01 Jan 2014

An Online Outlier Identification and Removal Scheme for Improving Fault Detection Performance

Hasan Ferdowsi

Sarangapani Jagannathan
Missouri University of Science and Technology, sarangap@mst.edu

Maciej Jan Zawodniok
Missouri University of Science and Technology, mjzx9c@mst.edu

Recommended Citation

H. Ferdowsi et al., "An Online Outlier Identification and Removal Scheme for Improving Fault Detection Performance," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 5, pp. 908 - 919, article no. 6658905, Institute of Electrical and Electronics Engineers, Jan 2014.
The definitive version is available at https://doi.org/10.1109/TNNLS.2013.2283456

An Online Outlier Identification and Removal Scheme for Improving Fault Detection Performance

Hasan Ferdowsi, Sarangapani Jagannathan, and Maciej Zawodniok, *Member, IEEE*

***Abstract*— Measured data or states for a nonlinear dynamic system is usually contaminated by outliers. Identifying and removing outliers will make the data (or system states) more trustworthy and reliable since outliers in the measured data (or states) can cause missed or false alarms during fault diagnosis. In addition, faults can make the system states nonstationary needing a novel analytical model-based fault detection (FD) framework. In this paper, an online outlier identification and removal (OIR) scheme is proposed for a nonlinear dynamic system. Since the dynamics of the system can experience unknown changes due to faults, traditional observer-based techniques cannot be used to remove the outliers. The OIR scheme uses a neural network (NN) to estimate the actual system states from measured system states involving outliers. With this method, the outlier detection is performed online at each time instant by finding the difference between the estimated and the measured states and comparing its median with its standard deviation over a moving time window. The NN weight update law in OIR is designed such that the detected outliers will have no effect on the state estimation, which is subsequently used for model-based fault diagnosis. In addition, since the OIR estimator cannot distinguish between the faulty or healthy operating conditions, a separate model-based observer is designed for fault diagnosis, which uses the OIR scheme as a preprocessing unit to improve the FD performance. The stability analysis of both OIR and fault diagnosis schemes are introduced. Finally, a three-tank benchmarking system and a simple linear system are used to verify the proposed scheme in simulations, and then the scheme is applied on an axial piston pump testbed. The scheme can be applied to nonlinear systems whose dynamics and underlying distribution of states are subjected to change due to both unknown faults and operating conditions.**

***Index Terms*— Data analysis, fault diagnosis, neural networks, nonlinear systems.**

NOMENCLATURE

x	Actual system states.
y	Measured system states.
x_s	Estimated outlier-free states.
u	System input.
ω	Known dynamics of the system.
η	Modeling uncertainties.
Π	Time profile of the fault.
h	Fault function.
ν	Noise and outliers.
c_g	Lipschitz constant for the function ω.
η_M	Upper bound on modeling uncertainty.
p	Window length.
W	Unknown target weights for outlier removal NN.
$\hat{W}$	Estimated weights.
$\tilde{W}$	Weight estimation error.
μ	Approximation error.
μ_M	Upper bound on μ.
ψ	Basis function for outlier removal NN.
β	Learning rate for updating $\hat{W}$.
e	Difference between the measured and estimated states, i.e., $e = y - x_s$.
$\tilde{x}$	State estimation error, i.e., $\tilde{x} = x - x_s$.
M	Median.
σ	Standard deviation.
$\hat{x}$	Estimated states from the fault detection (FD) observer.
$\hat{h}_d$	Estimated fault function.
θ_d	Unknown parameter matrix of the FD online approximator.
$\hat{\theta}_d$	Estimated parameters.
$\tilde{\theta}_d$	Parameter estimation error.
ε	Approximation error.
ε_M	Upper bound on ε.
ϕ	Basis function for FD online approximator in discrete time (OLAD).
α	Learning rate for updating $\hat{\theta}$.
$\bar{e}$	FD residual, i.e., $\bar{e} = x_s - \hat{x}$.

Manuscript received July 29, 2012; accepted September 3, 2013. Date of publication November 8, 2013; date of current version April 10, 2014. This work was supported in part by NSF I/UCRC on Intelligent Maintenance Systems and in part by the Intelligent Systems Center.

The authors were with the University of Missouri-Rolla, Rolla, MO 65401 USA. They are now with the Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO 65401 USA (e-mail: hfqk6@mst.edu; sarangap@mst.edu; mjzx9c@mst.edu).

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

Digital Object Identifier 10.1109/TNNLS.2013.2283456

I. INTRODUCTION

OUTLIERS are present in data sets of practical industrial systems. By definition [1], an outlier is an observation, which deviates sufficiently from other observations thus creating suspicion that it was from a different system. In industrial systems, outliers can appear in the measured data. If measured data from a system is contaminated by outliers, processing the data becomes difficult since these outliers can render inaccurate decisions during fault diagnosis. In many cases, the underlying distribution of the measured data can change due to outliers.

On the other hand, due to the high risk of component and system failures, reliable fault diagnosis schemes are required to guarantee safe system operation even in the presence of uncertainties, outliers, and faults. A reliable FD scheme is the one that can detect faults at an early stage, without missed or false alarms before the root cause analysis.

Recently, the topic of FD and diagnosis has attracted a number of researchers around the world. FD is performed data driven [2], model-based [3], or a combination of both [4], [5]. Several model-based FD techniques have been developed in the past decade [6]–[10]. However, even the best FD schemes can become unreliable in the presence of data corrupted with outliers since outliers can cause false alarms.

Several outlier identification and removal (OIR) schemes have been proposed in the literature such as distribution-based [11], distance-based [12], [13], clustering [14], [15], and density-based methods [16]. In addition, surveys of different outlier detection methods are given in [17] and [18]. However, these methods are data driven and work offline. An online outlier detection scheme is a prerequisite for improving the performance of the model-based FD scheme where system states are used. Therefore, several online outlier removal methods have also been developed for linear systems with known dynamics using Kalman filter and its variations [19]–[21] by assuming that the system states and measurement noise belong to the Gaussian distribution. However, due to changes in the operating conditions and presence of faults, the underlying distribution of states is not necessarily fixed and therefore nonstationary. Therefore, a novel scheme to detect the presence and the removal of outliers from system states is needed for both linear and nonlinear systems in nonstationary environments.

Since the system states are considered to be contaminated with noise and outliers, the main objective of this paper is to develop an online OIR scheme for system states prior to FD stage in contrast with traditional model-based FD framework where outlier-free state assumption is made [22]. The robustness and reliability of the model-based FD scheme is evaluated using detection rate, missed and false alarms with and without outliers. In model-based FD, data points are system states or outputs.

Since in practice outlier-free system states are not available, they need to be estimated and then compared with the measured states to detect outliers. Here, traditional observers cannot be used because the system dynamics are not known and subjected to unknown faults. Therefore, a two-layer feed-forward neural network (NN) is used to estimate the actual system states and at the same time to identify and remove the outliers. The NN outputs are the estimated states filtered for noise and outliers. At each time instant, the estimated state vector from the OIR scheme is calculated for the next instant of time and compared with the measured states to generate the state estimation error.

Next, median and standard deviation of the state estimation error in a limited time window are found. If the state estimation error and the calculated median are both higher than three standard deviations, an outlier is detected. A novel NN weight update law is derived using the state estimation error. To prevent an update of the NN weights in the state estimator in response to an outlier, a variable learning rate is selected such that it takes on a zero value when an outlier is detected. The stability of the state observer used for OIR is discussed in this paper.

For the purpose of FD, a different observer with known nominal dynamics of the system is introduced. Since the OIR scheme estimates the known system dynamics, uncertainties, and fault function, it cannot be used as an FD observer. Moreover, using a second observer for FD, the fault function can be approximated for isolation and prognostics. Therefore, an observer-based FD scheme is introduced that uses the estimated outlier-free state vector instead of the measured state vector. A fault is detected by comparing the observed states with the outlier-free system state vector. Upon detection, an online approximator is activated to estimate the fault dynamics. The performance of the FD scheme is evaluated with and without the proposed outlier scheme. Again note that the state estimation for the OIR scheme is different than the one used for fault diagnosis.

Therefore, the contributions of this paper involve the development of an OIR Scheme, which can operate online in contrast with data-based methods [11]–[16], and can be applied to both linear and nonlinear systems in nonstationary environments in contrast with existing Kalman filter-based schemes [19]–[21]. Since the proposed NN estimator is quite generic and does not use the system representation or model, it is useful even when the system dynamics are not known. In other words, the OIR scheme can be used both for data driven and model-based fault diagnosis schemes.

Moreover, a model-based FD scheme, which uses the estimated outlier-free states instead of the actual measured states of the system is presented and the stability analysis of the proposed fault diagnosis scheme is included when the underlying distribution of states are nonstationary in contrast with all the available model-based fault diagnosis [3]–[8]. This requires a complete novel analytical framework.

To verify the performance and effectiveness of the proposed outlier removal technique and observe its effect on FD process, a three-tank water system is used. A fault is seeded in one of the tanks and outlier removal is performed on both healthy and faulty data. It is shown that FD can only provide reliable results when the outliers are removed from the measured data. In addition, a linear example is used to compare the proposed scheme with a Kalman filter-based method. Simulations have been repeated for a significant number of times to evaluate the proposed scheme in different cases of random noise and outliers. Furthermore, an experimental study has been conducted on an axial piston pump testbed in healthy operating conditions and it is shown that the measured outlet pressure involves several outliers, which will trigger false alarms during FD. The outliers are shown to be removed successfully from the measured outlet pressure using this online OIR scheme.

This paper is organized as follows. Section II introduces the system description and required assumptions. Section III presents the outlier detection and removal technique while FD scheme is introduced in Section IV. Section V discusses the simulation and experimental results.

