The Sliding Innovation Filter

S. A. Gadsden[1], Senior Member, IEEE, and M. Al-Shabi[2], Senior Member, IEEE

[1]University of Guelph, Guelph, Ontario, Canada, N1G 2W1
[2]University of Sharjah, Sharjah, United Arab Emirates, 27272

Corresponding author: S. Andrew Gadsden (e-mail: gadsden@uoguelph.ca).

ABSTRACT In this paper, a new filter referred to as the sliding innovation filter (SIF) is presented. The SIF is an estimation strategy formulated as a predictor-corrector that makes use of a switching gain and innovation term. In estimation theory, a trade-off exists between robustness to disturbances and optimality in terms of estimation error. Unlike the Kalman filter (KF), the SIF is a sub-optimal filter in the sense that it does not provide the optimal solution to the linear estimation problem. However, the switching gain provides an inherent amount of robustness to estimation problems that may be ill-conditioned or contain modeling uncertainties and disturbances. The paper includes the proof of stability and explanation of the SIF gain. Furthermore, the SIF is extended to nonlinear estimation problems using a Jacobian matrix, resulting in the extended sliding innovation filter (ESIF). The methods are applied to a linear and nonlinear aerospace actuator system under the presence of a leakage fault. The results of the simulation demonstrate the improved performance of the SIF and ESIF strategies over popular KF-based methods.

INDEX TERMS Estimation theory, Kalman filters, observers, robustness, sliding modes, state space methods, modeling uncertainty

I. INTRODUCTION

Estimation strategies extract useful information from sensors with noisy measurements. This process is critical for developing and implementing accurate and reliable control systems. The most popular and well-studied estimation strategy is the Kalman filter (KF) which was introduced nearly 60 years ago by Rudolph Kalman [1]. The KF is formulated as a predictor-corrector estimator. State estimates are first predicted using knowledge of the system, controller input, and state values from the previous time step. These estimates are then updated using a gain which is calculated based on the state error covariance and measurement errors. The state error covariance is a function of the expected state error values squared. The KF gain provides the optimal solution to the linear estimation problem assuming that the system is well-known and the system and measurement noise are zero-mean and Gaussian distributed, or known as white [1]. The KF is derived such that the trace of the state error covariance is minimized, thereby providing an optimal solution. The trace is used because it represents a sum of the estimation error squared, or in other words, is the sum of the diagonal elements of the state error covariance [2].

The KF has been extended to nonlinear systems, where it no longer provides an optimal solution. The most common nonlinear estimators based on the KF are the extended KF

(EKF), the unscented KF (UKF), and the cubature KF (CKF) [1, 2]. The EKF utilizes a first-order Taylor series expansion (or Jacobian) that approximates the nonlinearities about the states of interest [2]. The EKF is formulated similarly to the KF with the exception that the nonlinearities are linearized by forming Jacobian matrices. The UKF provides improved estimation results in terms of accuracy by utilizing an unscented transform and 'sigma points' to approximate the nonlinearities [3]. The UKF is slightly more complicated than the EKF in terms of derivation and computation, however it does not require any linearization. The CKF was also introduced to try and improve upon the estimation accuracy compared to the UKF [4]. It has been shown in literature that the CKF, which uses a cubature rule instead of an unscented transform to approximate the nonlinearities, is a special case of the UKF [1].

Depending on the application, KF-based solutions can provide good estimates, however most KF strategies lack robustness to modeling errors, uncertainties, and disturbances. Other issues include numerical instability and matrix inversions [2]. A number of strategies have been implemented in an attempt to improve the robustness and stability of the KF [5, 6, 7]. Square-root formulations of the KF utilize QR decomposition and Cholesky factoring to ensure the state error

covariance matrix is symmetric, which improves numerical stability [8, 9]. Another strategy involves imposing boundaries on the state estimate [10]. For example, given the upper bounds on the level of modeling uncertainty, the KF gain may be bounded to help improve estimation stability [10]. Other methodologies include the addition of fictitious system noise, and adding a fading memory strategy [11, 12]. Modifying the system noise matrix is done when less confidence is placed in the system model used by the filter, or in other words, when there is a lot of system uncertainty [11]. Doing this causes the filter to place more emphasis on the measurements, and less on the system model which may be incorrect resulting in inaccurate estimates and instabilities [11, 13, 14].

In an effort to overcome robustness issues with the KF, the H_∞ filter was introduced [15, 16]. This method utilizes boundaries based on the worst-case uncertainties to regulate the filtering gain which ensures the state estimates are bounded to within a region of the true state trajectory. However, a trade-off exists between optimality and robustness. Other methods have extended the H_∞ filtering strategy further by combining it with the KF and its nonlinear variants [17, 18].

Another strategy for improving estimation performance in the presence of modeling uncertainties and disturbances is to utilize variable structure techniques [19, 20, 21]. Variable structure systems are designed to include discontinuity hyperplanes. These hyperplanes divide the state space into regions. Inside these regions, the system dynamics are continuous [20, 22]. Sliding mode observers were introduced based on sliding mode and variable structure theory [12, 23]. The observer gain is calculated based on the innovation, and the error surface moves towards zero (ideally) [24]. Sliding mode observers define a hyperplane (i.e., a sliding surface) and apply a discontinuous switching force on the estimate to keep the estimate bounded within an area of the hyperplane [12, 25].

Based on variable structure theory and sliding mode observers, the smooth variable structure filter (SVSF) was introduced [26]. Similar to the KF, the SVSF is formulated as a predictor-corrector estimator but utilizes a different gain structure [26, 27]. The SVSF gain is a function of the measurement errors and a switching term [26]. The switching structure of the gain brings an inherent amount of stability to the estimation process as its bounds the estimates to the trajectory of the true state values [2, 28]. However, the SVSF as presented in [26] did not contain a state error covariance derivation. The SVSF was expanded in [2, 27, 29] to contain a covariance function which increased the number of useful applications for the filter. Other improvements were also presented that included the use of a chattering function for fault detection, higher-order solutions, and multi-target tracking formulations [12, 30, 31, 32, 33].

In this paper, the sliding innovation filter (SIF) is proposed. The SIF is based on variable structure techniques similar to the SVSF and sliding mode observers, however its gain structure

is simpler and it provides more accurate results while maintaining robustness. Note also that the SIF may be combined with control strategies for improved tracking performance and robustness to uncertainties and disturbances.

The paper is organized as follows. For completeness, the standard KF equations are provided in Section II, followed by the proposed SIF equations in Section III. Section IV provides the proof of stability for the SIF gain and a discussion on the sliding boundary layer. The simulation results are provided and discussed in Section V, followed by concluding remarks.

