here is just about the battery model identification problem, and the comparison between the algorithms’ performance cannot be extended to other problems. For example, GD algorithm is not recommended for problems in which lots of local minima exist. Here, GD method is just selected as an option based on the battery identification surfaces that were obtained in the first part of this paper.

D. Investigating the Tradeoff Between Accuracy and Simplicity for Different Battery Model Structures

The results of this paper demonstrate that in addition to the identification algorithm selection, the battery model structure selection can also be a bit tricky especially for online applications. In order to address this issue, we have focused on the question, how well can the identification algorithms handle an additional complexity in the model?

Here, the battery model’s complexity is increased by adding more RC networks to the model. Referring to the results
presented in Fig. 7 and Table IV, it is concluded that GD algorithm’s speed decreases significantly by adding model’s parameters is proportional to $2^k$. For example, the time needed to identify 1RC model (which has four parameters) is 4 $(2^4/2^2 = 4)$ times more than $R$ model (which has two parameters), and the time needed to identify 2RC model (which has six parameters) is 16 $(2^6/2^2 = 16)$ times more than $R$ model and so on. So, GD method is not suggested for very complex models, however, it is a good choice when a simple and easily embeddable identification algorithm is needed for online applications.

On the other hand, GA algorithm may not be successful in a short period of time, however, it can handle additional complexity if there is no time limitation such as in offline applications. In this case, GA algorithm outperforms GD because of its additional capabilities like jumping out from the local minima. So, GA method is suggested to be used for very complex models when time is not the matter. In addition, GA algorithm is the most difficult one to be embedded because of its programming complexity.

Referring to the results of PEM method presented in Table V, there is an improvement in RMSE by adding the first RC network. However, this achievement is obtained by increasing the identification time to four times more. Adding the second RC network causes a small improvement of error while the identification time becomes two times more approximately. So, it is concluded that 2RC battery model structure is not necessary in this case. All in all, PEM method is suggested when battery model’s complexity is needed in online applications, however, its embeddable program is not as simple as GDs.

**E. Online Battery SoC Estimation Using System Identification and ANFIS**

In a case study, the proposed battery model identification framework was used for online battery SoC estimation. The battery parameterization curves, presented in Fig. 5, demonstrate that NiMH battery OCV curve is a smooth and consistent curve that can be used individually for SoC estimation. Based on this result, the proposed analytical method is used to obtain OCV from online measurements. The same procedure can be applied in all similar cases in which OCV and/or IOR is enough for that particular target. In other cases that more complexity is needed, the other identification algorithms can be applied as discussed before.

ANFIS was utilized as an estimator here in order to find the relationship between battery parameters and SoC. However, other techniques/tools can be used as well to build such a mapping between the identification results and battery SoC. Consequently, SoC estimation error depends on the identification precision and the estimator’s performance both as demonstrated in Fig. 10 in which an identification error directly caused an error in SoC estimation. For NiMH battery, an identification error in the middle SoC range, between 30% and 80%, would cause much estimation error since SoC is more sensitive to OCV variation in this area.

**VII. Conclusion**

In this paper, the following contributions were presented to the advancement of the battery model identification problem particularly for online applications.

1) A new battery model identification procedure by focusing on simplicity (speed) for online applications. In this concept, the battery model would be only as complex as really needed and not more than that. The idea is to simplify the battery model identification procedure by reducing the number of unknown parameters in an application-oriented way.

2) Battery identification surfaces. For a deeper understanding about the battery model identification problem, the identification surfaces were presented and discussed as a novelty of this paper. A battery identification surface demonstrates how the identification error changes as a function of the unknown battery parameters. It should be noted that the proposed approach is limited to the simplest battery model ($R$ model) with two parameters.

3) A novel analytical formula for battery model identification. The proposed formula enables us to obtain battery’s OCV and IOR from current and voltage measurements in time domain. The proposed method is computationally inexpensive and accurate. It can be used for different targets such as battery SoC or SOH estimation if battery OCV and/or IOR is enough for that purpose.

4) Tradeoff between accuracy and speed. The proposed approach takes both of the battery model structures and the identification algorithms into consideration at the same time. Pareto optimal sets are obtained which allow a designer to easily select the desired levels of accuracy and speed. In other words, our results demonstrate that a combination of the both factors should be considered with respect to the needed accuracy and speed for each particular application.

**REFERENCES**