II. System Description

Consider the nonlinear discrete-time system described by the following state space representation:

$$x(k+1) = \omega(x(k), u(k)) + \eta(x(k), u(k))$$

where $u \in \mathbb{R}^m$ is the control input vector, $x \in \mathbb{R}^n$ is the system state vector, $\omega : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is the known nonlinear system dynamics, and $\eta : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is the system uncertainties.

Now, consider the nonlinear system with a fault as

$$x(k+1) = \omega(x(k), u(k)) + \eta(x(k), u(k)) + \Pi(k-k_0)\, h(x(k), u(k)) \quad (1)$$

where $h(x(k), u(k))$ is a vector of possible fault dynamics. The time profile of a fault is given by $\Pi(k-k_0)$. The time profile $\Pi(k-k_0)$ is modeled by $\Pi(k-k_0) = \text{diag}\{\Omega_1(k-k_0), \Omega_2(k-k_0), \ldots, \Omega_n(k-k_0)\}$ where

$$\Omega_i(k-k_0) = \begin{cases} 0, & \text{if } \tau < 0 \\ 1 - e^{-\bar{\kappa}_i \tau}, & \text{if } \tau \geq 0 \end{cases} \quad \text{for } i = 1, \ldots, n$$

is the time profile variable and $\bar{\kappa}_i$ is an unknown constant that represents the rate at which a fault occurs. A larger value of $\bar{\kappa}_i$ indicates that it is an abrupt fault. The use of such time profiles is common in fault diagnosis literature [6].

Note that this fault will definitely change the system dynamics and might even change the underlying distribution of system states. Normally in the literature [6]–[9], it is assumed that the states are free with noise and does not change its underlying distribution, which is not practical. Therefore, Kalman filters [23] cannot be used to eliminate noise and outliers from the measured states because they require system states and noise to have a fixed distribution and also require the exact system representation, which is not available in our case due to unknown fault. In this paper, a NN-based approach will be taken with an appropriate selection of NN weights such that this assumption is relaxed.

For the purpose of monitoring the system and performing FD, state measurements are required. Usually, the measured system states are contaminated with noise and outliers. The measured state vector $y(k)$ can be represented by

$$y(k) = x(k) + v(k)$$

where $v(k)$ includes the measurement noise and outliers, which is considered as bounded above such that $\|v(k)\| \leq v_M$. The distribution of the measurement noise can change over time. According to the definition [1], if the measured states $y(k)$ deviates significantly from the actual system states $x(k)$, the data point is said to be an outlier. Model-based FD schemes cannot distinguish a residual increasing due to a fault or an outlier. Therefore, outliers in the measured states can cause false alarms during FD and diagnosis. This fact clearly emphasizes the importance of detection and removal of outliers before FD. The following standard assumptions are needed to proceed.

Assumption 1: The modeling uncertainty is bounded, i.e., $\|\eta(x(k), u(k))\| \leq \eta_M$, $\forall (x, u) \in (\chi \times U)$, where η_M is a positive known constant.

Remark 1: Assumption 1 is needed to distinguish between the faults and the system uncertainties and to analytically determine the FD threshold.

Assumption 2: The nonlinear system dynamics $\omega(x, u)$ is Lipschitz in x, i.e., $\|\omega(x_1, u) - \omega(x_2, u)\| \leq c_g \|x_1 - x_2\|$, where $c_g > 0$ is the Lipschitz constant.

Remark 2: This assumption is only required for the FD part, mainly because the estimated outlier-free states are used in the proposed FD estimator instead of the actual system states. This assumption has been used in other papers on fault diagnosis [6], [24], where the entire state vector is not available and estimated states have to be used in the estimator dynamics instead of actual system states.

Assumption 3: The functions $\omega(x(k), u(k))$, $\eta(x(k), u(k))$, and $h(x(k), u(k))$ can be expressed as nonlinear in the unknown parameters (NLIP), thus can be approximated by two-layer NNs with bounded weights and approximation errors.

Next, the proposed outlier detection scheme is introduced.

III. OIR Scheme

The main objective of this paper is to design an outlier scheme, which can detect, identify, and remove the outliers online, before an outlier triggers a false alarm during fault diagnosis. Therefore, the outlier detection must be performed online and prior to FD and root cause analysis.

According to Chebyshev's theorem and outlier detection method [25], almost all the observations in a data set of system states will fall into the interval $[\mu - 3\sigma, \mu + 3\sigma]$, where μ and σ are the mean and standard deviation of the data set, respectively, and the data points outside this interval are declared outliers. If the distribution of the actual system states was fixed over time, traditional outlier detection methods [25], [26] can be employed whereas for the present scenario, these methods cannot be used. Now initially assume that the measured system state vector y has following fixed distribution:

$$y(k) \sim \mathcal{N}(x(k), \sigma^2)$$

where $\mathcal{N}(\mu, \sigma^2)$ is a Gaussian distribution with mean μ and variance σ^2. In this case, an outlier can be defined as a point, where $|y(k) - x(k)| > 3\sigma$ where $|\cdot|$ is the absolute value operator. If the mean value of the actual states is equal to μ, then this definition can be rephrased as $|y(k) - \mu| > 3\sigma$. However, this method is offline and also it cannot be used when the system states do not have a fixed distribution. In this paper, we have assumed that the system is subjected to a fault, which can change the nominal dynamics of the system as well as the underlying distribution of the states.

To develop a method of online outlier detection for a system with changing dynamics, we will investigate the measured state vector in a fixed time window, assuming that the measurement noise has a fixed distribution over each of these small time windows. Suppose that the state vector at time instant k is being investigated and consider a finite window of time with length p in which the measured state vector is available, i.e., $\{y(k-p+1), \ldots, y(k-1), y(k)\}$.

If the outlier-free state vector in the current window is available, then the difference between the actual and the measured state vector can be calculated by $\delta = y - x$ and its mean and variance over the selected time interval can be

found by

$$\mu(k)=\frac{1}{p}\sum_{j=0}^{p-1}(\delta(k-j)),\ \ \sigma^2(k)=\frac{1}{p}\sum_{j=0}^{p-1}(\delta(k-j)-\mu(k))^2.$$

Since the distribution of measurement noise is normal, assuming that its variance is constant within the considered time window, an outlier can be detected at time k, when $|\delta(k)-\mu(k)| > 3\sigma(k)$. Although this is an online outlier detection method, which can also handle the changes in the system dynamics, it is impossible to implement since the outlier-free system states are not available in practice and the only available data would be the measured states contaminated with noise and outliers.

To overcome this issue, an estimator will be proposed to estimate the unknown system states by assuming that the states are available for measurement. In the literature, the outlier removal is traditionally done without an estimator while an observer is normally used for model-based FD and not for outlier removal. In contrast, using an observer, we are estimating the actual states and also performing outlier removal.

If the system dynamics was known, it could be used to construct an observer to estimate the system states, similar to Kalman filter-based outlier detection methods. However, this is only possible when the system is working in healthy conditions with known dynamics. In our case, the system is subjected to unknown changes like faults. Therefore, an estimator, which is able to approximate the system states without using the system dynamics is required. To construct this online approximator and its learning mechanism, initially we consider the case when the measured data does not have outliers, and then the general case will be investigated. Since $\omega(x(k), u(k))$, $\eta(x(k), u(k))$, $h(x(k), u(k))$ are all smooth functions, $x(k+1)$ in (1) can be approximated by a two-layer NN, if x_m does not involve outliers. Therefore, $x(k+1)$ can be written as

$$x(k+1)=W^T(k)\psi(x(k), u(k))+\mu(k)$$

where $W\in\mathbb{R}^{q\times n}$ is the unknown parameter matrix, which will change when a fault occurs in the system or the model parameters change due to shift in the operating conditions, which can also change the distribution of the states, $\psi(x(k), u(k))$ is a basis function-like sigmoid, and $\mu(k)$ is the approximation error, which is bounded by μ_M [27].

Now, let the estimated states be denoted as x_s and consider the NN output as

$$x_s(k+1)=\hat{W}^T(k)\psi(x_s(k), u(k)) \tag{2}$$

where $\hat{W}\in\mathbb{R}^{q\times n}$ represents the unknown weights of the output layer of NN. Now, an update law for training $\hat{W}$ is required. Define the state estimation error $\tilde{x}(k)=x(k)-x_s(k)$ and parameter estimation error $\tilde{W}(k)=W(k)-\hat{W}(k)$. When there is no noise and outliers in the measured states, which means y is equal to x at all times, the weight update law can be selected as

$$\hat{W}(k+1)=\hat{W}(k)+\beta\psi(x_s(k), u(k))e^T(k+1) -\gamma\|I-\beta\psi(x_s(k), u(k))\psi^T(x_s(k), u(k))\|\hat{W}(k) \tag{3}$$

where $\beta > 0$ is a constant learning rate, $0 < \gamma < 1$, and $e(k+1)=y(k+1)-x_s(k+1)$. Note that the NN weights are updated by the difference between the measured and estimated states, because the actual system states are not available. Then, the state estimation error can be written as

$$\tilde{x}(k+1)=W^T(k)\tilde{\psi}+\hat{W}^T(k)\psi(x_s(k), u(k))+\mu(k)$$

where $\tilde{\psi}(k)=\psi(x(k), u(k))-\psi(x_s(k), u(k))$.

Remark 3: Instead of the measured state vector y, the delayed output of the NN ($x_s(k)$) along with the input vector $u(k)$ are used as NN inputs, to prevent the outliers in measured data from affecting the state estimates, whereas y is only used for updating the NN weights. Later on, the weight update law in (3) will also be modified to cancel the effect of outliers on the NN weights.

