Received April 6, 2019, accepted July 9, 2019, date of publication July 16, 2019, date of current version August 9, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2929186

A New Observer for Nonlinear Systems With Application to Power Systems

MOHAMED F. HASSAN AND MOHAMMAD HAMMUDA
Electrical Engineering Department, Kuwait University, Kuwait City 13060, Kuwait

Corresponding author: Mohamed F. Hassan (m.f.hassan47@gmail.com)

This work was supported by Kuwait University under Grant EE 01/16.

ABSTRACT In this paper, a new observer is presented for discrete-time nonlinear dynamical systems. The designed observer is a modified version of the recently developed regularized least square (RLS) observer. It leads to an estimator with almost zero or very small overshoot while keeping the settling time of the estimation error very short, a case which cannot be satisfied by the well-known observers available in the literature. The predicted estimate of the developed estimator is calculated from a weighted average of a set of predicted estimate of a set of points to be generated around the filtered estimate of the state vector at a given sampling instant of time. Through this approach, we get a highly accurate result of the predicted state vector, which intuitively leads to highly accurate filtered estimates of the states. The developed estimator can deal with highly nonlinear systems, does not need any state transformation, has no restrictions on the output measurement model, leads to a unique solution, and last but not least avoids the computation of the Jacobian matrices. The convergence of the proposed observer is analyzed, and the results show its superior performance when compared with the RLS observer. Moreover, a modified version of the developed observer is proposed to reduce the computational time while maintaining its main features. Illustrative examples of highly nonlinear power systems are presented to show the effectiveness of the proposed approach and its superiority when compared with other well-known observers.

INDEX TERMS Discrete-time nonlinear systems, state estimation, stability analysis, power systems.

I. INTRODUCTION

The design of observers is one of the main pillars of estimation theory as they play a very important role in systems monitoring and/or control, especially when some of the states are unmeasurable or expensive to measure. Observers are used in many applications such as knee stiffness [1], charge state in batteries [2], inertial navigation systems [3], glucose regulation in diabetic patients [4], temperature and emissivity in strip annealing furnaces [5], chaos synchronization in secure communication systems [6], brake clutch in hybrid electrical vehicles [7], rotor crack detection in rotating machinery [8] and others. Therefore, developing linear and nonlinear observers with improved performance is always of interest.

For linear systems, whether deterministic or stochastic, estimation theory is well developed and the convergence of any new developed observer can be directly related to the observability and detectability of the system. On the other hand, the design of new observers for nonlinear systems is a challenging task as there is no general systematic framework to follow, and it is challenging to prove their convergence.

In the literature, different techniques have been developed to address the challenging problem of nonlinear state estimation. These approaches can be classified into two classes. In the first class, the model is transformed to a new form close enough to a linear system and hence an observer is designed using linear system estimation theory [9]–[23]. Observers designed in this class are greatly dependent on system's nonlinearities which may lead them to be inapplicable for certain nonlinear systems. However, once the system is transformed to the new form, the implementation of the designed observer is straight forward and often with relatively small computational time. In the second class, observers are designed using systematic methodologies regardless of the system's nonlinearities being strong or weak [24]–[32]. The simplicity of the implementation of the second class observers compared to the first class makes them more preferable in case of highly nonlinear systems.

The associate editor coordinating the review of this manuscript and approving it for publication was Ramazan Bayindir.

Observers designed for nonlinear systems satisfying Lipschitz condition can be classified in the aforementioned first class. Among these observers are Thau's observer [9], Linear Matrix Inequalities (LMI)-based observer [10], H_∞-based observer [11], Geometric observer [12]–[15], observers based on sliding mode approach and its variants [16]–[19], and High Gain (HG) observer [20]–[23]. The main disadvantage of the observers belonging to this class is that they are not applicable for all nonlinear systems. For example, some of these observers necessitate the transformation of the nonlinear system to a specific form which can be either complicated or even impossible for highly nonlinear systems. Others require some conditions and/or restrictions to be satisfied such as the global, local, or at least one-sided Lipschitz conditions, linear measurement model, and others which cannot be easily fulfilled in many practical applications. Overcoming these drawbacks has motivated the development of new approaches to be classified in the second class as mentioned earlier.

On the other hand, examples of the well-known observers belonging to the second class are the Extended Luenberger Observer (ELO) [24]–[27], the nonlinear State Dependent Riccati Equation (SDRE) observer [28]–[30], and the recently developed RLS observer [31], [32]. Unlike the SDRE observer and the ELO for which the results of the estimated states are not unique since they depend on the chosen State Dependent Matrix (SDM) for the former, and the chosen gain matrix for the later, the result of the RLS observer is unique since the mathematical structure of the algorithm was achieved through the formulation and solution of a regularized least-square optimization problem at each sampling point

Although there is a variety of state estimation techniques for nonlinear systems as mentioned earlier, there is always a need to develop new state estimation approaches which offer improved performance and overcome the drawbacks of existing estimators. In this paper, a new approach for nonlinear state estimation belonging to the aforementioned second class, namely the Smoothed Regularized Least Square (SRLS) observer, is developed. As usual, the final result is achieved in a two-step procedure, namely, the prediction step and the filtering step. It is well known that the better we get an estimate for the predicted state vector, the better the result of filtered estimate of the state vector and hence the final result. Having this in mind, the proposed SRLS estimator introduces a new approach for the estimation of the predicted state vector, while the filtered estimate of the state vector is that of the RLS observer.

The proposed approach uses the exact model of the nonlinear system in the prediction phase of the algorithm rather than the linearized model as used in the RLS observer. Our objective is to finally achieve a smooth estimator with a very small or no overshoot as well as a very short transient period of the estimation error. The main idea in our approach can be summarized as follows. It is well known that the performance of nonlinear systems is highly dependent on the starting point. Therefore, at the current instant of sampling time, instead of calculating the predicted estimate of the state vector from its present filtered estimate, a set of predetermined points around the present filtered estimate is calculated. This set of points is propagated through the system model to get a set of predicted estimates of the state vectors. This set is then used to calculate a set of predicted outputs through the measurement model. The weighted averages of these two sets are calculated to get the final predicted estimates of the state and output vectors. The weighted averages of the deviation of the elements of the sets of the predicted states and predicted outputs from their weighted averages are used to calculate the state dependency matrix $\boldsymbol{P}_{xx_{k+1|k}}$, the output dependency matrix $\boldsymbol{P}_{yy_{k+1|k}}$ and the cross dependency matrix $\boldsymbol{P}_{xy_{k+1|k}}$. The calculated dependency matrices are in turn used to calculate the gain matrix to be used in the filter step. Using such a procedure, it is expected to get much better estimates of the predicted state and output vectors which obviously lead to much better filtered estimate of the state vector as we receive the $(k+1)^{th}$ measurement. It is obviously expected that, better filtered estimates of the states reduce their overshoots at the start of the estimation process and hence lead to shorter transient period of the estimation errors. The convergence analysis of the proposed observer is presented. From this analysis, it is shown that the SRLS estimator converges to the zero steady state error faster than the RLS estimator, which is also verified through our simulation.

In the next section, we further highlight the features and performance goals that we aim to achieve using the SRLS observer. Also, numerical simulations are presented to demonstrate that the proposed approach leads to a superior performance when compared with the RLS observer and other well-known nonlinear observers.

The rest of the paper is organized as follows. Section II presents the motivation and the goals of our research work. Section III presents the problem formulation for the state estimation of discrete-time deterministic nonlinear systems. The mathematical structure of the proposed observer is presented in section IV. Section V is devoted to the convergence analysis of the developed approach. In section VI, different examples are demonstrated to show the effectiveness of the proposed approach compared with others such as the HG, SDRE, and RLS observers. Finally, the paper is concluded in section VII.

II. MOTIVATION AND GOALS OF THIS WORK

Observers are used either to monitor the behavior of a given system or to generate feedback control strategies in observer-based control systems. Consequently, as the result of the estimator is improved, the system monitoring and/or the performance of the feedback observer-based controlled system are improved. In the following, the main drawbacks of some of the well-known observers are reviewed. Our aim is to alleviate these drawbacks and hence achieve better state estimation results.

1- Some of the well-known observers, such as the HG and Thau's observer, require state transformation and/or special form of the output models. Therefore, the type of system nonlinearities, whether weak or strong, dictates the applicability of these techniques. For some systems, it can be easily applied while for others, it can be either difficult or impossible to apply.

2- For Luenberger based-observers, HGO and others, there is a trade-off between the settling time of the estimation error and the overshoot at the start of the estimation process. In other words, to get short transient period of the estimation error, large overshoot is observed at the start of the estimation process, and vice versa. In observer-based control systems, this might pose a problem as large overshoots may lead to undesirable control signals for some applications. On the other hand, long transient period of the estimation process delays the achievement of the desired performance of the controlled system.

3- The SDRE observer and others using similar approaches necessitate the observability or the detectability of the system at each point in the sampling space in order to achieve a solution for the static Riccati equation of the dual problem. Moreover, the result of the estimation process is not unique since it depends on the chosen SDM. In some applications, it may be difficult and even impossible to select a SDM that leads to a convergent state estimator.

4- Some techniques, such as the RLS estimator, require the computation of the Jacobian matrix of the system model which may be difficult for some applications.

Based on the aforementioned drawbacks, the goal of our research is to develop an observer with the following features:

1- It can be applied to any nonlinear system irrespective of the type of nonlinearities, output model,…etc.

2- It avoids the undesired overshoot at the start of the estimation process while maintaining a very short transient period of the estimation error.

3- It is directly applicable to the original system rather than a transformed one.

4- It does not require the computation of the system's Jacobian matrix.

5- It has no restrictive conditions and leads to a unique solution.

6- The solution has to be achieved within a reasonable average CPU time per sample in order to be applicable to a wider class of practical applications.

III. PROBLEM FORMULATION

Consider the following discrete-time deterministic nonlinear system:

$$\begin{aligned} \boldsymbol{x}_{k+1} &= f_k(\boldsymbol{x}_k) \\ \boldsymbol{y}_{k+1} &= h_{k+1}(\boldsymbol{x}_{k+1}) \end{aligned} \tag{1}$$

where $\boldsymbol{x}_k \in R^n$, $\boldsymbol{y}_k \in R^m$ represent, respectively, the state and the output measurement vectors; $f_k(\boldsymbol{x}_k) \in R^n : R^n \to R^n$ is the nonlinear vector function of the states while $h_{k+1}(\boldsymbol{x}_{k+1}) \in R^m : R^n \to R^m$ is the nonlinear vector function of the measurements. The vector functions $f_k(\boldsymbol{x}_k)$ and $h_{k+1}(\boldsymbol{x}_{k+1})$ are assumed to be differentiable with respect to their arguments $\boldsymbol{x}_k, \boldsymbol{x}_{k+1}$. The vector $\boldsymbol{x}_0 = \boldsymbol{x}_{k_0} \in R^n$ represents the initial conditions and finally $k \in \{0, 1, 2, \ldots\}$ is the discrete time instant.

IV. THE MATHEMATICAL STRUCTURE

In this section, the mathematical structure of the RLS estimator is reviewed and that of the SRLS estimator is presented. Before proceeding with the details, it is of great importance to clarify the following. The estimation methodologies and theories for deterministic and stochastic systems are totally different and cannot be used alternatively. In that regard, the RLS and SRLS observers are developed for deterministic systems and are not suitable for state estimation of stochastic systems. The opposite is also true as the techniques developed for stochastic systems like the Extended Kalman Filter (EKF) are not suitable for state estimation of deterministic systems. Based on this fact, we have never seen in the literature that, for example EKF has been used to estimate the states of deterministic nonlinear systems, and the HGO has been used to estimate the states of stochastic nonlinear systems.

From the first glance, it happened that the mathematical structure of the RLS observer (and later on, the SRLS observer) looks similar to that of the EKF. However, they are entirely different as the EKF is applicable only for stochastic systems whereas the RLS and SRLS observers are applicable only for deterministic systems. Moreover, the RLS observer has been developed in [31], [32] through the formulation and the solution of RLS estimation problem while those of KF or EKF rely on the properties of conditional Gaussian distribution. For interested readers as well as to avoid any ambiguities, the development of the mathematical structure of the RLS estimator is presented in appendix A.

A. THE REGULARIZED LEAST SQUARE ESTIMATOR

Since the SRLS estimator is a modified version of the recently developed RLS estimator [31], [32], then we start firstly with the presentation of the mathematical structure of the RLS estimator.

1) PREDICTION STEP OF THE RLS ESTIMATOR

Let us assume that at the sampling instant k^{th} we know the filtered estimate of the state vector $\hat{\boldsymbol{x}}_{k|k}$, and its associated Riccati-like matrix, $\boldsymbol{P}_{k|k} \in \boldsymbol{R}^{nxn}$. Then, the predicted estimate of the state vector $\hat{\boldsymbol{x}}_{k+1|k}$, its associated Riccati-like matrix $\boldsymbol{P}_{k+1|k} \in \boldsymbol{R}^{nxn}$, and the predicted estimate of the measurement vector $\hat{\boldsymbol{y}}_{k+1|k}$ are given by:

$$\hat{\boldsymbol{x}}_{k+1|k} = f_k(\hat{\boldsymbol{x}}_{k|k}) \tag{2a}$$

$$\hat{\boldsymbol{y}}_{k+1|k} = h_{k+1}(\hat{\boldsymbol{x}}_{k+1|k}) \tag{2b}$$

$$P_{k+1|k} = \hat{A}_k P_{k|k} \hat{A}_k^T + N \tag{2c}$$

where:

$$\hat{A}_k = \left. \frac{\partial f_k}{\partial x_k} \right|_{\hat{x}_{k|k}} \tag{3}$$

2) FILTER STEP OF THE RLS ESTIMATOR

As we receive the measurement vector y_{k+1}, the filtered estimate of the state vector $\hat{x}_{k+1|k+1}$, and its associated Riccati-like matrix, $P_{k+1|k+1}$, are given by:

$$\hat{\boldsymbol{x}}_{k+1|k+1} = \hat{\boldsymbol{x}}_{k+1|k} + K_{k+1}(\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1|k}) \quad (4)$$

$$P_{k+1|k+1} = [I - K_{k+1}\hat{H}_{k+1}]P_{k+1|k} \quad (5)$$

where the gain matrix K_{k+1} is given by:

$$K_{k+1} = P_{k+1|k}\hat{H}_{k+1}^T[S + \hat{H}_{k+1}P_{k+1|k}\hat{H}_{k+1}^T]^{-1} \quad (6)$$

and

$$\hat{H}_{k+1} = \left.\frac{\partial h_{k+1}}{\partial \boldsymbol{x}_{k+1}}\right|_{\hat{\boldsymbol{x}}_{k+1|k}} \quad (7)$$

B. THE SMOOTHED REGULARIZED LEAST SQUARE ESTIMATOR

The main idea of the proposed SRLS estimator is simply as follows. For both linear and nonlinear systems, it is obvious that the behavior of any state variable depends on the others through the system model. Let us define a matrix $\boldsymbol{P}_{xx_{k|k}} \in \boldsymbol{R}^{nxn}$ in which the element $\boldsymbol{P}_{xx_{k|k}}(i,j)$ gives a measure of the dependency of the behavior of the state variable x_i on the state variable x_j. Intuitively, at $k = 0$; the initial conditions are either due to disturbances or as specified by any other means. Therefore, $x_i(0), x_j(0)$ are independent of each other and hence $P_{xx_{0|0}}(i,j) = 0$ for $i \neq j$. On the other hand, at $k = 0$; $P_{xx_{0|0}}(i,i) > 0$, which represents the possible range in which the initial value of x_i can vary around such a specified value. With the progress of time (i.e. $k > 0$) the dependency between the different states of the system exists and hence $\boldsymbol{P}_{xx_{k|k}}(i,j) \neq 0$.

