Sign-regressor Wilcoxon and Sign-Sign Wilcoxon

Upendra Kumar Sahoo
Electronics and Communication Engineering Department
National Institute of Technology
Rourkela,India
Email: uksahoo01@gmail.com

Ganapati Panda
School of Electrical Sciences
IIT Bhubaneswar,India
Email: ganapati.panda@gmail.com

Bernard Mulgrew
IDCOM
University of Edinburgh,UK
Email: B.Mulgrew@ed.ac.uk

***Abstract*—It is known that sign sign LMS and sign regressor LMS are faster than LMS. Inspiring from this idea we have proposed sign regressor Wilcoxon and sign-sign wilcoxon which are robust against the outlier present in the desired data and also faster than Wilcoxon and sign Wilcoxon norm. It had applied to varities of linear and nonlinear system identification problems with Gaussian noise and impulse noise present in the desired. The simulation results are compared among Wilcoxon,sign Wilcoxon and proposed sign-sign Wilcoxon and sign-regressor Wilcoxon. From simulation results it has proved that the proposed techniques are robust against outlier in the desired data and convergence speed are faster compared to other two norms.**

***Index Terms*—Sign-regressor Wilcoxon, sign-sign Wilcoxon, sign Wilcoxon, Wilcoxon**

I. INTRODUCTION

Wilcoxon norm is a robust norm used by statistician for regression analysis [1], [2]. In [3] Wilcoxon learning machine is designed which is robust against the outlier present in the desired. In [4] wilcoxon LMS is proposed to estimate the parameter of system in presence of outlier in desired data. It is known that the convergence speed of sign-regressor LMS and sign-sign LMS [5] are faster than the LMS but it's performance decreases with respect to LMS.

Inspiring form the above [5] literature we have proposed sign-sign Wilcoxon and sign-regressor Wilcoxon which are robust against the outlier in the desired data and convergence speed are faster than the wilcoxon norm and sign Wilcoxon norm. These are simulated for varieties of the system identification problems in presence of outlier. The performance of the proposed techniques are giving the less performance than the Wilcoxon norm and sign Wilcoxon norm in some systems but in case of some other systems the performance are better than the Wilcoxon norm and sign Wilcoxon norm.

The notation of the paper is same as used in [6]. The organization of the paper is as follows. In section II a short description of the wilcoxon norm and sign wilcoxon norm is given. The proposed techniques are given in section III. In section IV the proposed techniques are applied for system identification and it's simulation results are discussed. The paper is cocluded in section V.

II. WILCOXON NORM

To define the Wilcoxon norm a score function is required. The score function is $\varphi \begin{bmatrix} 0 & 1 \end{bmatrix} \to \Re$ which is non decresing and bounded

$\int \varphi^2(u)du < \infty$

The score value is

$$a(i) = \varphi\left(i/l + 1\right) \tag{1}$$

Where i is a fixed positive integer. The following norm can be shown as a pseuso norm [1].

$$\|v\|_w = \sum_{i=1}^{l} a(R(v_i))v_i \tag{2}$$

Where $v = \begin{bmatrix} v_1 & v_2 & \cdots & v_l \end{bmatrix}$ Here l is the size of the vector and $R(v_i)$ is the rank of v_i among $v_1, v_2, \cdots, v_l$.

$v^{(1)} \le v^{(2)} \le \cdots \le v^{(l)}$ is the ordered value of the vector.

$$a(i) = \varphi\left(R(v_i)/\left(l+1\right)\right) \tag{3}$$

and

$$\varphi(u) = \sqrt{12}\left(u - 0.5\right) \tag{4}$$

For the sign Wilcoxon norm only the score function of the preveious norm (4) is changed to

$$\varphi(u) = sign(u - 0.5) \tag{5}$$

III. PROPOSED TECHNIQUES

A. problem formulation

Let the model of a system is

$$\begin{array}{l} d_i = \mathbf{u}_i^T \mathbf{w} + e_i + v_i \\ i = 1, 2, \cdots, n \end{array} \tag{6}$$

$\mathbf{w} \in \Re^p, \mathbf{u}_i \in \Re^p$ are column matrix. Here p is the order of the system.