By choosing the following Lyapunov function candidate:

$$V=\tilde{x}^T(k)\tilde{x}(k)+\frac{1}{\beta}\mathrm{tr}\{\tilde{W}^T(k)\tilde{W}(k)\}$$

it can easily be shown that state and parameter estimation errors will be uniformly ultimately bounded (UUB). However, measured states involve outliers, so this approach cannot be used since the outlier-free state vector is not available. In other words, when y is contaminated with outliers and this measured data is used to update the NN weights in (3), the actual states will not be estimated correctly while the outlier detection will also be unreliable.

To solve this problem, the outlier detection and state estimation processes will be combined to properly detect the outliers and design a new weight update law that is not affected by the outliers. The parameter update law in (3) is modified using a variable learning rate whose value will be zero when an outlier is detected at time $(k+1)$ and not zero otherwise. Suppose that $y(k+1)$ is an outlier. In this case, $e(k+1)$, which is used to update the parameters will be large, even if the weights are close to their desired values and $x_s(k+1)$ is close to its desired value $x(k+1)$. To prevent the NN weights to be updated by an outlier at this time instant, the variable leaning rate $\beta(k+1)$ used in the update law should take zero value.

Since $\hat{W}(k)$ is available at the time instant k, $x_s(k+1)$ can be calculated and used for outlier detection before updating the weights. To perform the outlier detection on $y(k+1)$, again consider a finite window of time with length p. The median value of $\|e\|=\|y-x_s\|$ (where $\|\cdot\|$ is the norm operator) in a window ending at time $(k+1)$ is defined by

$$M(k+1)=\mathrm{Median}\{\|e(k-p+2)\|, \ldots, \|e(k)\|, \|e(k+1)\|\}$$

and the standard deviation in the same time window is defined by

$$\sigma(k+1)=\sqrt{\mathrm{Var}\{\|e(k-p+2)\|, \ldots, \|e(k)\|, \|e(k+1)\|\}}.$$

Similar to the first case, we assume that the variance of the measurement noise is constant within the time window.

Therefore, a threshold value of three times the standard deviation is used to detect the outliers. Because of the limited time window, the mean value of the data set inside a window might be significantly affected even by a single outlier, which might increase the probability of a false or missed alarm. In contrast, median value is not easily affected by a single outlier. Therefore, if the mean value is used for outlier detection, the unwanted change in the mean value can definitely degrade the performance of outlier detection process. This simple example clarifies the reason why median is used in this outlier detection scheme instead of mean value.

Thus, median value is used instead of mean value to overcome this problem. Finally, the data point at time $(k+1)$ is considered an outlier if

$$|\,\|e\,(k+1)\| - M(k+1)| > 3\sigma\,(k+1)\,.$$

Now that the outlier detection is performed for the data set comprised of system states at time $(k+1)$, we need to construct an analytical formula to find the variable learning rate $\beta(k+1)$ based on whether or not an outlier exists at this time instant. The idea is to reduce the learning rate, preferably to zero, when an outlier is detected at time $(k+1)$, to prevent the NN weights from getting updated by an outlier. In addition, the learning rate needs to be small when outliers are not present with a relatively large amplitude noise.

For this purpose, define the function $S\,(\cdot)$ as

$$S\,(z) = \begin{cases} \left(1-z^2\right)^2, & \text{for} \quad |z| < 1 \\ 0, & \text{otherwise.} \end{cases}$$

This bell-shaped function achieves its maximum at $z = 0$ and takes zero value when $z \geq 1$. This function is used to construct the robust variable learning rate given by

$$\beta\,(k+1) = \beta_M S\left(\frac{1}{3\sigma\,(k+1)}\,|\,\|e\,(k+1)\| - M\,(k+1)|\right)$$

where β_M is the maximum possible learning rate parameter, which keeps the estimator stable. Larger noise amplitude will result in larger values for $|\|e(k+1)\| - M(k+1)|$, thus smaller values for the learning rate. Particularly, when $|\|e(k+1)\| - M(k+1)| > 3\sigma(k+1)$ [which means an outlier is detected at time $(k+1)$] $\beta(k+1)$ will automatically be set to zero, so the weights will not be updated upon detecting an outlier. Furthermore, considering the definition of outliers, it can be inferred that an outlier is detected at time k if $\beta(k) = 0$.

Finally, the proposed parameter update law can be represented by

$$\begin{aligned}\hat{W}(k+1) = {} & \hat{W}(k) + \beta(k+1)\psi(x_s(k), u(k))e^T(k+1) \\ & -\gamma\ \|\ I - \beta(k+1)\psi(x_s(k), u(k))\psi^T(x_s(k), u(k))\ \|\ \hat{W}(k).\end{aligned} \tag{4}$$

The definition of the learning rate implies that, if $e\,(k+1)$ is relatively close to $M(k+1)$, then the corresponding learning rate $\beta(k+1)$, which appears in the parameter update law, will be close to maximum possible learning rate. Whereas, if $e\,(k+1)$ is largely deviated from $M(k+1)$ then the corresponding learning rate, will be zero or close to zero. Thus, the measurement noise cannot make significant change on the NN weights while the effect of outliers on the weight update law is completely eliminated. In the following theorem, the performance of state estimation with the proposed outlier detection and removal scheme is discussed.

Theorem 1: Let an adaptive observer in (2) be used to estimate the state vector of system (1) when the measured state vector y is contaminated with outliers. Then, the state estimation error, $\tilde{x}\,(k) = x\,(k) - x_s(k)$, and the NN weight estimation error $\tilde{W}(k)$ are UUB in the mean if the user-defined variables are selected such that

$$\begin{aligned}3W_{\max}^2 c_\psi^2(1+2\beta_M^2\psi_{\max}^2) &< 12\gamma\,\|I - \beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\| \\ &> 2\beta_M^2\psi_{\max}^4 + 3\psi_{\max}^2 + 2/3 \\ &\quad + \gamma^2\|I - \beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\|^2.\end{aligned}$$

Proof: Consider the following Lyapunov function:

$$V = E(\mathrm{tr}\{\tilde{x}^T(k)I\tilde{x}(k)\}) + \frac{1}{3}E(\mathrm{tr}\{\tilde{W}^T(k)\tilde{W}(k)\})$$

where E$(\cdot)$ is the expectation operator. The first difference of this function is given by

$$\begin{aligned}\Delta V = {} & E(\mathrm{tr}\{\tilde{x}^T(k+1)I\tilde{x}(k+1) - \tilde{x}^T(k)I\tilde{x}(k)\}) \\ & + \frac{1}{3}E(\mathrm{tr}\{\tilde{W}^T(k+1)\tilde{W}(k+1) - \tilde{W}^T(k)\tilde{W}(k)\}).\end{aligned}$$

By substituting $\tilde{x}(k+1)$ from the state estimation error equation and $\tilde{W}(k+1)$ from the update law in the above equation and applying Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}\Delta V \leq {} & E(\mathrm{tr}\{3\hat{\psi}^T(k)\tilde{W}(k)\tilde{W}^T(k)\hat{\psi}(k) \\ & + 3\tilde{\psi}^T(k)W(k)W^T(k)\tilde{\psi}(k) + 3\mu^T(k)\mu(k) - \tilde{x}^T(k)\tilde{x}(k)\}) \\ & + \frac{1}{3}E(\mathrm{tr}\{3(1-\gamma\,\|I - \beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\|)^2\tilde{W}^T(k)\tilde{W}(k) \\ & + 3\beta^2(k+1)e(k+1)\hat{\psi}^T(k)\hat{\psi}(k)e^T(k+1) \\ & + 3\gamma^2\|I - \beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\|^2 W^T(k)W(k) \\ & - \tilde{W}^T(k)\tilde{W}(k)\}) \\ \leq {} & 3\psi_{\max}^2 E(\|\tilde{W}(k)\|^2) + 3E(\mathrm{tr}\{\tilde{\psi}^T(k)W(k)W^T(k)\tilde{\psi}(k)\}) \\ & + 3E(\|\mu(k)\|^2) - E(\|\tilde{x}(k)\|^2) \\ & + \frac{1}{3}E(\mathrm{tr}\{-6\gamma\,\|I - \beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\|\tilde{W}^T(k)\tilde{W}(k) \\ & + 3\gamma^2\|I - \beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\|^2\tilde{W}^T(k)\tilde{W}(k) \\ & + 6\beta^2(k+1)\tilde{x}(k+1)\hat{\psi}^T(k)\hat{\psi}(k)\tilde{x}^T(k+1) \\ & + 6\beta^2(k+1)v(k+1)\hat{\psi}^T(k)\hat{\psi}(k)v^T(k+1) \\ & + 3\gamma^2\|I - \beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\|^2 W^T(k)W(k) \\ & + 2\tilde{W}^T(k)\tilde{W}(k)\})\end{aligned}$$

where $\hat{\psi}(k) = \psi(x_s(k), u(k))$. Note that $E(v^T(k+1)v(k+1)) = E(\|v(k+1)\|^2) \leq v_M^2$. Assuming that the basis function $\psi(\cdot)$ is a Lipschitz function with the Lipschitz constant c_ψ, in the difference of Lyapunov function, we

arrive at

$$\begin{aligned}\Delta V \le\ & 3\psi_{\max}^2 E(\|\tilde{W}(k)\|^2) + 3W_{\max}^2 c_\psi^2 E(\|\tilde{x}(k)\|^2) + 3\mu_M^2 \\ & -E(\|\tilde{x}(k)\|^2) + \frac{2}{3}E(\|\tilde{W}(k)\|^2) \\ & -2\gamma\,\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\|E(\|\tilde{W}(k)\|^2) \\ & +\gamma^2\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\|^2 E(\|\tilde{W}(k)\|^2) \\ & +\gamma^2\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\|^2 W_{\max}^2 + 2\beta_M^2\psi_{\max}^2 v_M^2 \\ & +2\beta_M^2\psi_{\max}^2\big(\psi_{\max}^2 E(\|\tilde{W}(k)\|^2) \\ & +3W_{\max}^2 c_\psi^2 E(\|\tilde{x}(k)\|^2) + 3\mu_M^2\big)\end{aligned}$$

where $W_{\max}$ and $\psi_{\max}$ are the maximum norm values of W and ψ. By substituting $\tilde{x}(k+1)$ from the state estimation error equation and combining similar terms to obtain

$$\begin{aligned}\Delta V \le\ & -\big(1-3W_{\max}^2 c_\psi^2\big(1+2\beta_M^2\psi_{\max}^2\big)\big)E(\|\tilde{x}(k)\|^2) \\ & -(2\gamma\,\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\| - 2\beta_M^2\psi_{\max}^4 - 3\psi_{\max}^2 \\ & -\frac{2}{3}-\gamma^2\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\|^2)E(\|\tilde{W}(k)\|^2) \\ & +\gamma^2\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\|^2 W_{\max}^2 + 2\beta_M^2\psi_{\max}^2 v_M^2 \\ & +3\mu_M^2\big(1+2\beta_M^2\psi_{\max}^2\big).\end{aligned}$$

