

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

Detection of False Data Injection Attacks in Battery Stacks Using Input Noise-Aware Nonlinear State Estimation and Cumulative Sum Algorithms

Victoria O'Brien, Vittal S. Rao, and Rodrigo D. Trevizan

Abstract—Grid-scale battery energy storage systems (BESSs) are vulnerable to false data injection attacks (FDIAs), which could be used to disrupt state of charge (SoC) estimation. Inaccurate SoC estimation has negative impacts on system availability, reliability, safety, and the cost of operation. In this paper a combination of a Cumulative Sum (CUSUM) algorithm and an improved input noise-aware extended Kalman filter (INAEKF) is proposed for the detection and identification of FDIAs in the voltage and current sensors of a battery stack. The series-connected stack is represented by equivalent circuit models, the SoC is modeled with a charge reservoir model and the states are estimated using the INAEKF. The root mean squared error of the states' estimation by the modified INAEKF was found to be superior to the traditional EKF. By employing the INAEKF, this paper addresses the research gap that many state estimators make asymmetrical assumptions about the noise corrupting the system. Additionally, the INAEKF estimates the input allowing for the identification of FDIA, which many alternative methods are unable to achieve. The proposed algorithm was able to detect attacks in the voltage and current sensors in 99.16% of test cases, with no false positives. Utilizing the INAEKF compared to the standard EKF allowed for the identification of FDIA in the input of the system in 98.43% of test cases.

Index Terms— anomaly detection, cumulative sum, equivalent circuit model, false data injection attacks, noisy input, smart grid.

I. INTRODUCTION

BATTERY Energy Storage Systems (BESSs) tied to the electric power grid are composed of several battery cells, which are connected to form battery stacks to meet power requirements [2]-[4]. BESS are cyberphysical systems (CPSs) that require several electronic control and protection devices equipped with computation and communications capabilities for their safe and efficient operation [5]. One of them, the Battery management system (BMS) is required to

control the charging and discharging of the cells by monitoring sensor readings, estimating system states, and ensuring safe operation of the BESS [2], [6], [7]. BMSs typically include stack current sensors, voltage sensors on each of the battery cells, and an additional voltage sensor for the battery stack, all of which could be susceptible to false data injection attacks (FDIAs). FDIAs corrupt sensor data, which could result in incorrect BMS protection, operation, and inaccurate state of charge (SoC) estimation [2]. Incorrect SoC or voltage estimates can cause cell overcharge or deep discharge, which accelerate battery degradation and can cause battery thermal runaway, power outages, and damage to costly power grid equipment [2], [8]-[13]. The goal of attackers could range from practical reduction in usable SoC to accelerated battery degradation to system malfunctions or to battery failure, in extreme cases.

To compromise system reliability and damage equipment, the attacker must carefully construct attack vectors which may be stealthy or agnostic to the targeted system. Stealthy FDIAs are constructed to evade bad data detection (BDD) techniques [14], [15]. As discussed in [15], to elude BDD algorithms the measurement residuals (following the injection of FDIA) must remain undisturbed. Therefore, designing stealthy FDIAs requires extensive information about the system's topology and dynamic models to obtain carefully constructed attack vectors. Liu et al. [14] discussed multiple methods, including brute-force and heuristic methods, for designing stealthy attack vectors. In this paper, we consider FDIA vectors that are agnostic to the system's topology and parameters. These simpler, non-stealthy attacks are more straightforward and inexpensive to launch and are more likely to be encountered in real systems.

For the reliable and safe operation of the grid, it is critical that FDIAs targeting grid-scale BESSs are detected and miti-

This paragraph of the first footnote will contain the date on which you submitted your paper for review, which is populated by IEEE. This work was supported by the U.S. Department of Energy, Office Electricity, Energy Storage program. This article has been authored by an employee of National Technology & Engineering Solutions of Sandia, LLC under Contract No. DE-NA0003525 with the U.S. Department of Energy (DOE). The employee owns all right, title and interest in and to the article and is solely responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan <https://www.energy.gov/downloads/doe-public-access-plan>. (Corresponding author: V. O'Brien).

Victoria O'Brien is with the Department of Electrical & Computer Engineering at Texas Tech University, Lubbock, TX USA (e-mail: Victoria.obrien@ttu.edu) and the Energy Storage Technology & Systems group at Sandia National Laboratories, Albuquerque, NM, USA (e-mail: vaobrie@sandia.gov).

Vittal S. Rao is with the Department of Electrical & Computer Engineering at Texas Tech University, Lubbock, TX USA (e-mail: vittal.rao@ttu.edu)

Rodrigo D. Trevizan is with the Energy Storage Technology & Systems group at Sandia National Laboratories, Albuquerque, NM, USA (e-mail: rdtrevi@sandia.gov).

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>

This article has been authored by an employee of National Technology & Engineering Solutions of Sandia, LLC under Contract No. DE-NA0003525 with the U.S. Department of Energy (DOE). The employee owns all right, title and interest in and to the article and is solely responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan <https://www.energy.gov/downloads/doe-public-access-plan>.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

gated as quickly as possible. Methods for detecting data integrity attacks, including FDIAs, targeting CPSs have been studied extensively in literature [2], [14], [16] - [25]. The FDIAs are detected using either model-based or data-driven methods. Model-based methods typically utilize statistical approaches to detect anomalies on the residuals obtained by state estimators [1], [18], [20], [26]. On the other hand, Data-driven methods [23] - [25] do not require a model of the physical system and focus on learning system dynamics by processing sensor data streams. Common statistical methods used to detect FDIAs and other anomalies are the chi-squared test [18] and variations of the cumulative sum (CUSUM) algorithm [1], [19], [20], [26]. The chi-squared test is a popular choice to discover abnormalities in CPSs [18], [27] but is known to have higher false positive rates. The chi-squared test was used to detect outliers in the measurements of a battery during SoC estimation and was found to trigger multiple false positives [27]. Bombarding the BMS with warnings about anomalies when nothing is truly wrong could hinder system operation, therefore it is critical to minimize false positives while maximizing true positives. One benefit of the CUSUM algorithm used in this paper is that the false positive rate can be adjusted to zero by appropriately selecting the parameters of the algorithm, most commonly the value of the upper and lower thresholds.

A significant limitation of residual-based FDIA detectors is their inability to detect errors in the input variables [1], [2]. Most state estimation methods, including variants of the Kalman Filter (KF) [28], assume that inputs are accurately known and deterministic. Therefore, residual-based error detection cannot be applied to the input since their residuals cannot be calculated. The dynamic models used in SoC estimation utilize the stack current as the system input [3], which is obtained from current sensors. Considering the architecture of BMSs [5], [29], it is clear that, similar to voltage measurements, current sensors are also subject to disturbances from noise, faults, or attacks. Nevertheless, a gap exists in the battery SoC estimation literature as uncertainty in the input variable is not accounted for. Variants of the KF that consider noisy inputs, such as [23], [24], have not been applied to battery SoC estimation nor FDIA detection. State estimators based on the unknown input observer or robust state estimation (e.g. [21], [22], [30]) could potentially mitigate errors introduced by FDIA in the input variable or even detect input attacks in SoC estimation. However, those require the solution of relatively large optimization problems involving time-series of sensor data, making them ill-suited for implementation in real-time embedded systems with limited memory and computational capacity like BMSs.

Data-driven anomaly identification methods are used more frequently in literature but have their own set of limitations. Authors in [24] and [25] use deep learning methods, specifically convolutional neural networks (CNNs), to detect errors in battery systems' operation which may include faults or cyberattacks. In the survey paper [23], Sayghe et al. discuss supervised, semi-supervised, unsupervised, and deep learning methods to detect FDIAs in power systems. To train data-driven methods, hundreds of datasets may be required. For example, to train the

CNN presented in [24], 100 control datasets and 300 fault datasets were used. Battery data used to train data-based methods is often proprietary data that may be challenging or expensive to obtain, giving residual-based battery modeling methods an advantage over data driven methods. In addition, battery models can be adjusted to account for changing battery parameters by updating the model parameters [8], [31], whereas data driven methods rely highly on the training data and algorithm setup to yield accurate detection [23].

A. Overview of the Method and Paper Organization

This paper extends the work [1] and addresses many of the aforementioned limitations of the state-of-the-art by developing an input noise-aware extended KF (INAEKF) to perform SoC estimation. The INAEKF is better suited for the SoC estimation framework than the extended KF (EKF), as it is well known that battery stack current measurements are susceptible to noise and may be subject to spoofing attacks [32]. The proposed method extends the EKF utilized in [1] by considering additive Gaussian noise in the input of the state estimator. Accounting for additive noise in the input variable has been previously proposed for the KF [33] and the unscented KF (UKF) [34], but none of those methods have been applied to SoC estimation and we were unable to find literature that provided the framework for an EKF that considered input noise, which is accomplished in this paper. These methods allow for estimation of the input variable and, consequently, an input residual, which allows for postprocessing of the input signal for faults, bad data, and attack detection. Additionally, the INAEKF can mitigate the error introduced by noisy input data, thus providing more accurate state estimates than the traditional EKF when the input variable is subject to zero-mean random noise. That is especially relevant given that SoC estimation of batteries is known to be susceptible to measurement errors [35].

