Multiple Sensor Fusion using Adaptive Divided Difference Information Filter

Aritro Dey, Smita Sadhu and Tapan Kumar Ghoshal
Department of Electrical Engineering, Jadavpur University, Kolkata, 700032, India

Keywords: Sensor Fusion, Information Filter, Process Noise Covariance, Adaptive Filter, State Estimation.

Abstract: This paper addresses the problem of multiple sensor fusion in situations where the system dynamics suffers from unknown parameter variation. An adaptive nonlinear information filter has been proposed for such multi sensor estimation problems where the process noise covariance becomes unknown as a consequence of unknown parameter variation. The proposed filter, based on the Divided Difference interpolation formula, ensures satisfactory estimation performance by online adaptation of the unknown process noise covariance and makes sensor fusion successful. Efficacy of the proposed filter is demonstrated with the help of a tracking problem in a sensor fusion configuration. Results from Monte Carlo simulation indicate that though the process noise covariance is unknown, the performance of the proposed filter is demonstrably superior to its non adaptive version in the context of joint estimation of parameter and states.

1 INTRODUCTION

Sensor fusion is a conventional process of integration of information from multiple sensors (homogeneous as well as heterogeneous sensors) to provide sufficiently reliable and enriched knowledge of the unmeasured states of the system under observation. Sensor fusion is extensively employed method which may find many real life applications, like, target tracking in collaborative sensor networks (Vercauteren, 2005), sensor fusion in the fields of robot navigation, intelligent vehicle, surveillance (Lee, 2008). Formulation of an estimation problem and its solution is one of the central aspects of successful sensor fusion. The Information filter variant of state estimators is widely recommended for multiple sensor estimation (Jia, 2013), (Liu, 2011), (Ge, 2014) and plays a significant role in sensor fusion. Because of simple computation methodology and easy initialization (Anderson, 1979) Information filters are preferred over the traditional estimators with the standard error covariance form.

For multiple sensor estimation several nonlinear information filters viz., Unscented information filters (Lee, 2008) , Central Difference information filters (Liu, 2011), Cubature and higher order cubature information filters (Jia, 2013), (Ge, 2014) have been reported in literature where the task of multi sensor estimation is found to be satisfactory only when the process noise and the measurement noise covariances are precisely known. Improper choice of noise covariance deteriorates estimation results as it is also observed for nonlinear estimation problem using the traditional nonlinear filters.

Unavailability of the knowledge of process noise covariance because of unknown parameter variation or the process noise statistics is, therefore, a serious issue of multiple sensor fusion which needs attention. An arbitrary choice of process noise covariance degrades the estimation results and the nonlinear information filter may even face divergence.

This paper presents a new algorithm for Adaptive Divided Difference Information filter (ADDIF) which is intended for situations when enough knowledge of system dynamics is unavailable due to parametric uncertainty. In such cases the process noise covariance becomes unknown. The proposed filter based on divided difference interpolation formula (Nørgaard, 2000) ensures satisfactory estimation performance by adapting online the unknown process noise covariance with ensured positive definiteness. Because of the unavailability of the proof for convergence like other information filters the superiority of the proposed filter is established with the help of an extensive Monte Carlo simulation.

Dey A., Sadhu S. and Ghoshal T..
Multiple Sensor Fusion using Adaptive Divided Difference Information Filter.
DOI: 10.5220/0005537303980406
In *Proceedings of the 12th International Conference on Informatics in Control, Automation and Robotics* (ICINCO-2015), pages 398-406
ISBN: 978-989-758-122-9

In this paper the adaptation rule is mathematically established and incorporated in the algorithm of Divided Difference Information filter, alias, Central Difference Information filter (Liu, 2011) to circumvent the problem of the unknown process noise covariance. The method for adaptation of process noise covariance ($\boldsymbol{Q}$) in the proposed filter is based on 'Maximum Likelihood Estimation' which is inspired from the early works on adaptive filters (Maybeck, 1982), (Mohamed, 1999) for linear signal models. The adaptive sigma point filters with standard error covariance form have been reported in the literatures which are developed extending the concept of adaptation for linear signal models. Adaptive UKF (Das, 2013), (Hajiyev, 2014), (Lee, 2005), (Soken, 2014) adaptive DDF (Lee, 2005) have been cited in literatures. A $\boldsymbol{Q}$ adaptive first order DDF is presented by (Lee, 2005) where the state residual has been used for adaptation. A robust adaptive second order DDF is presented by (Karlgaard, 2011) where, in lieu of adaptation, focus is on robustness. However, formulation of adaptive nonlinear information filter has not yet been reported in literature to the best knowledge of the authors.