Assuming that at the sampling instant k we know the filtered estimate $\hat{\boldsymbol{x}}_{k|k}$ and its associated dependency matrix $\boldsymbol{P}_{xx_{k|k}}$. As we receive the measurement $\boldsymbol{y}_{k+1}$, it is desired to get the filtered estimate $\hat{\boldsymbol{x}}_{k+1|k+1}$ and its associated dependency matrix $\boldsymbol{P}_{xx_{k+1|k+1}}$. Since for nonlinear systems the performance of the system changes dramatically with the starting point, the following approach is conducted to improve the estimation accuracy. We firstly predetermine a set of $2n + 1$ vectors (to be denoted by $\hat{\mathrm{X}}_{k|k}$) around the filtered estimate $\hat{\boldsymbol{x}}_{k|k}$ within the range specified by $\boldsymbol{P}_{xx_{k|k}}$. The i^{th} element of this set is denoted by $\hat{\boldsymbol{x}}_{ik|k}$. The system model is then used to calculate the predicted value $\hat{\boldsymbol{x}}_{ik+1|k}$ for each element $i \in \{1, 2, \ldots, 2n+1\}$ of the predetermined set $\hat{\mathrm{X}}_{k|k}$. Finally, the weighted average of the set of estimated predicted vectors is calculated to get the final value of the predicted estimate of the state vector $\hat{\boldsymbol{x}}_{k+1|k}$, and hence its associated dependency matrix $\boldsymbol{P}_{xx_{k+1|k}}$. Using such an approach, it is expected to get much better result for $\hat{\boldsymbol{x}}_{k+1|k}$ than that achieved while using (2-a).

The proposed estimator is implemented in two steps, namely the prediction step and the filter step. The details of each are given in the following:

1) PREDICTION STEP OF THE SRLS ESTIMATOR

1- The filtered estimate $\hat{\boldsymbol{x}}_{k|k}$ and its associated dependency matrix $\boldsymbol{P}_{xx_{k|k}}$ are used to generate the set $\hat{\mathrm{X}}_{k|k}$ of the $2n + 1$ predetermined vectors around the estimated value $\hat{\boldsymbol{x}}_{k|k}$. This is given by:

$$\begin{aligned}\hat{\mathrm{X}}_{k|k} &= \{\hat{\boldsymbol{x}}_{1k|k} \quad \hat{\boldsymbol{x}}_{2k|k} \quad \hat{\boldsymbol{x}}_{3k|k} \quad \cdots \quad \hat{\boldsymbol{x}}_{(2n+1)k|k}\} \\ &= \{\hat{\boldsymbol{x}}_{k|k} \quad \hat{\boldsymbol{x}}_{k|k}+\boldsymbol{\zeta}_1 \quad \hat{\boldsymbol{x}}_{k|k}-\boldsymbol{\zeta}_1 \quad \hat{\boldsymbol{x}}_{k|k}+\boldsymbol{\zeta}_2 \ldots \hat{\boldsymbol{x}}_{k|k}+\boldsymbol{\zeta}_n\}\end{aligned} \quad (8)$$

where $\boldsymbol{\zeta}_i$ is the i^{th} column $(i = 1, 2, \ldots, n)$ of the matrix $\mathbf{Z}$ defined by:

$$\mathbf{Z} = \sqrt{(n+\lambda)P_{xx_{k|k}}} \quad (9)$$

where λ is a scaling parameter to be specified later on.

The matrix $\mathbf{Z}$ can be computed using one of the efficient factorization techniques such as Cholesky method. Then:

$$\begin{aligned}\mathbf{Z}(\mathbf{Z})^T &= \sqrt{(n+\lambda)P_{xx_{k|k}}}\left(\sqrt{(n+\lambda)P_{xx_{k|k}}}\right)^T \\ &= (n+\lambda)P_{xx_{k|k}}\end{aligned} \quad (10)$$

2- By propagating the elements of the set $\hat{\mathrm{X}}_{k|k}$ through the function f_k defined by (1), we get:

$$\begin{aligned}\hat{\mathrm{X}}_{k+1|k} &= \left\{\hat{\boldsymbol{x}}_{1k+1|k} \quad \hat{\boldsymbol{x}}_{2k+1|k} \quad \hat{\boldsymbol{x}}_{3k+1|k} \quad \cdots \quad \hat{\boldsymbol{x}}_{(2n+1)k+1|k}\right\} \\ &= \left\{f_k(\hat{\boldsymbol{x}}_{k|k}) \quad f_k(\hat{\boldsymbol{x}}_{k|k}+\boldsymbol{\zeta}_1) \quad f_k(\hat{\boldsymbol{x}}_{k|k}-\zeta_1)\right. \\ &\quad \left. f_k(\hat{\boldsymbol{x}}_{k|k}+\boldsymbol{\zeta}_2) \quad f_k(\hat{\boldsymbol{x}}_{k|k}-\zeta_2)\ldots f_k(\hat{\boldsymbol{x}}_{k|k}-\zeta_n)\right\}\end{aligned} \quad (11)$$

where $\hat{\mathrm{X}}_{k+1|k}$ is the set of the $2n + 1$ predicted vectors corresponding to the set $\hat{\mathrm{X}}_{k|k}$ of the predetermined vectors.

3- The weighted average is then computed using the formula:

$$\hat{\boldsymbol{x}}_{k+1|k} = \sum_{j=1}^{2n+1} w_j \hat{\boldsymbol{x}}_{j_{k+1|k}} \quad (12)$$

where the weight w_j is given by:

$$w_j = \begin{cases} \dfrac{\lambda}{n+\lambda} & j = 1 \\ \dfrac{1}{2(n+\lambda)} & j \neq 1 \end{cases} \quad (13)$$

4- The dependency matrix $P_{xx_{k+1|k}}$ is calculated through the weighted sum of the squared variations of the elements of the set $\hat{\mathrm{X}}_{k+1|k}$ around the calculated weighted average $\hat{\boldsymbol{x}}_{k+1|k}$ while using the same weights as defined above. This leads to:

$$P_{xx_{k+1|k}} = \sum_{j=1}^{2n+1} w_j \left\{(\hat{\boldsymbol{x}}_{j_{k+1|k}} - \hat{\boldsymbol{x}}_{k+1|k})(\hat{\boldsymbol{x}}_{j_{k+1|k}} - \hat{\boldsymbol{x}}_{k+1|k})^T\right\} \quad (14)$$

5- The elements of the set $\hat{X}_{k+1|k}$ are then used to get the set of predicted outputs $\hat{Y}_{k+1|k}$ through the function h_{k+1}, as defined by (1). This is given by:

$$\begin{aligned}\hat{Y}_{k+1|k} &= \{\hat{\boldsymbol{y}}_{1k+1|k} \quad \hat{\boldsymbol{y}}_{2k+1|k} \quad \hat{\boldsymbol{y}}_{3k+1|k} \quad \cdots \quad \hat{\boldsymbol{y}}_{(2n+1)k+1|k}\} \\ &= \{h_{k+1}(\hat{\boldsymbol{x}}_{1k+1|k}) \quad h_{k+1}(\hat{\boldsymbol{x}}_{2k+1|k})\ldots h_{k+1}(\hat{\boldsymbol{x}}_{(2n+1)k+1|k})\}\end{aligned} \quad (15)$$

6- The weighted average of the predicted output is computed using the same weights as defined above. This leads to:

$$\hat{\boldsymbol{y}}_{k+1|k} = \sum_{j=1}^{2n+1} w_j \hat{\boldsymbol{y}}_{j_{k+1|k}} \tag{16}$$

7- Using these results, the dependency matrix $P_{yy_{k+1|k}}$ is calculated from the weighted sum of the squared variations of the elements of the set $\hat{\Upsilon}_{k+1|k}$ around their average $\hat{\boldsymbol{y}}_{k+1|k}$. Finally, the elements of the cross dependency matrix $P_{xy_{k+1|k}}$ are estimated from the weighted sum of the variations of the elements of $\hat{X}_{k+1|k}$ around their average $\hat{\boldsymbol{x}}_{k+1|k}$ multiplied by the variations of the elements of the set $\hat{\Upsilon}_{k+1|k}$ around their average $\hat{\boldsymbol{y}}_{k+1|k}$. As a result we get:

$$P_{yy_{k+1|k}} = \sum_{j=1}^{2n+1} w_j \left\{ (\hat{\boldsymbol{y}}_{j_{k+1|k}} - \hat{\boldsymbol{y}}_{k+1|k})(\hat{\boldsymbol{y}}_{j_{k+1|k}} - \hat{\boldsymbol{y}}_{k+1|k})^T \right\} \tag{17}$$

$$P_{xy_{k+1|k}} = \sum_{j=1}^{2n+1} w_j \left\{ (\hat{\boldsymbol{x}}_{j_{k+1|k}} - \hat{\boldsymbol{x}}_{k+1|k})(\hat{\boldsymbol{y}}_{j_{k+1|k}} - \hat{\boldsymbol{y}}_{k+1|k})^T \right\} \tag{18}$$

To avoid singularity, equation (17) is modified by adding a very small extra term ($V \simeq 10^{-6} I_m$) to ensure the existence of $P^{-1}_{yy_{k+1|k}}$. Therefore, (17) takes the form:

$$P_{yy_{k+1|k}} = V + \sum_{j=1}^{2n+1} w_j \left\{ (\hat{\boldsymbol{y}}_{j_{k+1|k}} - \hat{\boldsymbol{y}}_{k+1|k})(\hat{\boldsymbol{y}}_{j_{k+1|k}} - \hat{\boldsymbol{y}}_{k+1|k})^T \right\} \tag{19}$$

2) FILTERING STEP OF THE SRLS ESTIMATOR

The mathematical structure of the filtering stage of the RLS observer is used in this step. Therefore, we have:

$$K_{k+1} = P_{xy_{k+1|k}} P^{-1}_{yy_{k+1|k}} \tag{20a}$$
$$\hat{\boldsymbol{x}}_{k+1|k+1} = \hat{\boldsymbol{x}}_{k+1|k} + K_{k+1}(\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1|k}) \tag{20b}$$
$$P_{xx_{k+1|k+1}} = P_{xx_{k+1|k}} - K_{k+1} P_{yy_{k+1|k}} K^T_{k+1} \tag{20c}$$

C. COMPARISON BETWEEN THE PREDICTION STEP OF THE SRLS AND RLS OBSERVERS

Although the procedure used to calculate the predicted estimate of the state vector using the SRLS and the RLS observers are different, relationships between the results of these two estimators exist. The details of our analysis to achieve these relationships are presented in Appendix B.

By expanding (12) while using (11), getting Taylor series expansion of the expanded equation up to the second order approximation, and finally after simple mathematical manipulation, $\hat{\boldsymbol{x}}_{k+1|k}$ can be written in the form:

$$\hat{\boldsymbol{x}}_{k+1|k} = f_k(\hat{\boldsymbol{x}}_{k|k}) + \boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k}) + H.O.T \tag{21}$$

where:

$$\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k}) = \frac{1}{2} \left[\nabla^T P_{xx_{k|k}} \nabla \right] f_k(\boldsymbol{x}_k) \Big|_{\hat{\boldsymbol{x}}_{k|k}} \tag{22}$$

and ∇ donates for the gradient operator with respect to the vector x, and $H.O.T.$is the higher order terms.

Equation (21) is not more than the predicted estimate of the RLS estimator in addition to the extra terms $\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})$ and the $H.O.T$. As will be demonstrated later on, these extra terms will not only improve the results of the prediction step and hence the overall performance of the SRLS observer, but also improves its convergence behavior.

Similarly, by expanding (14) while using (11) and (21), then after simple algebraic manipulations, one gets:

$$P_{xx_{k+1|k}} = \hat{A}_k P_{xx_{k|k}} \hat{A}^T_k - \boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k}) \boldsymbol{D}^T_k(\hat{\boldsymbol{x}}_{k|k}) + H.O.T \tag{23}$$

Or:

$$P_{xx_{k+1|k}} = \hat{A}_k P_{xx_{k|k}} \hat{A}^T_k + \Phi_k(\hat{\boldsymbol{x}}_{k|k}) + H.O.T \tag{24}$$

where:

$$\Phi_k(\hat{\boldsymbol{x}}_{k|k}) = -\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k}) \boldsymbol{D}^T_k(\hat{\boldsymbol{x}}_{k|k}) \tag{25}$$

and $\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})$ is as defined by (22).

One can notice that (24) is equivalent to (2-c) of the RLS observer except for the constant matrix $N$is replaced by $\Phi_k(\hat{x}_{k|k})$, in addition to the $H.O.T.$

The similarities between (16), (17), and (18) and those of the RLS observer can be achieved by using the same procedure as presented in Appendix B. This leads to following results:

$$\hat{\boldsymbol{y}}_{k+1|k} = h_{k+1}(\hat{\boldsymbol{x}}_{k+1|k}) + H.O.T \tag{26a}$$
$$P_{yy_{k+1|k}} = \hat{H}_{k+1} P_{xx_{k+1|k}} \hat{H}^T_{k+1} + \Psi_k(\hat{\boldsymbol{x}}_{k|k}) + H.O.T \tag{26b}$$
$$P_{xy_{k+1|k}} = P_{xx_{k+1|k}} \hat{H}^T_{k+1} + \Omega_k(\hat{\boldsymbol{x}}_{k|k}) + H.O.T \tag{26c}$$

where $\hat{H}_{k+1}$ is as given by (7), while:

$$\begin{aligned} \Psi_k(\hat{\boldsymbol{x}}_{k|k}) &= \hat{H}_{k+1} \Phi_k(\hat{\boldsymbol{x}}_{k|k}) \hat{H}^T_{k+1} + H.O.T \\ \Omega_k(\hat{\boldsymbol{x}}_{k|k}) &= \Phi_k(\hat{\boldsymbol{x}}_{k|k}) \hat{H}^T_{k+1} + H.O.T \end{aligned} \tag{27}$$

From this analysis, although there are similarities between the outputs of the prediction step resulting from the SRLS and the RLS observers, two main factors highly improve the performance of the SRLS estimator. Firstly, the computation of the Jacobian matrices, especially that of f_k, is not required to get (14), (17), and (18). Secondly, higher order terms in Taylor series expansion, which are not included in the RLS observer, will obviously improve the estimation results of the SRLS observer.

From our detailed analysis of $P_{xx_{k+1|k}}$, $P_{yy_{k+1|k}}$ and $P_{xy_{k+1|k}}$, it has been noticed that higher-order terms in $\Phi_k(\hat{\boldsymbol{x}}_k)$, $\Psi_k(\hat{\boldsymbol{x}}_{k|k})$, $\Omega_k(\hat{\boldsymbol{x}}_{k|k})$ are multiplied by $(n+\lambda)^i$ where the exponent i increases as the order increases. Therefore, if $(n+\lambda) > 1$, higher order terms will increase and lead to misleading results. Therefore, to avoid this problem, the parameter λ is chosen such as:

$$\lambda = -n + \alpha^2 n \tag{28}$$

where α is an additional scaling parameter to be chosen such that $(n+\lambda) << 1$ (e.g. $\alpha \simeq 10^{-3}$).

Moreover, the terms $\Phi_k(\hat{\boldsymbol{x}}_k)$, $\Psi_k(\hat{\boldsymbol{x}}_{k|k})$, $\Omega_k(\hat{\boldsymbol{x}}_{k|k})$ in (24) and (26) contain higher-order terms of Taylor series expansion of f_k which can affect the state estimation results. Therefore, the weighting factor w_1 in the summation of the squared variations around the mean can be modified to the form $w_1 = \frac{\lambda}{n+\lambda} + (1 + \beta - \alpha^2)$, where β is another scaling parameter which permits the control of the higher order weights. A reasonable value of this parameter is 2. Therefore, (25) is modified to take the form:

$$\Phi_k(\hat{\boldsymbol{x}}_k) = -(\beta - \alpha^2)\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})\boldsymbol{D}_k^T(\hat{\boldsymbol{x}}_{k|k}) \tag{29}$$

V. CONVERGENCE ANALYSIS

This section is devoted to the convergence analysis of the proposed algorithm. In order to facilitate the follow-up of our analysis, the procedure to be used is summarized in the following steps:

Step 1: The error equation of the estimated state vector resulting from the SRLS observer is firstly derived.

Step 2: After getting the relationship between the estimation errors resulting from the SRLS and the RLS estimators, we transform the equation derived in step 1 to that of the RLS estimator.

Step 3: We prove the convergence of the estimation error resulting from the RLS estimator to the desired zero steady state.

Step 4: Through the relation derived in step 2, we show that the estimation error of the SRLS estimator converges to the desired zero steady state faster than that resulting from the RLS estimator.

Before starting, the system described by (1) is assumed to fulfill the following assumptions:

Assumption 1: The system (1) is affine on the neighborhoods of $\boldsymbol{x}_k, \hat{\boldsymbol{x}}_{k|k}$ and $\boldsymbol{x}_{k+1}, \hat{\boldsymbol{x}}_{k+1|k}$.