In this linear model $\mathbf{u}_i$ is the input to the system at i^{th} time which is a tap delay system of order p. Where e_i is the additive white Gaussian noise present in the system. The outlier is indicated by v_i which can think of as the impulse noise present in the system. d_i is the output of the system at i^{th} time.

The model (6) can be written in vector form like below

$$\mathbf{d} = \mathbf{U}^T\mathbf{w} + \mathbf{e} + \mathbf{v} \tag{7}$$

So in adaptive system identification problem we have to estimate the parameter $\mathbf{w}$ from the input $\mathbf{U}$ and output $\mathbf{d}$. From geometric point of view we can formulate the problem as below. We have to find a point on the span of the input space for which the the distance between the desired point and the point on the input space will be minimum, that is called the projection of the desired to the input space. Let it is $\mathbf{d}_u$. Now the optimum parameter is $\mathbf{w} = \mathbf{d}_u^{-1}\mathbf{U}$.

In order to find the point on the span space of the input that will minimize the distance a norm is used. In conventional technique L_2 norm is used but this norm is very sensitive to the impulse noise present in the desired. But Wilcoxon norm and sign wilcoxon norm are robust to the impulse noise. For estimating the parameter of a system using Wilcoxon norm as cost function gradient based technique has been using [1], [3], [4]. The update equation is

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu \left(\sum_{j=1}^{L} \varphi \left(R \left(e_{iL+j} \right) \right) \mathbf{u}_{iL+j} \right) \tag{8}$$

For sign wilcoxon norm based estimation the update equation (8) will changed to

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu \left(\sum_{j=1}^{L} \varphi \left(R \left(e_{iL+j} \right) \right) \mathbf{u}_{iL+j} \right) \tag{9}$$

The update equation for LMS case is

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu \mathbf{u}_i e_i \tag{10}$$

Equation (9),(9) can modify to matrix vecor multiplication form as given below

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu \begin{bmatrix} \mathbf{u}_{iL+1}^T \\ \mathbf{u}_{iL+2}^T \\ \cdots \\ \mathbf{u}_{(i+1)L}^T \end{bmatrix}^T \begin{bmatrix} \varphi \left(R \left(e_{iL+1} \right) \right) \\ \varphi \left(R \left(e_{iL+2} \right) \right) \\ \cdots \\ \varphi \left(R \left(e_{(i+1)L} \right) \right) \end{bmatrix} \tag{11}$$

Which can be written like

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu \mathbf{U}_i^T \mathbf{s}_i \tag{12}$$

Where $\mathbf{U}_i = \begin{bmatrix} \mathbf{u}_{iL+1} & \mathbf{u}_{iL+2} & \cdots & \mathbf{u}_{iL+n} \end{bmatrix}$ and $\mathbf{s}_i^T = \begin{bmatrix} \varphi \left(R \left(e_{iL+1} \right) \right) & \varphi \left(R \left(e_{iL+2} \right) \right) & \cdots & \varphi \left(R \left(e_{iL+n} \right) \right) \end{bmatrix}$

In case of sign Wilcoxon norm we can write simillar matrix vector multiplication form like (12) as

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu \mathbf{U}_i^T sign \left(\mathbf{s}_i \right) \tag{13}$$

B. Sign-regressor Wilcoxon

Comparing (10),(12) and (13), we can find that $\mathbf{U}_i^T$ and $\mathbf{s}_i$ in wilcoxon update equation are acting like input and error in LMS respectively. It is known that sign regressor LMS and sign-sign LMS are faster than the LMS.

The update equation for sign regressor LMS is

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu sign \left(\mathbf{u}_i \right) e_i \tag{14}$$

Comparing (14) with (12) and taking sign of the $\mathbf{U}_i^T$ of (12) we designed the update equation for sign regressor Wilcoxon like below

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu sign \left(\mathbf{U}_i^T \right) \mathbf{s}_i \tag{15}$$

Changing the matrix vector multiplication part of (15) into summation of multiplication form we can get

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu \left(\sum_{j=1}^{L} \varphi \left(R \left(e_{iL+j} \right) \right) sign \left(\mathbf{u}_{iL+j} \right) \right) \tag{16}$$

This is the final equation for sign-regressor Wilcoxon.