Define b_1, b_2, and D as

$$\begin{aligned}B_1 &= 1-3W_{\max}^2 c_\psi^2\big(1+2\beta_M^2\psi_{\max}^2\big) \\ B_2 &= 2\gamma\,\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\| - 2\beta_M^2\psi_{\max}^4 - 3\psi_{\max}^2 \\ &\quad -2/3-\gamma^2\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\|^2 \\ D &= \gamma^2\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\|^2 W_{\max}^2 \\ &\quad +2\beta_M^2\psi_{\max}^2 v_M^2 + 3\mu_M^2\big(1+2\beta_M^2\psi_{\max}^2\big).\end{aligned}$$

If the design parameters are selected such that $B_1 > 0$ and $B_2 > 0$, the state and weight estimation errors will be UUB in the mean with the following bounds given by:

$$E(\|\tilde{x}(k)\|^2) \le \frac{D}{B_1}, \quad E(\|\tilde{W}(k)\|^2) \le \frac{D}{B_2}.$$

Remark 4: The conditions stated in Theorem 1 can be satisfied by proper selection of the basis function ψ and user-defined parameters including β_M, $\psi_{\max}$, c_ψ, and γ.

In summary, the proposed NN and weight update law can both detect and remove outliers from the measured data or system state or output vector y. First of all, if $\beta(k)$ is zero then an outlier is detected in $y(k)$, which will have no impact on NN weight update. Therefore, outliers will automatically be removed in the state estimates x_s and measurement noise will be moderated. Furthermore, by reducing the effect of outliers on the weight update law, the state estimation issue is resolved and the boundedness of state and parameter estimation errors can be obtained similar to the case without outliers in measured data.

After detection and removal of outliers, the estimated outlier-free state vector x_s can be used for FD without the risk of having false alarms. Fig. 1 shows an overview of the combined online outlier detection/removal and FD scheme. The following section briefly discusses about the fault diagnosis after outlier removal.

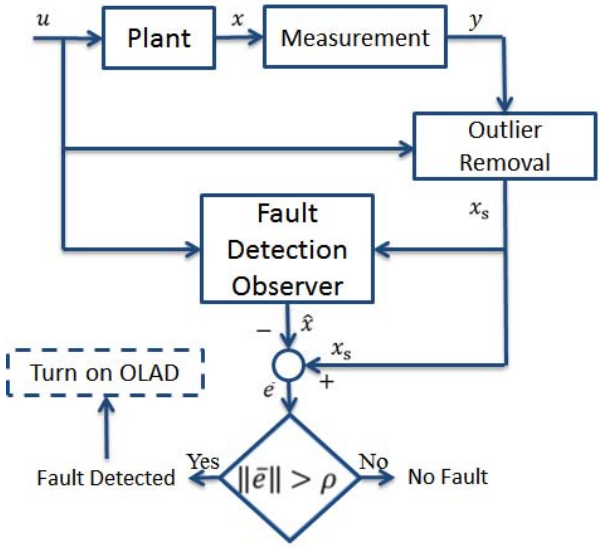

Fig. 1. Overview of the combined outlier removal and FD scheme.

IV. Fault Diagnosis Scheme

Model-based FD schemes require an observer to estimate the system states. Then an FD residual will be generated by comparing the actual and estimated system states [28]. Traditional model-based FD schemes use the measured states in the FD observer and then compare them with the observer states to detect faults. However, when the measured states are not reliable and involve outliers, false alarms could be triggered. Therefore, in this section, the outlier-free state vector x_s is used for the purpose of fault diagnosis, instead of the actual measured state vector y. As mentioned in previous section, it can be shown that $\tilde{x} = x - x_s$ is bounded in the mean, i.e., $E(\|\tilde{x}(k)\|) < \sigma_M = \sqrt{D/B_1}$.

Consider the nonlinear FD estimator

$$\begin{aligned}\hat{x}(k+1) = A_d\hat{x}(k) + \omega(x_s(k),u(k)) \\ + \hat{h}d(x_s(k),u(k);\hat{\theta}_d(k)) - A_d x_s(k)\end{aligned} \quad (5)$$

where $\hat{x}(k) \in \mathbb{R}^n$ is the estimated state vector, $\hat{h}_d$: $\mathbb{R}^n \times \mathbb{R}^{p\times n} \to \mathbb{R}^n$ is the output of the OLAD with $\hat{\theta}_d \in \mathbb{R}^{p\times n}$ being its set of unknown parameters, and A_d is a user-defined diagonal matrix, which must be selected in a way that the eigenvalues of the closed-loop system lie within the unit circle [16]. Initial values of the FD estimator are taken to be $\hat{x}(0) = \hat{x}_0$, $\hat{\theta}_d(0) = \hat{\theta}_{d_0}$, such that $\hat{h}(x_s,u,\hat{\theta}_{d_0}) = 0$.

In the proposed FD estimator, NNs are used as the OLADs. NN-based OLAD is off prior to the detection of a fault and thus its output is zero. Upon detection of a fault, the OLAD is turned on to estimate the fault dynamics.

Define the detection residual as $\bar{e} = x_s - \hat{x}$. Prior to the detection of a fault, the residual dynamics are given by

$$\begin{aligned}\bar{e}(k+1) &= x_s(k+1) - \hat{x}(k+1) \\ &= x(k+1) - \tilde{x}(k+1) - \hat{x}(k+1) \\ &= A_d\bar{e}(k) + \omega(x(k),u(k)) - \omega(x_s(k),u(k)) \\ &\qquad + \eta(x(k),u(k)) - \tilde{x}(k+1) \\ &= A_d\bar{e}(k) + \tilde{\omega}(k) + \eta(x(k),u(k)) - \tilde{x}(k+1)\end{aligned}$$

where $\tilde{\omega}(k) = \omega(x(k), u(k)) - \omega(x_s(k), u(k))$. As mentioned earlier, $\tilde{x}$, which is the difference between the actual system states x and the estimated outlier-free states x_s, is bounded. In addition, from assumptions 1 and 2, we know that $\tilde{\omega}$ and η are bounded. Therefore, with the appropriate selection of A_d, the detection residual $\bar{e}$ will remain bounded in healthy operating conditions of the system.

Now, consider a dead-zone operator

$$D[z] = \begin{cases} 0, & \text{if } |z| \le \rho \\ z, & \text{if } |z| > \rho \end{cases}$$

where ρ is the FD threshold. When the detection residual exceeds the detection threshold, a fault is declared active through the dead-zone operator and the OLAD that generates $\hat{h}_d(k)$ is initiated and tuned online using the following update law:

$$\hat{\theta}(k+1) = \hat{\theta}_d(k) + \alpha\hat{\phi}(k)D[\bar{e}^T(k)|] - \gamma\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\|\hat{\theta}_d(k) \quad (6)$$

where $\alpha > 0$ is the learning rate, $0 < \gamma < 1$ is the forgetting factor, and $\hat{\phi}(k) = \phi(x_s(k), u(k))$ is a basis function like sigmoid or RBF. Moreover, the output of the OLAD will be given by

$$\hat{h}_d(k) = \hat{\theta}_d^T(k)\phi(x_s(k), u(k)).$$

After detection, the residual dynamics can be described by

$$\bar{e}(k+1) = A_d\bar{e}(k) + \tilde{\omega}(k) + \eta(x(k), u(k)) - \tilde{x}(k+1) + h(x(k), u(k)) - \hat{h}_d(x_s(k), u(k); \hat{\theta}_d(k)).$$

Asserting the NLIP assumption on the fault function, the above equation can be rewritten as

$$\bar{e}(k+1) = A_d\bar{e}(k) + \tilde{\omega}(k) + \eta(x(k), u(k)) - \tilde{x}(k+1) + \tilde{\theta}_d^T(k)\phi(x_s(k), u(k)) + \varepsilon(k) \quad (7)$$

where $\tilde{\theta}_d(k) = \theta_d - \hat{\theta}_d(k)$ is the OLAD parameter estimation error, and $\varepsilon(k)$ is the approximation error, which is bounded by ε_M.

The stability of the proposed scheme will be investigated in the following theorem.