II. THE KALMAN FILTER

As described earlier, the KF provides the optimal solution to the linear estimation problem which is described by (2.1) and (2.2). The goal of any estimator is to obtain the true state value x_{k+1} using noisy measurements z_{k+1}.

$$x_{k+1} = Ax_k + Bu_k + w_k \qquad (2.1)$$

$$z_{k+1} = Cx_{k+1} + v_{k+1} \qquad (2.2)$$

The system and measurement noises are represented by w_k and v_k, respectively. A, B, and C represent the system (or process) matrix, input gain matrix, and measurement matrix, respectively. In (2.1) and (2.2), it is assumed that these matrices are fixed and do not change with time. The input to the system is defined as u_k. For the KF and most estimation methods, it is assumed that the system and measurement noises are statistically zero mean with Gaussian distribution [2]. The system and measurement noise are generated using the covariance matrices Q and R, respectively.

It is well established in literature, however the main KF equations are summarized here for completeness [1]. The KF is formulated as a predictor-corrector estimator and is an iterative process. The prediction stage involves calculating the state estimates based on the previous state values and knowledge of the system, as per (2.3). The corresponding state error covariance matrix is calculated in (2.4) and is used in the update stage to calculate the KF gain in (2.5), and is also used to update the state error covariance as per (2.7).

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_k \qquad (2.3)$$

$$P_{k+1|k} = AP_{k|k}A^T + Q_k \qquad (2.4)$$

The update stage is summarized by (2.5) through (2.7). The gain calculated in (2.5) is used to update the state estimates in (2.6) based on the measurement error (or innovation). The gain is also used along with the predicted state error covariance to update the state error covariance in (2.7).

$$K_{k+1} = P_{k+1|k}C^T\big(CP_{k+1|k}C^T + R_{k+1}\big)^{-1} \qquad (2.5)$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\big(z_{k+1} - C\hat{x}_{k+1|k}\big) \qquad (2.6)$$

$$P_{k+1|k+1} = (I - K_{k+1}C)P_{k+1|k}(I - K_{k+1}C)^T \dots$$
$$\dots + K_{k+1}R_{k+1}K_{k+1}^T \qquad (2.7)$$

Note that k refers to the time step, $k|k$ refers to the updated values at the previous iteration, and $k+1|k$ refers to the predicted values at time $k+1$ based on information at time k. Equations (2.3) through (2.7) represent the KF estimation process for linear systems and measurements defined by (2.1) and (2.2), respectively. The process is iterative and repeats every time step k. Note that (2.7) is known as the Joseph covariance form, and is considered to be numerically stable.

III. THE SLIDING INNOVATION FILTER

This section describes the main sliding innovation filter (SIF) equations used for linear systems and measurements. Similar to the KF, the SIF is formulated as a predictor-corrector estimation method. The state estimates and state error covariances are first predicted using values obtained at the previous time step (or initialization), and then the state estimates and state error covariance are updated based on the measurements and correction term at the current time step. The correction term in this case is referred to as the SIF gain.

A. Linear Systems and Measurements

Similar to the KF, the prediction stage includes calculating the predicted or *a priori* ('before the fact') state estimates $\hat{x}_{k+1|k}$, the predicted state error covariance $P_{k+1|k}$, and the predicted innovation $\tilde{z}_{k+1|k}$ as per the following three equations, respectively:

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_k \tag{3.1}$$

$$P_{k+1|k} = AP_{k|k}A^T + Q_k \tag{3.2}$$

$$\tilde{z}_{k+1|k} = z_{k+1} - C\hat{x}_{k+1|k} \tag{3.3}$$

The update stage includes calculating the SIF gain K_{k+1}, the updated or *a posteriori* ('after the fact') state estimates $\hat{x}_{k+1|k+1}$, and the updated state error covariance $P_{k+1|k+1}$ as per the following three equations, respectively:

$$K_{k+1} = C^+ \overline{sat}(|\tilde{z}_{k+1|k}|/\delta) \tag{3.4}$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\tilde{z}_{k+1|k} \tag{3.5}$$

$$P_{k+1|k+1} = (I - K_{k+1}C)P_{k+1|k}(I - K_{k+1}C)^T \dots$$
$$\dots + K_{k+1}R_{k+1}K_{k+1}^T \tag{3.6}$$

Note that C^+ refers to the pseudoinverse of the measurement matrix, $\overline{sat}$ refers to the diagonal of the saturation term, sat refers to the saturation of a value (yields a result between -1 and +1), $|\tilde{z}_{k+1|k}|$ refers to the absolute value of the innovation, δ refers to the sliding boundary layer width, and I refers to the identity matrix (of dimension n-by-n where n is the number of states). Equations (3.1) through (3.6) represent the SIF estimation process for linear systems and measurements defined by (2.1) and (2.2), respectively.

The main difference between the KF and SIF strategies is in the structure of the gain. For the KF, the gain is derived as a function of the state error covariance, which offers optimality [1, 2]. However, for the SIF, the gain is based on the measurement matrix, the innovation, and a sliding

boundary layer term. Although the state error covariance is not used to calculate the SIF gain, it still provides useful information as it represents the amount of estimation error in the filtering process. Figure 1 provides an overview of the SIF estimation concept. An initial estimate is pushed towards the sliding boundary layer which is defined based on the amount of uncertainties in the estimation process. Once inside the sliding boundary layer, the estimates are forced to switch about the true state trajectory by the SIF gain.
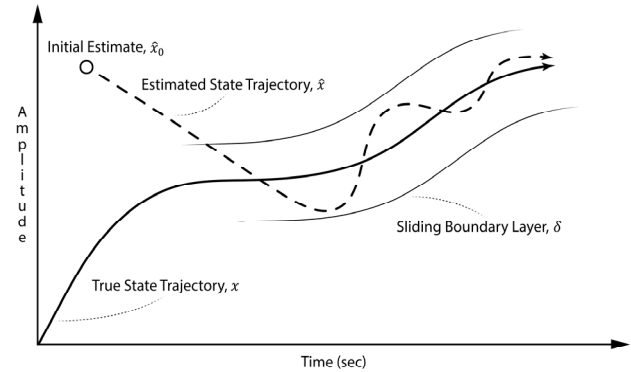

FIGURE 1. The sliding innovation filter (SIF) concept illustrating the effects of the switching gain and sliding boundary layer.