C. Sign-sign Wilcoxon

The update equation for sign-sign LMS is

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu sign \left(\mathbf{u}_i \right) sign \left(e_i \right) \tag{17}$$

Comparing (17) with (12) and taking sign of the $\mathbf{U}_i^T$ and $\mathbf{s}_i$ in (12) we have designed the update equation of sign-sign wilcoxon

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu sign \left(\mathbf{U}_i^T \right) sign \left(\mathbf{s}_i \right) \tag{18}$$

Changing the matrix vector multiplication part of (18) into summation of multiplication term we will get

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu \left(\sum_{j=1}^{L} sign \left(\varphi \left(R \left(e_{iL+j} \right) \right) \right) sign \left(\mathbf{u}_{iL+j} \right) \right) \tag{19}$$

IV. Application for adaptive system Identification problem

In this section we have applied proposed techniques for adaptive system identification problem. The parameters of the system are

$$\mathbf{w} = \begin{bmatrix} 0.26 \\ 0.93 \\ 0.26 \end{bmatrix} \tag{20}$$

.

The 20000 input data is taken which is a random value between $(-0.5, 0.5)$. The output is generated passing the input data through a tap delay system like (6). Here e is the additive white Gaussian noise having SNR 30dB.The impulse noise v is a random value between $(-R, R)$ where R is a fixed maximum value. The number of data added with impulse noise per 100 data of the desired is called the percentage of the additive impulse noise. Simulation is done taking combination of random value between (-30,30) magnitude with 10% and 40% of total desired data. The simulation result is plotted between iteration and normalized MSD in dB. The given plot is obtained by averaging it over 20 independent experiments. The block of the data is 40 and the step size is 0.001.

Normalized MSD in dB is $10 \log \frac{\|\mathbf{w} - \hat{\mathbf{w}}\|_2^2}{\|\mathbf{w}\|_2^2}$. In addition to the linear system (6), following nonlinear system are also tested with the proposed techniques system-II

$$d_i = \tan\left(\mathbf{u}_i^T \mathbf{w} + e_i + v_i\right) \tag{21}$$

system-III

$$d_i = \tan\left(\mathbf{u}_i^T \mathbf{w}\right) + e_i + v_i \tag{22}$$

system-IV

$$d_i = \mathbf{u}_i^T \mathbf{w} + 0.2 \times \left(\mathbf{u}_i^T \mathbf{w}\right)^2 - 0.1 \times \left(\mathbf{u}_i^T \mathbf{w}\right)^3 + e_i + v_i \tag{23}$$

system-V

$$d_i = \mathbf{u}_i^T \mathbf{w} - 0.9 \times \left(\mathbf{u}_i^T \mathbf{w}\right)^3 + e_i + v_i \tag{24}$$

Fig.1-2 are for linear system (6). In this case we have given outlier 10% and 40% with magnitude between(-30,30). From the plot we can found that the proposed technique is robust and convergence speed is faster than the other two norm. The same result we can also found for other impulse noise percentage upto 60%.

Fig 3-4 are for nonlinear system-II. In this case the performance is degrading with comparision to linear system but in comparision the proposed technique is showing good performance with faster convergence than the other two. Fig 5-6 are for nonlinear system-III. In this case the performance is degrading but in comparision the proposed technique is showing good performance with faster convergence than the other two techniques. Fig 7-8 is for nonlinear system-IV. In this case the performance is degrading but in comparision the proposed technique is showing less performance with faster convergence than the other two technique. Fig 9-10 is for nonlinear system-V. In this case the performance is degrading more but in comparision the proposed technique is showing less performance with faster convergence.