Theorem 2 (Fault Diagnosis Observer Performance): Let the proposed observer in (5) be used to monitor the system in (1), with the OLAD turned on upon detection of a fault. Let the update law in (6) be used to update the unknown parameter set $\hat{\theta}_d$. Then, the FD residual, $\bar{e}(k)$, and the parameter estimation errors, $\tilde{\theta}_d(k)$ are UUB in the mean.

Proof: Consider the following Lyapunov function candidate:

$$V = E(\mathrm{tr}\{\bar{e}^T(k)\bar{e}(k)\}) + E\left(\mathrm{tr}\left\{\tilde{\theta}_d^T(k)\tilde{\theta}_d(k)\right\}\right).$$

Then, the first difference of the Lyapunov function is given by

$$\Delta V = \underbrace{E(\mathrm{tr}\{\bar{e}^T(k+1)\bar{e}(k+1) - \bar{e}^T(k)\bar{e}(k)\})}_{\Delta V_1} + \underbrace{E\left(\mathrm{tr}\left\{\tilde{\theta}_d^T(k+1)\tilde{\theta}_d(k+1) - \tilde{\theta}_d(k+1) - \tilde{\theta}_d^T(k)\tilde{\theta}_d(k)\right\}\right)}_{\Delta V_2}$$

By substituting $\bar{e}(k+1)$ from (7) in ΔV_1 and applying the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}\Delta V_1 \le\ & E\big(\mathrm{tr}\big\{6\bar{e}^T(k)A_d^TA_d^{\bar{e}}(k) + 6\tilde{\omega}^T(k)\tilde{\omega}(k) \\ & +6\eta^T(x(k),u(k))\eta(x(k),u(k)) \\ & +6\tilde{x}^T(k+1)\tilde{x}(k+1) + 6\hat{\phi}^T(k)\tilde{\theta}_d(k)\tilde{\theta}_d^T(k)\hat{\phi}(k) \\ & +6\varepsilon^T(k)\varepsilon(k) - \bar{e}^T(k)\bar{e}(k)\big\}\big) \\ \le\ & 6E\big(\mathrm{tr}\big\{A_d^TA_d\bar{e}(k)\bar{e}^T(k)\big\}\big) + 6E\big(\tilde{\omega}^T(k)\tilde{\omega}(k)\big) \\ & +6E\big(\eta^T(x(k),u(k))\eta(x(k),u(k))\big) \\ & +6E\big(\tilde{x}^T(k+1)\tilde{x}(k+1)\big) + 6E\big(\varepsilon^T(k)\big) \\ & +6E\big(\mathrm{tr}\big\{\hat{\phi}(k)\hat{\phi}^T(k)\tilde{\theta}_d(k)\tilde{\theta}_d^T(k)\big\}\big) - E\big(\bar{e}^T(k)\bar{e}(k)\big).\end{aligned}$$

Now, substitute $\tilde{\theta}_d(k+1)$ from (6) in ΔV_2 and use the Cauchy–Schwarz inequality

$$\begin{aligned}\Delta V_2 =\ & E(\mathrm{tr}\{\tilde{\theta}_d(k) - \alpha\hat{\phi}(k)\bar{e}^T(k) - \gamma\|I - \alpha\hat{\phi}^T(k)\|\tilde{\theta}_d(k) \\ & +\gamma\|I-\alpha\hat{\phi}(k)\hat{\phi}^T(k)\|\theta_d)^T(\tilde{\theta}_d(k) - \alpha\hat{\phi}(k)\bar{e}^T(k) \\ & -\gamma\|I-\alpha\hat{\phi}(k)\hat{\phi}^T(k)\|\tilde{\theta}_d(k) + \gamma\|I-\alpha\hat{\phi}(k)\hat{\phi}^T(k)\|\theta_d) \\ & -\tilde{\theta}_d^T(k)\tilde{\theta}(k)\}) \\ \le\ & 3\alpha^2\hat{\phi}^T(k)\hat{\phi}(k)E(\bar{e}^T(k)\bar{e}(k)) \\ & +3\gamma^2\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\|^2\ E\left(\mathrm{tr}\left\{\theta_d^T\theta_d\right\}\right) \\ & +2E(\mathrm{tr}\{\tilde{\theta}_d^T(k)\tilde{\theta}_d(k)\}) \\ & -6\gamma\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\|E(\mathrm{tr}\{\tilde{\theta}_d^T(k)\tilde{\theta}_d(k)\}) \\ & +3\gamma^2\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\|^2\ E(\mathrm{tr}\{\tilde{\theta}_d^T(k)\tilde{\theta}_d(k)\}).\end{aligned}$$

By combining ΔV_1 and ΔV_2, taking Frobenius norm, and using Assumption 2, we arrive at

$$\begin{aligned}\Delta V \le\ & 6A_{d_{\max}}^2E(\|\bar{e}(k)\|^2) + 6c_g^2E(\|\tilde{x}(k)\|^2) + 6\eta_M^2 + 6\sigma_M^2 \\ & +6\varepsilon_M^2 + 6\phi_{\max}^2E(\tilde{\theta}_d(k)\|^2) - E(\|\bar{e}(k)\|^2) \\ & +3\alpha^2\phi_{\max}^2E(\|\bar{e}(k)\|^2) + 3\gamma^2\ \|I-\alpha\hat{\phi}(k)\hat{\phi}^T(k)\|^2\ \theta_{d_{\max}}^2 \\ \Delta V \le\ & 6A_{d_{\max}}^2E(\|\bar{e}(k)\|^2) + 6c_g^2E(\|\tilde{x}(k)\|^2) + 6\eta_M^2 + 6\sigma_M^2 \\ & +6\varepsilon_M^2 + 6\phi_{\max}^2E(\|\tilde{\theta}_d(k)\|^2) - E(\|\bar{e}(k)\|^2) \\ & +3\alpha^2\phi_{\max}^2E(\|\bar{e}(k)\|^2) + 3\gamma^2\|I-\alpha\hat{\phi}(k)\hat{\phi}^T(k)\|^2\theta_{d_{\max}}^2 \\ & -6\gamma\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\|E(\|\tilde{\theta}_d(k)\|^2) \\ & +(3\gamma^2\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\|^2 + 2)E(\|\tilde{\theta}_d(k)\|^2) \\ \le\ & -\big(1 - 6A_{d_{\max}}^2 - 3\alpha^2\phi_{\max}^2\big)E(\|\bar{e}(k)\|^2) \\ & -(6\gamma\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\| - 6\phi_{\max}^2 \\ & -3\gamma^2\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\|^2 - 2)E(\|\tilde{\theta}_d(k)\|^2) \\ & +6c_g^2\sigma_M^2 + 6\eta_M^2 + 6\sigma_M^2 + 6\varepsilon_M^2 \\ & +3\gamma^2\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\|^2\ \theta_{d_{\max}}^2.\end{aligned}$$

Therefore, $\bar{e}$ and $\tilde{\theta}_d$ are UUB in the mean if the following conditions are satisfied:

$$6A_{d_{\max}}^2 + 3\alpha^2\phi_{\max}^2 < 1$$
$$3\gamma\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\|(2 - \gamma\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\|) > 6\phi_{\max}^2 + 2.$$

Moreover, the bounds are given by

$$E(\|\bar{e}(k)\|^2) \le \frac{F}{C_1}, \qquad E(\|\tilde{\theta}_d(k)\|^2) \le \frac{F}{C_2}$$

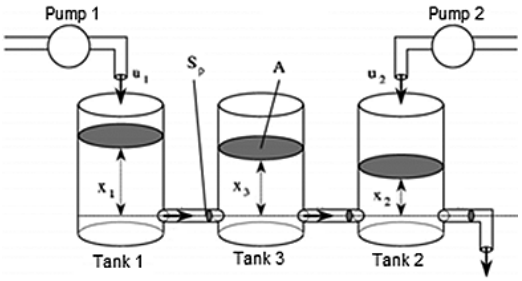

Fig. 2. Schematic view of the three-tank system.

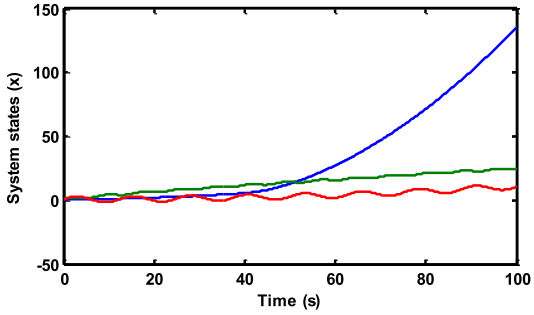

Fig. 3. Actual system states x.

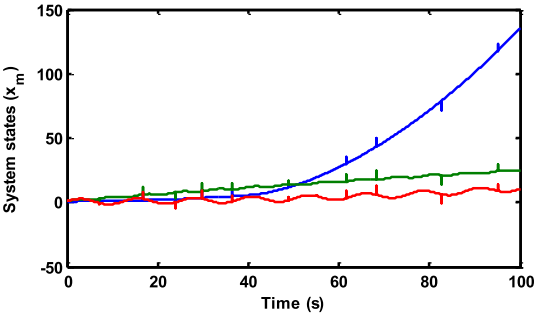

Fig. 4. Measured system states.