To help illustrate the SIF gain, consider a system with two measurements (and $C = I$), such that the saturation term in (3.4) could be expanded further:

$$K_{k+1} = \overline{sat}(|\tilde{z}_{k+1|k}|/\delta) = \begin{bmatrix} sat\left(\dfrac{|\tilde{z}_1|}{\delta_1}\right) & 0 \\ 0 & sat\left(\dfrac{|\tilde{z}_2|}{\delta_2}\right) \end{bmatrix} \tag{3.7}$$

When multiplied with the innovation $\tilde{z}_{k+1|k}$, as in (3.5), the state estimates $\hat{x}_{k+1|k}$ are updated with the following term:

$$K_{k+1}\tilde{z}_{k+1|k} = \begin{bmatrix} sat\left(\dfrac{|\tilde{z}_1|}{\delta_1}\right)\tilde{z}_1 \\ sat\left(\dfrac{|\tilde{z}_2|}{\delta_2}\right)\tilde{z}_2 \end{bmatrix} \tag{3.8}$$

As shown in (3.8), the state estimates are updated with their corresponding innovation and sliding boundary layer term. The SIF gain effectively acts as a switching term, which forces the measurement errors to be bounded towards the true state trajectory. The sliding boundary layer δ is defined as a function of the modeling uncertainty and noise present in the estimation process. The width can be tuned to obtain the desired estimation result. A method to set the width is explained later in Section IV using the maximum uncertainties in the estimation process (4.23). Another starting point for tuning is to use the values of the measurement noise covariance. For example, $\delta = 10 diag(R)$. The values can then be tuned by trial-and-error, grid search methods, or optimization techniques to reduce the estimation error.

For the cases when there are fewer measurements than states ($m < n$), artificial measurements can be created based on existing measurements to create a full measurement matrix. The structure could also be modified as per a Luenberger observer or other strategies as per [12, 34]. This process would be required to estimate parameters of the system matrix using the SIF.

B. Nonlinear Systems and Measurements

Similar to the extended Kalman filter (EKF), the proposed extended sliding innovation filter (ESIF) makes use of the linearized form of the nonlinear system and measurement functions. For example, consider the nonlinear system function $f(\hat{x}_{k|k}, u_k)$ and the nonlinear measurement function $h(\hat{x}_{k+1|k})$. Linearized forms of these nonlinearities may be calculated respectively using the following partial derivatives:

$$F_{k+1} = \frac{\partial f}{\partial x}\Big|_{\hat{x}_{k|k}, u_k} \tag{3.9}$$

$$H_{k+1} = \frac{\partial h}{\partial x}\Big|_{\hat{x}_{k+1|k}} \tag{3.10}$$

The structure of the nonlinear SIF estimation process is similar to the linear SIF case, with the main difference being the formulation of the gain. The prediction stage consists of three main equations, as follows:

$$\hat{x}_{k+1|k} = f(\hat{x}_{k|k}, u_k) \tag{3.11}$$

$$P_{k+1|k} = F_{k+1} P_{k|k} F_{k+1}^T + Q_{k+1} \tag{3.12}$$

$$\tilde{z}_{k+1|k} = z_{k+1} - h(\hat{x}_{k+1|k}) \tag{3.13}$$

Note that f refers to the nonlinear system function, F_k refers to the linearized version of the system (Jacobian matrix or first-order Taylor series expansion) at time k, and h refers to the nonlinear measurement function. The states are first predicted in (3.11) before being updated in (3.15) using the innovation defined in (3.13) and gain defined in (3.14). The state error covariance matrix is first predicted in (3.12) before being updated in (3.16). Note that the gain (3.14) is also used to update the state error covariance (3.16).

The update stage consists of three main equations, as follows:

$$K_{k+1} = H_{k+1}^+ \overline{sat}(|\tilde{z}_{k+1|k}|/\delta) \tag{3.14}$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\tilde{z}_{k+1|k} \tag{3.15}$$

$$P_{k+1|k+1} = (I - K_{k+1}H_{k+1})P_{k+1|k}(I - K_{k+1}H_{k+1})^T \\ \dots + K_{k+1}R_{k+1}K_{k+1}^T \tag{3.16}$$

Note that H_{k+1}^+ refers to the pseudoinverse of the linearized measurement matrix (first-order Taylor series expansion) at time $k+1$, and H_{k+1} refers to the linearized measurement matrix at time $k+1$. Equations (3.9) through (3.16) represent the ESIF estimation process for nonlinear systems and measurements.

IV. SIF PROOF OF STABILITY

The SIF gain as defined in (3.4) provides robustness to modeling uncertainties and disturbances. This section provides the proof of stability. Consider a Lyapunov function M_{k+1} defined by the updated innovation:

$$M_{k+1} = |\tilde{z}_{k+1|k+1}| \tag{4.1}$$

According to Lyapunov stability theory, the estimation process is considered stable if the following is held true:

$$M_{k+1} \circ \Delta M_{k+1} < 0 \tag{4.2}$$

Equation (4.2) states that the rate of change of the Lyapunov function (4.1) must be negative, or in other words, the absolute magnitude of the innovation must decrease with time. Based on (4.2) and the definition in (4.1), we are able to find the following condition for stability:

$$|\tilde{z}_{k+1|k+1}|\left(\frac{|\tilde{z}_{k+1|k+1}| - |\tilde{z}_{k|k}|}{T}\right)^T < 0 \tag{4.3}$$

where T refers to the sample rate or time step. Equation (4.3) can be rewritten as follows:

$$|\tilde{z}_{k+1|k+1}||\tilde{z}_{k+1|k+1}|^T < |\tilde{z}_{k+1|k+1}||\tilde{z}_{k|k}|^T \tag{4.4}$$

Simplifying (4.4) yields:

$$|\tilde{z}_{k+1|k+1}| < |\tilde{z}_{k|k}| \tag{4.5}$$

Therefore, the estimation process is considered stable if the innovation is decreasing over time. To explore the stability further, we need to find the expectation of $\tilde{z}_{k+1|k+1}$ as per (4.5). First, consider the definition for the updated state error $\tilde{x}_{k+1|k+1}$ which is the difference between x_{k+1} in (2.1) and $\hat{x}_{k+1|k+1}$ in (3.5):

$$\tilde{x}_{k+1|k+1} = \tilde{x}_{k+1|k} - C^+\overline{sat}(|\tilde{z}_{k+1|k}|/\delta)\tilde{z}_{k+1|k} \tag{4.6}$$

Furthermore, note that the innovation may be defined as a function of the state error and measurement noise:

$$\tilde{z}_{k+1|k} = C\tilde{x}_{k+1|k} + v_{k+1} \tag{4.7}$$

Multiplying (4.6) by C and utilizing (4.7) transforms (4.6) into the following:

$$\tilde{z}_{k+1|k+1} = \tilde{z}_{k+1|k} - \overline{sat}(|\tilde{z}_{k+1|k}|/\delta)\tilde{z}_{k+1|k} \tag{4.8}$$

Simplifying (4.8) yields:

$$\tilde{z}_{k+1|k+1} = \left(I - \overline{sat}(|\tilde{z}_{k+1|k}|/\delta)\right)\tilde{z}_{k+1|k} \tag{4.9}$$