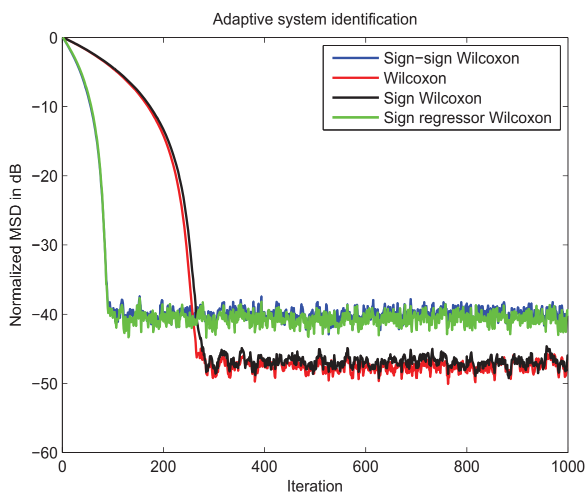

Fig. 1. 10% impulse noise is added in desired having magnitude a random value between (-30,30)

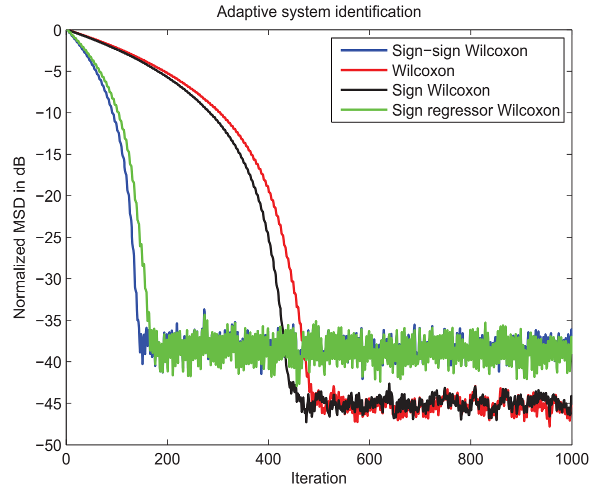

Fig. 2. 40% impulse noise is added in desired having magnitude a random value between (-30,30)

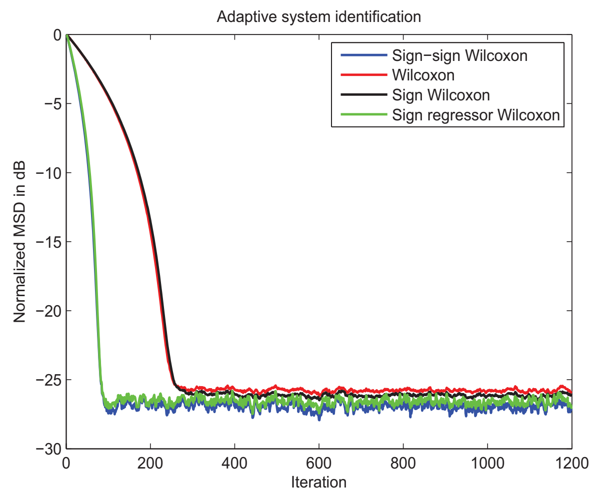

Fig. 3. 10% impulse noise is added in desired having magnitude a random value between (-30,30)

V. Conclusion

From simulation results we can conclude that the convergence speed of proposed techniques Sign-regressor Wilcoxon and sign-sign Wilcoxon are robust against the impulse noise present in desired data and convergence speed are faster than the sign Wilcoxon and Wilcoxon. The performance of the proposed techniques are dependent on the system we are using. In this paper only simulation results are shown but there are large work to be done like convergence analysis, stability and breakdown point of the algorithms with respect to impulse noise present in the desired data. It can consider as the future work of the algorithms. Since the convergence speed of the proposed techniques are very fast it can apply to the fast varying system.