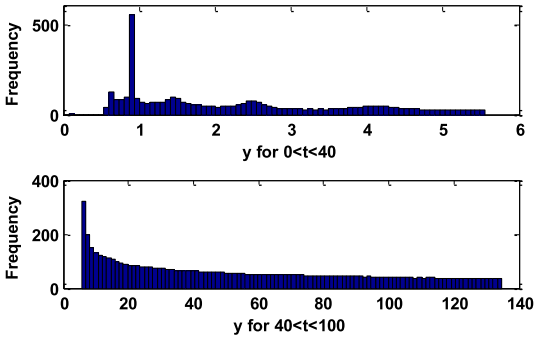

Fig. 5. Distribution of the measured data y_1.

where F, C_1, and C_2 are defined by

$$\begin{aligned} F &= 6c_g^2\sigma_M^2 + 6\eta_M^2 + 6\sigma_M^2 + 6\varepsilon_M^2 \\ &\quad +3\gamma^2 \parallel I - \alpha\hat{\phi}(k)\hat{\phi}^T(k) \parallel^2 \theta_{d_{\max}}^2 \\ C_1 &= 1 - 6A_{d_{\max}}^2 - 3\alpha^2\phi_{\max}^2 \\ C_2 &= 6\gamma \parallel I - \alpha\hat{\phi}(k)\hat{\phi}^T(k) \parallel -6\phi_{\max}^2 - 2 \\ &\quad -3\gamma^2 \parallel I - \alpha\hat{\phi}(k)\hat{\phi}^T(k) \parallel^2 . \end{aligned}$$

V. Simulation Results

In this section, a three-tank water system is selected to verify the performance of the proposed schemes in simulations and then an axial piston pump testbed is used as an experimental study to show the effectiveness of the proposed outlier removal technique in practice.

Example 1: A schematic view of the three-tank benchmarking system [29] is shown Fig. 2. This system consists of three tanks connected to each other, two input pumps on tanks 1 and 2 and one water outlet on tank 2.

The three-tank system dynamics are described by

$$x(k+1) = \omega(x(k), u(k)) + \eta(x(k))$$

where $x = [x_1, x_2, x_3]^T$ is the state vector and $\omega(x(k), u(k))$ is the known nonlinear dynamics of the system [17] given by

$$\begin{aligned} &\omega x(k, u(k)) \\ &= \begin{bmatrix} \frac{T}{A}\{-c_1 S_p \text{sign}(x_1(k) - x_3(k))\sqrt{2g|x_1(k) - x_3(k)|} \\ \quad +u_1(k)\} + x_1(k) \\ \frac{T}{A}\{-c_3 S_p \text{sign}(x_2(k) - x_3(k))\sqrt{2g|x_2(k) - x_3(k)|} \\ \quad -c_2 S_p \sqrt{2_g x_2(k)} + u_2(k)\} + x_2(k) \\ \frac{T}{A}\{-c_1 S_p \text{sign}(x_1(k) - x_3(k))\sqrt{2g|x_1(k) - x_3(k)|} \\ -c_3 s_p \text{sign}(x_3(k) - x_2(k))\sqrt{2_g|x_3(k)} + x_2(k)\} - x_3(k) \end{bmatrix} \end{aligned}$$

where T is the sampling time chosen to be 0.01 s, $A = 0.0154$ m^2 is the cross section of the tanks, $S_p = 5 \times 10^{-5}$ m^2 is the cross section of the connecting pipes, $c_1 = 1, c_2 = 0.8$, and $c_3 = 1$ are the outflow coefficients, and $g = 9.8$ m/s^2 is the standard gravity. Moreover, $\eta(x(k)) = [10^{-3}\sin(0.7\ \text{kT})10^{-2}\cos(0.8\ \text{kT})10^{-1.65}\cos(0.5kT)]^T$ represents the modeling uncertainty.

This system is subjected to a fault, which is given in terms of leakage in tank 1 and occurs at time $t_0 = 40$ s. The fault function is described by

$$h(x(k)) = [0.0154(1 - e^{-0.5T(k-k_0)})\sqrt{2gx_1(k)}\,,\ \ 0\,,\ \ 0]^T.$$

The FD estimator in (5) is used to detect the faults, where $A_d = 0.99I_{3\times3}$. The OLAD output is given by $\hat{h}_d(k) = \hat{\theta}_d^T(k)\phi(Vx_s(k) + B)$, where $\hat{\theta}_d \in \mathbb{R}^{8\times3}$ is the estimated parameters while $\phi \in \mathbb{R}^8$ is a vector of sigmoid functions. Moreover, V and B are selected randomly and the update law parameters are $\alpha = 0.5, \gamma = 10^{-4}$. The detection threshold, ρ, is selected to be 2.5.

Fig. 3 shows the actual system states, while Fig. 4 shows the measured system states involving a number of outliers. The state distribution for healthy and faulty periods is shown in Fig. 5. The mean and variance of the distribution in healthy period are 2.33 and 2.05, respectively, while mean and variance values for the faulty period are 54.24 and 1479.53, respectively. If the measured data is used for FD, the outliers will cause false alarms to be triggered. This can be observed in Fig. 6, where the detection residual is plotted along with detection threshold. Fault is seeded at $t = 40$ s, but a false alarm will be triggered at about $t = 8$ s. It is obvious that in

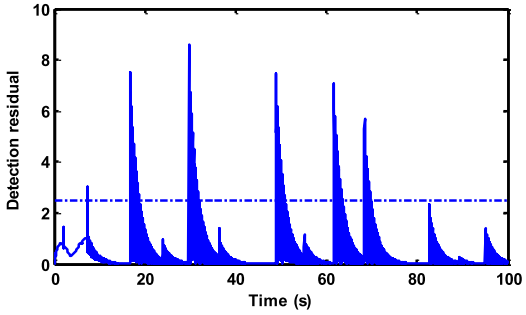

Fig. 6. Detection residual without outlier removal.

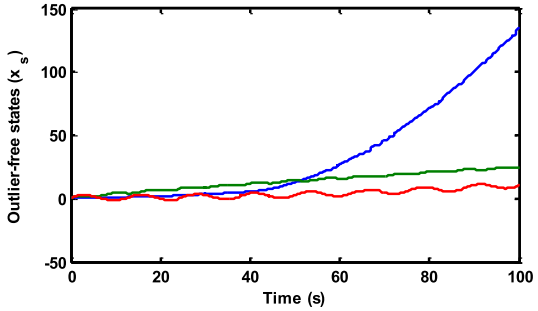

Fig. 7. Estimated outlier-free states x_s.

this case, the estimated fault given by the OLAD cannot be close to the actual fault at all.

To fix this problem, the proposed outlier removal scheme is first used to remove the outliers from the measured data. An NN with 12 hidden layer neurons and sigmoid activation functions is used to estimate x_s and p, which is the window size is selected as 100 (that means 1 s). The estimated x_s is shown in Fig. 7. It can be observed that the outliers are removed from the measured states. Table I shows the outlier detection results for several cases with different number of outliers at randomly selected times. It is observed that the proposed scheme has been able to detect 100% of the outliers in most of the cases, with low number of false positives.

When x_s is used in the FD observer, no false alarm will be triggered and the actual fault is detected at $t = 74$ s (Fig. 8). Fig. 9 shows the estimated and actual fault magnitudes in this case. Unlike the previous case, when no outlier removal was performed, the fault can be estimated with a small error. The simulation results clearly show the effectiveness of the proposed outlier detection/removal scheme. Furthermore, the importance of removing the outliers before performing the FD is clarified.

Example 2: A linear system has been selected as the second example to compare the proposed scheme with a Kalman filter-based approach. The system is described by $x(k+1) = Ax(k)$, where A is defined by

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad \text{for } t \le 50$$

$$A = 1.003\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad \text{for } t > 50$$

with $\theta = \pi/500$. In fact, the system dynamics are slightly changed during the simulation, to test the proposed scheme

TABLE I
OUTLIER DETECTION STATISTICS

Case #	Total number of outliers	Number of True positives	Number of False positives
1*	12	12	1
2	8	8	0
3	14	14	2
4	18	17	1
5	9	9	0
6	11	11	0

* The presented plots correspond to case 1

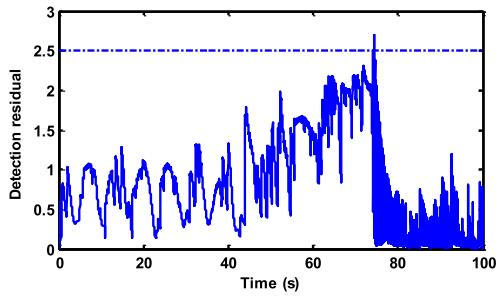

Fig. 8. Detection residual with outlier removal.

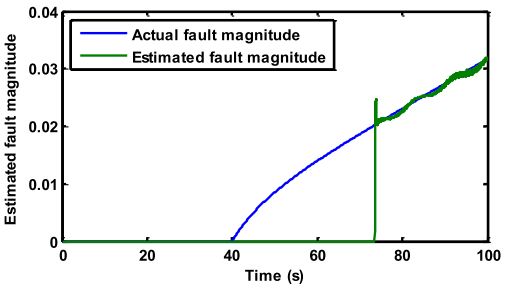

Fig. 9. Detection residual with outlier removal.

and compare its performance with a Kalman filter-based method under nonstationary operating conditions. The entire simulation time is taken to be 100 s and the sampling time is 0.1 s, and the states are measured as follows:

$$y(k) \sim \mathcal{N}\left(x(k), \begin{bmatrix} .25 & 0 \\ 0 & 0.25 \end{bmatrix} + \begin{bmatrix} r_1(k) & 0 \\ 0 & r_2(k) \end{bmatrix}\right)$$

where $r_1(k)$ is set to 1 with probability κ and to zero with probability $1-\kappa$.

Fig. 10 shows the actual system states x, along with measured states y when $\kappa = 0.05$ and Fig. 11 shows the distribution of $y_2 - x_2$ for $t < 50$ and $t > 50$. In the first half of the simulation, the mean and variance of $y_2 - x_2$ are 0.049 and 0.82, respectively, whereas mean and variance in the second half of the simulation are -0.008 and 1.37. The proposed outlier detection and removal scheme is applied on the measured data, with $p = 10$, $\gamma = 0.01$, and randomly selected W matrix. The estimated states are shown in Fig. 12. It is worth mentioning that the performance of the scheme is not degraded after the change in the system dynamics at $t = 50$ s.