Assumption 2: The pair of matrices $\hat{A}_{ss} = \left.\frac{\partial f_k^T}{\partial \boldsymbol{x}_k}\right|_{\hat{\boldsymbol{x}}_{ss}}^T$, $\hat{H}_{ss} = \left.\frac{\partial h_{k+1}^T}{\partial \boldsymbol{x}_{k+1}}\right|_{\hat{\boldsymbol{x}}_{ss}}^T$ is observable at the equilibrium point $\boldsymbol{x}_{ss}$. Moreover, the system (1) is locally observable at the equilibrium point $\boldsymbol{x}_{ss}$, i.e. there exists an entire non-empty domain $\Omega(\boldsymbol{x}_{ss}) \subseteq \chi_r(\boldsymbol{x}_{ss})$ in the neighborhood of the equilibrium point $\boldsymbol{x}_{ss}$ such that for every $\boldsymbol{x}_k$ in that neighborhood other than $\boldsymbol{x}_{ss}$ the system is distinguishable from $\boldsymbol{x}_{ss}$.

Assumption 3: The vector functions f_k, h_{k+1}; the matrices $A_k = \left(\frac{\partial f_k^T}{\partial \boldsymbol{x}_k}\right)^T$, $H_{k+1} = \left(\frac{\partial h_{k+1}^T}{\partial \boldsymbol{x}_{k+1}}\right)^T$ and their estimates are bounded in $\boldsymbol{x}_k, \hat{\boldsymbol{x}}_{k|k}$ for all $\|\boldsymbol{x}_k\| \leq \bar{r}$ and $\left\|\hat{\boldsymbol{x}}_{k|k}\right\| \leq \bar{r}$ where $\bar{r}$ is a positive scalar.

Assumption 4: The pairs of matrices (A_k, H_{k+1}) and their estimates are observable or detectable for all $\boldsymbol{x}_k \in \Omega$ and $\hat{\boldsymbol{x}}_{k|k} \in \Omega$ except for a finite number of points.

Step 1 (State Estimation Error of the SRLS Estimator): As the states and the estimated states converge to $\boldsymbol{x}_{ss}$, then $\hat{A}_k, \hat{H}_{k+1}$ converge to A_{ss}, H_{ss} respectively. Let:

$$\hat{A}_{\mathrm{k}} = A_{ss} + \Delta\hat{A}_{\mathrm{k}}, \quad \hat{H}_{\mathrm{k}+1} = H_{ss} + \Delta\hat{H}_{\mathrm{k}+1}, \quad K_{k+1} = K_{ss} + \Delta K_{k+1} \tag{30}$$

where K_{ss} represents the gain matrix at the equilibrium. Now, we define the following:

$$\boldsymbol{x}_k = \hat{\boldsymbol{x}}_{k|k} + \tilde{\boldsymbol{x}}_{k|k} \tag{31}$$
$$\boldsymbol{x}_{k+1} = \hat{\boldsymbol{x}}_{k+1|k} + \tilde{\boldsymbol{x}}_{k+1|k} \tag{32}$$
$$\boldsymbol{x}_{k+1} = \hat{\boldsymbol{x}}_{k+1|k+1} + \tilde{\boldsymbol{x}}_{k+1|k+1} \tag{33}$$

Substituting from (31) into (1), and according to assumption 1, f_k can be expanded using Taylor series to the first order approximation. Therefore, we get:

$$\boldsymbol{x}_{k+1} = f_k\left(\hat{\boldsymbol{x}}_{k|k} + \tilde{\boldsymbol{x}}_{k|k}\right) \tag{34}$$
$$\boldsymbol{x}_{k+1} = f_k\left(\hat{\boldsymbol{x}}_{k|k}\right) + \hat{A}_k\tilde{\boldsymbol{x}}_{k|k} \tag{35}$$

Using (21), then (34) takes the form:

$$\boldsymbol{x}_{k+1} = \hat{\boldsymbol{x}}_{k+1|k} - \boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k}) + \hat{A}_k\tilde{\boldsymbol{x}}_{k|k} \tag{36}$$

From which:

$$\tilde{\boldsymbol{x}}_{k+1|k} = \boldsymbol{x}_{k+1} - \hat{\boldsymbol{x}}_{k+1|k} = \hat{A}_k\tilde{\boldsymbol{x}}_{k|k} - \boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k}) \tag{37}$$

Again, according to assumption 1, we can expand h_{k+1} using Taylor series to the first order approximation. Hence, the output $\boldsymbol{y}_{k+1}$ while using (32) is such that:

$$\boldsymbol{y}_{k+1} = h_{k+1}\left(\hat{\boldsymbol{x}}_{k+1|k} + \tilde{\boldsymbol{x}}_{k+1|k}\right) \tag{38}$$
$$\boldsymbol{y}_{k+1} = h_{k+1}\left(\hat{\boldsymbol{x}}_{k+1|k}\right) + \hat{H}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k} \tag{39}$$

Using (26-a) while ignoring the $H.O.T$, then (39) is such that:

$$\boldsymbol{y}_{k+1} = \hat{\boldsymbol{y}}_{k+1|k} + \hat{H}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k} \tag{40}$$
$$\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1|k} = \hat{H}_{k+1}\tilde{\boldsymbol{x}}_{k+1|k} \tag{41}$$

Substituting from (20-b) into (33) we get:

$$\begin{aligned}\tilde{\boldsymbol{x}}_{k+1|k+1} &= \boldsymbol{x}_{k+1} - \hat{\boldsymbol{x}}_{k+1|k+1} \\ &= \boldsymbol{x}_{k+1} - \left[\hat{\boldsymbol{x}}_{k+1|k} + K_{k+1}\left(\boldsymbol{y}_{k+1} - \hat{\boldsymbol{y}}_{k+1|k}\right)\right]\end{aligned} \tag{42}$$

Using (36) and (40) into (42), one gets:

$$\tilde{\boldsymbol{x}}_{k+1|k+1} = \hat{A}_{\mathrm{k}}\tilde{\boldsymbol{x}}_{k|k} - \boldsymbol{D}_{\mathrm{k}}(\hat{\boldsymbol{x}}_{k|k}) - K_{k+1}\hat{H}_{\mathrm{k}+1}\tilde{\boldsymbol{x}}_{k+1|k} \tag{43}$$

Again, substituting from (36) into (43), we have:

$$\begin{aligned}\tilde{\boldsymbol{x}}_{k+1|k+1} &= \hat{A}_k\tilde{\boldsymbol{x}}_{k|k} - \boldsymbol{D}_{\mathrm{k}}(\hat{\boldsymbol{x}}_{k|k}) \\ &\quad - K_{k+1}\hat{H}_{\mathrm{k}+1}\left[\hat{A}_k\tilde{\boldsymbol{x}}_{k|k} - \boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})\right] \\ \tilde{\boldsymbol{x}}_{k+1|k+1} &= \left[I - K_{k+1}\hat{H}_{\mathrm{k}+1}\right]\left[\hat{A}_k\tilde{\boldsymbol{x}}_{k|k} - \boldsymbol{D}_{\mathrm{k}}(\hat{\boldsymbol{x}}_{k|k})\right]\end{aligned} \tag{44}$$

Now using (30), equation (44) can be written in the form:

$$\begin{aligned}\tilde{\boldsymbol{x}}_{k+1|k+1} = &\left[I - (K_{ss} + \Delta K_{k+1})\left(H_{ss} + \Delta\hat{H}_{\mathrm{k}+1}\right)\right]^* \\ &\left[\left(A_{ss} + \Delta\hat{A}_{\mathrm{k}}\right)\tilde{\boldsymbol{x}}_{k|k} - \boldsymbol{D}_{\mathrm{k}}(\hat{\boldsymbol{x}}_{k|k})\right]\end{aligned} \tag{45}$$

Collecting terms and simplifying:

$$\tilde{\boldsymbol{x}}_{k+1|k+1} = \phi_{ss}\tilde{\boldsymbol{x}}_{k|k} + \psi_k(\tilde{\boldsymbol{x}}_{k|k})\tilde{\boldsymbol{x}}_{k|k} - \mathbf{E}_{\mathrm{k}} \tag{46}$$

where:

$$\begin{aligned}\phi_{ss} &= (I - K_{ss}H_{ss})A_{ss} \\ \psi_k(\hat{\boldsymbol{x}}_{k|k}) &= -\left(K_{ss}\Delta\hat{H}_{\mathrm{k}+1} + \Delta K_{k+1}\hat{H}_{\mathrm{k}+1}\right)A_{ss} \\ &\quad + \left(I - K_{k+1}\hat{H}_{\mathrm{k}+1}\right)\Delta\hat{A}_{\mathrm{k}} \\ \mathbf{E}_{\mathrm{k}} &= \left(I - K_{k+1}\hat{H}_{\mathrm{k}+1}\right)\boldsymbol{D}_{\mathrm{k}}(\hat{\boldsymbol{x}}_{k|k})\end{aligned} \tag{47}$$

Step 2 (Transformation of the Estimation Error to That of the RLS Estimator):

Now define:

$$\tilde{x}_{k|k_{RLS}} = \tilde{x}_{k|k_{SRLS}} + \Delta\tilde{x}_{k|k} \tag{48}$$

where $\tilde{x}_{k|k_{RLS}}, \tilde{x}_{k|k_{SRLS}}$ are, respectively, the errors of the RLS and SRLS observers.

Substituting from (48) into (46), and assuming that $\psi_k(\tilde{x}_{k|k})$ is affine on a neighborhood of $\tilde{x}_{k|k_{RLS}}, \tilde{x}_{k|k_{SRLS}}$. Then, by expanding ψ_k using Taylor series to the first order approximation, we get:

$$\tilde{x}_{k+1|k+1_{SRLS}} = \phi_{ss}\left[\tilde{x}_{k|k_{RLS}} - \Delta\tilde{x}_{k|k}\right] + \psi_k(\tilde{x}_{k|k_{RLS}} - \Delta\tilde{x}_{k|k})\left[\tilde{x}_{k|k_{RLS}} - \Delta\tilde{x}_{k|k}\right] - \mathrm{E_k} \tag{49}$$

OR

$$\begin{aligned}\tilde{x}_{k+1|k+1_{SRLS}} &= \phi_{ss}\tilde{x}_{k|k_{RLS}} - \phi_{ss}\Delta\tilde{x}_{k|k} \\ &+ \left[\psi_k(\tilde{x}_{k|k_{RLS}}) - \hat{\Psi}_k\Delta\tilde{x}_{k|k}\right] \\ &\times \left[\tilde{x}_{k|k_{RLS}} - \Delta\tilde{x}_{k|k}\right] - \mathrm{E_k}\end{aligned} \tag{50}$$

where $\hat{\Psi}_k = \frac{\partial\psi_k^T}{\partial x_k}\Big|_{\tilde{x}_{k|k_{RLS}}}^T$

Equation (50) can also be rewritten as:

$$\begin{aligned}&\tilde{x}_{k+1|k+1_{SRLS}} + \phi_{ss}\Delta\tilde{x}_{k|k} + \psi_k(\tilde{x}_{k|k_{RLS}})\Delta\tilde{x}_{k|k} \\ &\quad + \hat{\Psi}_k\Delta\tilde{x}_{k|k}\left[\tilde{x}_{k|k_{RLS}} - \Delta\tilde{x}_{k|k}\right] + \mathrm{E_k} \\ &= \phi_{ss}\tilde{x}_{k|k_{RLS}} + \psi_k(\tilde{x}_{k|k_{RLS}})\tilde{x}_{k|k_{RLS}}\end{aligned} \tag{51}$$

Since the right hand side (R.H.S) of (51) is the estimation error resulting from the RLS estimator, then the left hand side (L.H.S) is obviously $\tilde{x}_{k+1|k+1_{RLS}}$. Therefore, the relation between the estimation errors resulting from the two observers is such that:

$$\begin{aligned}&\tilde{x}_{k+1|k+1_{SRLS}} + \phi_{ss}\Delta\tilde{x}_{k|k} + \psi_k(\tilde{x}_{k|k_{RLS}})\Delta\tilde{x}_{k|k} \\ &\quad + \hat{\Psi}_k\Delta\tilde{x}_{k|k}\left[\tilde{x}_{k|k_{RLS}} - \Delta\tilde{x}_{k|k}\right] + \mathrm{E_k} = \tilde{x}_{k+1|k+1_{RLS}}\end{aligned} \tag{52}$$

Using (52), we can write (51) as:

$$\tilde{x}_{k+1|k+1_{RLS}} = \phi_{ss}\tilde{x}_{k|k_{RLS}} + \psi_k(\tilde{x}_{k|k_{RLS}})\tilde{x}_{k|k_{RLS}} \tag{53}$$

Step 3 (Convergence Analysis of Estimation Error While Using RLS Estimator):

Let: $\mathbf{Z}_{k+1} = \tilde{x}_{k+1|k+1_{RLS}}, \mathbf{Z}_0 = \tilde{x}(0)$, then (53) can be rewritten in the form:

$$\mathbf{Z}_{k+1} = \phi_{ss}^{k+1}\mathbf{Z}_0 + \sum_{j=0}^{k}\phi_{ss}^{k-j}\psi_j\mathbf{Z}_j \tag{54}$$

Theorem 1: If the solutions of the homogeneous equation:

$$\mathbf{Z}_{k+1} = \phi_{ss}\mathbf{Z}_k \quad \text{for } k \geq 0 \tag{55}$$

remain bounded as $k \to \infty$, then:

a) The same is true for all the solutions of the homogeneous system:

$$\mathbf{Z}_{k+1} = \phi_k\mathbf{Z}_k \quad \text{for } k \geq 0 \text{ where } \phi_k = \phi_{ss} + \psi_k \tag{56}$$

provided that:

$$\sum_{k=0}^{\infty}\|\psi_k\| < \infty \tag{57}$$

b) $\lim_{k\to\infty}\|\mathbf{Z}_{k+1}\| = 0$, and the system (56) is locally exponentially stable.

Proof of Theorem 1:

a) Equation (55) can be rewritten as:

$$\mathbf{Z}_{k+1} = \phi_{ss}^{k+1}\mathbf{Z}_0 \tag{58}$$

As a result of the assumption that the solution of (58) remains bounded as $k \to \infty$, there exist a constants c_1 such that for $k+1 \geq 0$:

$$\left\|\phi_{ss}^{k+1}\right\| \leq c_1 \tag{59}$$

Since $\phi_k = \phi_{ss} + \psi_k$, then (56) can be written in the form:

$$\mathbf{Z}_{k+1} = \phi_{ss}^{k+1}\mathbf{Z}_0 + \sum_{j=0}^{k}\phi_{ss}^{k-j}\psi_j\mathbf{Z}_j \tag{60}$$

Taking the norm of (60), we get:

$$\|\mathbf{Z}_{k+1}\| \leq \left\|\phi_{ss}^{k+1}\right\|\|\mathbf{Z}_0\| + \sum_{j=0}^{k}\left\|\phi_{ss}^{k-j}\right\|\left\|\psi_j\right\|\left\|\mathbf{Z}_j\right\| \tag{61}$$

Substituting from (59) into (61), one gets:

$$\|\mathbf{Z}_{k+1}\| \leq c_1\|\mathbf{Z}_0\| + \sum_{j=0}^{k}c_1\left\|\psi_j\right\|\left\|\mathbf{Z}_j\right\| \tag{62}$$

Let $g_j = c_1\left\|\psi_j\right\|$, then we have:

$$\|\mathbf{Z}_{k+1}\| \leq c_1\|\mathbf{Z}_0\| + \sum_{j=0}^{k}g_j\left\|\mathbf{Z}_j\right\| \tag{63}$$

From discrete-time Gronwall lemma [33], we have:

$$\|\mathbf{Z}_{k+1}\| \leq c_1\|\mathbf{Z}_0\|\prod_{j=0}^{k}(1+g_j) \leq c_1\|\mathbf{Z}_0\|\exp\left(\sum_{j=0}^{k}g_j\right) \tag{64}$$

Using assumption (57), there exists a positive constant c_2 such that $\sum_{k=0}^{\infty}\left\|\psi_j\right\| \leq c_2$. Therefore, we have:

$$\|\mathbf{Z}_{k+1}\| \leq c_1\|\mathbf{Z}_0\|\exp(c_1c_2) \tag{65}$$

Thus, the solution of the homogeneous system (56) is bounded.