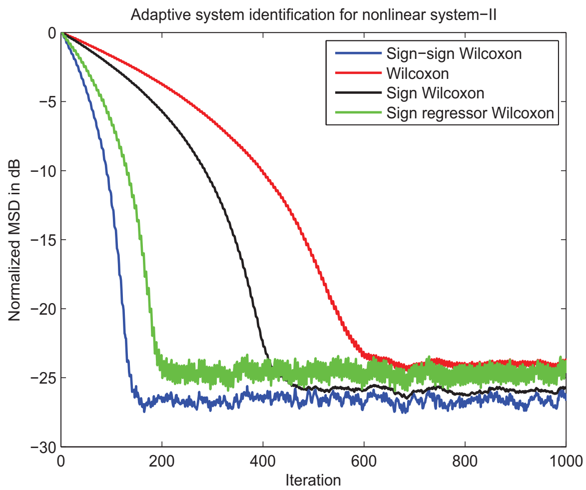

Fig. 4. 40% impulse noise is added in desired having magnitude a random value between (-30,30)

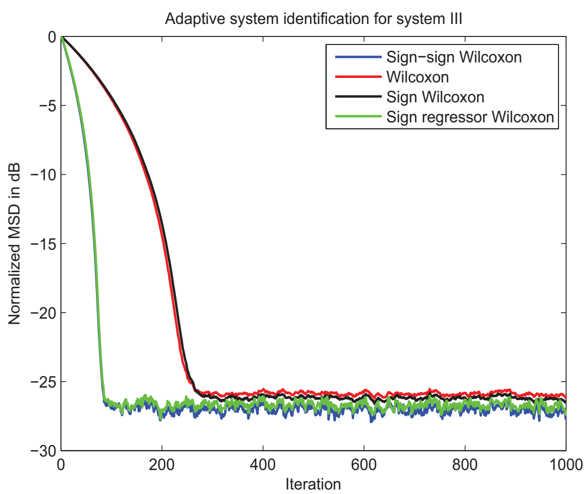

Fig. 5. 10% impulse noise is added in desired having magnitude a random value between (-30,30)

REFERENCES

[1] Robust nonparametric statistical methods, Thomas P Hettmansperger and Joseph W McKean, Kendall's Library of Statistics 5, 1998

[2] Mckean, J.W.: 'Robust analysis of linear models', Stat. Sci, 2004, 19, (4), pp. 562-570

[3] Preliminary study of Wilcoxon Learning Machines, Jer-Guang Hsieh, Yih-Lon Lin, and Jyh-Horng Jeng, Senior Member, IEEE, IEEE transactions on Neural Networks, VOL. 19, NO. 2, February 2008

[4] B. Majhi, G. Panda and B. Mulgrew,"'Robust identification using new Wilcoxon least mean square algorithm"'ELECTRONICS LETTERS, 12th March, 2009, Vol. 45, No. 6

[5] B.Farhang-Boroujeny, Adaptive Filters Theory and Applications,Wiley (February 2, 1999)

[6] S. Haykin, Adaptive Filter Theory. Englewood Cliffs, NJ: Prentice- Hall, 2001.

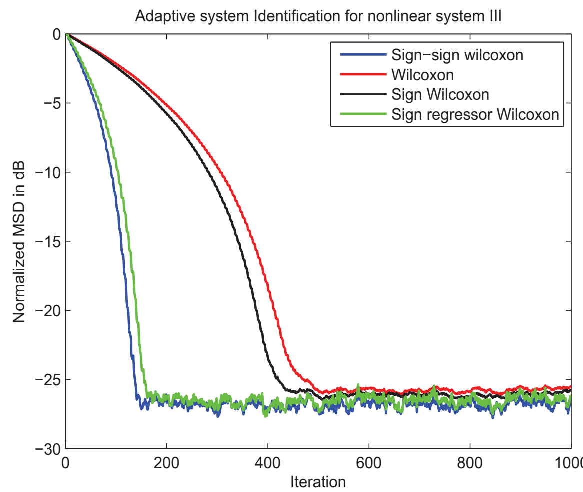

Fig. 6. 40% impulse noise is added in desired having magnitude a random value between (-30,30)