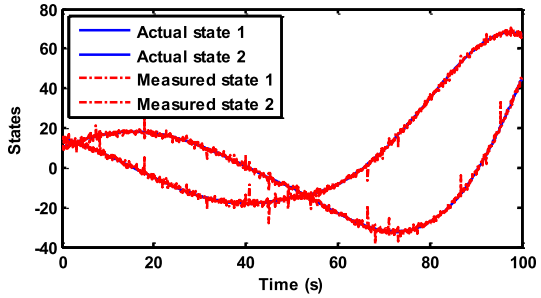

Fig. 10. Actual and measured system states.

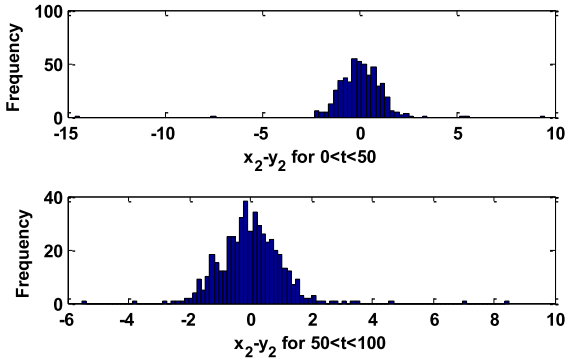

Fig. 11. Distribution of the difference between actual and measured states.

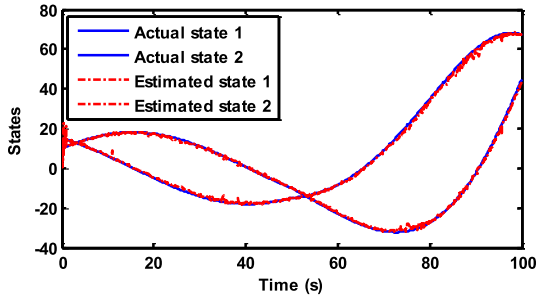

Fig. 12. Estimated states using the proposed NN-based scheme.

Next an outlier-robust Kalman filter whose parameters are fitted according to the maximum-likelihood criterion [19], is applied on the same measured data and the result is shown in Fig. 13. Although this method has a good performance in the first 50 s of the simulation, its performance is extremely degraded when the dynamics of the system changes. Mean squared error of state estimation for both of the methods are presented in Table II for comparison. This simulation clearly shows that unlike Kalman filter-based approaches, our outlier removal method is robust to changes in the system dynamics (which could be due to the faults or changes in the operating conditions). Although the state estimation error of the Kalman filter method after the occurrence of fault seems to be useful for FD, the large error in estimation makes this method useless for outlier removal in the presence of fault.

The simulations have been repeated for 4000 times using the proposed method. The noise and outliers are random, thus vary from one simulation to the other. The average

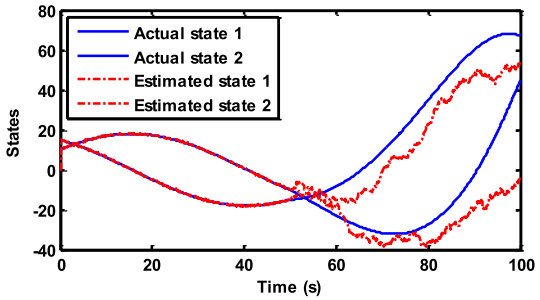

Fig. 13. Estimated states using an outlier robust Kalman filter.

TABLE II
OUTLIER REMOVAL PERFORMANCE COMPARISON

Method	Mean squared error
Proposed scheme	2.14
Outlier robust Kalman filter	110.29

TABLE III
RESULTS OF THE REPEATED SIMULATIONS USING NN-BASED SCHEME

Total number of simulations	Average MSE	Maximum MSE	Average percentage of detected outliers
4000	1.98	3.56	%97

Fig. 14. Picture of the axial piston pump testbed.

and maximum mean squared error of state estimation and the average percentage of detected outliers are shown in Table III. The results imply that the proposed method is able to detect and remove the outliers with consistently high performance. The important point is that the average percentage of detected outliers is as high as 97% and the maximum mean squared error in all the simulations is less than four, which is still very low compared with the Kalman filter method [19].

Example 3: An axial piston pump testbed is used to test the performance of the proposed outlier removal scheme and observe its effect on FD in an experimental study. A picture of this testbed is shown in Fig. 14. The nonlinear dynamics

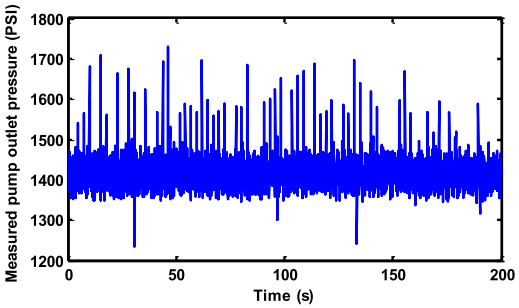

Fig. 15. Measured pump outlet pressure.

of this system is described by

$$x_i(k+1) = x_i(k) + \frac{BT}{C - A_p S_{pi}}\left(Q_{kpi}(k) - Q_{pi}(k) - Q_{lpi}(k)\right) \quad \text{for } i = 1, \ldots, 9$$

$$x_{10}(k+1) = x_{10}(k) + \frac{CBT}{V_c}\left(\sum_{i=1}^{9} Q_{pi}(k) - C_{d2} A_v \sqrt{\frac{2x_{10}(k)}{\rho_c}}\right)$$

where $[x_1(k), \ldots, x_9(k)]$ represent the pressures in the nine pistons, $x_{10}(k)$ is the pump outlet pressure, B is the bulk modulus of the hydraulic fluid, V_c is the theoretical volume of flow, A_p is the piston area, C_{d2} is the discharge coefficient of needle valve orifice, A_v is the orifice area of the needle valve, ρ_c is the flow density, and T is the sampling time. Furthermore, $Q_{kpi}(k)$, $Q_{lpi}(k)$, $Q_{pi}(k)$, $S_{pi}(k)$ are obtained by

$$Q_{kpi}(k) = \frac{\omega \pi d^2 R_p}{4} \tan(\beta_c) \sin(\omega k - (i-1)\alpha_p)$$

$$Q_{lpi}(k) = \frac{\pi r h_g^3}{6\mu L}(x_i(k) - P_c)$$

$$Q_{pi}(k) = C_{d1} A_{d1} \sqrt{\frac{2\,|x_i(k) - x_{10}(k)|}{\rho_c}}$$

$$S_{pi}(k) = R_p \tan(\beta_c)\,(1 - \cos(\omega k - (i-1)\alpha_c))$$

where ω is the angular velocity of the pump drive shaft, d is the diameter of the piston, R_p is the piston radius on barrel, β_c is the angle of swash plate, α_p is the phase delay, r is the radius of piston, μ is the absolute fluid velocity, and L is the length of leakage passage.

In this system, only one of the states, namely the pump outlet pressure is measurable. Therefore, for the purpose of FD, an output observer [24] is constructed using the model of the system. The dynamics of the output observer is slightly different from the full state observer presented in this paper, in that it uses the output of the system instead of the entire state vector. The sampling time for measuring the data is 0.1 s. The output of the system is measured in healthy operating conditions for 200 s and is shown in Fig. 15. Mean and variance of the whole data set are 1428.79 and 59.34, respectively.

If the measured data is directly used for FD, several false alarms will be triggered. This can be clearly observed in Fig. 16, which shows the FD residual and threshold without any outlier removal performed. To solve this problem, we

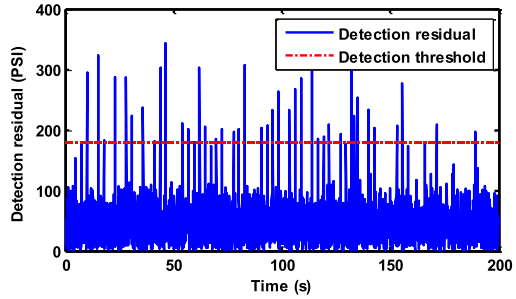

Fig. 16. Detection residual without outlier removal.

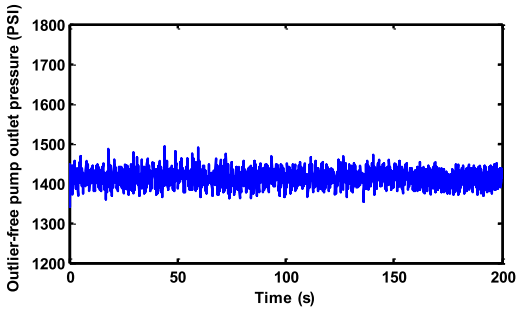

Fig. 17. Estimated outlier-free pump outlet pressure.

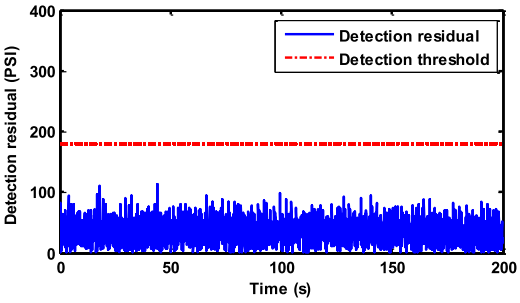

Fig. 18. Detection residual with outlier removal.

use our proposed outlier removal scheme. The user-defined parameters of the update law are selected as $\tau = 0.2$ and $p = 20$. The estimated state x_s is shown in Fig. 17 and the FD residual when x_s is used for FD is shown in Fig. 18. It is clearly observed that the outliers are removed from the data and no false alarm is triggered.