Since the homogeneous system is stable at the equilibrium point x_{ss}, and according to assumption 4, the pair $(\hat{A}_k, \hat{H}_{k+1})$ is observable or detectable for all $\hat{x}_{k|k} \in \Omega$ except for a finite number of points; then there exist $0 \leq \varepsilon(k+1) < 1$ such that $\left\|\phi_{ss}^{k+1}\right\| = \varepsilon(k+1)^{k+1}$ for $k+1 \geq K$.

From Gronwall lemma while using (57), equation (65) can be written as:

$$\|\mathbf{Z}_{k+1}\| \leq \varepsilon(k+1)^{k+1}\|\mathbf{Z}_0\|\exp(c_1c_2) \leq \bar{\varepsilon}^{k+1}\|\mathbf{Z}_0\|\exp(c_1c_2) \quad (66)$$

where $\bar{\varepsilon} = \sup\{\varepsilon(j), j \in (K, K+1, \ldots, \infty)\}$ and $0 \leq \bar{\varepsilon} < 1$. Therefore, $\lim_{k\to\infty} \|\mathbf{Z}_{k+1}\| = 0$ as $\lim_{k\to\infty} \bar{\varepsilon}^{k+1} = 0$, and hence the system is locally exponentially stable.

Step 4 (Comparison Between the Speed of Convergence of the RLS and SRLS Estimators):

From (62), since the term:

$$\{\phi_{ss}\Delta\tilde{\boldsymbol{x}}_{k|k} + \psi_k(\tilde{\boldsymbol{x}}_{k|k_{RLS}})\Delta\tilde{\boldsymbol{x}}_{k|k} + \hat{\Psi}_k\Delta\tilde{\boldsymbol{x}}_{k|k}\left[\tilde{\boldsymbol{x}}_{k|k_{RLS}} - \Delta\tilde{\boldsymbol{x}}_{k|k}\right] + \mathbf{E}_k\}$$

is a vector with a finite value, it can be written in the form:

$$\phi_{ss}\Delta\tilde{\boldsymbol{x}}_{k|k} + \psi_k(\tilde{\boldsymbol{x}}_{k|k_{RLS}})\Delta\tilde{\boldsymbol{x}}_{k|k} + \hat{\Psi}_k\Delta\tilde{\boldsymbol{x}}_{k|k}\left[\tilde{\boldsymbol{x}}_{k|k_{RLS}} - \Delta\tilde{\boldsymbol{x}}_{k|k}\right] + \mathbf{E}_k = \theta_k\tilde{\boldsymbol{x}}_{k+1|k+1_{SRLS}} \quad (67)$$

where $\theta_k \in R^{nxn}$ is a matrix with adjustable parameters to satisfy (67).

As a result, (51) can be written as:

$$(\mathrm{I} + \theta_k)\tilde{\boldsymbol{x}}_{k+1|k+1_{SRLS}} = \phi_{ss}\tilde{\boldsymbol{x}}_{k|k_{RLS}} + \psi_k(\tilde{\boldsymbol{x}}_{k|k_{RLS}})\tilde{\boldsymbol{x}}_{k|k_{RLS}} \quad (68)$$

Let $\left\|(\mathrm{I}+\theta_k)\tilde{\boldsymbol{x}}_{k+1|k+1_{SRLS}}\right\| = (1+\mu_k)\left\|\tilde{\boldsymbol{x}}_{k+1|k+1_{SRLS}}\right\|$, then the norm of (68) is given by:

$$\left\|\tilde{\boldsymbol{x}}_{k+1|k+1_{SRLS}}\right\| \leq \frac{1}{(1+\mu_k)}\{\|\phi_{ss}\|\left\|\tilde{\boldsymbol{x}}_{k|k_{RLS}}\right\| + \left\|\psi_k(\tilde{\boldsymbol{x}}_{k|k_{RLS}})\right\|\left\|\tilde{\boldsymbol{x}}_{k|k_{RLS}}\right\|\} \quad (69)$$

One can notice that the R.H.S of (69) is that of the RLS observer divided by $(1+\mu_k)$. This means that the convergence rate of the SRLS is a scaled version of the RLS observer with a scaling factor $\frac{1}{1+\mu_k} < 1$. As it is made clear from (67), such a scaling factor is due to the extra term $\theta_k\tilde{\boldsymbol{x}}_{k+1|k+1_{SRLS}}$ resulting from $\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})$ in (21). Since $\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})$ represents the higher order terms in Taylor series expansion of f_k, then, as expected, the convergence of $\tilde{\boldsymbol{x}}_{k+1|k+1_{SRLS}}$ to the desired zero steady state value will be faster than that of $\tilde{\boldsymbol{x}}_{k+1|k+1_{RLS}}$.

VI. SIMULATION EXAMPLES

In this section, two illustrative examples of highly nonlinear power systems are presented to show the effectiveness of the developed SRLS estimator in handling discrete-time nonlinear estimation problems. For the sake of comparison, the states of the two systems are also estimated using the HG, the SDRE, and the RLS observers.

A. EXAMPLE I: SYNCHRONOUS GENERATOR

Consider the system of an uncontrolled synchronous generator connected to an infinite bus through a transmission line. The model of the system is given by [34], (70), as shown at the top of the next page, where $x_1 = \delta$ is the angular position (rad); $x_2 = \Delta\omega = \omega - \omega_0$ is the change in angular speed $\omega(rad/s)$ from its nominal value $\omega_o(rad/s)$; $x_3 = E'_q$ is the q-axis transient EMF of the armature; $x_4 = E_f$ is the field voltage; ΔT is the sampling rate; $\bar{P}$ is the mechanical power; x'_d is the generator direct-axis transient reactance; x_d, x_q are the direct and quadrature-axis current; x_e is the transmission line reactance; T'_{d_0} is the d-axis open circuit field time constant (sec); H is the inertia constant (sec); K_E is the exciter gain constant; T_E is the exciter time constant (sec); ν is the infinite bus voltage; and ν_{ref} is the reference voltage. The values of the parameters are given in Table 1.

TABLE 1. System parameters of example I.

Par.	Val.	Par.	Val.	Par.	Val.
x_d	1.6	ν	1	K_E	25
x'_d	0.32	ω_0	$100\pi\ rad/s$	T_E	0.05 sec
x_q	1.55	T'_{d_0}	6 sec	x_e	0.4
H	5	ν_{ref}	1.2553	ΔT	0.001 sec

1) THE HG OBSERVER

To be able to apply this observer, the system should be transformed to the form [35]:

$$F_k(\boldsymbol{x}_k) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & c_1c_2\dfrac{x_q - x_e}{x'_d + x_e} & 0 \\ 0 & 0 & 0 & \dfrac{1}{T_E} \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}^T \quad (71a)$$

$$\Lambda(\boldsymbol{x}_k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c_1c_2\dfrac{x_q - x_e}{x'_d + x_e} & 0 \\ 0 & 0 & 0 & \dfrac{c_1c_2}{T_E}\dfrac{x_q - x_e}{x'_d + x_e} \end{bmatrix}, \quad (71b)$$

$$S^{-1}C^T = \begin{bmatrix} 4 & 6 & 4 & 1 \end{bmatrix}^T$$

$$\Delta_\theta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \dfrac{1}{\theta} & 0 & 0 \\ 0 & 0 & \dfrac{1}{\theta^2} & 0 \\ 0 & 0 & 0 & \dfrac{1}{\theta^3} \end{bmatrix} \quad (71c)$$

$$
\begin{aligned}
x_{1_{k+1}} &= x_{1_k} + \Delta T \left(x_{2_k}\right) \\
x_{2_{k+1}} &= x_{2_k} + \Delta T \frac{\omega_o}{2H}\left(\bar{P} - \frac{v \sin x_{1_k}}{x_q + x_e} x_{3_k}\right) \\
&\quad - \Delta T \frac{\omega_o}{2H}\left(\frac{v \sin x_{1_k}}{x_q + x_e}\left[x_q \frac{x_{3_k} - v\cos x_{1_k}}{x_d' + x_e} - x_d' \frac{x_{3_k} - v\cos x_{1_k}}{x_d' + x_e}\right]\right) \\
x_{3_{k+1}} &= x_{3_k} + \Delta T \frac{1}{T_{do}'}\left(x_{4_k} - x_{3_k} - x_d \frac{x_{3_k} - v\cos x_{1_k}}{x_d' + x_e}\right) \\
x_{4_{k+1}} &= x_{4_k} + \Delta T \frac{1}{T_E}\left(-x_{4_k} + K_E v_{ref}\right) \\
&\quad - \Delta T \frac{1}{T_E}\left(K_E \sqrt{\left[\frac{x_q v \sin(x_{1_k})}{x_q + x_e}\right]^2 + \left[x_{3_k} - x_d' \frac{x_{3_k} - v\cos x_{1_k}}{x_d' + x_e}\right]^2}\right) \\
y_{k+1} &= x_{1_{k+1}}
\end{aligned} \tag{70}
$$

$$
\begin{aligned}
a_{21} &= \frac{\Delta T}{x_{1_k}} c_1 \left(\bar{P} - c_2\left[x_q - x_d'\right]\frac{x_{3_k} - v\cos x_{1_k}}{x_d' + x_e}\right) \\
a_{31} &= \frac{\Delta T}{x_{1_k}}\left(\frac{\left[x_d' - x_d\right]}{T_{do}'}\right)\left(\frac{x_{3_k} - v\cos x_{1_k}}{x_d' + x_e}\right) \\
a_{41} &= \left\{\frac{\Delta T\,(K_E)}{T_E x_{1_k}} v_{ref}\right. \\
&\quad \left. - \frac{\Delta T\,(K_E)}{T_E x_{1_k}}\sqrt{\left[\frac{x_q v \sin(x_{1_k})}{x_q + x_e}\right]^2 + \left[x_{3_k} - x_d' \frac{x_{3_k} - v\cos x_{1_k}}{x_d' + x_e}\right]^2}\right\} \\
C_k(\boldsymbol{x}_k) &= \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}
\end{aligned} \tag{73}
$$

where θ is chosen to be equal to 2, and:

$$
c_1 = \frac{\omega_o}{2H}, \quad c_2 = \frac{v \sin(\hat{x}_{1_k})}{x_q + x_e} \tag{72}
$$

2) THE SDRE OBSERVER

By writing the system in the state dependent linear form [36], we get:

$$
A_k(\boldsymbol{x}_k) = \begin{bmatrix} 1 & \Delta T & 0 & 0 \\ a_{21} & 1 & -\Delta T\,(c_1 c_2) & 0 \\ a_{31} & 0 & 1 - \frac{\Delta T}{T_{d_0}'} & \frac{\Delta T}{T_{d_0}'} \\ a_{41} & 0 & 0 & 1 - \frac{\Delta T}{T_E} \end{bmatrix}
$$

where, (73) as shown at the top of this page. The parameters c_1, c_2 are as given by (72). The weighting matrices of the dual infinite regulator problem, Q; R, are chosen as follows:

$$
\begin{aligned}
Q &= diag\,([0.001; 0.001; 0.001; 0.001]) \\
R &= 0.001
\end{aligned} \tag{74}
$$

3) THE RLS OBSERVER

The Jacobian matrix A_k of the system is given by:

$$
A_k = \begin{bmatrix} 1 & \Delta T & 0 & 0 \\ -\Delta T\,(c_1 d_1) & 1 & -\Delta T\,(c_1 c_2)(1 + c_3) & 0 \\ -\Delta T(c_4 v \sin x_{1_k}) & 0 & 1 - \frac{\Delta T}{T_{d_0}}(1 + c_4) & \frac{\Delta T}{T_{d_0}} \\ -\Delta T \frac{K_E d_2}{T_E} & 0 & -\Delta T \frac{K_E}{T_E} d_3 & 1 - \frac{\Delta T}{T_E} \end{bmatrix}
$$

$$
C_{k+1} = \begin{bmatrix} 1 \; 0 \; 0 \; 0 \end{bmatrix} \tag{75}
$$

where c_1, c_2 are as given by (72).

$$
\begin{aligned}
c_3 &= \frac{x_q - x_d'}{x_d' + x_e}, \quad c_4 = \frac{x_d - x_d'}{x_d' + x_e}, \quad c_5 = \frac{x_q v}{x_q + x_e}, \\
c_6 &= \frac{x_d'}{x_d' + x_e}, \\
c_7 &= \sqrt{\left[c_5 \sin x_{1_k}\right]^2 + \left[x_{3_k} - c_6\left(x_{3_k} - v\cos x_{1_k}\right)\right]^2} \\
d_1 &= \frac{v \cos x_{1_k}}{x_q + x_e}\left[(1 + c_3) x_{3_k} - c_3 v \cos x_{1_k}\right] \\
&\quad + c_3 v \sin x_{1_k}\left[\frac{v\cos x_{1_k}}{x_q + x_e}\right]
\end{aligned}
$$

$$d_2 = \left[\left(c_5^2 - v^2 c_6^2\right)\cos x_{1_k} - vc_6 x_{3_k}(1-c_6)\right]\sin x_{1_k}/c_7$$
$$d_3 = \left[c_6 x_{3_k} - x_{3_k} - vc_6 \cos x_{1_k}\right][c_6 - 1]/c_7 \quad (76)$$

The matrices $P_{0|0}; N; S$, as given in (2-c) and (6), are chosen as follows:

$$P_{0|0} = diag([1;\ 1;\ 1;\ 1])$$
$$N = diag([0.001;\ 0.001;\ 0.001;\ 0.001])$$
$$S = 0.001 \quad (77)$$

4) THE SRLS OBSERVER

This observer does not need any state transformation or the computation of the Jacobian matrix. The algorithm is directly applied while using the matrix $P_{xx_{0|0}} = diag([1;\ 1;\ 1;\ 1])$.

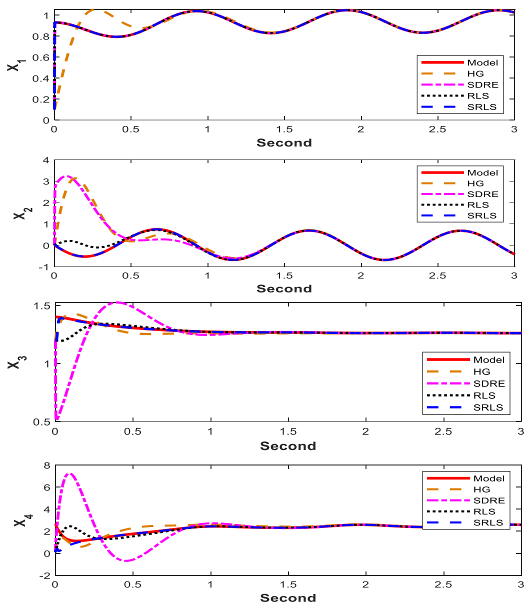

FIGURE 1. Estimated states for the synchronous generator in example I.

5) SIMULATION RESULTS OF EXAMPLE I

The initial condition of the state vector and the estimator are, respectively, given by $\boldsymbol{x}_0 = [0.9309 \quad 1.4001 \quad 2.6907]^T$ and $\hat{\boldsymbol{x}}_{0|0} = \left[0.1 \ \ 0 \ \ 1.2 \ \ 0.1\right]^T$. Fig. 1 shows the estimated states of the system using the four estimation approaches, whereas Fig. 2 shows their corresponding absolute state estimation errors (taken as the absolute difference between the estimated and actual states). For each observer, the maximum overshoot, the settling time of the estimation error of each state, and the average CPU time are presented in Table 2. The reported settling times are taken at the instant the error responses reaches and remain within ±0.02.

From the demonstrated results, it is clear that the SRLS observer leads to the shortest settling time with the least overshoot when compared to the other three observers. More preciously, and as shown in Table 2, the settling times of the estimation errors resulting from the SRLS observer (except for the first state) is less than 8% of the HG observer, 12% of the SDRE observer, and 22% of the RLS observer. For the first state, the settling time of the RLS and the SRLS observers are the same and 25% less than that of the SDRE observer and 0.17% less than the HG observer. The maximum error overshoot of the SRLS observer is less than 5% of the

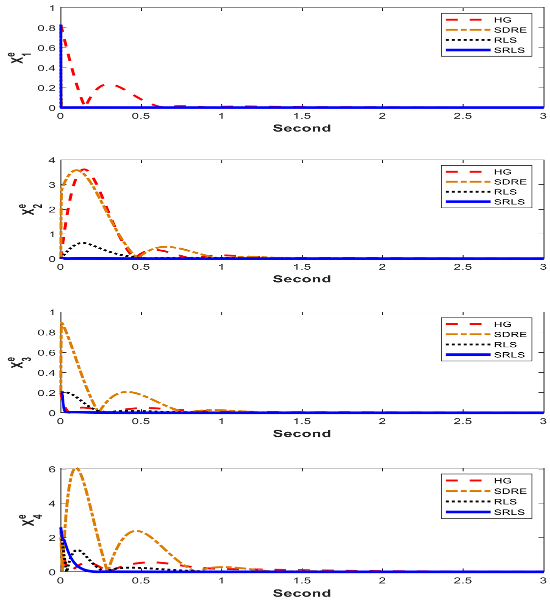

FIGURE 2. Absolute estimation error in the estimated states of example I.