To conform to the new state estimation framework, the battery model proposed in this paper presents improvements upon the model utilized in [1]. These enhancements include removing the stack current from the vector of measurements and modeling it in the state estimator as a single noisy input variable, i.e., its signal is not split into two variables representing the charging and discharging currents. Instead, a nonlinear function is used to account for the stack current-dependent Coulombic efficiency. The case studies demonstrate the superior performance of the INAEKF when compared with the EKF in terms of estimation accuracy and FDIA detection. We also compare the CUSUM with the chi-squared test for attack detection.

In summary, the key contributions of the paper are:

- 1) The proposed estimation method, the INAEKF, which is an extension of the EKF and considers uncertainties in the input variables;
- 2) Modifications to the nonlinear equivalent circuit models (ECM) of the battery stack to accommodate the noisy input current variables.
- 3) A residual-based input error estimation method which is accomplished with the INAEKF; and
- 4) The application of the combination of the CUSUM algorithm with the INAEKF and its input residual estimator

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

for the identification of FDIA targeting battery stack SoC estimation.

In this paper, a three-pronged method is presented to detect and identify simple FDIAs that are injected to the voltage and/or current sensors of a series connected battery stack. First, a battery model is used to describe the batteries' dynamics accurately within the state estimator. Then, the residuals are obtained by comparing the INAEKF-estimated voltages and current of the battery stack with their respective measurements. Finally, the residuals are post processed using a tuned CUSUM algorithm to identify anomalies, indicating the presence of an attack. The overall process is outlined in Fig 1.

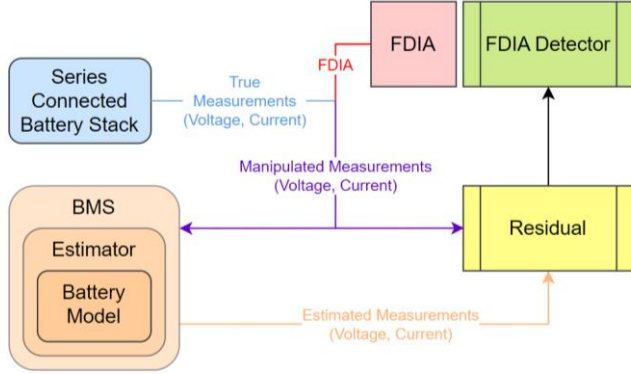


Fig. 1. Framework for detection and identification of FDIA in SoC estimation.

The remainder of the paper is organized as follows. Section II. presents the battery models and governing equations. The INAEKF state estimation method and the input noise aware filter are introduced in Section III. Section IV. presents the detection of FDIAs using the Cumulative Sum (CUSUM) algorithm. The simulation setup used to assess the performance of the proposed methods is presented in Section V. The results of the case studies are presented in Section VI. The conclusion of the paper is presented in Section VII.

II. BATTERY MODEL

Thevenin ECMs (Fig. 2.) are commonly used to model the dynamics of single battery cells when subject to changes in the stack current [16], [26]. The Thevenin ECM used in this paper includes a voltage source (V_{oc}) that is a function of the SoC of the battery cell. The dynamics of the SoC for each battery cell are represented by a charge reservoir model (CRM) (Fig. 3.), which defines capacity in units of electric charge [3]. The battery stack is modeled by considering that multiple battery cells are connected in series, therefore subject to the same current.

We consider that the stack of batteries is equipped with voltage sensors for each cell and an additional stack voltage sensor. The stack voltage can be found by adding each cell's voltage, so including the stack sensor (V_{stack}) allows for the redundancy of measurements. The input to the system is the stack current ($i_{bat}[k]$), which is used to control the charging and discharging of the battery stack and is measured using a current sensor.

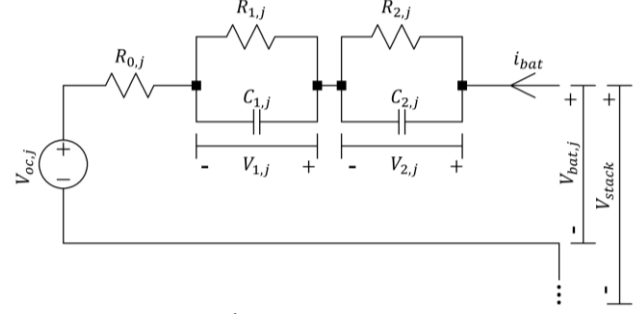


Fig. 2. ECM for the j^{th} cell in a stack of N batteries.

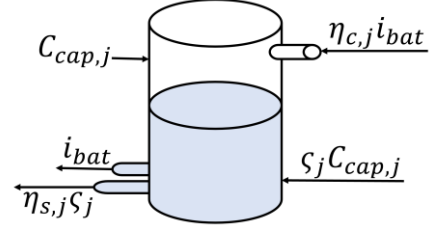


Fig. 3. CRM for the j^{th} cell in a stack of N batteries.

Instead of considering two input currents to represent the Coulombic SoC efficiency as in [1], [3], [36] we have decided to leverage the INAEKF's ability to handle nonlinear equations and utilized the nonlinear function $\eta_{c,j}(i_{bat}[k])$ to represent the current-dependent cell efficiency. This formulation eliminates the need to represent the stack current as two separate variables by introducing a function whose value varies between 1, when the battery is discharging ($i_{bat} < 0$), and η_c , when the battery is charging ($i_{bat} > 0$). The discrete-time governing equations describing the physics of the battery stack for ECM and CRM are summarized in Table I.

TABLE I
GOVERNING EQUATIONS FOR THE J^{TH} BATTERY CELL

State transition:
$\zeta_j[k+1] = e^{\eta_s \Delta t} \zeta_j[k] + \frac{\eta_{c,j}(i_{bat}[k]) \Delta t}{C_{cap,j}} i_{bat}[k]$
$v_{i,j}[k+1] = e^{\frac{-\Delta t}{R_{i,j} C_{i,j}}} v_{i,j}[k] + \frac{\Delta t}{C_{i,j}} i_{bat}[k]$
Charge and discharge efficiency:
$\eta_{c,j}(i_{bat}[k]) = 1 + \eta_{c,j} + (1 - \eta_{c,j}) \tanh(-a i_{bat}[k])$
Output:
$v_{bat,j}[k] = v_{oc,j}(\zeta_j[k]) + v_{1,j}[k] + v_{2,j}[k] + R_{0,j} i_{bat}[k]$
$V_{stack}[k] = v_{bat,1}[k] + \dots + v_{bat,N}[k]$

In the battery equations j is a subscript representing the j^{th} battery in the stack, i is a subscript that denotes the i^{th} RC pair in the ECM, Δt is the sampling time, k is the timestep, v_i is the voltage drop across the i^{th} RC pair, R_i and C_i are the resistance and capacitance of the i^{th} RC pair, respectively, R_0 is the battery ohmic resistance, i_{bat} is the battery current, C_{cap} is the battery capacity, η_c is the charge efficiency, η_s is the self-discharge coefficient, a is a parameter for the step function $\eta_{c,j}(i_{bat}[k])$,

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

v_{bat} is the battery cell voltage, V_{stack} is the battery stack voltage, v_{oc} is the open circuit voltage (OCV), and ζ is the SoC.

III. NONLINEAR SOC ESTIMATION METHODS

The nonlinear relationship between OCV and SoC in the output equations and the input-based nonlinearities in the state transition equation of the model require the use of a nonlinear filter to perform state estimation. This section discusses nonlinear estimation methods, specifically the EKF in Section III.A and its extension considering additive noise in the input variables, the INAEKF, in Section III.B.

The basic framework of each estimator was the same as [1] regardless of the estimator used. For each timestep (k), the state vector ($x[k]$) is comprised of the states: the SoC ($\zeta_1 \dots \zeta_N$) for each cell and the RC pairs' voltage drops ($v_{1,1}, v_{2,1}, \dots, v_{1,N}, v_{2,N}$) for each cell. The input ($u[k]$) is the battery current (i_{bat}). The vector of system outputs ($y[k]$) contains: the battery voltage for each cell ($v_{bat,1}, \dots, v_{bat,N}$) and the total voltage of the battery stack (V_{stack}). The process ($w[k]$) and measurement ($e[k]$) noise was also considered. The noise vectors are assumed to be Gaussian white noise where $w[k] \sim \mathcal{N}(0, Q)$ and $e[k] \sim \mathcal{N}(0, R)$, with covariances Q and R . For the INAEKF, an additional noise vector was considered where $n[k] \sim \mathcal{N}(0, \Sigma)$ with covariance Σ .

A. Extended Kalman Filter

The EKF requires the calculation of Jacobian matrices, so the derivatives of the state transition (A), state input (B), and output (C) matrices must be found, which is simple to compute for this battery model. The standard EKF equations are used in this formulation and are explained in detail in [1].

B. Input Noise-Aware EKF

In many state estimation applications, such as SoC estimation of batteries, the dynamic system input for state estimation purposes (e.g., battery current) is a measurement obtained in the same way as any other system output (e.g., battery terminal voltages). Therefore, it is expected that the input would be subject to noise in the same way as the system outputs, and not as a deterministic and known signal as assumed by most state estimation frameworks. In this paper, we assume there is a mismatch between the input to the plant (the battery stack) and the input signal obtained by the state estimator, $s[k]$, which is the input signal corrupted by additive Gaussian noise $n[k]$. Fig. 4 provides a visual description of the problem.

Accounting for a noisy input signal in the equations of the EKF follows a similar development applied to the KF and UKF provided in [33] and [34], respectively, so a summarized version of the derivation will be presented in this paper.