Here we define $D = \overline{sat}(|\tilde{z}_{k+1|k}|/\delta)$. We need to solve for $\tilde{z}_{k+1|k}$ to expand and simplify (4.9) further. Based on the definitions for state and measurement error (innovation), we find the following for the predicted state error and predicted innovation, respectively:

$$\tilde{x}_{k+1|k} = A\tilde{x}_{k|k} + w_k \tag{4.10}$$

$$\tilde{z}_{k+1|k} = C(A\tilde{x}_{k|k} + w_k) + v_{k+1} \tag{4.11}$$

Substituting (4.11) into (4.9) yields:

$$\tilde{z}_{k+1|k+1} = (I - D)\big(C(A\tilde{x}_{k|k} + w_k) + v_{k+1}\big) \quad (4.12)$$

We can rewrite (4.12) such that:

$$\tilde{z}_{k+1|k+1} = (I - D)CAC^+\tilde{z}_{k|k} + \eta \quad (4.13)$$

where η is defined as an uncertainty vector:

$$\eta = -(I - D)C(AC^+v_k - w_k) + (I - D)v_{k+1} \quad (4.14)$$

Equation (4.14) is unknown, however its expectation or expected value is zero provided the noise is zero mean (one of the main assumptions for standard estimation theory). In other words, $E[w_k] = E[v_k] = 0$. Taking the expectation of (4.13) yields:

$$E[\tilde{z}_{k+1|k+1}] = (I - D)CAC^+E[\tilde{z}_{k|k}] \quad (4.15)$$

where E refers to the expectation or expected value. This means that the estimation process is stable as long as $|(I - D)CAC^+|$ is less than unity. This ensures that (4.5) satisfies the Lyapunov condition for stability defined by (4.1) and (4.2) provided that:

$$-1 < (I - D)CAC^+ < 1 \quad (4.16)$$

For (4.16), there are two cases to consider. The first case is when CAC^+ is positive:

$$I + CA^{-1}C^+ > D > I - CA^{-1}C^+ \quad (4.17)$$

The second case is when CAC^+ is negative:

$$I + CA^{-1}C^+ < D < I - CA^{-1}C^+ \quad (4.18)$$

Based on the fact that D should be larger than zero and less than 1, D is contained and defined by:

$$D \in [max(I - |CA^{-1}C^+|, 0), I] \quad (4.19)$$

Based on the conditions of (4.5), (4.16), and (4.19), the SIF is considered stable as the Lyapunov function defined by (4.1) and (4.2) is satisfied.

Finally, note that the uncertainty vector defined by (4.14) may be used to help set the sliding boundary layer width. In order to reduce the effects of chattering (or high-frequency switching), the width should be set larger than or equal to the following:

$$\delta \geq |\tilde{z}|_{max} + |\eta|_{max} \quad (4.20)$$

The sliding boundary layer width can be set using (4.20), where $|\tilde{z}|_{max}$ refers to the largest innovation and η_{max} represents the largest uncertainty vector. To find the largest uncertainty vector, we set $D = 0$ which occurs when the innovation is zero or the boundary layer is infinitely large. In this case, we modify (4.20) to consider maximum values:

$$|\eta|_{max} = |C|_{max}(|A|_{max}|C^+|_{max}|v|_{max} \dots$$
$$\dots + |w|_{max}) + |v|_{max} \quad (4.21)$$

Note that (4.21) may be simplified further to yield:

$$|\eta|_{max} = (|C|_{max}|A|_{max}|C^+|_{max} + I)|v|_{max} \dots$$
$$\dots + |C|_{max}|w|_{max} \quad (4.22)$$

Based on (4.20) and (4.22), we can now define a condition for the sliding boundary layer:

$$\delta \geq |\tilde{z}_{max}| + (|C|_{max}|A|_{max}|C^+|_{max} + I)|v|_{max} \dots$$
$$\dots + |C|_{max}|w|_{max} \quad (4.23)$$

The above condition can be used as a starting point for tuning the sliding boundary layer width. Note that if the sliding boundary layer is set wider than the uncertainties, the estimates will be smoothed. If the width is set smaller, chattering will occur which represents high-frequency switching due to the SIF gain.

V. SIMULATION SETUP AND RESULTS

A. LINEAR AEROSPACE SYSTEM

To demonstrate the robustness of the proposed estimation strategy, the KF and SIF are applied on a linear system with noise. The studied system is a type of aerospace flight surface actuator, referred to as the electrohydrostatic actuator (EHA). It has been well-studied and presented in literature [35, 36, 37]. A simplified linear EHA model was formulated in state space where the states of interest refer to position, velocity, and acceleration [2, 26]. The model parameters were found through experimentation of an EHA [26, 38]. The linear form of the system and measurements are described using the following state space equations [26]:

$$x_{k+1} = \begin{bmatrix} 1 & T & 0 \\ 0 & 1 & T \\ -557 & -28.6 & 0.94 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \\ 557 \end{bmatrix} u_k + w_k \quad (5.1)$$

$$z_{k+1} = Cx_{k+1} + v_{k+1} \quad (5.2)$$

where the sample rate T is defined as $1\ ms$, k is the time step, C refers to the measurement matrix which in this case is an identity matrix of dimension $m \times m$ or 3×3, and u is the controller input for the system (defined in Fig. 2) that drives the desired trajectory. The system and measurement noises (w and v) are normally distributed with zero mean and covariance's Q and R defined by (5.3) and (5.4), respectively.
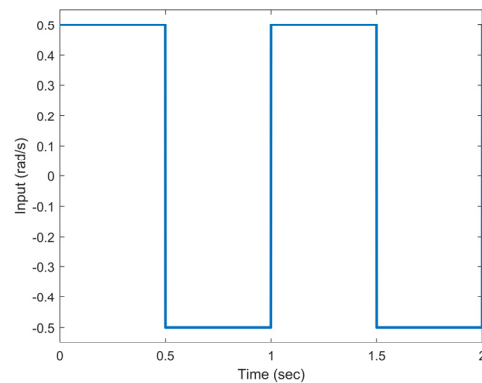

FIGURE 2. Control signal (rad/s) used as the input for the linear EHA simulation described by (5.1) to (5.4).

$$Q = diag([10^{-5} \quad 10^{-3} \quad 0.1]) \tag{5.3}$$

$$R = diag([10^{-4} \quad 10^{-2} \quad 1]) \tag{5.4}$$

The three states in (5.1) represent the position, velocity, and acceleration of the linear actuator, respectively. The initial state values, measurements, and estimates were set to zero. The initial state error covariance values were set to $P_{0|0} = 10Q$. The sliding boundary layer width was manually tuned to yield the smallest estimation error, and was found for this simulation to be $\delta = [0.05 \quad 1 \quad 0.5]$. The simulation was coded in MATLAB.