TABLE 2. Comparison between performance the observers in example I*.

	State	HG	SDRE	RLS	SRLS
Set. time (sec)	x_1	0.589 (0.17%)	0.004 (25%)	0.001 (100%)	0.001
	x_2	1.907 (0.89%)	1.417 (1.2%)	1.006 (1.69%)	0.017
	x_3	0.837 (2.87%)	1.067 (2.25%)	0.229 (10.48 %)	0.024
	x_4	2.566 (7.72%)	1.732 (11.43%)	0.899 (22%)	0.198
Max O.S	x_1	0.234	---	---	---
	x_2	3.613 (0.83%)	3.579 (0.84%)	0.633 (4.74%)	0.03
	x_3	0.0516	0.207	0.0175	---
	x_4	0.544 (1.84%)	6.053 (0.17%)	1.272 (0.79%)	0.01
CPU time (msec)		0.05 (800%)	0.6 (67%)	0.1 (400%)	0.4**

* The results in parentheses represent the ratio between the SRLS observer results to that of the respective observer.
** The MSRLS observer CPU time is 0.26 msec.

RLS observer and much less than that of the HG and the SDRE observers.

Furthermore, no estimation error overshoot is observed in the error responses of the first and third states while using the SRLS observer. The improvements introduced by the SRLS observer came on the expense of increasing the average CPU time (0.4 msec) when compared to the RLS observer (0.1 msec) and the hg observer (0.05 msec). However, the computational times of the four observers are within the 1 msec sampling period as used in our simulation. In the next subsection, we propose a procedure to decrease the computational time of the SRLS observer with negligible effects on the achieved results.

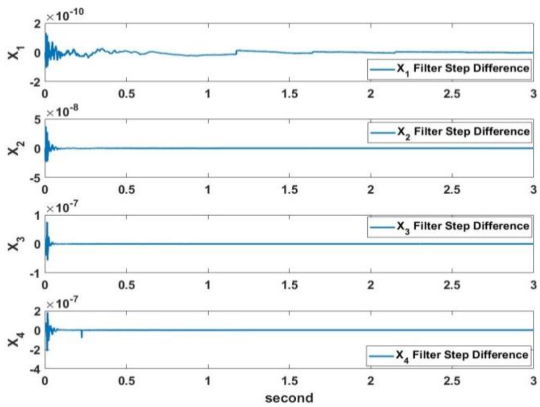

FIGURE 3. Difference of the filter step of the SRLS and MSRLS observers.

6) REDUCING THE COMPUTATIONAL TIME OF THE SRLS OBSERVER

As have been stated in the convergence analysis, the achieved improvements in the estimated states resulting from the SRLS observer are contributed to its prediction step due to the extra terms included in (21). To reduce the computational time of the SRLS observer without affecting the estimation results, a modified version of the SRLS observer (MSRLS) is proposed. In this version, equations (2-b), (4), (5), (6), (7) of the RLS observer are used to calculate $\hat{y}_{k+1|k}, K_{k+1}, P_{k+1|k+1}$ instead of (16), (18), (19), (20-a), (20-c) which usually need more computational time. By applying the MSRLS estimator to the problem at hand, the average CPU time was reduced by almost 35% of that of the SRLS observer. The difference between the filtered state estimates of the SRLS and the MSRLS observers are shown in Fig. 3. From the achieved results, it is very clear that the differences are negligible since they are less than 3×10^{-7}. Fig. 4 demonstrates the estimation errors from both observers which are almost identical. One has to notice that the estimation errors resulting from the HG and SDRE observers (shown in Fig. 2) are not displayed in Fig. 4 due to the large differences in their settling times and overshoots when compared with the SRLS observer that would have made it difficult for us to clearly compare the behavior of the SRLS and MSRLS observers. The preceding results validate the applicability of the MSRLS observer as a computationally efficient variant of the SRLS observer.

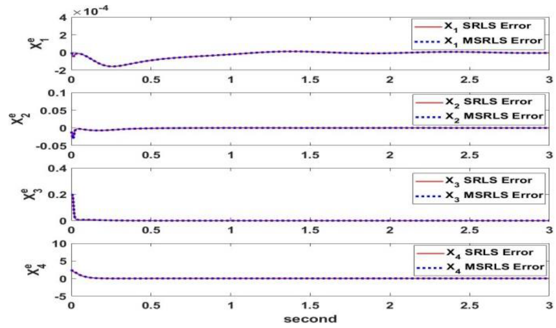

FIGURE 4. Estimation error of SRLS and MSRLS observers.

B. EXAMPLE II: INDUCTION MACHINE

In this example, the aim is to monitor the behavior of the nonlinear induction machine for which the mathematical model is given by [37]:

$$\begin{aligned} x1_{k+1} &= x1_k + \Delta T\left(k_1 x1_k + z_1 x2_k + k_2 x3_k + z_2\right) \\ x2_{k+1} &= x2_k + \Delta T\left(-z_1 x1_k + k_1 x2_k + k_2 x4_k\right) \\ x3_{k+1} &= x3_k + \Delta T\left(k_3 x1_k + k_4 x3_k + \left[z_1 - x5_k\right] x4_k\right) \\ x4_{k+1} &= x4_k + \Delta T\left(k_3 x2_k + k_4 x4_k - \left[z_1 - x5_k\right] x3_k\right) \\ x5_{k+1} &= x5_k + \Delta T\left(k_6 z_3 + k_5\left[x1_k x4_k - x2_k x3_k\right]\right) \\ y_{k+1} &= \left[k_7 x1_{k+1} + k_8 x3_{k+1} \quad k_7 x2_{k+1} + k_8 x4_{k+1}\right]^T \end{aligned} \tag{78}$$

In the above system, x_1, x_2 are the components of the stator fluxes; x_3, x_4 are the components of the rotor fluxes; and x_5 is the angular velocity. The frequency and the amplitude of the stator's voltage are denoted by z_1 and z_2, respectively; while z_3 represents the load torque. The parameters $k_1, k_2, \ldots, k_8$ depend on the considered drive. The values of these parameters are given in Table 3.

TABLE 3. System parameters of example II.

Par.	Val.	Par.	Val.	Par.	Val.
k_1	-0.186	k_2	0.178	k_3	0.225
k_4	-0.234	k_5	-0.081	k_6	-0.018
k_7	4.643	k_8	-4.448	z_1	1
z_2	1	z_3	0	ΔT	0.001 sec

The HG observer cannot be applied to estimate the states of this system since the output is a linear combination of the states.

1) THE SDRE OBSERVER

Although, there are several choices for the A, C matrices, we could not achieve any combination that leads to a convergent estimator. Therefore, we failed to apply this observer to the system at least with the several trails we made.

2) THE RLS OBSERVER

The Jacobian matrix of the system is given by, (79), as shown at the bottom of this page. The matrices $P_{0|0}$, N, S are chosen as follows:

$$\begin{aligned} P_{0|0} &= diag\,([10;\,10;\,10;\,10;\,10]) \\ N &= diag\,([0.0001;\,0.0001;\,0.0001;\,0.0001;\,0.0001]) \\ S &= diag\,([0.01;\,0.01]) \end{aligned} \tag{80}$$

3) THE SRLS AND THE MSRLS OBSERVERS

Both the SRLS and the MSRLS observers are directly applicable to the system model while the matrix $P_{xx_{0|0}}$ is as given in (80).

4) SIMULATION RESULTS OF EXAMPLE II

The initial conditions of the state vectors of the model and the observer are chosen as: $\boldsymbol{x}_0 = [0.2 - 0.6 - 0.4\ 0.1\ 0.3]^T$ and $\hat{\boldsymbol{x}}_{0|0} = \begin{bmatrix} 0.5 & 0.1 & 0.3 & -0.2 & 4 \end{bmatrix}^T$, respectively. The results of the RLS and the SRLS observers are shown in Figs. 5. The maximum overshoots and settling times of the estimation errors are presented in Table 4 in addition to the average CPU time per sample of each observer. The errors in the filtering step between the SRLS and the MSRLS are also shown in Fig. 6, while the estimation errors between the actual and the estimated states of the two filters are shown in Fig. 7.

From the achieved results, it can be concluded that the SRLS observer leads to much better estimated states when compared with the RLS observer. Moreover, the settling times of the SRLS observer are less than 10% of the RLS observer, and its maximum overshoots are less than 41% of the RLS observer. It can also be noticed that no overshoot is observed in the error responses of the first, third, and fifth states while using the SRLS observer. The average CPU time per sample of the SRLS observer is 0.52 msec while that of the RLS observer is 0.15 msec. However, both are within the chosen sampling period (1 msec).

It can also be observed from Fig. 7 that the results of the SRLS and the MSRLS observers are almost identical. The suggested modification reduced the average CPU time per sample to 0.33 msec which is almost 60% of the CPU time per sample for the SRLS observer. This again supports the applicability of the MSRLS estimator in practice.

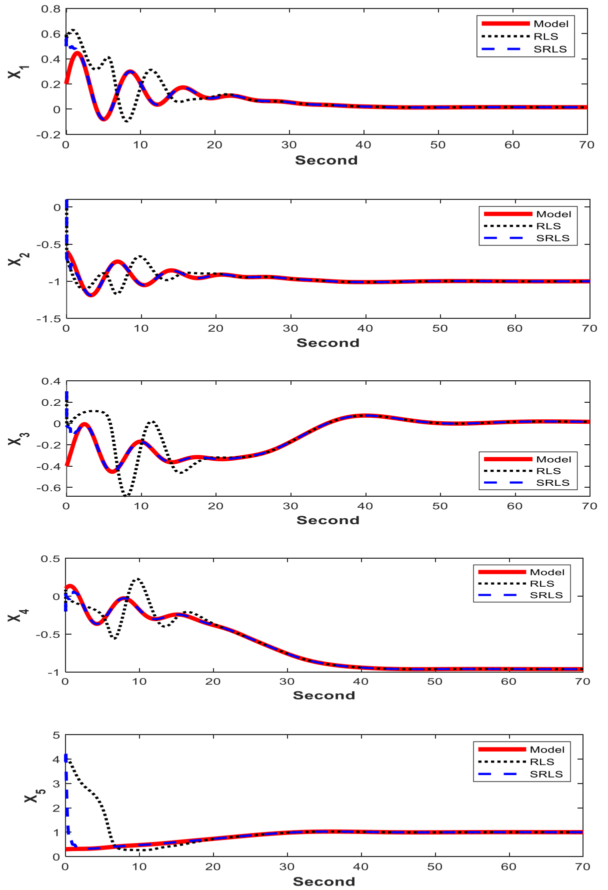

FIGURE 5. The estimated states (X_1-X_5) of the Induction Machine in example II.

C. DISCUSSION OF THE RESULTS

1- The proposed SRLS estimator and its modified version (MSRLS) lead to much better estimation results compared with those achieved from the HG, the SDRE, and the RLS estimators.

2- Unlike other available estimators in the literature, the developed SRLS estimator and its modified version did not show any undesired large overshoots at the start of the estimation process. More specifically, it is clear from Tables 2,4 and Figs. 1,5 that the overshoot of the SRLS and the MSRLS observers are almost negligible

$$A_k = \begin{bmatrix} 1+\Delta T k_1 & \Delta T z_1 & \Delta T k_2 & 0 & 0 \\ -\Delta T z_1 & 1+\Delta T k_1 & 0 & \Delta T k_2 & 0 \\ \Delta T k_3 & 0 & 1+\Delta T k_4 & \Delta T(z_1 - x_{5_k}) & -\Delta T x_{4_k} \\ 0 & \Delta T k_3 & \Delta T(x_{5_k} - z_1) & 1+\Delta T k_4 & \Delta T x_{3_k} \\ \Delta T k_5 x_{4_k} & -\Delta T k_5 x_{3_k} & -\Delta T k_5 x_{2_k} & \Delta T k_5 x_{1_k} & 1 \end{bmatrix}$$

$$C_{k+1} = \begin{bmatrix} k_7 & 0 & k_8 & 0 & 0 \\ 0 & k_7 & 0 & k_8 & 0 \end{bmatrix} \tag{79}$$

TABLE 4. Performance comparison between the observers in example II*.

	Settling time (sec)		Max. O.S	
State	RLS	SRLS	RLS	SRLS
x_1	18.19 (7.81%)	1.42	0.475	---
x_2	19.89 (7.54%)	1.5	0.4365 (40.78%)	0.178
x_3	18.22 (7.85%)	1.43	0.4963	---
x_4	19.94 (7.62%)	1.52	0.4556 (40.23%)	0.1833
x_5	19.23 (9.15%)	1.76	0.2218	---
CPU Time (msec)	0.15 (347%)	0.52**		

* The results in parentheses represent the ratio between the SRLS observer results to that of the RLS observer.
** The MSRLS observer CPU time is 0.33 msec.

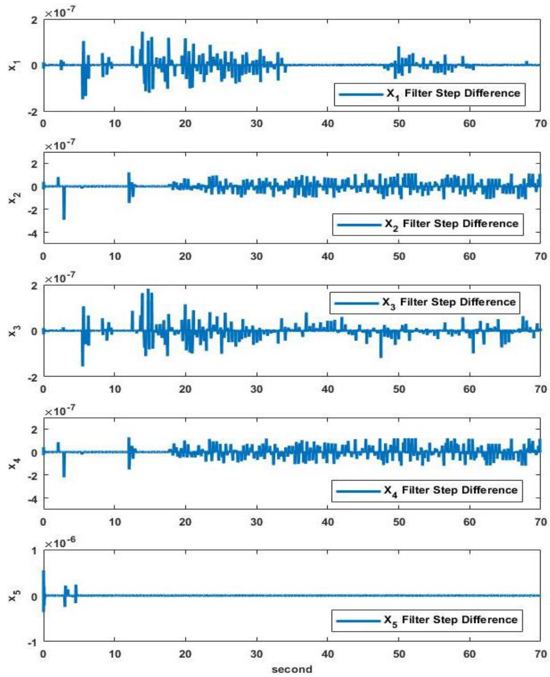

FIGURE 6. **Differences of the filter step of the SRLS, the MSRLS observers.**

when compared with the other observer(s). Therefore, they do not lead to any harmful control strategies when used to generate control signals in observer-based controlled systems.

3- The transient period of the estimation error is very short compared to that of the HG, the SDRE and the RLS estimators. This fact can also be justified from Tables 2,4 which clearly indicate that the transmission

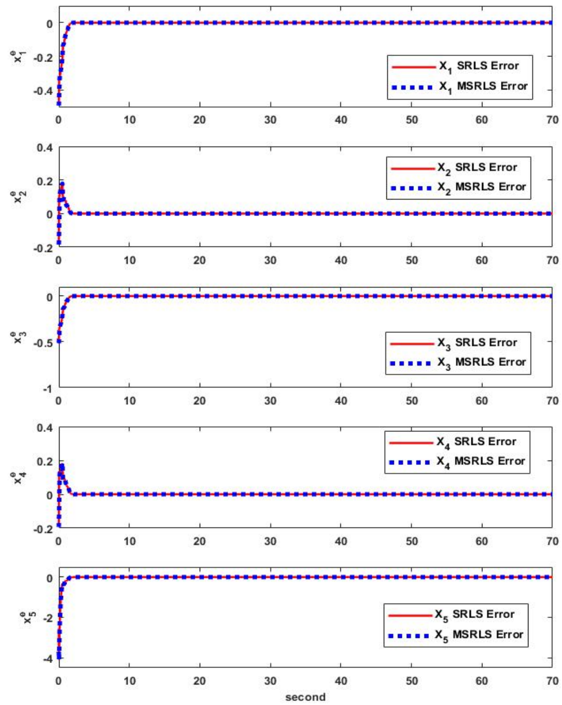

FIGURE 7. **Estimation Errors of the SRLS and the MSRLS observers.**

period of the estimation error of the SRLS estimator is very short relative to the other estimators. This significantly enhances the applicability of this observer whether in monitoring and/or generating state feedback control strategies.