The dynamics of the system are defined by (1) and (2). The state transition function (1) and the output function (2) are affected by process and measurement noise, respectively.

$$x[k+1] = f(x[k], u[k]) + w[k] \quad (1)$$

$$y[k] = g(x[k], u[k]) + e[k], \quad (2)$$

$$s[k] = u[k] + n[k]. \quad (3)$$

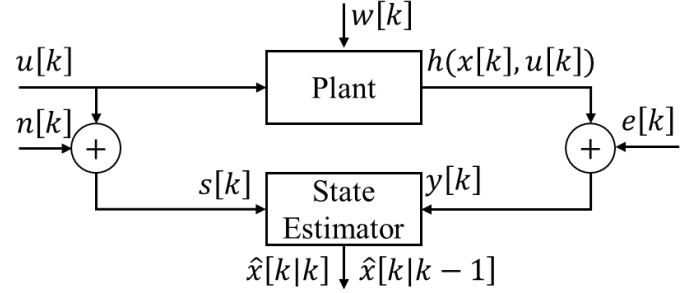


Fig. 4. The framework for SoC estimation developed in this paper considers that the input signal $s[k]$ observed by the state estimator is corrupted by the noise signal $n[k]$.

The noiseless variable $u[k]$ cannot be observed directly by the state estimator. Therefore, it is necessary to estimate $u[k]$ based on the system observed noisy outputs and inputs so it is possible to track the system states more accurately. By replacing $u[k]$ with $s[k] - n[k]$ in the above equations and aggregating the noise term $n[k]$ to the state transition and output measurement signals, we obtain the equations

$$x[k+1] = f(x[k], s[k]) + \tilde{w}[k], \quad (4)$$

$$y[k] = g(x[k], s[k]) + \tilde{e}[k], \quad (5)$$

where $\tilde{w}[k] = w[k] - B[k]n[k]$ and $\tilde{e}[k] = e[k] - D[k]n[k]$, respectively, are correlated noise terms considering the linearized approximations $B[k] = \frac{\partial f(x[k], s[k])}{\partial s[k]}$ and $D[k] = \frac{\partial g(x[k], s[k])}{\partial s[k]}$.

For the dynamic system model considered in this paper, $D[k] = D$ will be constant and, for a well-tuned parameter a , $B[k]$ is dependent on the sign of the stack current.

Following a similar approach as described in [37], the noise terms can be made uncorrelated if we apply the transformation

$$l[k] = \tilde{w}[k] - L[k](y[k] - g(x[k], s[k])), \quad (6)$$

$$L[k] = B[k]\Sigma D^T (D\Sigma D^T + R)^{-1}. \quad (7)$$

With uncorrelated noise terms it is possible to apply the improved INAEKF framework to the equations above.

1) Initialization

Similar to the standard EKF, it is necessary to initialize the recursive algorithm with initial estimates for the states and covariance matrix

$$\hat{x}[0|0] = \mathbb{E}[x[0]], \quad (8)$$

$$P[0|0] = P[0], \quad (9)$$

where $\hat{x}[0|0]$ is the initial guess for the state and $P[0]$ is the initial estimate of the state covariance.

2) State prediction equations

After rearranging the terms and calculating the conditional expected value with respect to the sequence of outputs $Y[k] = \{y[0], y[1], \dots, y[k]\}$, $\mathbb{E}[x[k+1]|Y[k]] = \hat{x}[k+1|k]$, we then obtain the state prediction equations

$$\hat{x}[k+1|k] = f(\hat{x}[k|k], s[k]) + L[k]\tilde{y}[k|k] \quad (10)$$

$$\tilde{y}[k|k] = y[k] - g(\hat{x}[k|k], s[k]) \quad (11)$$

$$P[k+1|k] = (A[k] - L[k]C[k|k])P[k|k](A[k] - L[k]C[k|k])^T + B[k]\Sigma B[k]^T - L[k]D\Sigma B[k]^T + Q \quad (12)$$

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

where $\hat{x}[k+1|k]$ is the state prediction at $k+1$, $P[k+1|k]$ is its covariance matrix, $P[k|k]$ is the covariance of the state correction vector, $A[k] = \left. \frac{\partial f(x[k], s[k])}{\partial x[k]} \right|_{x[k]=\hat{x}[k|k]}$, $C[k|k] = \left. \frac{\partial g(x[k], s[k])}{\partial x[k]} \right|_{x[k]=\hat{x}[k|k]}$, and $\tilde{y}[k|k]$ is the vector of output residuals.

3) State correction equations

Similarly, we can leverage the identity from [37], $\mathbb{E}[x[k]|Y[k]] = \mathbb{E}[x[k]|Y[k-1]] + cov(x[k]|y[k])\tilde{y}[k|k-1]$ to obtain the state correction equations

$$\hat{x}[k|k] = \hat{x}[k|k-1] + K[k]\tilde{y}[k|k-1], \quad (13)$$

$$\tilde{y}[k|k-1] = y[k] - g(\hat{x}[k|k-1], s[k]), \quad (14)$$

$$K[k] = P[k|k-1]C[k|k-1]^T(C[k|k-1]P[k|k-1]C[k|k-1]^T + D\Sigma D^T + R)^{-1}, \quad (15)$$

$$P[k|k] = P[k|k-1](I - C[k|k-1]^TK[k]^T), \quad (16)$$

where $\mathbb{E}[x[k]|Y[k]] = \hat{x}[k|k]$ is the vector of corrected states,

$C[k|k-1] = \left. \frac{\partial g(x[k], s[k])}{\partial x[k]} \right|_{x[k]=\hat{x}[k|k-1]}$, I is an identity matrix,

$K[k]$ is the Kalman gain, and $\tilde{y}[k|k-1]$ is the innovation.

4) Input estimation

This framework allows the estimation of the input signal $u[k]$. Similar to how the state correction is obtained, we can estimate the value of the input using the linear estimator $\mathbb{E}[u[k]|Y[k]] = \mathbb{E}[u[k]|Y[k-1]] + cov(u[k]|y[k])\tilde{y}[k|k-1]$, which can be developed to obtain

$$\hat{u}[k] = s[k] + U[k]\tilde{y}[k|k-1], \quad (17)$$

$$U[k] = \Sigma D^T(C[k]P[k|k-1]C[k]^T + R + D\Sigma D^T)^{-1}, \quad (18)$$

where $\hat{u}[k] = \mathbb{E}[u[k]|Y[k]]$ is the input estimate and $U[k]$ is the covariance between input and the innovations.

5) Residual calculations

The residuals of the measurement (19) and the input (20) are required to be postprocessed by the CUSUM algorithm for FDIA detection and identification, and must be calculated

$$z[k|k-1] = y[k] - \hat{y}[k|k-1], \quad (19)$$

$$S[k] = s[k] - \hat{u}[k], \quad (20)$$

where $z[k|k-1]$ is the a priori measurement residual and $S[k]$ is the input residual.

A summary of the equations of the INAEKF considering noisy inputs is given in Table II.

IV. DETECTION OF FDIA IN BATTERY STACKS USING CUSUM

A. False Data Injection Attack Model

Bad actors could use FDIAs to corrupt the system's measurements to disturb state estimation. In the case of an attacker targeting SoC estimation, inaccurate estimation could result in the overcharging or overdischarging of battery cells. For attacks targeting the output, we assume the FDIA could be injected into one or more voltage sensors and that all voltage sensors in the battery stack could be susceptible to FDIAs. Like the study in [1], the FDIAs injected into the voltage sensors were small magnitude bias attacks (21). This paper extends the work done in [1] by also injecting small magnitude bias attacks into the input of the system (22).

$$y_a[k] = y[k] + \Delta y_a[k] \quad (21)$$

TABLE II
SUMMARY OF THE EQUATIONS OF THE INAEKF

System model: $x[k+1] = f(x[k], u[k]) + w[k]$ $y[k] = g(x[k], u[k]) + e[k]$ $s[k] = u[k] + n[k]$ $w[k] \sim \mathcal{N}(0, Q)$, $e[k] \sim \mathcal{N}(0, R)$, $n[k] \sim \mathcal{N}(0, \Sigma)$
Initialization: $\mathbb{E}[x[0]] = \hat{x}[0 0] = \hat{x}[0]$ $Cov(x[0] - \hat{x}[0 0]) = P[0 0] = P[0]$
State prediction (prior estimation): $\hat{x}[k+1 k] = f(\hat{x}[k k], s[k]) + L[k](y[k] - g(\hat{x}[k k], \hat{u}[k]))$ $L[k] = B[k]\Sigma D^T(D\Sigma D^T + R)^{-1}$ $P[k+1 k] = (A[k] - L[k]C[k])P[k k](A[k] - L[k]C[k])^T + B[k]\Sigma B[k]^T - L[k]D\Sigma B[k]^T + Q$ $\hat{y}[k k-1] = g(\hat{x}[k k-1], s[k])$ $A[k] = \left. \frac{\partial f(x[k], s[k])}{\partial x[k]} \right _{x[k]=\hat{x}[k k]}$ $C[k k] = \left. \frac{\partial g(x[k], s[k])}{\partial x[k]} \right _{x[k]=\hat{x}[k k]}$
State correction (posterior estimation): $\hat{x}[k k] = \hat{x}[k k-1] + K[k](y[k] - g(\hat{x}[k k-1], s[k]))$ $\tilde{y}[k k-1] = y[k] - g(\hat{x}[k k-1], s[k])$ $K[k] = P[k k-1]C[k k-1]^T(C[k k-1]P[k k-1]C[k k-1]^T + D\Sigma D^T + R)^{-1}$ $P[k k] = P[k k-1](I - C[k k-1]^TK[k]^T)$ $C[k k-1] = \left. \frac{\partial g(x[k], s[k])}{\partial x[k]} \right _{x[k]=\hat{x}[k k-1]}$
Input estimation: $\hat{u}[k] = s[k] + U[k]\tilde{y}[k k-1]$ $U[k] = \Sigma D^T(C[k k-1]P[k k-1]C[k k-1]^T + R + D\Sigma D^T)^{-1}$
Residual calculation: $z[k k-1] = y[k] - \hat{y}[k k-1]$ $S[k] = s[k] - \hat{u}[k]$

where $\Delta y_a[k]$ denotes a small magnitude attack vector (where each element may be positive, negative, or zero) added to the measurement vector ($y[k]$) and $y_a[k]$ denotes the manipulated measurement vector that is used during state estimation.

$$s_a[k] = s[k] + \Delta s_a[k] \quad (22)$$

where $\Delta s_a[k]$ is a small magnitude FDIA added to the noisy input ($s[k]$) and $s_a[k]$ denotes the corrupted input used during state estimation.