The results of applying the KF and SIF strategies on the linear EHA are shown in Fig. 3. As expected, since the system is linear and well-known, the KF yields better results in terms of root mean square error (RMSE) under normal operating conditions. Note that RMSE is defined by (5.5) where n is the number of time steps. The results are summarized in Tab. I.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{n}} \tag{5.5}$$
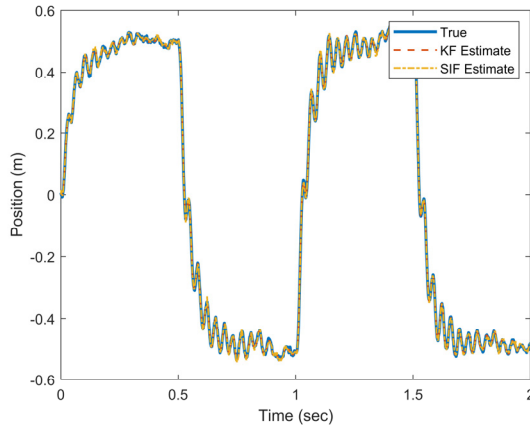

FIGURE 3. True and estimated state values for the linear EHA system example under normal operating conditions. The results are nearly identical so the lines appear over-lapping.

TABLE I
RMSE RESULTS FOR THE LINEAR SIMULATION: NORMAL CASE

State	Kalman Filter	Sliding Innovation Filter
Position (m)	3.80×10^{-3}	5.89×10^{-3}
Velocity (m/s)	4.95×10^{-2}	5.56×10^{-2}
Acceleration (m/s^2)	9.21×10^{-1}	9.87×10^{-1}

In an effort to demonstrate the robustness of the SIF strategy, consider the case when the system has a fault injected half-way through the simulation (at $t = 1 \sec$). In this case, the linear system state equation used by the filters is changed:

$$x_{k+1} = \begin{bmatrix} 1 & T & 0 \\ 0 & 1 & T \\ -240 & -28 & 0.94 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \\ 557 \end{bmatrix} u_k + w_k \tag{5.6}$$

The results of the modeling uncertainty and its effects on the filters are shown in Fig. 4. The model mismatch at 1 second causes the KF to deviate from the true state trajectory, yielding poor estimates of the true position. The SIF was still able to perform relatively well, and was bounded to the true state trajectory due to the switching effects in the SIF gain.
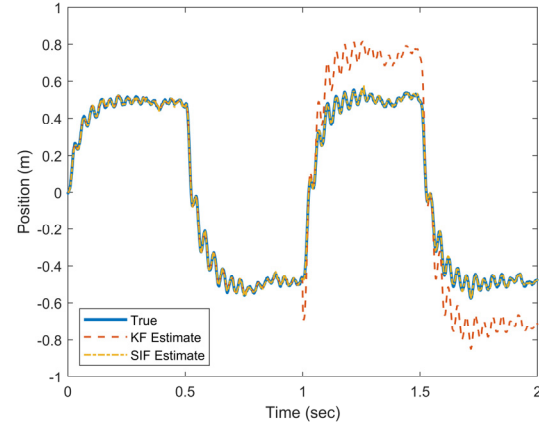

FIGURE 4. True and estimated state values for the linear EHA system example under faulty operating conditions.

TABLE II
RMSE RESULTS FOR THE LINEAR SIMULATION: FAULTY CASE

State	Kalman Filter	Sliding Innovation Filter
Position (m)	1.54×10^{-1}	5.95×10^{-3}
Velocity (m/s)	1.74	5.97×10^{-2}
Acceleration (m/s^2)	8.91	1.02

The RMSE results for the faulty case are shown in Tab. II. The SIF performs only slightly worse than the normal case. However, the KF is unable to recover from the modeling uncertainty and yields worse performance. This was expected as the KF is derived based on the assumption that the system is known. Figure 5 further illustrates the presence of the modeling uncertainty. The KF was unable to recover from the system change, whereas the SIF provided a good estimate.
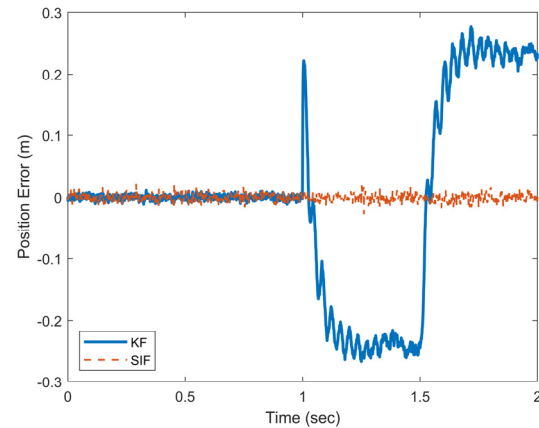

FIGURE 5. Position estimation error for the KF and SIF for the linear EHA faulty case. This further illustrates the robustness of the SIF gain.

B. NONLINEAR AEROSPACE SYSTEM

A nonlinear form of the EHA system is studied to further demonstrate the robustness of the SIF [2]. In this case, the extended sliding innovation filter (ESIF) is applied and compared with the EKF. The ESIF is summarized by (3.9) through (3.16). The nonlinear EHA can be accurately described by the following state space (related to its position, velocity, and acceleration) equations [2]:

$$x_{1,k+1} = x_{1,k} + Tx_{2,k} + w_{1,k} \tag{5.7}$$

$$x_{2,k+1} = x_{2,k} + Tx_{3,k} + w_{2,k} \tag{5.8}$$

$$x_{3,k+1} = \left[1 - T\frac{a_2V_0 + M\beta_e L}{MV_0}\right]x_{3,k} - T\frac{(A_E^2 + a_2L)\beta_e}{MV_0}x_{2,k}$$

$$\dots - T\frac{2a_1V_0x_{2,k}x_{3,k} + \beta_e L(a_1x_{2,k}^2 + a_3)}{MV_0}sgn(x_{2,k})$$

$$\dots + T\frac{A_E\beta_e}{MV_0}u_k + w_{3,k} \tag{5.9}$$

The differential pressure $x_{4,k}$ may also be modelled using the velocity of the actuator and a number of friction constants:

$$x_{4,k} = A_E^{-1}\left(a_2x_{2,k} + (a_1x_{2,k}^2 + a_3)sgn(x_{2,k})\right)\dots$$

$$\dots + w_{4,k} \tag{5.10}$$

The sample rate of the system is $T = 1\ ms$, and the input to the EHA system is defined by:

$$u_k = D_p\omega_{p,k} - sgn(x_{4,k})Q_{L0} \tag{5.11}$$

As described in [35], A_E refers to the piston cross-sectional area, β_e is the effective bulk modulus (i.e., the 'stiffness' in the hydraulic circuit), D_p refers to the pump displacement, L represents the leakage coefficient, M is the load mass (i.e., weight of the cylinders), V_0 is the initial cylinder volume, and ω_p refers to the pump angular velocity. Two main parameters that are affected by induced fault conditions in the EHA are the leakage coefficients L and Q_{L0}, and the friction coefficients a_1, a_2, and a_3. The following table lists the known parameter values that were experimentally determined in [38].