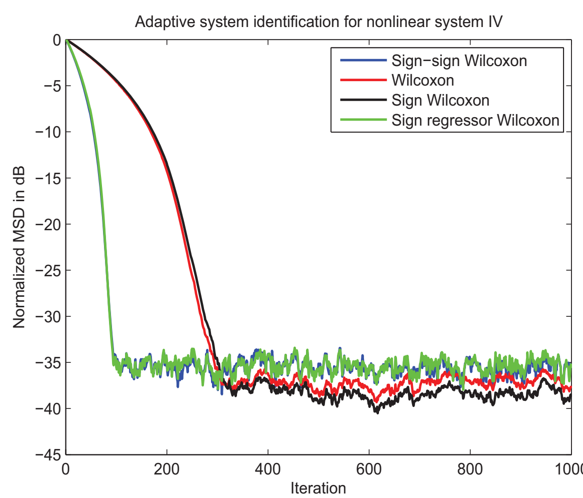

Fig. 7. 10% impulse noise is added in desired having magnitude a random value between (-30,30)

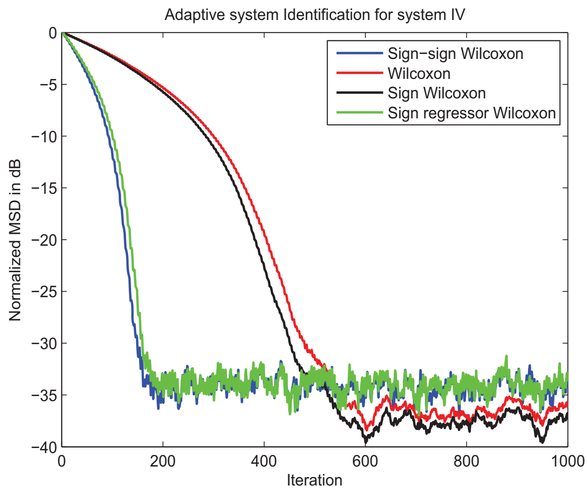

Fig. 8. 40% impulse noise is added in desired having magnitude a random value between (-30,30)

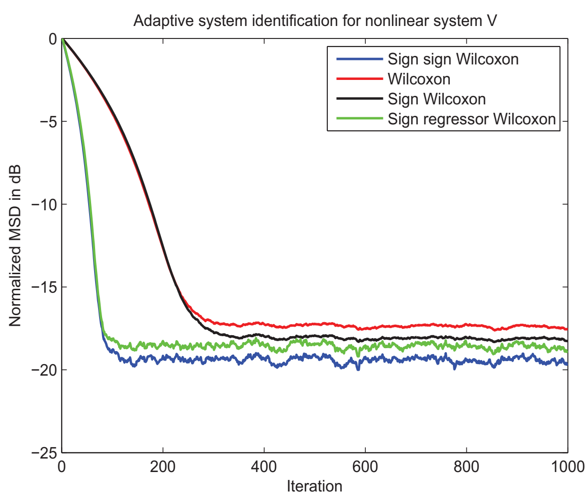

Fig. 9. 10% impulse noise is added in desired having magnitude a random value between (-30,30)

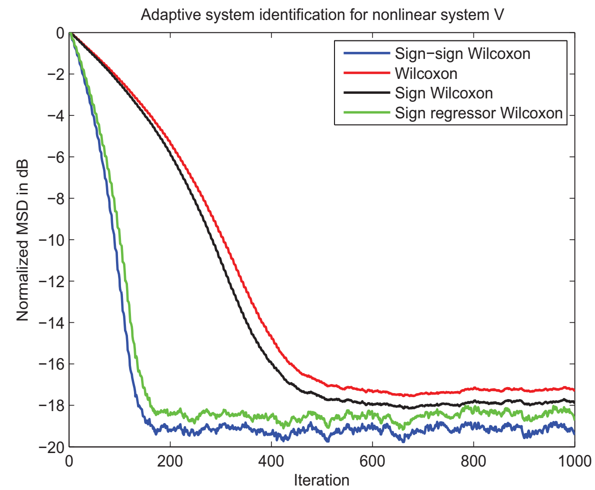

Fig. 10. 40% impulse noise is added in desired having magnitude a random value between (-30,30)