B. CUSUM Algorithm

The CUSUM algorithm is a statistical method that can be used to detect a shift in the mean of a timeseries data set [38]. In this paper the a priori measurement residual data and input residual data are used as inputs to the CUSUM algorithm. The a priori residual was chosen over the a posteriori residual since papers such as [26] found the a priori residual to result in more sensitive detection.

Variations of the CUSUM algorithm are used for change detection and the algorithm can be adjusted to suit the application it is needed for. An effective CUSUM detector requires the calculation of an upper (UCL) and lower (LCL) control limit, the

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

definition of parameters, data preprocessing (including separating data into samples), and the tuning of parameters. A high (SH) and low (SL) recursive cumulative sum is used to gauge if there is a shift in the mean of the random process. The sums are initialized to zero and then for each subsequent sample the SH and SL are calculated by using the equations found in Table III. The system is considered out-of-control if the SL exceeds the LCL or the SH exceeds the UCL for any sample. In the context of FDIA detection, an out-of-control system indicates that an attack is present in at least one of the sensors. The h parameter controls the width of the UCL and LCL and can be tuned experimentally to adjust the sensitivity of the detector. For example, increasing the h parameter could reduce or eliminate the number of false positives while lowering the h parameter could create a more sensitive detector.

The CUSUM algorithm equations are discussed in detail in [1] and are summarized in Table III. The parameters for the CUSUM algorithm are defined in Table IV and are derived in [1], [26], [39]. Since the timeseries used in the CUSUM algorithm is the residual data, the expected value is assumed to be zero [40] and therefore the population mean (μ) is equal to zero. The actual population standard deviation is unknown, but the population standard deviation (σ_z) can be estimated with the equation in Table III. The reader should refer to [1] to see how the parameters are defined in the scope of this algorithm.

TABLE III
SUMMARY OF CUSUM ALGORITHM EQUATIONS

Parameter selection:
$\gamma = \frac{\delta}{2}$
$h = d\gamma$
$d = \frac{2}{\delta^2} \ln\left(\frac{1-\beta}{\alpha}\right)$
$\sigma_z = \frac{A_3 s}{3}$
$\mu = \mathbb{E}[z]$
Control limits:
$UCL = h\sigma_z$
$LCL = -h\sigma_z$
Initialization:
$SH_0 = 0$
$SL_0 = 0$
Recursive cumulative sums:
$SH_i = \max(0, \bar{z}_i - \mu - \gamma\sigma_z + SH_{i-1})$
$SL_i = \min(0, \bar{z}_i - \mu + \gamma\sigma_z + SL_{i-1})$

TABLE IV
CUSUM PARAMETERS [1]

Parameter	h	γ	m	n_{samp}	α	β	δ
Value	5.9045	0.5	86	12	0.0027	0.01	1

V. CASE STUDY AND SIMULATION SETUP

A. Battery Characteristics

The battery system discussed in the case studies is modeled using Fig. 2. and Fig. 3. for a stack of three series-connected battery cells. The battery cells are LiFePO₄ and have a nominal

voltage of 3.3 V, consistent with [1]. In this extension i_{bat} was used as the system input, which was affected by noise, and the voltage sensors ($v_{bat,1} \dots v_{bat,N}$, V_{stack}) were used as output measurements.

The parameters for the battery cells were taken from [16] and are listed in Table V. The v_{oc} curves for each cell can be seen in Fig. 5. and their functions are described by (23) – (25).

TABLE V
BATTERY MODEL PARAMETERS [16]

Parameter	Cell 1	Cell 2	Cell 3
$R_{0,j}$	0.0043 Ω	0.0045 Ω	0.0042 Ω
$R_{1,j}$	0.00032 Ω	0.00031 Ω	0.00034 Ω
$C_{1,j}$	629.7 F	632.4 F	605.3 F
$R_{2,j}$	0.0028 Ω	0.0031 Ω	0.0030 Ω
$C_{2,j}$	2247.7 F	2367.5 F	2453.6 F
$C_{cap,j}$	4.369 Ah	4.130 Ah	4.270 Ah
$\eta_{c,j}$	0.99	0.99	0.99
$\eta_{s,j}$	0.000010	0.000010	0.000010

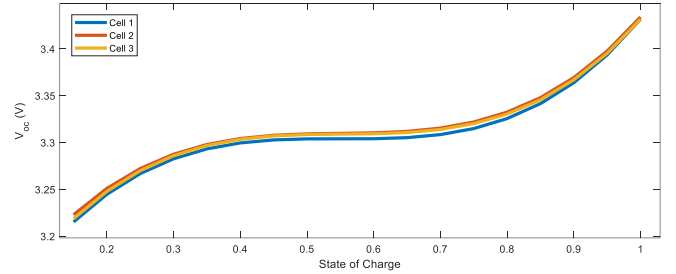


Fig. 5. Average OCV for each battery cell as done in [1] and [16].

$$v_{oc,1} = 1.404\zeta^3 - 2.314\zeta^2 + 1.2693\zeta + 3.0723 \quad (23)$$

$$v_{oc,2} = 1.3184\zeta^3 - 2.1722\zeta^2 + 1.2\zeta + 3.0875 \quad (24)$$

$$v_{oc,3} = 1.35\zeta^3 - 2.2404\zeta^2 + 1.2437\zeta + 3.0781 \quad (25)$$

B. Battery Cycling Simulations

The effectiveness of the proposed methods was evaluated using simulations of battery charging and discharging cycles implemented in MATLAB/Simulink. Each simulation ran for one charge/discharge cycle, which was simulated for 8100 s, with a sampling time of 0.1 s. Fig. 6 shows Cell 1 charging and discharging following the application of the stack current.

We assume that the input (i_{bat}) is affected by noise and that the battery current (i_{bat}) is the same throughout each cell due to their series-connected configuration. So, the simulations consider process, measurement, and input noise. The standard deviation values of the noise parameters are listed in Table VI, and all considered noise parameters were applied to both the sensor measurements and the INAEKF. Then the CUSUM algorithm was run on each sensor's residual data to determine if an attack was present in the system.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

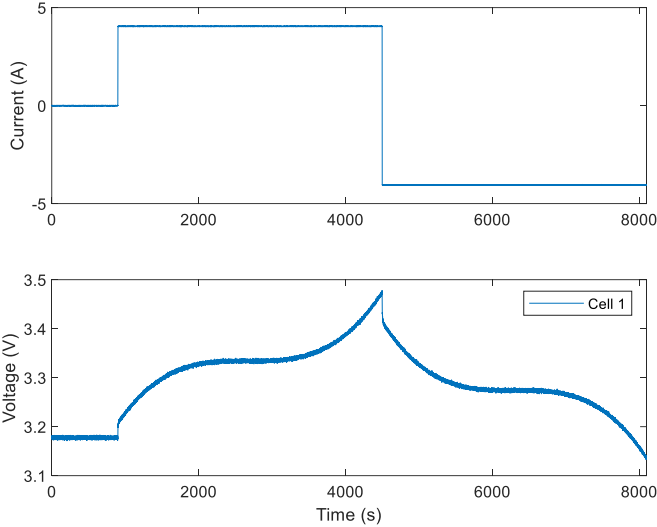


Fig. 6. Battery stack current (top) and $v_{bat,1}$ measurement of Cell 1 (bottom).

TABLE VI
SIMULATION PARAMETERS

Parameter	Value
Simulation time	8100 s
Number of battery cells	3
Sampling time of state estimators	0.1 s
Number of FDIA detection tests run	4800
Standard deviation of cell voltage noise	0.0017 V
Standard deviation of stack voltage noise	0.005 V
Standard deviation of stack current noise	0.0067 A
Standard deviation of SoC process noise	$1 \cdot 10^{-5}$
Standard deviation of RC voltage process noise	$5 \cdot 10^{-4}$ V

C. FDIA Protocol

It has been assumed that the attacker could elect to attack one or more sensors and it was also of interest to evaluate the cases where no attacks were present to assess the occurrence of false positives. Consequently, there were 32 possible permutations of sensors being attacked, ranging from no sensors being attacked to all the sensors ($v_{bat,1}, v_{bat,2}, v_{bat,3}, V_{stack}, i_{bat}$) being injected with FDIA. To assess the robustness of the FDIA detection and identification method 4800 simulations were run. At each simulation run, the combination of attack magnitude, the number of sensors affected (from 0 to 5), the time where the FDIA first affected the sensor, and the magnitude of each attack was varied randomly following uniform distributions. The attack injection time for any sensor was selected randomly from a uniform distribution with a minimum value of 2000 s and a maximum value of 7000 s. Once a FDIA was injected, its bias remained until the end of the simulation.