TABLE III
EHA PARAMETER VALUES FOR NONLINEAR CASE

Parameter	Physical Significance	Model Values
A_E	Piston Area	$1.52 \times 10^{-3}\ m^2$
D_p	Pump Displacement	$5.57 \times 10^{-7}\ m^3/rad$
M	Load Mass	$7.376\ kg$
V_0	Initial Cylinder Volume	$1.08 \times 10^{-3}\ m^3$
x_0	Maximum Stroke	$0.14335\ m$
β_e	Effective Bulk Modulus	$2.07 \times 10^8\ Pa$
a_1	Friction Parameter 1	6.589×10^4
a_2	Friction Parameter 2	2.144×10^3
a_3	Friction Parameter 3	436
L	Leakage: Normal Case	$4.78 \times 10^{-12}\ m^3/sPa$
	Leakage: Fault Case	$6.01 \times 10^{-11}\ m^3/sPa$
Q_{L0}	Flow Rate: Normal Case	$2.41 \times 10^{-6}\ m^3/s$
	Flow Rate: Fault Case	$1.47 \times 10^{-5}\ m^3/s$

There are four important states and parameters: actuator position, velocity, acceleration, and differential pressure. It is assumed that measurements are available for each state, such that the measurement matrix C is full or $C = I$:

$$z_{k+1} = Cx_{k+1} + v_{k+1} \tag{5.12}$$

To properly implement the EKF and ESIF, the first-order Taylor series expansion (Jacobian) of (5.7) through (5.10) must be solved, as per (3.9). The linearized system matrix F_k is defined as follows:

$$F_k = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & T & 0 \\ 0 & F_{32} & F_{33} & 0 \\ 0 & F_{42} & 0 & 0 \end{bmatrix} \tag{5.13}$$

where F_{32}, F_{33}, and F_{42} are as follows:

$$F_{32} = \frac{-T\left(2V_0a_1x_{3,k}sign(x_{2,k})\right)}{MV_0}\dots$$

$$\dots - \frac{\beta_e T\left(A_E^2 + La_2 + 2La_1x_{2,k}sign(x_2,k)\right)}{MV_0}$$

$$F_{33} = 1 - \frac{2Ta_1x_{2,k}sign(x_{2,k})}{M} - \frac{T(V_0a_2 + \beta_e LM)}{MV_0}$$

$$F_{42} = A_E^{-1}\left(a_2 + 2a_1x_{2,k}sign(x_{2,k})\right)$$

The system noise covariance matrix Q was defined based on tuning the experimental setup in [35], and in this simulation was set to:

$$Q = \begin{bmatrix} 10^{-12} & 0 & 0 & 0 \\ 0 & 10^{-10} & 0 & 0 \\ 0 & 0 & 10^{-9} & 0 \\ 0 & 0 & 0 & 10^6 \end{bmatrix} \tag{5.14}$$

The measurement noise covariance R was set to 10^3Q. The system noise w and measurement noise v were defined as normal distributions with covariances Q and R, respectively. For both strategies, the initial state estimate $\hat{x}_0$ and error covariance matrix $P_{0|0}$ are respectively defined as follows:

$$\hat{x}_0 = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}^T \tag{5.15}$$

$$P_{0|0} = 10Q \tag{5.16}$$

The sliding boundary layer was manually tuned to yield the smallest estimation error, and was set to $\delta = 10^{-3} \times \begin{bmatrix} 1 & 5 & 10 & 10^9 \end{bmatrix}$. The angular pump velocity ω_p was set to a square wave with $\pm 100\ rad/s$, and was used in (5.11) to generate the input used by the EHA system (as shown in Fig. 6). Note that the input signal refers to volumetric flow rate within the hydraulic circuit of the EHA. Under normal operating conditions, both the EKF and ESIF were able to successfully track the three kinematic states and the differential pressure.
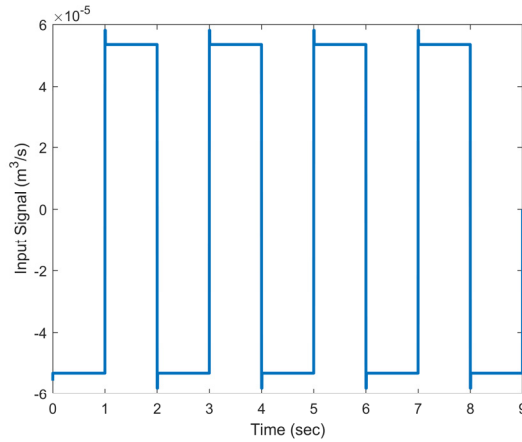
FIGURE 6. Input signal used for the nonlinear EHA scenario. Note that the input signal (5.10) was calculated based on a square wave representing the pump angular velocity.

The EHA velocity and corresponding EKF and ESIF estimates are shown in Fig. 7. The differential pressure is shown in Fig. 8. The RMSE results are summarized in Tab. IV and show comparable performance between the two strategies. The ESIF provided a slightly better position estimate in terms of error, whereas the EKF yielded slightly better estimates for the velocity and acceleration. The ESIF provided a significantly better estimate for the differential pressure (about three-times lower error).
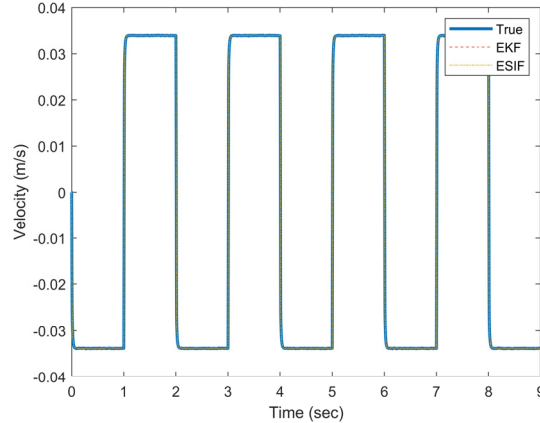

FIGURE 7. EHA velocity trajectory and corresponding EKF and ESIF estimates. Note that the results are nearly identical, so appear to be over-lapping.