4- The applicability of the proposed technique has no limitation. In other words, it can deal with any type of system nonlinearities (whether strong or weak), needs no state transformation, can deal with any measurement model (linear or nonlinear), does not have any restrictions (such as that in the SDRE approach), and finally does not need the derivation of the Jacobian matrix which may be difficult to achieve in some cases.

5- The proposed MSRLS estimator reduces the average CPU time per sample with a reasonable value (the CPU time of the MSRLS estimator is more or less 60% of that of the SRLS estimator) without affecting the estimation results (the error between the two estimators is within the value of 10^{-7}). It is obvious that this increases the domain of its applicability.

6- In principle, the proposed observer can be applied to estimate the states of large-scale nonlinear systems. However, in this case it is expected that the numerical accuracy of the results will not be as good as for small-scale nonlinear systems. This is simply due to rounding off resulting from the inversion of relatively large matrices. For this reason, we are now in the process of developing a decomposed RLS estimator. The preliminary achieved results are very encouraging and we hope to finalize this research in the near future.

From the preceding results, it is clear that the objectives of this work, as stated in section 2, have been fully satisfied.

VII. CONCLUSION

In this paper, the SRLS observer is developed as a new approach for discrete-time nonlinear estimation problems. Through this approach, the results of the prediction step are greatly improved which in turn lead to very precise estimation results. Unlike other available techniques in the literature, undesired large overshoots, usually taking place at the start of the estimation process, are avoided while keeping the transient period of the estimation error very short. In the proposed observer, a set of predetermined points around the filtered estimate of the state vector are used to get highly precise predicted estimates of the states. This obviously leads to highly improved filtered estimates of the states. The developed observer overcomes the main drawbacks of the well-known observers in the literature. It can deal with general nonlinear systems irrespective of their type of nonlinearities, leads to a unique solution, and avoids the computation of the Jacobian matrices which may be difficult to achieve in some nonlinear systems. The mathematical structure of the developed approach is presented and compared with the RLS observer. The convergence of the SRLS observer is analyzed and the results show that it has better performance when compared with the RLS observer. Illustrative examples of highly nonlinear power systems are presented to show the effectiveness and the superiority of the proposed approach when compared with the HG, the SDRE, and the RLS observers. Moreover, a modified version of the developed observer is proposed to reduce the computational time while maintaining the same quality of the estimation results.

APPENDIX A
THE REGULARIZED LEAST SQUARE ESTIMATOR

The idea is to formulate and solve at each sampling point a regularized least-square estimation problem for nonlinear discrete-time dynamical systems. Therefore, let us consider the following nonlinear discrete-time dynamical system:

$$\boldsymbol{x}_{k+1} = \boldsymbol{f}_k(\boldsymbol{x}_k) + \boldsymbol{\gamma}_k \tag{A-1}$$
$$\boldsymbol{y}_{k+1} = \boldsymbol{h}_{k+1}(\boldsymbol{x}_{k+1}) \tag{A-2}$$

where $\boldsymbol{x}_k$, $\boldsymbol{y}_{k+1}$, $\boldsymbol{f}_k(\boldsymbol{x}_k)$ are as defined in section II, while $\boldsymbol{\gamma}_k \in \boldsymbol{R}^n$ is a vector representing model inaccuracy.

Let $\hat{\boldsymbol{x}}_{k|k} \in \boldsymbol{R}^n$ denote for the estimate of $\boldsymbol{x}_k$ at the sampling instant $\boldsymbol{k}$. Fix the time instant $\boldsymbol{k}$, and assume that the filtered estimate $\hat{\boldsymbol{x}}_{k|k}$ has been computed as well as the recursive solution of the nonlinear matrix discrete Riccati like equation $\boldsymbol{P}_{k|k}$ (assumed positive definite).

Given a new measurement $\boldsymbol{y}_{k+1}$, we pose the problem of estimating x_k again along with $\boldsymbol{\gamma}_k$ by solving the RLS problem given by:

$$\min_{x_k,\gamma_k} \boldsymbol{J} = \frac{1}{2}\{\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k}\|^2_{\boldsymbol{P}^{-1}_{k|k}} + \|\boldsymbol{\gamma}_k\|^2_{\boldsymbol{N}^{-1}} + \|\boldsymbol{y}_{k+1} - \boldsymbol{h}_{k+1}(\boldsymbol{x}_{k+1})\|^2_{\boldsymbol{W}^{-1}}\} \tag{A-3}$$

where $\boldsymbol{N} \in \boldsymbol{R}^{nxn}$, $\boldsymbol{W} \in \boldsymbol{R}^{mxm}$, $\boldsymbol{P}_{k|k} \in \boldsymbol{R}^{nxn}$ are positive definite symmetric weighting matrices and $\boldsymbol{P}_{k|k}$ is as defined above.

Theorem 1: 1- Given a new measurement $\boldsymbol{y}_{k+1}$, the estimate of the vectors $\boldsymbol{\gamma}_k$ (to be denoted by $\hat{\boldsymbol{\gamma}}_{k|k+1}$) and $\boldsymbol{x}_k$ (to be denoted by $\hat{\boldsymbol{x}}_{k|k+1}$) which minimize (A-3) while taking into consideration (A-1) are given by:

$$\hat{\boldsymbol{\gamma}}_{k|k+1} = \boldsymbol{N}\hat{\boldsymbol{H}}^T_{k+1}\boldsymbol{W}^{-1}\left[\boldsymbol{y}_{k+1} - \hat{\boldsymbol{h}}_{k+1}(\hat{\boldsymbol{x}}_{k+1|k+1})\right] \tag{A-4}$$
$$\hat{\boldsymbol{x}}_{k|k+1} = \hat{\boldsymbol{x}}_{k|k} + \boldsymbol{P}_{k|k}\hat{\boldsymbol{A}}^T_k\hat{\boldsymbol{H}}^T_{k+1}\boldsymbol{W}^{-1}\left[\boldsymbol{y}_{k+1} - \hat{\boldsymbol{h}}_{k+1}(\hat{\boldsymbol{x}}_{k+1|k+1})\right] \tag{A-5}$$

where $\hat{\boldsymbol{x}}_{k+1|k+1}$ is the filtered estimate of $\boldsymbol{x}_{k+1}$ given the set of measurements $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k, \boldsymbol{y}_{k+1}\}$; $\hat{\boldsymbol{A}}_k$ is the estimate of the Jacobean matrix $\boldsymbol{f}_k$using $\hat{\boldsymbol{x}}_{k|k}$; $\hat{\boldsymbol{h}}_{k+1}(\hat{\boldsymbol{x}}_{k+1|k+1})$ is the estimate of $\boldsymbol{h}_{k+1}$ using $\hat{\boldsymbol{x}}_{k+1|k+1}$; and $\hat{\boldsymbol{H}}_{k+1}$ is the estimate of the Jacobean matrix of $\boldsymbol{h}_{k+1}$ using $\hat{\boldsymbol{x}}_{k+1|k}$.

2- The filtered estimate $\hat{\boldsymbol{x}}_{k+1|k+1}$ is given by:

$$\hat{\boldsymbol{x}}_{k+1|k+1} = \hat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{G}_{1k+1}\left[\boldsymbol{y}_{k+1} - \hat{\boldsymbol{h}}_{k+1}(\hat{\boldsymbol{x}}_{k+1|k})\right] \tag{A-6}$$

where, $\hat{\boldsymbol{x}}_{k+1|k}$ the predicted estimate of $\boldsymbol{x}_{k+1}$, and the gain matrix $\boldsymbol{G}_{1k+1}$ are given by:

$$\hat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{f}_k(\hat{\boldsymbol{x}}_{k|k}) \tag{A-7}$$
$$\boldsymbol{G}_{1k+1} = \boldsymbol{P}_{k+1|k}\hat{\boldsymbol{H}}^T_{k+1}\left[\hat{\boldsymbol{H}}_{k+1}\boldsymbol{P}_{k+1|k}\hat{\boldsymbol{H}}^T_{k+1} + \boldsymbol{W}\right]^{-1} = \boldsymbol{P}_{k+1|k+1}\hat{\boldsymbol{H}}^T_{k+1}\boldsymbol{W}^{-1} \tag{A-8}$$

with

$$\boldsymbol{P}_{k+1|k} = \hat{\boldsymbol{A}}_k\boldsymbol{P}_{k|k}\hat{\boldsymbol{A}}^T_k + \boldsymbol{N} \tag{A-9}$$

and the matrix Riccati like equation $\boldsymbol{P}_{k+1|k+1}$ is given by:

$$\boldsymbol{P}_{k+1|k+1} = \left[\boldsymbol{P}^{-1}_{k+1|k} + \hat{\boldsymbol{H}}^T_{k+1}\boldsymbol{W}^{-1}\hat{\boldsymbol{H}}_{k+1}\right]^{-1} = [\boldsymbol{I} - \boldsymbol{G}_{1k+1}\boldsymbol{H}_{k+1}]\boldsymbol{P}_{k+1|k} \tag{A-10}$$

Proof: 1- The necessary conditions of optimality, taking into consideration (A-1), lead to:

$$\frac{\partial \boldsymbol{J}}{\partial \boldsymbol{\gamma}_k} = 0 \Rightarrow \boldsymbol{N}^{-1}\boldsymbol{\gamma}_k + \left(\frac{\partial x^T_{k+1}}{\partial \boldsymbol{\gamma}_k}\right)\left[-\frac{\partial \boldsymbol{h}^T_{k+1}(\boldsymbol{x}_{k+1})}{\partial x_{k+1}}\boldsymbol{W}^{-1}\boldsymbol{y}_{k+1}\right] + \left(\frac{\partial \boldsymbol{x}^T_{k+1}}{\partial \boldsymbol{\gamma}_k}\right)\left[\frac{\partial \boldsymbol{h}^T_{k+1}(\boldsymbol{x}_{k+1})}{\partial \boldsymbol{x}_{k+1}}\boldsymbol{W}^{-1}\boldsymbol{h}_{k+1}(\boldsymbol{x}_{k+1})\right] = 0 \tag{A-11}$$

Denoting the minimizing argument of $\boldsymbol{x}_{k+1}$ by $\hat{\boldsymbol{x}}_{k+1|k+1}$ we get:

$$\hat{\boldsymbol{\gamma}}_{k|k+1} = \boldsymbol{N}\left.\frac{\partial \boldsymbol{h}^T_{k+1}(\boldsymbol{x}_{k+1})}{\partial \boldsymbol{x}_{k+1}}\right|_{\hat{x}_{k+1|k+1}} \times \boldsymbol{W}^{-1}\left[\boldsymbol{y}_{k+1} - \hat{\boldsymbol{h}}_{k+1}(\hat{\boldsymbol{x}}_{k+1|k+1})\right] \tag{A-12}$$

Defining $\hat{x}_{k+1|k}$, the predicted estimate of x_{k+1} by:

$$\hat{x}_{k+1|k} = f_k(\hat{x}_{k|k}) \tag{A-13}$$

and let $\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + \Delta\hat{x}_{k+1|k}$ where $\Delta\hat{x}_{k+1|k}$ is the difference between the filtered estimate $\hat{x}_{k+1|k+1}$ and the predicted estimate $\hat{x}_{k+1|k}$. Then, the Jacobean matrix $\hat{H}_{k+1}(\hat{x}_{k+1|k+1})$ of h_{k+1} is given by:

$$\hat{H}^T_{k+1}(\hat{x}_{k+1|k+1}) = \hat{H}^T_{k+1}(\hat{x}_{k+1|k} + \Delta\hat{x}_{k+1|k}) = \left.\frac{\partial h^T_{k+1}(x_{k+1})}{\partial x_{k+1}}\right|_{\hat{x}_{k+1|k+1}} \tag{A-14}$$

Assume that $\hat{x}_{k+1|k+1}$ is close to $\hat{x}_{k+1|k}$ that we can approximate $\hat{H}_{k+1}(\hat{x}_{k+1|k}+\Delta\hat{x}_{k+1|k})$ by $\hat{H}_{k+1}(\hat{x}_{k+1|k})$. For simplicity of writing and without introducing any ambiguities, let us drop the argument of $\hat{H}_{k+1}(\cdot)$, hence:

$$\hat{H}_{k+1} = \hat{H}^T_{k+1}(\hat{x}_{k+1|k+1}) \cong \hat{H}_{k+1}(\hat{x}_{k+1|k}) \tag{A-15}$$

Using (A-13) into (A-11) we get (A-4).

Also;

$$\frac{\partial J}{\partial x_k} = 0$$

$$\Rightarrow P^{-1}_{k|k}[x_k - \hat{x}_{k|k}] - \frac{\partial f^T_k(x_k)}{\partial x_k}\left[\frac{\partial h^T_{k+1}(x_{k+1})}{\partial x_{k+1}} W^{-1} y_{k+1}\right] + \frac{\partial f^T_k(x_k)}{\partial x_k}\left[\frac{\partial h^T_{k+1}(x_{k+1})}{\partial x_{k+1}} W^{-1} h_{k+1}(x_{k+1})\right] = 0$$

From which we get:

$$\hat{x}_{k|k+1} = \hat{x}_{k|k} + P_{k|k}\left.\frac{\partial f^T_k(x_k)}{\partial x_k}\right|_{\hat{x}_{k|k+1}} \hat{H}^T_{k+1} W^{-1} y_{k+1} - P_{k|k}\left.\frac{\partial f^T_k(x_k)}{\partial x_k}\right|_{\hat{x}_{k|k+1}} \hat{H}^T_{k+1} W^{-1} \hat{h}_{k+1}(\hat{x}_{k+1|k+1}) \tag{A-16}$$

Again, let $\hat{x}_{k|k+1} = \hat{x}_{k|k} + \Delta\hat{x}_{k|k}$, where $\Delta\hat{x}_{k|k+1}$ is the difference between the filtered estimate $\hat{x}_{k|k}$ and $\hat{x}_{k|k+1}$ resulting from the minimization of (A-3) with respect to (w.r.t.) x_k, then the Jacobean matrix $\hat{A}^T_k(\hat{x}_{k|k+1})$ of f_k is such that:

$$\hat{A}^T_k(\hat{x}_{k|k+1}) = \hat{A}^T_k(\hat{x}_{k|k} + \Delta\hat{x}_{k|k}) = \left.\frac{\partial f^T_k(x_k)}{\partial x_k}\right|_{\hat{x}_{k|k+1}}$$

Again, assuming that $\hat{x}_{k|k+1}$ is close to $\hat{x}_{k|k}$ that we can approximate $\hat{A}_{k+1}(\hat{x}_{k|k} + \Delta\hat{x}_{k|k})$ by $\hat{A}_{k+1}(\hat{x}_{k|k})$. Therefore, and after dropping the argument, we get:

$$\hat{A}_k \cong \hat{A}_k(\hat{x}_{k|k}) \tag{A-17}$$

Using (A-15) into (A-14) we get (A-5).