We have modeled additive FDIAs on sensors as a modification of the measurements obtained from discrete-time sampled values of the voltage and current sensor readings. As such, the resolution of the attack follows the resolution of the analog-to-digital conversion used by the BMS. A short survey of commercial BMSs found that a 15-bits resolution is common. Also, cell

voltages are often sampled between 0 V to 5 V, so the measurement and the FDIA resolution is $153 \mu\text{V}$.

Assuming large magnitude attacks are simpler to detect, the maximum attack value used in simulations was limited to ± 20 mV, since it has been previously shown that attacks larger than that could be detected by less sensitive FDIA detection methods like the chi-squared test [26]. The injected attack on each voltage sensor was selected randomly between $\pm 153 \mu\text{V}$ and ± 20 mV, excluding 0 V as that would indicate no attack was injected. For the voltage stack measurement, it has been considered that both its measurement range and resolution are three times that of the individual cell voltages.

A similar approach was applied to the FDIAs on the current sensor. We have considered that the stack current sensor measurement range is from -20 A to 20 A, which results in 1.22 mA of current measurement resolution considering 15-bit sampling. To exclude overly large FDIA in the current sensor, the maximum attack value was selected as ± 500 mA. The parameters utilized in the attack protocol are summarized in Table VII.

TABLE VII
SENSOR ATTACK PROTOCOL

Parameter	Value
Cell voltage attack magnitude range*	-20 mV to 20 mV
Cell voltage attack resolution	153 μV
Stack voltage attack magnitude range*	-60 mV to 60 mV
Stack voltage attack resolution	459 μV
Stack current attack magnitude range*	-500 mA to 500 mA
Stack current magnitude resolution	1.22 mA
Attack time range	2000 s to 7000 s

*Excluding 0.

D. Comparison with the Chi-Squared Error Detector

The well-accepted chi-squared detector [41] was used to provide a benchmark for the effectiveness of the CUSUM algorithm in detecting FDIA. The chi-squared detector followed a similar implementation as [26] but considers four degrees of freedom when processing the output innovations. A second chi-squared detector was applied to the input variable residuals, with a covariance equal to $U[k]D\Sigma$ and one degree of freedom. To minimize false positives, we have chosen a confidence level of 99.999%, the equivalent of 1 false positive in 100,000 time steps.

VI. RESULTS

This section contains the results of the tests performed, which include: an observability study to determine if the estimators could perform in the event of sensor failures, an accuracy study comparing the state estimate errors of EKF and INAEKF, and multiple attack scenarios to determine if the CUSUM was consistent in detecting the presence of FDIAs in one or more voltage and current sensors. To allow for a more detailed analysis of the simulation results, we have divided the sections that present results. One section contains a summary of all results, and the remaining sections discuss attacks on single voltage sensors, the input current sensor, and multiple sensors.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

A. Observability Study

To perform state estimation, the system must be observable. Since the system described in this paper is time-variant, an observability check must be performed at each timestep. An observability study was done on the stack of three series-connected batteries by calculating an observability matrix for each timestep k . The system is considered observable if the rank of the observability matrix is equal to the number of states at each every timestep. The observability study was performed using the standard observability matrix, which is shown in [1].

The observability study was first done with all the sensors active to determine if the system was observable when using the EKF and INAEKF. To determine if the system remained observable in the event of sensor failures, the sensors were disconnected from the system and the observability study was performed. A disconnected sensor is defined as the sensor being completely zeroed out from the system.

When all four sensors were online the system remained observable regardless of the nonlinear estimator used. The redundancy of voltage measurements added by the V_{stack} sensor allowed the system to remain observable in the event of a single sensor failure. If two or more sensors failed, the system became unobservable and state estimation could not be relied on.

B. State Estimation Accuracy of EKF vs. INAEKF

To compare the state estimation accuracy of the traditional EKF and the INAEKF, we have elected to use the root mean squared error (RMSE) (24) of the state estimates as the comparison metric. The RMSE is found by comparing the a priori ($\hat{x}[k] = \hat{x}[k|k-1]$) or a posteriori ($\hat{x}[k] = \hat{x}[k|k]$) state estimates with the true system state ($x[k]$) as obtained by the simulation:

$$\hat{x}_{RMSE}(\hat{x}[k]) = \sqrt{\frac{\sum_{k=1}^T (x[k] - \hat{x}[k])^2}{T}} \quad (26)$$

where T is the total number of timesteps, x is the actual state, and \hat{x} is the estimated state which could be estimated by the EKF or the INAEKF.

The RMSE of each state variable (for the a priori and a posteriori state) was calculated using both estimation methods. Then, the percent change between the INAEKF and EKF results was calculated following:

$$\hat{x}_{RMSE}^{pred} = 100\% \cdot \left(1 - \frac{\hat{x}_{RMSE}^{INAEKF}(\hat{x}[k|k-1])}{\hat{x}_{RMSE}^{EKF}(\hat{x}[k|k-1])}\right) \quad (27)$$

$$\hat{x}_{RMSE}^{corr} = 100\% \cdot \left(1 - \frac{\hat{x}_{RMSE}^{INAEKF}(\hat{x}[k|k])}{\hat{x}_{RMSE}^{EKF}(\hat{x}[k|k])}\right) \quad (28)$$

Throughout this study all battery parameters and initial conditions were the same regardless of whether the EKF or adjusted INAEKF was used. In this specific accuracy study, no FDIAs were present.

The percent change of the RMSE results for each system state can be seen in Table VIII. Table VIII presents the reduction of the RMSE (as a percent) for the a priori (top) and a posteriori (bottom) states when using the INAEKF. For example, when estimating the v_{11} state, the INAEKF had a 0.039%

and 0.073% estimation accuracy improvement over the EKF for the a priori and a posteriori state, respectively. The results show that when considering input noise in the system, the INAEKF was more accurate in estimating the states than the traditional EKF in all cases.

TABLE VIII

PERCENT REDUCTION OF RMSE OF INAEKF VERSUS EKF

State RMSE	ζ_1	ζ_2	ζ_3	v_{11}	v_{21}	v_{12}	v_{22}	v_{13}	v_{23}
\hat{x}_{RMSE}^{pred}	0.047	0.057	0.056	0.039	0.048	0.042	0.057	0.037	0.056
\hat{x}_{RMSE}^{corr}	0.047	0.057	0.056	0.073	0.048	0.076	0.057	0.071	0.056

C. Summary of FDIA Detection

After running 4800 simulations, the FDIA detection results can be summarized in the confusion matrix shown in Table IX. In this case a true positive indicated an attack was injected and detected, a true negative meant that no attack was present and none of the CUSUM sums diverged, a false positive meant that there was no attack injected but the CUSUM falsely flagged one, and a false negative meant that the CUSUM algorithm missed an attack that was injected to the system. The CUSUM algorithm had a correct detection rate of 99.18%, with no false positives, and a true positive rate of 99.16%.

TABLE IX

CONFUSION MATRIX FOR FDIA DETECTION

Total = 4800	Attack Injected	No Attack Injected
Attack Flagged	4611	0
No Attack Flagged	39	150

It was common for multiple output CUSUM charts to diverge, regardless of the sensor(s) attacked. Therefore, in these studies the CUSUM algorithm was not suitable to identify the corrupted voltage sensor(s). On the other hand, the CUSUM algorithm was able to identify attacks on the input (current sensor) very consistently. The input attack confusion matrix is summarized in Table X, where true positive indicated an attack was injected into the current sensor and the input CUSUM chart diverged, true negative indicated that no attack was injected to the current sensor and the input CUSUM chart did not diverge, false positive indicated that the input CUSUM chart flagged an attack although no attack was injected to the input, and false negative meant that the attack in the input was not captured by the input CUSUM chart.

TABLE X

CONFUSION MATRIX FOR INPUT FDIA IDENTIFICATION

Total = 4800	Attack on Current Sensor	No Current Attack
Flagged Current Sensor	2510	241
Current Sensor Not Flagged	40	2009

Based on the confusion matrix in Table X, the true positive rate was found to be 98.43% and the true negative rate was found to be 89.29%, making the CUSUM algorithm an effective method to identify attacks targeting the input of the system. The tuned CUSUM algorithm presented in this paper was suitable for FDIA detection in the voltage and current sensors of

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

the studied battery stack and was suitable for FDIA identification in the current sensor of the system.

When the CUSUM algorithm was run on the attack-free output and input residual data there were no false positives detected in the system for both cases (with the EKF or adjusted INAEKF as the nonlinear estimator). That is, the SH did not exceed the UCL, and the SL did not breach the LCL for any sample. The CUSUM algorithm had a false positive rate of zero, that is, the CUSUM chart did not diverge when FDIA was not injected into any sensor (Fig. 7 and Fig. 8.).