TABLE IV
RMSE RESULTS FOR THE NONLINEAR SIMULATION: NORMAL CASE

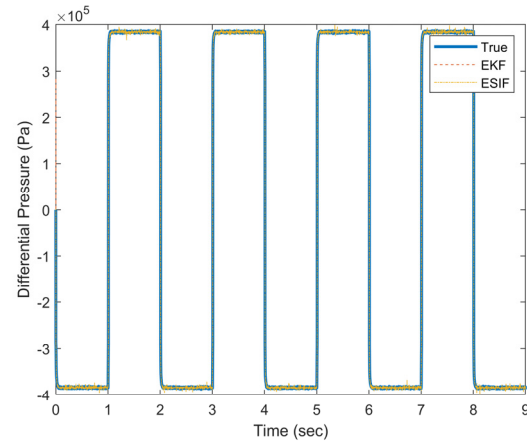| State	Extended Kalman Filter	Extended Sliding Innovation Filter
Position (m)	5.77×10^{-5}	7.43×10^{-6}
Velocity (m/s)	1.59×10^{-5}	5.24×10^{-5}
Acceleration (m/s^2)	7.14×10^{-4}	9.96×10^{-4}
Diff. Pressure (Pa)	6.11×10^3	2.01×10^3

FIGURE 8. EHA differential pressure trajectory and corresponding EKF and ESIF estimates. Initially, the EKF gave a poor estimate but quickly corrected (due to the initial conditions).

In an effort to demonstrate the robustness of the ESIF strategy, consider the case when the system has a fault injected half-way through the simulation (at $t = 4.5$ sec). In this case, the leakage parameters L and Q_{L0} are changed as per Tab. III (to the faulty values). This simulates a major leak occurring in the hydraulic circuit.

The results of the modeling uncertainty and its effects on the filters are shown in Fig. 9. At 4.5 seconds, the fault is injected into the system. The model mismatch causes the EKF to deviate from the true state trajectory, yielding poor estimates of the true differential pressure. The ESIF was still able to perform relatively well, and was bounded to the true state trajectory. Similar to the linear case, this is due to the inherent stability caused by the switching effects in the ESIF gain. The RMSE results for the faulty case are summarized in Tab. V.
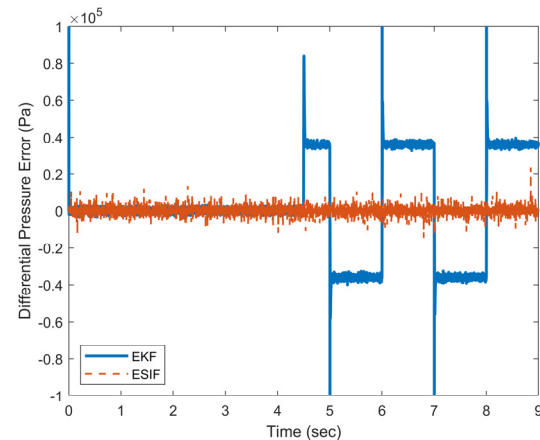

FIGURE 9. EHA differential pressure error based on corresponding EKF and ESIF estimates. The leakage fault is injected mid-way through the simulation. The ESIF is able to accurately compensate for the modeling uncertainty, whereas the EKF has difficulty overcoming the model mismatch.

TABLE V
RMSE RESULTS FOR THE NONLINEAR SIMULATION: FAULTY CASE

State	Extended Kalman Filter	Extended Sliding Innovation Filter
Position (m)	2.35×10^{-3}	7.94×10^{-6}
Velocity (m/s)	8.08×10^{-3}	6.12×10^{-5}
Acceleration (m/s^2)	2.01×10^{-2}	1.04×10^{-3}
Diff. Pressure (Pa)	2.90×10^4	2.48×10^3

The RMSE results further highlight the robustness of the proposed ESIF strategy. The sliding boundary layer δ and the switching effect of the ESIF gain keep the estimated states bounded to the true state trajectory. The sliding boundary layer δ may be tuned manually to improve estimation results, with initial values designed based on knowledge of the noise or uncertainties present in the estimation problem. In this scenario, it was initially tuned based on the amount of measurement noise present (for example, $\delta = 10 \times diag(R)$), and then optimized based on the estimation error.

VI. CONCLUSIONS

In this paper, a new estimation strategy called the sliding innovation filter (SIF) was proposed. The filter is called the SIF due to the fact that the innovation slides along a boundary layer close to the true states. The SIF gain structure is based on variable structure and sliding mode theory. It utilizes a switching term that allows the estimates to remain bounded to a region within the true state trajectory. This improves estimation robustness to modeling uncertainties, errors, and disturbances. A nonlinear formulation of the SIF, referred to as the extended sliding innovation filter (ESIF) was also presented. The SIF and ESIF were applied on an aerospace system and compared with the well-known Kalman filter and its nonlinear form, the extended KF. The simulation results demonstrated that the proposed SIF and ESIF strategies provide sub-optimal yet robust estimates for both linear and nonlinear systems under the presence of uncertainties and disturbances.

REFERENCES

[1] H. H. Afshari, S. A. Gadsden and S. R. Habibi, "Gaussian Filters for Parameter and State Estimation: A General Review and Recent Trends," *Signal Processing,* vol. 135, pp. 218-238, 2017.

[2] S. A. Gadsden, "Smooth Variable Structure Filtering: Theory and Applications," McMaster University (PhD Thesis), Hamilton, Ontario, 2011.

[3] S. J. Julier and J. K. Uhlmann, "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE,* vol. 92, no. 3, 2004.

[4] I. Arasaratnam and S. Haykin, "Cubature Kalman Filters," *IEEE Transactions on Automatic Control,* vol. 54, no. 6, pp. 1254-1269, 2009.

[5] J. E. Potter and R. G. Stern, "Statistical Filtering of Space Navigation Measurements," in *Proceedings of 1963 AIAA Guidance and Control Conference*, New York, 1963.

[6] N. A. Carlson, "Fast Triangular Formulation of the Square Root Filter," *AIAA Journal,* vol. 11, no. 9, pp. 1259-1265, 1973.

[7] G. J. Bierman, Factorization Methods for Discrete Sequential Estimation, New York: Academic Press, Inc., 1977.

[8] R. Van der Merwe and E. A. Wan, "The Square-Root Unscented Kalman Filter for State and Parameter-Estimation," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2001.

[9] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, Numerical Recipes: The Art of Scientific Computing, New York: Cambridge University Press, 2007.

[10] Y. Bar-Shalom, X. Rong Li and T. Kirubarajan, Estimation with Applications to Tracking and Navigation, New York: John Wiley and Sons, Inc., 2001.

[11] D. Simon, Optimal State Estimation: Kalman, H-Infinity, and Nonlinear Approaches, Wiley-Interscience, 2006.

[12] M. Al-Shabi, "The General Toeplitz/Observability SVSF," Hamilton, Ontario, 2011.

[13] A. Jazwinski, "Adaptive Filtering," *Automatica,* vol. 5, no. 4, pp. 475-485, July 1969.

[14] A. Jazwinski, Stochastic Processes and Filtering Theory, New York: Academic Press, 1970.

[15] D. S. Bernstein and W. H. Haddad, "Steady-state Kalman filtering with an H∞ error bound," *Syst. Contr. Lett.,* vol. 12, pp. 9-16, 1989.