2- Introducing the quantity in agreement with the state equation (A-1):

$$\hat{x}_{k+1|k+1} = f(\hat{x}_{k|k+1}) + \hat{\gamma}_{k|k+1} \tag{A-18}$$

Substituting from (A-4), (A-5) into (A-16), $\hat{x}_{k+1|k+1}$ is such that:

$$\hat{x}_{k+1|k+1} = f\left(\hat{x}_{k|k} + P_{k|k}\hat{A}^T_k\hat{H}^T_{k+1}W^{-1}\left[y_{k+1} - \hat{h}_{k+1}(\hat{x}_{k+1|k+1})\right]\right) + N\hat{H}^T_{k+1}W^{-1}\left[y_{k+1} - \hat{h}_{k+1}(\hat{x}_{k+1|k+1})\right] \tag{A-19}$$

Assume that $\hat{x}_{k|k}+P_{k|k}\hat{A}^T_k\hat{H}^T_{k+1}W^{-1}\left[y_{k+1}-\hat{h}_{k+1}(\hat{x}_{k+1|k+1})\right]$ is close to $\hat{x}_{k|k}$ that we can replace $f(\cdot)$ by its first order approximation, i.e. we assume $f(\cdot)$ affine on a neighborhood of $\hat{x}_{k|k}+P_{k|k}\hat{A}^T_k\hat{H}^T_{k+1}W^{-1}[y_{k+1}-\hat{h}_{k+1}(\hat{x}_{k+1|k+1})], \hat{x}_{k|k}$, then:

$$f\left(\hat{x}_{k|k} + P_{k|k}\hat{A}^T_k\hat{H}^T_{k+1}W^{-1}\left[y_{k+1} - \hat{h}_{k+1}(\hat{x}_{k+1|k+1})\right]\right) = f(\hat{x}_{k|k}) + \hat{A}_kP_{k|k}\hat{A}^T_k\hat{H}^T_{k+1}W^{-1}\left[y_{k+1} - , \hat{h}_{k+1}(\hat{x}_{k+1|k+1})\right]$$

As a result, the filtered estimate $\hat{x}_{k+1|k+1}$ takes the form:

$$\hat{x}_{k+1|k+1} = f(\hat{x}_{k|k}) + \hat{A}_kP_{k|k}\hat{A}^T_k\hat{H}^T_{k+1}W^{-1} \times \left[y_{k+1} - \hat{h}_{k+1}(\hat{x}_{k+1|k+1})\right] + N\hat{H}^T_{k+1}W^{-1} \times \left[y_{k+1} - \hat{h}_{k+1}(\hat{x}_{k+1|k+1})\right] \tag{A-20}$$

Let:

$$P_{k+1|k} = \hat{A}_kP_{k|k}\hat{A}^T_k + N \tag{A-21}$$

Then, by using (A-11), (A-19) and substituting for $\hat{x}_{k+1|k+1}$ by $\hat{x}_{k+1|k} + \Delta\hat{x}_{k+1|k}$ into (A-18), one gets:

$$\hat{x}_{k+1|k} + \Delta\hat{x}_{k+1|k} = \hat{x}_{k+1|k} + P_{k+1|k}\hat{H}^T_{k+1}W^{-1}y_{k+1} - P_{k+1|k}\hat{H}^T_{k+1}W^{-1}\hat{h}_{k+1}(\hat{x}_{k+1|k} + \Delta\hat{x}_{k+1|k})$$

From which $\Delta\hat{x}_{k+1|k}$ is such that:

$$\Delta\hat{x}_k + 1|k = P_{k+1|k}\hat{H}^T_{k+1}W^{-1} \times \left[y_{k+1} - \hat{h}_{k+1}(\hat{x}_{k+1|k} + \Delta\hat{x}_{k+1|k})\right] \tag{A-22}$$

Again, assume that $\hat{x}_{k+1|k+1}$ is close to $\hat{x}_{k+1|k}$, then we can replace $\hat{h}_{k+1}(\hat{x}_{k+1|k} + \Delta\hat{x}_{k+1|k})$ by its first order approximation, i.e. we assume $\hat{h}_{k+1}(\cdot)$ affine on a neighborhood $\hat{x}_{k+1|k+1}, \hat{x}_{k+1|k}$ therefore (A-20) reduces to:

$$\Delta\hat{x}_{k+1|k} = P_{k+1|k}\hat{H}^T_{k+1}W^{-1}\left[y_{k+1} - \hat{h}_{k+1}(\hat{x}_{k+1|k})\right] - P_{k+1|k}\hat{H}^T_{k+1}W^{-1}\hat{H}_{k+1}\Delta\hat{x}_{k+1|k}$$

where $\hat{h}_{k+1}(\hat{x}_{k+1|k})$ is the estimate of $h_{k+1}(x)$ using the predictive estimate $\hat{x}_{k+1|k}$.

Let

$$P_{k+1|k+1} = \left[P^{-1}_{k+1|k} + \hat{H}^T_{k+1}W^{-1}\hat{H}_{k+1}\right]^{-1} \tag{A-23}$$

Then, after simple mathematical manipulation, $\Delta\hat{x}_{k+1|k}$ is such that:

$$\Delta\hat{x}_{k+1|k} = P_{k+1|k+1}\hat{H}^T_{k+1}W^{-1}\left[y_{k+1} - \hat{h}_{k+1}(\hat{x}_{k+1|k})\right] \tag{A-24}$$

Using the matrix inversion lemma, it can be shown that:

$$\boldsymbol{P}_{k+1|k}\hat{\boldsymbol{H}}_{k+1}^{T}\left[\hat{\boldsymbol{H}}_{k+1}\boldsymbol{P}_{k+1|k}\hat{\boldsymbol{H}}_{k+1}^{T}+\boldsymbol{W}\right]^{-1} = \boldsymbol{P}_{k+1|k+1}\hat{\boldsymbol{H}}_{k+1}^{T}\boldsymbol{W}^{-1} \quad \text{(A-25)}$$

Hence, (A-22) takes the form:

$$\Delta\hat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{G}_{1k+1}\left[\boldsymbol{y}_{k+1}-\hat{\boldsymbol{h}}_{k+1}(\hat{\boldsymbol{x}}_{k+1|k})\right] \quad \text{(A-26)}$$

where:

$$\begin{aligned}\boldsymbol{G}_{1k+1} &= \boldsymbol{P}_{k+1|k}\hat{\boldsymbol{H}}_{k+1}^{T}\left[\hat{\boldsymbol{H}}_{k+1}\boldsymbol{P}_{k+1|k}\hat{\boldsymbol{H}}_{k+1}^{T}+\boldsymbol{W}\right]^{-1}\\ &= \boldsymbol{P}_{k+1|k+1}\hat{\boldsymbol{H}}_{k+1}^{T}\boldsymbol{W}^{-1}\end{aligned} \quad \text{(A-27)}$$

Finally, the filtered estimate $\hat{\boldsymbol{x}}_{k+1|k+1}$ is given by:

$$\hat{\boldsymbol{x}}_{k+1|k+1} = \hat{\boldsymbol{x}}_{k+1|k}+\boldsymbol{G}_{1k+1}\left[\boldsymbol{y}_{k+1}-\hat{\boldsymbol{h}}_{k+1}(\hat{\boldsymbol{x}}_{k+1|k})\right] \quad \text{(A-28)}$$

Again, by using the matrix inversion lemma, we get:

$$\begin{aligned}\boldsymbol{P}_{k+1|k+1} &= \left[\boldsymbol{P}_{k+1|k}^{-1}+\hat{\boldsymbol{H}}_{k+1}^{T}\boldsymbol{W}^{-1}\hat{\boldsymbol{H}}_{k+1}\right]^{-1}\\ &= [\boldsymbol{I}-\boldsymbol{G}_{1k+1}\boldsymbol{H}_{k+1}]\boldsymbol{P}_{k+1|k}\end{aligned} \quad \text{(A-29)}$$

APPENDIX B
THE RELATION BETWEEN THE PREDICTION STEP OF THE SRLS AND THE RLS ESTIMATORS

To get the relationships between the SRLS and the RLS observers, let us firstly expand (12) while using (11).

Therefore, we get:

$$\begin{aligned}\hat{\boldsymbol{x}}_{k+1|k} &= \frac{\lambda}{n+\lambda}\hat{\boldsymbol{x}}_{1k+1|k}+\frac{1}{2(n+\lambda)}[\hat{\boldsymbol{x}}_{2k+1|k}+\hat{\boldsymbol{x}}_{3k+1|k}\\ &\quad+\cdots+\hat{\boldsymbol{x}}_{(2n)k+1|k}+\hat{\boldsymbol{x}}_{(2n+1)k+1|k}]\end{aligned} \quad \text{(B-1)}$$

$$\begin{aligned}\hat{\boldsymbol{x}}_{k+1|k} &= \frac{\lambda}{n+\lambda}f_k(\hat{\boldsymbol{x}}_{k|k})+\frac{1}{2(n+\lambda)}[f_k(\hat{\boldsymbol{x}}_{k|k}+\boldsymbol{\zeta}_1)\\ &\quad+f_k(\hat{\boldsymbol{x}}_{k|k}-\boldsymbol{\zeta}_1)+\cdots+f_k(\hat{\boldsymbol{x}}_{k|k}+\boldsymbol{\zeta}_n)\\ &\quad+f_k(\hat{\boldsymbol{x}}_{k|k}-\boldsymbol{\zeta}_n)]\end{aligned} \quad \text{(B-2)}$$

Using Taylor series expansion up to the second order approximation, we have:

$$\begin{aligned}&\hat{\boldsymbol{x}}_{k+1|k}\\ &= \frac{\lambda}{n+\lambda}f_k(\hat{\boldsymbol{x}}_{k|k})+\frac{1}{2(n+\lambda)}\\ &\quad * \left\{\left[f_k(\hat{\boldsymbol{x}}_{k|k})+\hat{A}_k\boldsymbol{\zeta}_1+\frac{1}{2}\left[\nabla^T\Theta_1\nabla\right]f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}}\right]\right.\\ &\quad+\left[f_k(\hat{\boldsymbol{x}}_{k|k})-\hat{A}_k\boldsymbol{\zeta}_1+\frac{1}{2}\left[\nabla^T\Theta_1\nabla\right]f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}}\right]\\ &\quad+\cdots+\left[f_k(\hat{\boldsymbol{x}}_{k|k})+\hat{A}_k\boldsymbol{\zeta}_n+\frac{1}{2}\left[\nabla^T\Theta_n\nabla\right]f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}}\right]\\ &\quad+\left[f_k(\hat{\boldsymbol{x}}_{k|k})-\hat{A}_k\boldsymbol{\zeta}_n+\frac{1}{2}\left[\nabla^T\Theta_n\nabla\right]f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}}\right]\\ &\quad\left.+H.O.T\right\}\end{aligned} \quad \text{(B-3)}$$

where Θ_i is the nxn matrix: $\Theta_i = \boldsymbol{\zeta}_i\boldsymbol{\zeta}_i^T$, $\hat{A}_k = \frac{\partial f_k^T}{\partial \boldsymbol{x}_k}\Big|_{\hat{\boldsymbol{x}}_{k|k}}^T$, ∇ donates for the gradient operator with respect to the vector x, and $H.O.T.$ is the higher order terms.

Collecting terms and simplifying, we get:

$$\begin{aligned}\hat{\boldsymbol{x}}_{k+1|k} &= \frac{\lambda}{n+\lambda}f_k(\hat{\boldsymbol{x}}_{k|k})+\frac{n}{n+\lambda}f_k(\hat{\boldsymbol{x}}_{k|k})+\frac{1}{2(n+\lambda)}\\ &\quad * \left.\left[\nabla^T(\Theta_1+\cdots+\Theta_n)\nabla\right]f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}}+H.O.T\right\}\end{aligned} \quad \text{(B-4)}$$

Noting that $(\Theta_1+\cdots+\Theta_n) = Z(Z)^T = (n+\lambda)P_{xx_{k|k}}$, then (B-4) takes the form:

$$\hat{\boldsymbol{x}}_{k+1|k} = f_k(\hat{\boldsymbol{x}}_{k|k})+\frac{1}{2(n+\lambda)}\left[\nabla^T(n+\lambda)P_{xx_{k|k}}\nabla\right]f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}}+H.O.T \quad \text{(B-5)}$$

which can be written in the form:

$$\hat{\boldsymbol{x}}_{k+1|k} = f_k(\hat{\boldsymbol{x}}_{k|k})+\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})+H.O.T \quad \text{(B-6)}$$

where:

$$\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k}) = \frac{1}{2}\left[\nabla^T P_{xx_{k|k}}\nabla\right]f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}} \quad \text{(B-7)}$$

Equation (B-6) is not more than the predicted estimate of the RLS estimator in addition to the extra terms $\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})$ and the $H.O.T$.

Similarly, by expanding (14) we get:

$$\begin{aligned}&P_{xx_{k+1|k}}\\ &= \frac{\lambda}{n+\lambda}\left[\left(\hat{\boldsymbol{x}}_{1k+1|k}-\hat{\boldsymbol{x}}_{k+1|k}\right)\left(\hat{\boldsymbol{x}}_{1k+1|k}-\hat{\boldsymbol{x}}_{k+1|k}\right)^T\right]\\ &\quad+\frac{1}{2(n+\lambda)}\left\{\left(\hat{\boldsymbol{x}}_{2k+1|k}-\hat{\boldsymbol{x}}_{k+1|k}\right)\left(\hat{\boldsymbol{x}}_{2k+1|k}-\hat{\boldsymbol{x}}_{k+1|k}\right)^T\right.\\ &\quad+\left(\hat{\boldsymbol{x}}_{3k+1|k}-\hat{\boldsymbol{x}}_{k+1|k}\right)\left(\hat{\boldsymbol{x}}_{3k+1|k}-\hat{\boldsymbol{x}}_{k+1|k}\right)^T\\ &\quad+\cdots+\left(\hat{\boldsymbol{x}}_{(2n)k+1|k}-\hat{\boldsymbol{x}}_{k+1|k}\right)\left(\hat{\boldsymbol{x}}_{(2n)k+1|k}-\hat{\boldsymbol{x}}_{k+1|k}\right)^T\\ &\quad\left.+\left(\hat{\boldsymbol{x}}_{(2n+1)k+1|k}-\hat{\boldsymbol{x}}_{k+1|k}\right)\left(\hat{\boldsymbol{x}}_{(2n+1)k+1|k}-\hat{\boldsymbol{x}}_{k+1|k}\right)^T\right\}\end{aligned} \quad \text{(B-8)}$$

Let us consider the first term in (B-8). Substituting for $\hat{\boldsymbol{x}}_{k+1|k}$ from (B-6) and for $\hat{\boldsymbol{x}}_{1_{k+1|k}}$ from (11), we have:

$$\begin{aligned}&\left(\hat{\boldsymbol{x}}_{1k+1|k}-\hat{\boldsymbol{x}}_{k+1|k}\right)\left(\hat{\boldsymbol{x}}_{1k+1|k}-\hat{\boldsymbol{x}}_{k+1|k}\right)^T\\ &\quad= \left\{f_k(\hat{\boldsymbol{x}}_{k|k})-\left[f_k(\hat{\boldsymbol{x}}_{k|k})+\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})+H.O.T\right]\right\}\\ &\quad\quad * \left\{f_k(\hat{\boldsymbol{x}}_{k|k})-\left[f_k(\hat{\boldsymbol{x}}_{k|k})+\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})+H.O.T\right]\right\}^T\\ &\quad= \boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})\boldsymbol{D}_k^T(\hat{\boldsymbol{x}}_{k|k})+H.O.T\end{aligned} \quad \text{(B-9)}$$

Using the same procedure with the second term in (B-8) while expanding $\hat{\boldsymbol{x}}_{2_{k+1|k}}$ in (11) using Taylor series expansion up to the second order, we get:

$$\left(\hat{\boldsymbol{x}}_{2k+1|k}-\hat{\boldsymbol{x}}_{k+1|k}\right)\left(\hat{\boldsymbol{x}}_{2k+1|k}-\hat{\boldsymbol{x}}_{k+1|k}\right)^T = \Im\Im^T+H.O.T \quad \text{(B-10a)}$$

where:

$$\Im = \{f_k(\hat{\boldsymbol{x}}_{k|k}) + \hat{A}_k \boldsymbol{\zeta}_1 + \frac{1}{2}[\nabla^T \Theta_1 \nabla] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}} - [f_k(\hat{\boldsymbol{x}}_{k|k}) + \boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})]\}$$
$$= \hat{A}_k \boldsymbol{\zeta}_1 + \frac{1}{2}\left[\nabla^T \Theta_1 \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}} - \boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k}) \quad \text{(B-10b)}$$

Or:

$$\begin{aligned}
&\left(\hat{\boldsymbol{x}}_{2k+1|k} - \hat{\boldsymbol{x}}_{k+1|k}\right)\left(\hat{\boldsymbol{x}}_{2k+1|k} - \hat{\boldsymbol{x}}_{k+1|k}\right)^T \\
&= \hat{A}_k \Theta_1 \hat{A}_k^T \\
&\quad + \frac{1}{2}\hat{A}_k \boldsymbol{\zeta}_1 \left\{\left[\nabla^T \Theta_1 \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}}\right\}^T - \hat{A}_k \boldsymbol{\zeta}_1 \boldsymbol{D}_k^T(\hat{\boldsymbol{x}}_{k|k}) \\
&\quad + \frac{1}{2}\left[\nabla^T \Theta_1 \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}} \boldsymbol{\zeta}_1^T \hat{A}_k^T \\
&\quad + \frac{1}{4}\left[\nabla^T \Theta_1 \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}} \left\{\left[\nabla^T \Theta_1 \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}}\right\}^T \\
&\quad - \frac{1}{2}\left[\nabla^T \Theta_1 \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}} \boldsymbol{D}_k^T(\hat{\boldsymbol{x}}_{k|k}) - \boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})\boldsymbol{\zeta}_1^T \hat{A}_k^T \\
&\quad - \frac{1}{2}\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})\left\{\left[\nabla^T \Theta_1 \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}}\right\}^T \\
&\quad + \boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})\boldsymbol{D}_k^T(\hat{\boldsymbol{x}}_{k|k}) + H.O.T \qquad \text{(B-11)}
\end{aligned}$$