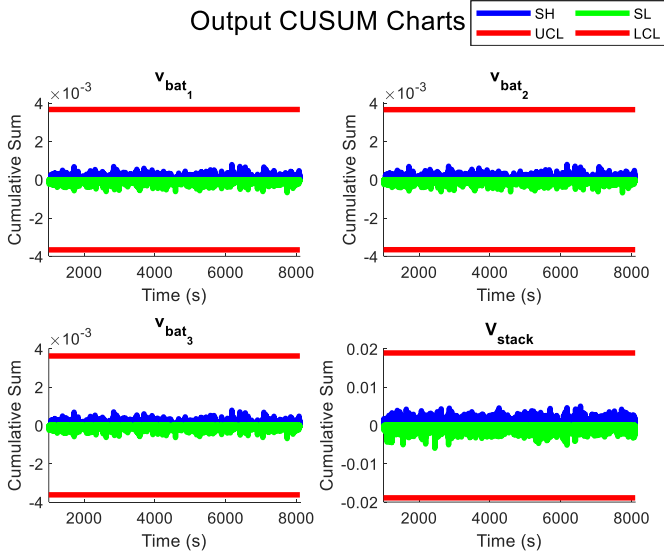


Fig. 7. Output CUSUM charts when no attack was injected, resulting in a false alarm rate of 0%.

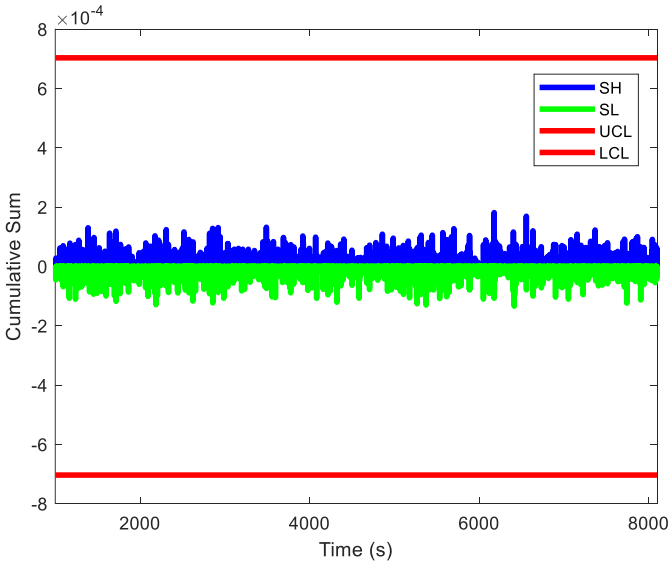


Fig. 8. Input CUSUM charts when no attack was injected, resulting in a false alarm rate of 0%.

D. Single Voltage Sensor Attacks

Single sensor voltage attacks are those in which a FDIA was injected into a single voltage sensor in the stack (modeled by Fig. 2. and Fig. 3.). A single-sensor attack is a likely scenario since the bad actor would want to target the minimum number

of sensors required to disturb state estimation. FDIAs are typically expensive to implement and require detailed knowledge of the system's configuration [14], so targeting the minimum number of sensors while causing the most damage would be desirable to an attacker. The targeted sensor was varied for each test, where either $v_{bat,1}$, $v_{bat,2}$, $v_{bat,3}$, or V_{stack} could be selected as the attacked sensor.

The a priori residual data generated by the estimator was used to calculate the SH and SL to determine if the system was in or out of control at each timestep. In the scope of this paper, detection is defined as any CUSUM chart diverging during a specific test case, thereby indicating the presence of FDIA somewhere in the system. An attack was detected when either the SH exceeded the UCL, or the SL exceeded the LCL for any sample in any of the CUSUM charts. Throughout this study, the divergent CUSUM chart did not consistently correspond to the affected sensor(s) and could not be relied on to identify the targeted sensor(s). The attack injection time did not appear to affect if the attack was detected or not. Using the traditional EKF resulted in the detection of slightly smaller magnitude FDIA than the INAEKF, but either estimator was found to be suitable in this application.

In the tests where the EKF was used as the nonlinear estimator the CUSUM was able to detect attacks as low as $\pm 500 \mu V$ in each individual voltage sensor with no false positives [1]. The INAEKF was slightly less sensitive to smaller-magnitude attacks, and attacks with values of $\pm 2mV$ were detected in each sensor, without triggering false positives. When attacks were injected to only the voltage sensors, the current sensor CUSUM chart did not diverge. The CUSUM charts for a +3 mV FDIA injected in the $v_{bat,1}$ sensor at 5500 s when the INAEKF was used is shown (Fig. 9.) where the SH clearly diverges from the UCL, indicating a FDIA was injected in the system.

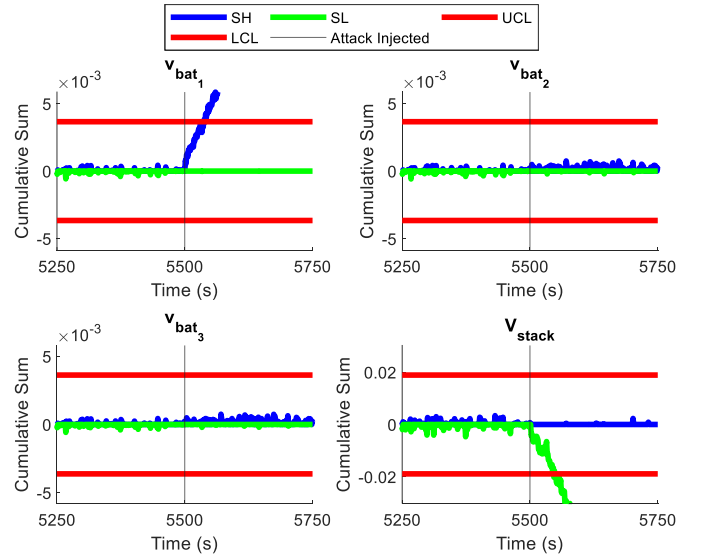


Fig. 9. Output CUSUM charts for a +3 mV attack injected in the $v_{bat,1}$ sensor at 5500 s with noisy inputs considered.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

E. Current Sensor Attacks

This paper extends the work in [1] by including a study where the input was injected with FDIA. When the EKF was used as the nonlinear estimation method, it is assumed that the input is deterministic and known, therefore an input residual cannot be calculated. Using the INAEKF allows the generation of an input residual, which could be run through the CUSUM algorithm as done with the voltage sensors' measurement residuals.

Attacks of ± 200 mA or greater on the input were able to be detected when either the EKF or INAEKF was used. In the case of the traditional EKF, the CUSUM chart for one or more voltage sensors diverged when an attack was added to the input of the system (in most cases the V_{stack} CUSUM chart was divergent) (Fig. 10). The implementation with the traditional EKF was unable to differentiate between attacks targeting the voltage sensors and the input, since it is not possible to generate an Input CUSUM Chart when the EKF is used as the estimator.

The INAEKF implementation allows for the generation of an input residual, which cannot be done with the traditional EKF. Therefore, an additional CUSUM chart can be generated for the input of the system when using the INAEKF. When the attack on the input was ± 300 mA or greater, the input CUSUM chart was divergent, allowing the input CUSUM chart to identify when the input is being attacked (Fig. 11.). In voltage sensor-only attacks, the input CUSUM chart of the implementation with the INAEKF did not diverge. So, in cases where the input CUSUM chart diverged, it can be concluded that the input has FDIA injected to it. In cases where only the output sensor CUSUM charts diverged it can be concluded that only the output sensors are being targeted.

F. Attacks on Multiple Sensors

Multi-sensor attacks are less likely than single sensor attacks due to the expense and expertise required to launch FDIAs, but for completeness this study was included. Every combination of attacked sensors was tested and the sensors susceptible to attack were the same as in the previous case study. The number of targeted sensors, selected corrupted voltage sensors, injected attack value, and injection time were randomly selected.

During some multi-sensor attacks, the CUSUM algorithm was less sensitive to small magnitude FDIA than during single sensor attacks. In the multi-sensor attack cases, all attacks of ± 6 mV or larger were able to be detected by the CUSUM algorithm in testcases with the INAEKF and ± 500 μ V in the testcases with the EKF [1]. Consistent with the results from Section V.D, the CUSUM algorithm was not able to determine how many sensors were attacked or identify the targeted sensors.

The CUSUM chart of a multi-voltage-sensor attack is shown in Fig. 12. In this example, a +3 mV attack was injected in the $v_{bat,2}$ sensor at 5722 s and a -3 mV attack was injected in the V_{stack} sensor at 3745 s. Following the injection of the attacks, the SHs and SLs diverged in the CUSUM charts indicating the presence of an attack in at least one of the voltage sensors. It is common for multiple CUSUM charts to diverge, although the FDIA may only be applied to a smaller sample of sensors.

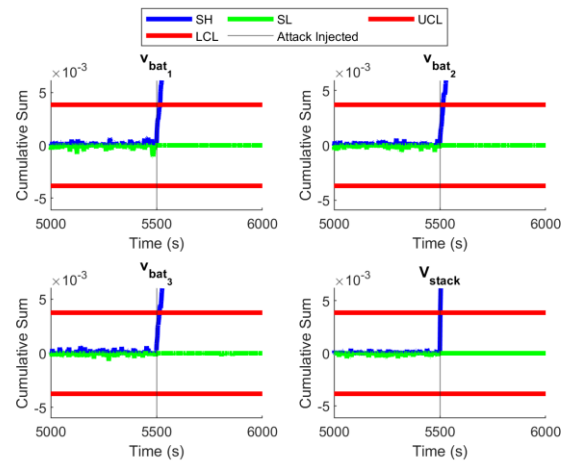


Fig. 10. Output CUSUM charts for a +300 mA attack injected in the input at 5500 s without noisy inputs considered (utilizing EKF as the estimator).