VI. Conclusion

In this paper, an NN-based online outlier detection and removal scheme was presented and combined with a model-based FD scheme. It was demonstrated that the underlying distribution of data in the case of a data driven scheme or states in the case of model-based FD is nonstationary due to the presence of changing dynamics, outliers, and noise. Moreover, it was shown that a separate OIR scheme is necessary prior to any FD and diagnosis. On the other hand, when the outliers are removed by the proposed scheme, FD can be performed successfully. The proposed observer-based method changes the learning rate to zero when an outlier is detected. The proposed OIR scheme can function even when the measured data, which is going to be used for monitoring and FD is contaminated

with outliers. Then, a data driven FD scheme can yield low detection rate and high false alarm rate similar to the model-based FD framework.

REFERENCES

[1] D. M. Hawkins, *Identification of Outliers*. London, U.K.: Chapman & Hall, 1980.

[2] S. M. Namburu, M. Wilcutts, S. Chigusa, Q. Liu, C. Kihoon, and K. Pattipati, "Systematic data-driven approach to real-time fault detection and diagnosis in automotive engines," in *Proc. IEEE Autotestcon*, Sep. 2006, pp. 59–65.

[3] R. Isermann, "Model-based fault-detection and diagnosis—Status and applications," *Annu. Rev. Control*, vol. 29, no. 1, pp. 71–85, 2005.

[4] B. J. Ohran, J. Liu, D. M. de la Peña, P. D. Christofides, and J. F. Davis, "Data-based fault detection and isolation using feedback control: Output feedback and optimality," *Chem. Eng. Sci.*, vol. 64, no. 10, pp. 2370–2383, 2009.

[5] R. Ghimire, C. Sankavaram, A. Ghahari, K. Pattipati, Y. Ghoneim, M. Howell, *et al.*, "Integrated model-based and data-driven fault detection and diagnosis approach for an automotive electric power steering system," in *Proc. IEEE AUTOTESTCON*, Sep. 2011, pp. 70–77.

[6] X. Zhang, M. M. Polycarpou, and T. Parisini, "Fault diagnosis of a class of nonlinear uncertain systems with Lipschitz nonlinearities using adaptive estimation," *Automatica*, vol. 46, no. 2, pp. 290–299, 2010.

[7] D. Blake and M. Brown, "Simultaneous, multiplicative actuator and sensor fault estimation using fuzzy observers," in *Proc. IEEE Int. Fuzzy Syst. Conf.*, Jul. 2007, pp. 1–6.

[8] H. A. Talebi and K. Khorasani, "A neural network-based actuator gain fault detection and isolation strategy for nonlinear systems," in *Proc. 46th IEEE Conf. Decision Control*, Dec. 2007, pp. 2614–2619.

[9] G. Jiaxing and T. Gang, "A feedback-based fault detection scheme for aircraft systems with damage," in *Proc. 8th ASCC*, 2011, pp. 1431–1436.

[10] H. Ferdowsi and S. Jagannathan, "A unified model-based fault diagnosis scheme for non-linear discrete-time systems with additive and multiplicative faults," *Trans. Inst. Meas. Control*, vol. 35, no. 6, pp. 742–752, Aug. 2013.

[11] P. J. Rousseeuw and K. V. Driessen, "A fast algorithm for the minimum covariance determinant estimator," *Technometrics*, vol. 41, no. 3, pp. 212–223, 1999.

[12] E. M. Knorr, R. T. Ng, and V. Tucakov, "Distance-based outliers: Algorithms and applications," *Int. J. Very Large Data Bases*, vol. 8, nos. 3–4, pp. 237–253, 2000.

[13] E. M. Knorr and R. T. Ng, "Algorithms for mining distance-based outliers in large datasets," presented at the Proc. 24rd Int. Conf. Very Large Data Bases, Vancouver, BC, Canada, 1998.

[14] G. Kollios, D. Gunopulos, N. Koudas, and S. Berchtold, "Efficient biased sampling for approximate clustering and outlier detection in large data sets," *IEEE Trans. Knowl. Data Eng.*, vol. 15, no. 5, pp. 1170–1187, Sep./Oct. 2003.

[15] Z. Jiang-She and L. Yiu-Wing, "Robust clustering by pruning outliers," *IEEE Trans. Syst., Man, Cybern., B, Cybern.*, vol. 33, no. 6, pp. 983–998, Dec. 2003.

[16] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and R. Sander, "LOF: Identifying density-based local outliers," presented at the Proc. ACM SIGMOD Int. Conf. Management of data, Dallas, TX, USA, 2000.

[17] V. J. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artif. Intell. Rev.*, vol. 22, no. 2, pp. 85–126, 2004.

[18] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surveys*, vol. 41, no. 3, pp. 1–58, 2009.

[19] G. Agamennoni, J. I. Nieto, and E. M. Nebot, "An outlier-robust Kalman filter," in *Proc. IEEE ICRA*, May 2011, pp. 1551–1558.

[20] S. C. Chan, Z. G. Zhang, and K. W. Tse, "A new robust Kalman filter algorithm under outliers and system uncertainties," in *Proc. IEEE ISCAS*, vol. 5. May 2005, pp. 4317–4320.

[21] J. A. Ting, E. Theodorou, and S. Schaal, "A Kalman filter for robust outlier detection," in *Proc. IEEE/RSJ Int. Conf. IROS*, Oct./Nov. 2007, pp. 1514–1519.

[22] J. J. Gertler, "Survey of model-based failure detection and isolation in complex plants," *IEEE Control Syst. Mag.*, vol. 8, no. 6, pp. 3–11, Dec. 1988.

[23] N. Tanabe, T. Furukawa, H. Matsue, and S. Tsujii, "Kalman filter for robust noise suppression in white and colored noises," in *Proc. IEEE ISCAS*, May 2008, pp. 1172–1175.

[24] B. T. Thumati and S. Jagannathan, "A model-based fault-detection and prediction scheme for nonlinear multivariable discrete-time systems with asymptotic stability guarantees," *IEEE Trans. Neural Netw.*, vol. 21, no. 3, pp. 404–423, Mar. 2010.

[25] B. G. Amidan, T. A. Ferryman, and S. K. Cooley, "Data outlier detection using the Chebyshev theorem," in *Proc. IEEE Aerosp. Conf.*, Mar. 2005, pp. 3814–3819.

[26] Z. Yue, L. Jie, and S. Bo, "A new algorithm for outlier detection based on offset," in *Proc. 5th Int. Conf. IAS*, 2009, pp. 3–6.

[27] S. Jagannathan, *Neural Network Control of Nonlinear Discrete-Time Systems*. Boca Raton, FL, USA: CRC Press, 2006.

[28] F. Caccavale, P. Cilibrizzi, F. Pierri, and L. Villani, "Actuators fault diagnosis for robot manipulators with uncertain model," *Control Eng. Pract.*, vol. 17, no. 1, pp. 146–157, 2009.

[29] D. Theilliol, H. Noura, and J.-C. Ponsart, "Fault diagnosis and accommodation of a three-tank system based on analytical redundancy," *ISA Trans.*, vol. 41, no. 3, pp. 365–382, 2002.

Hasan Ferdowsi received the B.S. and M.S. degrees in electrical engineering from the Isfahan University of Technology, Isfahan, Iran, in 2007 and 2010, respectively, and the Ph.D. degree in electrical engineering from the Missouri University of Science and Technology (formerly known as University of Missouri-Rolla), Rolla, MO, USA, in 2013.

He also contributed in research projects on fault diagnostics and prognostics, fault accommodation, and sensor data quality enhancement as a member of the Center for Inteligent Maintanance Systems. His current research interests include fault diagnostics and prognostics, adaptive control systems, nonlinear systems, large-scale interconnected systems, and neural networks.

Sarangapani Jagannathan is with the Missouri University of Science and Technology (former University of Missouri-Rolla), Rolla, MO, USA, where he is a Rutledge-Emerson Distinguished Professor and the Site Director for the NSF Industry/University Cooperative Research Center on Intelligent Maintenance Systems. He has a courtesy appointment with the Department of Computer Science. He has co-authored 114 peer reviewed journal articles mostly in IEEE Transactions, over 212 refereed IEEE conference articles, several book chapters and three books. He holds 20 patents with several pending. He supervised the completion of 17 doctoral students and 28 M.S. students. His research funding is in excess of $14 million dollars from NSF, NASA, AFRL, and Sandia. His current research interests include adaptive and neural network control, networked control systems and sensor networks, prognostics, and autonomous systems/robotics.

He is serving as the Co-Editor for the *IET Book Series on Control*. He has been on several IEEE Transaction editorial boards and served on various IEEE conference organizing committees. He has received a number of awards both from the IEEE Society and industrial entities.

Maciej J. Zawodniok (S'03–M'06) received the M.Sc. degree in computer science from the Silesian University of Technology, Katowice, Poland, in 1999, and the Ph.D. degree in computer engineering from the University of Missouri-Rolla, Rolla, MO, USA, in 2006.

He has been with the Missouri University of Science and Technology, since 2008, where currently he is an Assistant Professor of computer engineering and Assistant Director of NSF I/UCRC on intelligent maintenance systems. He has co-authored over 17 peer-reviewed journal articles, over 30 refereed IEEE conference articles, and two book chapters. His current research interests include adaptive and energy-efficient protocols for wireless networks, network-centric systems, network security, cyber-physical, and embedded systems with applications to manufacturing and maintenance.

Dr. Zawodniok received the Prestigious NSF CAREER Award in 2010.