Repeating the same procedure with the other terms in (28), adding the results of the expanded terms and simplifying, one gets:

$$\begin{aligned}
&P_{xx_{k+1|k}} \\
&= \frac{\lambda}{n+\lambda}\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})\boldsymbol{D}_k^T(\hat{\boldsymbol{x}}_{k|k}) \\
&\quad + \frac{1}{2(n+\lambda)}\Big\{2\hat{A}_k(\Theta_1 + \cdots + \Theta_n)\hat{A}_k^T \\
&\quad + 2n\left[\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})\boldsymbol{D}_k^T(\hat{\boldsymbol{x}}_{k|k})\right] \\
&\quad - \left[\nabla^T(\Theta_1 + \cdots + \Theta_n)\nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}} \boldsymbol{D}_k^T(\hat{\boldsymbol{x}}_{k|k}) \\
&\quad - \boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})\left\{\left[\nabla^T(\Theta_1 + \cdots + \Theta_n)\nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}}\right\}^T \\
&\quad + \frac{1}{2}\left[\nabla^T \Theta_1 \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}} \left\{\left[\nabla^T \Theta_1 \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}}\right\}^T \\
&\quad + \cdots + \frac{1}{2}\left[\nabla^T \Theta_n \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}} \\
&\quad \times \left\{\left[\nabla^T \Theta_n \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}}\right\}^T\Big\} + H.O.T \qquad \text{(B-12)}
\end{aligned}$$

Again, using the fact that $(\Theta_1 + \cdots + \Theta_n) = Z(Z)^T = (n+\lambda)P_{xx_{k|k}}$ while using (B-7) into (B-12) we have:

$$\begin{aligned}
&P_{xx_{k+1|k}} \\
&= \frac{\lambda}{n+\lambda}\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})\boldsymbol{D}_k^T(\hat{\boldsymbol{x}}_{k|k}) + \hat{A}_k P_{xx_{k|k}} \hat{A}_k^T \\
&\quad + \frac{n}{(n+\lambda)}\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})\boldsymbol{D}_k^T(\hat{\boldsymbol{x}}_{k|k}) - 2\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})\boldsymbol{D}_k^T(\hat{\boldsymbol{x}}_{k|k}) \\
&\quad + \frac{1}{4(n+\lambda)}\Big\{\left[\nabla^T \Theta_1 \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}} \\
&\quad \times \left\{\left[\nabla^T \Theta_1 \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}}\right\}^T + \cdots + \left[\nabla^T \Theta_n \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}} \\
&\quad \times \left\{\left[\nabla^T \Theta_n \nabla\right] f_k(\boldsymbol{x}_k)\Big|_{\hat{\boldsymbol{x}}_{k|k}}\right\}^T\Big\} + H.O.T \qquad \text{(B-13)}
\end{aligned}$$

Finally, after simple algebraic manipulations, one gets:

$$P_{xx_{k+1|k}} = \hat{A}_k P_{xx_{k|k}} \hat{A}_k^T - \boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})\boldsymbol{D}_k^T(\hat{\boldsymbol{x}}_{k|k}) + H.O.T \quad \text{(B-14)}$$

Or:

$$P_{xx_{k+1|k}} = \hat{A}_k P_{xx_{k|k}} \hat{A}_k^T + \Phi_k(\hat{\boldsymbol{x}}_{k|k}) + H.O.T \quad \text{(B-15)}$$

where:

$$\Phi_k(\hat{\boldsymbol{x}}_{k|k}) = -\boldsymbol{D}_k(\hat{\boldsymbol{x}}_{k|k})\boldsymbol{D}_k^T(\hat{\boldsymbol{x}}_{k|k}) \quad \text{(B-16)}$$

One can notice that (B-15) is equivalent to (2-c) of the RLS observer except for the constant matrix N is replaced by $\Phi_k(\hat{\boldsymbol{x}}_{k|k})$, in addition to the $H.O.T$ included in (B-15).

REFERENCES

[1] B. Misgeld, M. Lüken, R. Riener, and S. Leonhardt, "Observer-based human knee stiffness estimation," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 5, pp. 1033–1044, May 2017. doi: 10.1109/tbme.2016.2587841.

[2] F. Zhong, H. Li, S. Zhong, Q. Zhong, and C. Yin, "An SOC estimation approach based on adaptive sliding mode observer and fractional order equivalent circuit model for lithium-ion batteries," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 24, nos. 1–3, pp. 127–144, 2015. doi: 10.1016/j.cnsns.2014.12.015.

[3] J. M. Hansen, T. I. Fossen, and T. A. Johansen, "Nonlinear observer design for GNSS-aided inertial navigation systems with time-delayed GNSS measurements," *Control Eng. Pract.*, vol. 60, pp. 39–50, Mar. 2017. doi: 10.1016/j.conengprac.2016.11.016.

[4] D. Copp, R. Gondhalekar, and J. P. Hespanha, "Simultaneous model predictive control and moving horizon estimation for blood glucose regulation in Type 1 diabetes," *Optim. Control Appl. Methods*, vol. 39, no. 2, pp. 904–918, 2018. doi: 10.1002/oca.2388.

[5] S. Strommer, M. Niederer, A. Steinboeck, L. Jadachowski, and A. Kugi, "Nonlinear observer for temperatures and emissivities in a strip annealing furnace," *IEEE Trans. Ind. Appl.*, vol. 53, no. 3, pp. 2578–2586, May/Jun. 2017. doi: 10.1109/tia.2017.2669327.

[6] M. H. Abd, F. Tahir, G. Al-Suhail, and V.-T. Pham, "An adaptive observer synchronization using chaotic time-delay system for secure communication," *Nonlinear Dyn.*, vol. 90, no. 4, pp. 2583–2598, 2017. doi: 10.1007/s11071-017-3825-2.

[7] B. Zhou, J. Zhang, J. Gao, H. Yu, and D. Liu, "Clutch pressure estimation for a power-split hybrid transmission using nonlinear robust observer," *Mech. Syst. Signal Process.*, vol. 106, pp. 249–264, Jun. 2018. doi: 10.1016/j.ymssp.2018.01.001.

[8] D. Söffker, C. Wei, S. Wolff, and M.-S. Saadawia, "Detection of rotor cracks: Comparison of an old model-based approach with a new signal-based approach," *Nonlinear Dyn.*, vol. 83, no. 3, pp. 1153–1170, Feb. 2016. doi: 10.1007/s11071-015-2394-5.

[9] K. E. Starkov, L. N. Coria, and L. T. Aguilar, "On synchronization of chaotic systems based on the Thau observer design," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 17, no. 1, pp. 17–25, 2012. doi: 10.1016/j.cnsns.2011.04.020.

[10] A. Nath, R. Dey, and C. Aguilar-Avelar, "Observer based nonlinear control design for glucose regulation in type 1 diabetic patients: An LMI approach," *Biomed. Signal Process. Control*, vol. 47, pp. 7–15, Jan. 2019. doi: 10.1016/j.bspc.2018.07.020.

[11] Q. Zhu, L. Li, X. Hu, N. Xiong, and G.-D. Hu, "H∞-based nonlinear observer design for state of charge estimation of lithium-ion battery with polynomial parameters," *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 10853–10865, Dec. 2017. doi: 10.1109/tvt.2017.2723522.

[12] H. Hammouri, F. S. Ahmed, and S. Othman, "Observer design based on immersion technics and canonical form," *Syst. Control Lett.*, vol. 114, pp. 19–26, Apr. 2018. doi: 10.1016/j.sysconle.2018.02.003.

[13] C. Aguilar-Ibañez, E. García-Canseco, R. Martínez-García, J. Martínez-García, and M. Suarez-Castañon, "An I&I-based observer to solve the output-feedback synchronization problem for a class of chaotic systems," *Asian J. Control*, vol. 20, no. 4, pp. 1491–1503, 2018. doi: 10.1002/asjc.1650.

[14] J. Zhang, G. Feng, and H. Xu, "Observer design for nonlinear discrete-time systems: Immersion and dynamic observer error linearization techniques," *Int. J. Robust Nonlinear Control*, vol. 20, no. 5, pp. 504–514, Mar. 2010. doi: 10.1002/rnc.1443.

[15] C. Murguia, R. H. B. Fey, and H. Nijmeijer, "Immersion and invariance observers with time-delayed output measurements," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 30, nos. 1–3, pp. 227–235, 2016. doi: 10.1016/j.cnsns.2015.06.005.

[16] H. Yang, Y. Zhang, J. Liang, J. Gao, P. Walker, and N. Zhang, "Sliding-mode observer based voltage-sensorless model predictive power control of PWM rectifier under unbalanced grid conditions," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5550–5560, Jul. 2018. doi: 10.1109/tie.2017.2774730.

[17] M. Corradini, G. Ippoliti, and G. Orlando, "A sliding mode observer-based icing detection and estimation scheme for wind turbines," *J. Dyn. Syst., Meas., Control*, vol. 140, no. 1, 2018, Art. no. 014502. doi: 10.1115/1.4037387.

[18] H. Deng, Q. Li, W. Chen, and G. Zhang, "High-order sliding mode observer based OER control for PEM fuel cell air-feed system," *IEEE Trans. Energy Convers.*, vol. 33, no. 1, pp. 232–244, Mar. 2018. doi: 10.1109/tec.2017.2742144.

[19] Y. Huangfu, J. Xu, D. Zhao, Y. Liu, and F. Gao, "A novel battery state of charge estimation method based on a super-twisting sliding mode observer," *Energies*, vol. 11, no. 5, p. 1211, 2018. doi: 10.3390/n11051211.

[20] H. K. Khalil and L. Praly, "High-gain observers in nonlinear feedback control," *Int. J. Robust Nonlinear Control*, vol. 24, no. 6, pp. 993–1015, Apr. 2014.

[21] M. Farza, M. M'Saad, T. Ménard, A. Ltaief, and T. Maatoug, "Adaptive observer design for a class of nonlinear systems. Application to speed sensorless induction motor," *Automatica*, vol. 90, pp. 239–247, Apr. 2018. doi: 10.1016/j.automatica.2017.12.058.

[22] H. Liu and H. K. Khalil, "Output feedback stabilization using super-twisting control and high-gain observer," *Int. J. Robust Nonlinear Control*, vol. 29, no. 3, pp. 601–617, 2019. doi: 10.1002/rnc.4240.

[23] A. Minowa and M. Toda, "A high-gain observer-based approach to robust motion control of towed underwater vehicles," *IEEE J. Ocean. Eng.*, to be published. doi: 10.1109/joe.2018.2859458.

[24] S. Sriprang, B. Nahid-Mobarakeh, S. Pierfederici, N. Takorabet, N. Bizon, P. Kumam, P. Mungporn, and P. Thounthong, "Robust flatness control with extended Luenberger observer for PMSM drive," in *Proc. IEEE Transp. Electrific. Conf. Expo, Asia–Pacific*, Bangkok, Thailand, Jun. 2018, pp. 1–8.

[25] C. Afri, V. Andrieu, L. Bako, and P. Dufour, "State and parameter estimation: A nonlinear Luenberger observer approach," *IEEE Trans. Autom. Control*, vol. 62, no. 2, pp. 973–980, Feb. 2017. doi: 10.1109/tac.2016.2566804.

[26] N. Ding, W. Chen, Y. Zhang, G. Xu, and F. Gao, "An extended Luenberger observer for estimation of vehicle sideslip angle and road friction," *Int. J. Veh. Des.*, vol. 66, no. 4, pp. 385–414, 2014. doi: 10.1504/ijvd.2014.066071.

[27] V. Miskovic, V. Blasko, T. M. Jahns, A. H. C. Smith, and C. Romenesko, "Observer-based active damping of *LCL* resonance in grid-connected voltage source converters," *IEEE Trans. Ind. Appl.*, vol. 50, no. 6, pp. 3977–3985, Nov.-Dec. 2014. doi: 10.1109/TIA.2014.2317849.

[28] J. Jin, "Observer design for satellite angular rate estimation using the state-dependent riccati equation method," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 3, pp. 2496–2501, Jul. 2015. doi: 10.1109/taes.2015.140159.

[29] M. H. Korayem, N. Y. Lademakhi, and S. R. Nekoo, "Application of the state-dependent Riccati equation for flexible-joint arms: Controller and estimator design," *Optim. Control Appl. Methods*, vol. 39, no. 2, pp. 792–808, 2018. doi: 10.1002/oca.2377.

[30] M. Hassan and M. Zribi, "An observer-based controller for nonlinear systems: A gain scheduling approach," *Appl. Math. Comput.*, vol. 237, pp. 695–711, Jun. 2014. doi: 10.1016/j.amc.2014.03.130.

[31] M. F. Hassan, "A new approach for secure communication using constrained hyperchaotic systems," *Appl. Math. Comput.*, vol. 246, pp. 711–730, Nov. 2014. doi: 10.1016/j.amc.2014.08.029.

[32] M. F. Hassan, "Observer design for constrained nonlinear systems with application to secure communication," *J. Franklin Inst.*, vol. 351, no. 2, pp. 1001–1026, Feb. 2014. doi: 10.1016/j.jfranklin.2013.10.001.

[33] J. M. Holte, "Discrete Gronwall lemma and applications," in *Proc. MAA North Central Section Meeting Univ.*, North Dakota, ND, USA, 2009, pp. 1–7.

[34] P. Kundur, *Power System Stability and Control*. New York, NY, USA: McGraw-Hill, 1994.

[35] M. Farza, M. M'Saad, and L. Rossignol, "Observer design for a class of MIMO nonlinear systems," *Automatica*, vol. 40, no. 1, pp. 135–143, 2004. doi: 10.1016/j.automatica.2003.08.008.

[36] M. Hassan, "Observer-based controller for discrete-time systems: A state dependent Riccati equation approach," *Nonlinear Dyn.*, vol. 70, no. 1, pp. 693–707, 2012. doi: 10.1007/s11071-012-0488-x.

[37] K. Reif, F. Sonnemann, and R. Unbehauen, "An EKF-based nonlinear observer with a prescribed degree of stability," *Automatica*, vol. 34, no. 9, pp. 1119–1123, 1998. doi: 10.1016/s0005-1098(98)00053-3.

MOHAMED F. HASSAN was born in Cairo, Egypt, in 1947. He received the B.Sc. and M.Sc. degrees in electrical engineering from the Faculty of Engineering, Cairo University, Cairo, in 1970 and 1973, respectively, and the D.Sc. (Doctor d'Etat) degree in systems science and automatic control from Paul Sabatier University, Toulouse, France, in 1978.

He has been a Faculty Member with the Electronics and Communication Department, Faculty of Engineering, Cairo University, since 1978, where he has been a Professor, since 1988. He is currently on leave from Cairo University and joined Kuwait University, in 2002. He is the author or the coauthor of more than 170 technical papers in most of the international journals and conferences, two books, and seven book chapters. His current research interests include large-scale systems, optimization theory, stochastic control theory, estimation theory, and nonlinear systems.

Dr. Hassan is a member of the Egyptian Engineering Society and the Egyptian Society for Applied Operation research. He was a recipient of the 1981 Science State Prize for outstanding research in Engineering and the State Medal of Science and Arts (First Class), Egypt. He was nominated to receive the distinct "American Medal of Honor" from the American Biographical Institute, in 2002. He has been listed in the 1983 edition of Mathematics: Who's Who, the 2007 Academic Keys Who's Who in Engineering Higher Education, and the 2007 Who's Who in Science and Engineering.

MOHAMMAD HAMMUDA received the B.Sc. degree in electrical engineering and the M.Sc. degree in system and process control from Kuwait University, in 2014 and 2018, respectively. His research interests include nonlinear observers design, observer-based controllers, and nonlinear controllers design.