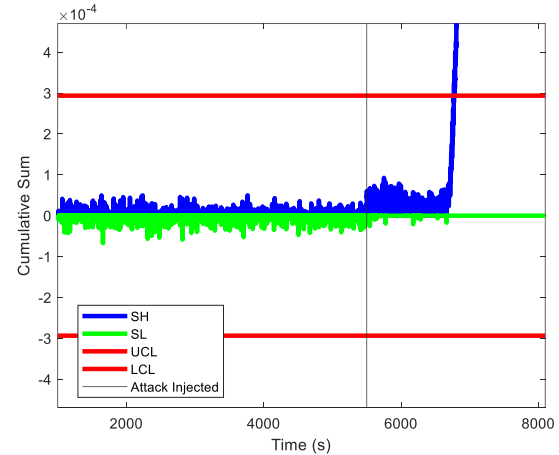


Fig. 11. Input CUSUM chart for a +300 mA attack injected in the input at 5500 s with noisy inputs considered (utilizing the INAEKF as the estimator).

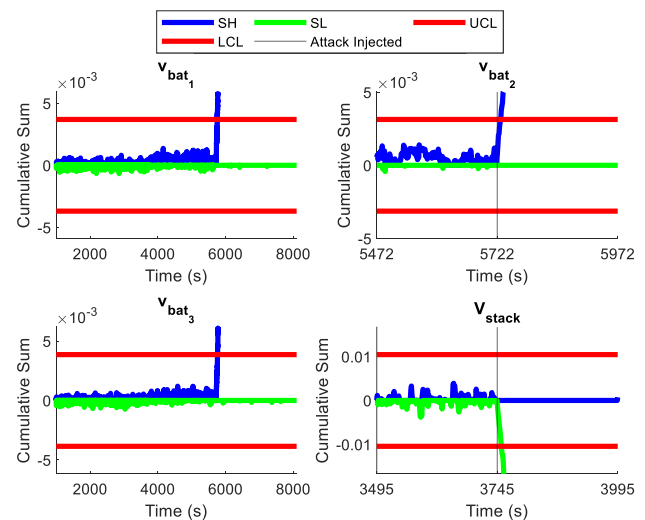


Fig. 12. Output CUSUM charts for a +3 mV attack injected in the $v_{bat,2}$ sensor at 5722 s and a -3 mV attack injected in the V_{stack} sensor at 3745 s with noisy inputs considered.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

The implementation with the traditional EKF [1] was able to detect smaller-magnitude attacks than when using the INAEKF, but due to the tiny magnitudes of the investigated attacks either estimator is suitable for FDIA detection in grid applications. The nominal voltage of each battery cell was 3.3 V [16], so a $\pm 500 \mu V$ and $\pm 6 mV$ attack was 0.015% and 0.18% of the nominal voltage, respectively. Since either method allowed for the detection of higher-magnitude attacks that would have a more destructive impact on the system, either implementation would work for the FDIA detection of BESSs.

G. Comparison with chi-squared error detector

In all 4800 simulations, the chi-squared detector applied to the output flagged errors at least 10 times over the 81001 time steps of the simulation. Given the confidence level chosen, this value is significantly larger than the expected value of 0.81 detections. That means that the chi-squared detector requires a procedure for processing false positives and processing transients. Following the filter initialization, its states and covariance are still far from the values to which they converge, which causes large residuals. If we consider only the last 61001 time steps of the simulation, we see a much different result, summarized in Table XI, reaching a correct classification rate of 87.85%, which is much smaller than the 99.18% rate of the CUSUM algorithm. There is also a correlation between the number of sensors attacked and the number of times the chi-squared detector has flagged data as anomalous. These results are summarized in Table XII.

TABLE XI

CONFUSION MATRIX FOR INPUT FDIA IDENTIFICATION USING THE CHI-SQUARED DETECTOR

Total = 4800	Attack Injected	No Attack
Attack Flagged	4067	0
Attack Not Flagged	583	150

TABLE XII

AVERAGE NUMBER OF DATA POINTS FLAGGED AS ANOMALOUS BY THE CHI-SQUARED DETECTOR (OUT OF 61001)

Attacked Sensors	0	1	2	3	4	5
Mean Flagged Data Points	0	4,225	7,867	12,161	15,568	19,074

The chi-squared detector applied to the input residual has proven insensitive to attacks, flagging 3 to 5 bad data points in all simulations regardless of the presence of an attack. It is important to highlight that, unlike CUSUM, the chi-squared test requires an accurate (or adaptive) estimate of the signal covariances so that a result matches the theoretical prediction.

VII. CONCLUSION

This paper presents a statistics-based CUSUM algorithm that was applied to a BESS consisting of three series-connected battery cells to detect FDIAs targeting the system input and voltage sensors. A noisy input environment is considered, where the input to the system is not assumed deterministic and known, and

a INAEKF is presented to account for the input noise. The BESS was modeled (Fig. 2. and Fig. 3.) and a priori residuals were generated using the nonlinear estimators described in Section III. This paper is an extension to [1] where an alternative nonlinear estimator, the INAEKF, is studied and compared to the nonlinear estimator used in the previous paper (the EKF). The INAEKF outperformed the EKF during state estimation, in terms of RMSE. Utilizing the INAEKF also allowed for the generation of an input residual which could be used to identify FDIAs targeting the system input. The observability study proved the extra stack sensor created a redundant voltage measurement that allowed estimation to be performed when a single sensor failed.

In single sensor attacks, attack vectors as low as $\pm 500 \mu V$ were able to be detected in experiments when using the traditional EKF and as low as $\pm 3 mV$ when using the INAEKF. The multi-sensor results for the implementation using the traditional EKF were the same as the results for single-sensor attack, where attacks of $\pm 500 \mu V$ or greater were able to be detected. When the INAEKF was used, attacks of $\pm 6 mV$ or greater were able to be detected in multiple sensors. Since the detectable attacks were a low percentage of the nominal voltage, either estimation method could be used to generate a priori measurement residuals to detect FDIAs in the voltage sensors of grid-scale battery stacks.

During input attacks, the implementation with the EKF and INAEKF were both able to detect attacks of $\pm 200 mA$ or larger. When the attack on the input was $\pm 300 mA$ or larger the implementation with the INAEKF was able to identify if an attack was targeting the input or not using the input CUSUM chart. Since the traditional EKF considers the input deterministic and known, it is not possible to generate input residuals and identify if FDIAs are targeting the input or just the output sensors. The INAEKF allows for the identification of FDIA in the input, which could help grid operators implement remedial actions, causing it to be an improvement over the traditional EKF implementation.

The proposed three-step approach utilizing the ECM and CRM for battery modeling, the INAEKF for state estimation, and a tuned CUSUM algorithm for FDIA detection was successful in detecting FDIA in the input and output sensors in 99.16% of the 4800 test cases, with a false positive rate of 0%. In addition, by utilizing the INAEKF an input estimation could be generated, allowing for the CUSUM algorithm to be run on the input residual. Post processing the input residual allowed for attacks to be identified as targeting the current sensor in 98.43% of the 2550 input attack testcases. Comparison with the popularly used chi-squared method has shown that the CUSUM method presents a superior anomaly detection rate.

A. Future Work

One of the research gaps identified during this work is that it is challenging to differentiate between cyberattacks and faults corrupting the sensor measurements used in FDIA detection mechanisms. This work focused on detecting corrupted sensor measurements in the voltage and current sensors of a series-connected battery stack. In the future we would like to expand

this work to determine if the sensor corruption was due to a cyberattack, fault, or another undesirable event. Another extension to this work is to construct control actions to mitigate FDIAs detected in the sensors of the battery stack.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Imre Gyuk, Director of the Energy Storage Program, for his continued support. We also acknowledge the support of the U.S Department of Education's program on Graduate Assistance in Areas of National Need (GAANN) grant to Texas Tech University. Thanks to Dr. Ujjwol Tamrakar for his technical advice.