[16] K. M. Nagpal and P. P. Khargonekar, "Filtering and smoothing in an Hinf setting," *IEEE Trans. Automat. Contr.,* vol. 36, pp. 152-166, 1991.

[17] G. A. Einicke and L. B. White, "Robust Extended Kalman Filtering," *IEEE Transactions on Signal Processing,* vol. 47, no. 9, pp. 2596-2599, 1999.

[18] Y. Li, L. Ding and Z. Jing, "Optimal-switched extended H∞ filter for nonlinear systems with stochastic uncertainties," *International Journal of Robust and Nonlinear Control,* vol. 30, no. 7, pp. 2850-2870, 2020.

[19] V. I. Utkin, "Variable Structure Systems With Sliding Mode: A Survey," *IEEE Transactions on Automatic Control,* vol. 22, pp. 212-222, 1977.

[20] V. I. Utkin, Sliding Mode and Their Application in Variable Structure Systems, English Translation ed., Mir Publication, 1978.

[21] A. F. Fillipov, "Differential Equation with Discontinuous Right Hand Side," *American Mathematical Society Transactions,* vol. 62, 1969.

[22] V. Utkin, Sliding Modes in Control Optimization, New York: Springer-Verlag, 1992.

[23] S. K. Spurgeon, "Sliding Mode Observers: A Survey," *International Journal of Systems Science,* pp. 751-764, 2008.

[24] S. H. Qaiser, A. I. Bhatti, M. Iqbal, R. Samar and J. Qadir, "Estimation of Precursor Concentration in a Research Reactor by Using Second Order Sliding Mode Observer," *Nuclear Engineering and Design,* vol. 239, pp. 2134-2140, 2009.

[25] J. J. E. Slotine and W. Li, Applied Nonlinear Control, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[26] S. R. Habibi, "The Smooth Variable Structure Filter," *Proceedings of the IEEE,* vol. 95, no. 5, pp. 1026-1059, 2007.

[27] S. A. Gadsden and S. R. Habibi, "Derivation of an Optimal Boundary Layer Width for the Smooth Variable Structure Filter," in *ASME/IEEE American Control Conference (ACC)*, San Francisco, California, 2011.

[28] S. A. Gadsden, S. R. Habibi and T. Kirubarajan, "Kalman and Smooth Variable Structure Filters for Robust Estimation," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 50, no. 2, pp. 1038-1050, 2014.

[29] S. A. Gadsden and S. R. Habibi, "A New Robust Filtering Strategy for Linear Systems," *ASME Journal of Dynamic Systems, Measurement, and Control,* vol. 135, no. 1, 2013.

[30] M. Al-Shabi, S. A. Gadsden and S. R. Habibi, "Kalman Filtering Strategies Utilizing the Chattering Effects of the Smooth Variable

Structure Filter," *Signal Processing,* vol. 93, no. 2, pp. 420-431, 2013.

[31] H. H. Afshari, "The 2nd-Order Smooth Variable Structure Filter (2nd-SVSF) for State Estimation: Theory and Applications," McMaster University (PhD Thesis), Hamilton, Ontario, 2014.

[32] S. A. Gadsden and A. S. Lee, "Advances of the Smooth Variable Structure Filter: Square-Root and Two-Pass Formulations," *Journal of Applied Remote Sensing,* vol. 11, no. 1, 2017.

[33] M. Al-Shabi and A. Elnady, "Recursive Smooth Variable Structure Filter for Estimation Processes in Direct Power Control Scheme Under Balanced and Unbalanced Power Grid," *IEEE Systems Journal,* vol. 14, no. 1, pp. 971-982, 2019.

[34] M. Attari, "SVSF Estimation for Target Tracking with Measurement Origin Uncertainty," McMaster University (PhD Thesis), Hamilton, Ontario, 2016.

[35] S. A. Gadsden, Y. Song and S. R. Habibi, "Novel Model-Based Estimators for the Purposes of Fault Detection and Diagnosis," *IEEE/ASME Transactions on Mechatronics,* vol. 18, no. 4, pp. 1237-1249, 2013.

[36] S. A. Gadsden, Y. Song and S. R. Habibi, "Mathematical Modeling and Fault Detection and Diagnosis of an Electrohydrostatic Actuator," in *ASME/IEEE American Control Conference (ACC)*, Montreal, Quebec, 2012.

[37] S. A. Gadsden, Y. Song, K. R. McCullough and S. R. Habibi, "Friction Fault Detection of an Electrohydrostatic Actuator," in *ASME/Bath Symposium on Fluid Power and Motion Control (FPMC)*, Washington DC, 2011.

[38] K. McCullough, "Design and Characterization of a Dual Electro-Hydrostatic Actuator," Hamilton, Ontario, 2011.

Mohammad Al-Shabi (M'10-SM'19) obtained his B.Sc. and M.Sc. in Mechatronics from the Mechanical Engineering Department at Jordan University for Science and Technology. He completed his PhD in Mechanical Engineering at McMaster University. His thesis was in the area of filtering and estimation on mechatronic systems. Mohammad is currently serving as an Assistant Professor in the Department of Mechanical and Nuclear Engineering at the University of Sharjah, UAE. His research areas are broad and consider nonlinear and robust control, filtering, estimation theory, fault detection, robotics, optimization, and artificial intelligence.

S. Andrew Gadsden (M'09–SM'19) earned his Bachelors in Mechanical Engineering and Management (Business) at McMaster University. He then completed his Ph.D. in Mechanical Engineering at McMaster in the area of estimation theory. Andrew worked as a postdoctoral researcher at the Centre for Mechatronics and Hybrid Technology (Hamilton, Ontario, Canada). He also worked as a Project Manager in the pharmaceutical industry (Apotex Inc.) for three years.

Andrew is currently an Associate Professor at the University of Guelph in Mechanical Engineering. Before joining Guelph, he was an Assistant Professor in the Department of Mechanical Engineering at the University of Maryland, Baltimore County, USA. Andrew has a number of colleagues in NASA, the US Army Research Laboratory (ARL), US Department of Agriculture (USDA), National Institute of Standards and Technology (NIST), and the Maryland Department of the Environment (MDE). He is an IEEE Senior Member, elected Fellow of ASME, and a Professional Engineer of Ontario. Andrew earned the *2019/2020 University Research Excellence Award* for the College of Engineering and Physical Sciences based on his research activities at the University of Guelph. In addition, he is a *2019 SPIE Rising Researcher* award winner based on his work in intelligent estimation theory, and is also a *2018 Ontario Early Researcher* (ERA) award winner based on his work in intelligent condition monitoring strategies. Andrew was also awarded the *2019 University of Guelph Faculty Association (UGFA) Distinguished Professor Award for Excellence in Teaching* in the College of Engineering and Physical Sciences.