REFERENCES

- [1] V. O'Brien, V. Rao and R.D. Trevizan, "Detection of False Data Injection Attacks in Battery Stacks Using Physics-Based Modeling and Cumulative Sum Algorithm," in *Proc. 2022 IEEE Power and Energy Conf. at Illinois (PECI)*, 2022, doi: 10.1109/PECI54197.2022.9744036.
- [2] R. Xiong, Q. Yu, W. Shen, C. Lin and F. Sun, "A Sensor Fault Diagnosis Method for a Lithium-Ion Battery Pack in Electric Vehicles," in *IEEE Trans. Power Electronics*, vol. 34, no. 10, pp. 9709-9718, Oct. 2019, doi: 10.1109/TPEL.2019.2893622.
- [3] D. M. Rosewater, D. A. Copp, T. A. Nguyen, R. H. Byrne and S. Santoso, "Battery Energy Storage Models for Optimal Control," in *IEEE Access*, vol. 7, pp. 178357-178391, 2019, doi: 10.1109/ACCESS.2019.2957698.
- [4] R. H. Byrne, T. A. Nguyen, D. A. Copp, B. R. Chalamala, and I. Gyuk, "Energy management and optimization methods for grid energy storage systems," *IEEE Access*, vol. 6, pp. 13 231–13 260, 2018.
- [5] R.D. Trevizan, J. Obert, V. De Angelis, T.A. Nguyen, V.S. Rao and B.R. Chalamala, "Cyberphysical Security of Grid Battery Energy Storage Systems," *IEEE Access*, vol. 10, pp. 59675-59722, 2022.
- [6] F. Han, K. See, Y. Feng, X. Yu and X. Yi, "Online SoC estimation for Li-ion batteries: A survey explore the distributed secure cloud management to battery packs," *2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2017, pp. 1838-1843, doi: 10.1109/ICIEA.2017.8283137.
- [7] A. B. Lopez, K. Vatanparvar, A. P. Deb Nath, S. Yang, S. Bhunia and M. A. Al Faruque, "A security perspective on battery systems of the Internet of Things", *J. Hardw. Syst. Secur.*, vol. 1, no. 2, pp. 188-199, Jun. 2017.
- [8] M. Zeng, P. Zhang, Y. Yang, C. Xie, and Y. Shi, "SOC and SOH joint estimation of the power batteries based on fuzzy unscented Kalman filtering algorithm," *Energies*, vol. 12, no. 16, 2019.
- [9] N. Kharlamova, S. Hashemi, and C. Træholt, "The cyber security of battery energy storage systems and adoption of data-driven methods," in *2020 IEEE Third Int. Conf. on Artificial Intelligence and Knowledge Eng. (AIKE)*, 2020, pp. 188–192.
- [10] S. Kumbhar, T. Faika, D. Makwana, T. Kim, and Y. Lee, "Cybersecurity for battery management systems in cyber-physical environments," in *2018 IEEE Transportation Electrification Conf. and Expo (ITEC)*, June 2018, pp. 934–938.
- [11] "Operational risk management in the U.S. energy storage industry: Lithium-ion fire and thermal event safety," Energy Storage Association, Tech. Rep., Sep 2019.
- [12] M. T. Lawder, B. Suthar, P.W. C. Northrop, S. De, C. M. Hoff, O. Leitermann, M. L. Crow, S. Santhanagopalan, and V. R. Subramanian, "Battery energy storage system (BESS) and battery management system (BMS) for grid-scale applications," *Proceedings of the IEEE*, vol. 102, no. 6, pp. 1014–1030, June 2014.
- [13] C. Miller, "Battery Firmware Hacking: Inside the Innards of a Smart Battery," *Proc. of BlackHat*, July 2011.
- [14] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. of the 16th ACM Conf. on Computer and Comm. Security, ser. CCS '09*. New York, NY, USA: ACM, 2009, pp. 21–32. [Online]. Available: <http://doi.acm.org/10.1145/1653662.1653666>
- [15] E. -N. S. Youssef and F. Labeau, "False Data Injection Attacks Against State Estimation in Smart Grids: Challenges and Opportunities," *2018 IEEE Canadian Conf. on Electrical & Computer Engineering (CCECE)*, Quebec, QC, Canada, 2018, pp. 1-5, doi: 10.1109/CCECE.2018.8447683.
- [16] Z. Liu and H. He, "Sensor fault detection and isolation for a lithium-ion battery pack in electric vehicles using adaptive extended Kalman filter", *Appl. Energy*, vol. 185, pp. 2033-2044, 2017.
- [17] D. Ye and T. -Y. Zhang, "Summation Detector for False Data-Injection Attack in Cyber-Physical Systems," in *IEEE Trans. Cybernetics*, vol. 50, no. 6, pp. 2338-2345, June 2020, doi: 10.1109/TCYB.2019.2915124.
- [18] Y. Mo and B. Sinopoli, "On the Performance Degradation of Cyber-Physical Systems Under Stealthy Integrity Attacks," in *IEEE Trans. Automatic Control*, vol. 61, no. 9, pp. 2618-2624, Sept. 2016.
- [19] K. Fang, Y. Huang, Q. Huang, S. Yang, Z. Li and H. Cheng, "An Event Detection Approach Based on Improved CUSUM Algorithm and Kalman Filter," *2020 IEEE 4th Conf. Energy Internet and Energy System Integration (EI2)*, 2020, pp. 3400-3403.
- [20] M. Severo and J. Gama, "Change Detection with Kalman Filter and CUSUM", *Proc. Int. Conf. Discovery Science*, pp. 243-254, 2006.
- [21] M. Pajic et al., "Robustness of attack-resilient state estimators," *2014 ACM/IEEE Int. Conf. on Cyber-Physical Systems (ICCP)*, Berlin, Germany, 2014, pp. 163-174, doi: 10.1109/ICCP.2014.6843720.
- [22] H. Fawzi, P. Tabuada and S. Diggavi, "Secure Estimation and Control for Cyber-Physical Systems Under Adversarial Attacks," in *IEEE Trans. Automatic Control*, vol. 59, no. 6, pp. 1454-1467, June 2014, doi: 10.1109/TAC.2014.2303233.
- [23] A. Sayghe, Y. Hu, I. Zografopoulos, X. Liu, R. G. Dutta, Y. Jin, et al., "Survey of machine learning methods for detecting false data injection attacks in power systems", *IET Smart Grid*, vol. 3, no. 5, pp. 581-595, Oct. 2020.
- [24] H. Lee, G. Bere, K. Kim, J. J. Ochoa, J. -h. Park and T. Kim, "Deep Learning-Based False Sensor Data Detection for Battery Energy Storage Systems," *2020 IEEE CyberPELS (CyberPELS)*, Miami, FL, USA, 2020, pp. 1-6, doi: 10.1109/CyberPELS49534.2020.9311542.
- [25] H. -J. Lee, K. -T. Kim, J. -H. Park, G. Bere, J. J. Ochoa and T. Kim, "Convolutional Neural Network-Based False Battery Data Detection and Classification for Battery Energy Storage Systems," in *IEEE Trans. Energy Conversion*, vol. 36, no. 4, pp. 3108-3117, Dec. 2021, doi: 10.1109/TEC.2021.3061493.
- [26] V. O'Brien, R. D. Trevizan and V. Rao, "Detecting False Data Injection Attacks to Battery State Estimation Using Cumulative Sum Algorithm," *53rd North American Power Symposium (NAPS)*, Nov. 2021 pp 1-6
- [27] H. Chen, E. Tian, L. Wang and S. Liu, "A Joint Online Strategy of Measurement Outliers Diagnosis and State of Charge Estimation for Lithium-Ion Batteries," in *IEEE Trans. Industrial Informatics*, vol. 19, no. 5, pp. 6387-6397, May 2023, doi: 10.1109/TII.2022.3202949.
- [28] G. L. Plett, "Sigma-point Kalman filtering for battery management systems of LiPb-based HEV battery packs: Part 1: Introduction and state estimation," *J. of Power Sources*, vol. 161, no. 2, pp. 1356 – 1368, 2006.
- [29] M. Lelie, T. Braun, M. Knips, H. Nordmann, F. Ringbeck, H. Zappen, and D. Sauer, "Battery management system hardware concepts: An overview," *Applied Sciences*, vol. 8, no. 4, p. 534, Mar 2018.
- [30] B. Alenezi, M. Zhang, S. Hui and S. H. Žak, "Simultaneous Estimation of the State, Unknown Input, and Output Disturbance in Discrete-Time Linear Systems," in *IEEE Transactions on Automatic Control*, vol. 66, no. 12, pp. 6115-6122, Dec. 2021, doi: 10.1109/TAC.2021.3061993.
- [31] N. Wassiliadis, J. Adermann, A. Frericks, M. Pak, C. Reiter, B. Lohmann, and M. Lienkamp, "Revisiting the dual extended Kalman filter for battery state-of-charge and state-of-health estimation: A use-case life cycle analysis," *Journal of Energy Storage*, vol. 19, pp. 73 – 87, 2018.
- [32] B. Bell, UCI Cyber-Physical Security Researchers Highlight Vulnerability of Solar Inverters, Aug. 2020, [online] Available: <https://news.uci.edu/2020/08/18/uci-cyber-physical-security-researchers-highlight-vulnerability-of-solar-inverters/>.
- [33] R. Diversi, R. Guidorzi and U. Soverini, "Kalman filtering in extended noise environments," in *IEEE Transactions on Automatic Control*, vol. 50, no. 9, pp. 1396-1402, Sept. 2005, doi: 10.1109/TAC.2005.854627.
- [34] Y. Zhou, J. Xu, Y. Jing, G.M. Dimirovski, "The unscented Kalman filtering in extended noise environments", *2009 American Control Conference*, pp.1865-1870, 2009.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

- [35] S. Zhao, S. R. Duncan and D. A. Howey, "Observability Analysis and State Estimation of Lithium-Ion Batteries in the Presence of Sensor Biases," in *IEEE Trans. Control Systems Technology*, vol. 25, no. 1, pp. 326-333, Jan. 2017, doi: 10.1109/TCST.2016.2542115.
- [36] D. Rosewater, S. Ferreira, D. Schoenwald, J. Hawkins and S. Santoso, "Battery Energy Storage State-of-Charge Forecasting: Models, Optimization, and Accuracy," in *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2453-2462, May 2019.
- [37] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, NJ, Eaglewood Cliffs:Prentice-Hall, 1979.
- [38] W. C. Navidi, *Statistics for Engineers and Scientists*, New York, NY, USA:McGraw-Hill Education, 2015
- [39] "e-Handbook of Statistical Methods," National Institute of Standards and Technology and SEMATECH, Jun 2012. [Online]. Available: <http://www.itl.nist.gov/div898/handbook/>
- [40] P. D. Hanlon and P. S. Maybeck, "Characterization of Kalman filter residuals in the presence of mismodeling," in *IEEE Trans. Aerospace and Electronic Syst.*, vol. 36, no. 1, pp. 114-131, Jan. 2000.
- [41] R. Mehra and J. Peschon, "An innovation approach to fault detection and diagnosis in dynamic systems", *Automatica*, vol. 7, no. 5, pp. 637-640, Sep. 1971.