The proposed $\boldsymbol{Q}$ adaptive information filter intended for sensor fusion has the following advantages: (i) Unlike Extended Information Filter and its higher order relatives computation of complex Jacobian and Hessian matrices are not required. (ii) As the proposed filter is based on Divided Difference Information filter it does not need tuning parameters like Unscented Information filter and can reportedly achieve same accuracy at a lower computational burden (Liu, 2011). (iii) Positive definiteness of the adapted process noise covariance is ensured. (iv) The proposed filter has the flexibility of multiple sensor estimation even in face of unknown noise covariance because of its dual aspect of information filter framework and $\boldsymbol{Q}$ adaptation algorithm.

2 ADAPTIVE DIVIDED DIFFERENCE INFORMATION FILTER

In this section the problem statement is provided followed by the solution methods which include the algorithm for the proposed filter.

2.1 Problem Statement

We consider an augmented nonlinear dynamic system as given below.

$$\boldsymbol{x}_k = \boldsymbol{f}(\boldsymbol{x}_{k-1}) + \boldsymbol{w}_k \tag{1}$$

$$\boldsymbol{y}_k^{\zeta} = \boldsymbol{g}^{\zeta}(\boldsymbol{x}_k) + \boldsymbol{v}_k^{\zeta} \tag{2}$$

Here $\boldsymbol{x}_k \in \boldsymbol{R}^n$ is an augmented state vector, By the term *augmented state vector*, it is meant that the unknown parameters have been concatenated with the state vector such that dimension of the augmented state vector is n. The difference equations corresponding to a typical unknown parameter ζ_k are considered to obey the random walk model, i.e., $\zeta_k = \zeta_{k-1} + \boldsymbol{w}_k^{\zeta}$, where $\boldsymbol{w}_k^{\zeta}$ is the noise term. $\boldsymbol{w}_k \in \boldsymbol{R}^n \sim (\boldsymbol{0}, \boldsymbol{Q}_k)$ indicates zero mean process noise (Gaussian) with unknown noise covariance. $\boldsymbol{y}_k^{\zeta} \in \boldsymbol{R}^m$ is the measurement available from the ζ^{th} sensor among M different sensors where $\zeta = 1, \cdots, M$. The measurement noise of each sensor is considered to be white (Gaussian) and denoted as, $\boldsymbol{v}_k^{\zeta} \in \boldsymbol{R}^m \sim (\boldsymbol{0}, \boldsymbol{R}_{\zeta})$. It is also considered the covariances of the sensors are known. The process noise covariance, $\boldsymbol{Q}_k$, however, remains unknown for parametric uncertainty and needs to be adapted.

2.2 Filtering Algorithm

Initialization:
Initialize $\hat{\boldsymbol{x}}_0, \hat{\boldsymbol{P}}_0, \hat{\boldsymbol{Q}}_0, \boldsymbol{R}_{\zeta}$
Time update steps:
calculate

$$\hat{\boldsymbol{S}}_x(\boldsymbol{k-1}) = CholeskyFactor(\hat{\boldsymbol{P}}_{k-1}) \tag{3}$$

Propagation of a-priori estimate of state:

$$\bar{x}_k = \tfrac{h^2-n}{h^2} f(\hat{x}_{k-1}) + \tfrac{1}{2h^2}\sum_{p=1}^{n}\left\{ \boldsymbol{f}(\hat{x}_{k-1} + h\hat{s}_{x,p}) + \boldsymbol{f}(\hat{x}_{k-1} - h\hat{s}_{x,p}) \right\} \tag{4}$$

where $\hat{\boldsymbol{s}}_{x,p}$ is p^{th} column of $\hat{\boldsymbol{S}}_x(\boldsymbol{k-1})$ and the interval length is chosen as, $h = \sqrt{3}$ for Gaussian distribution following (Nørgaard, 2000).
Propagation of a-priori error covariance:
The a-priori error covariance is

$$\bar{\boldsymbol{P}}_k = \left[\boldsymbol{S}_{x\hat{x}}^{(1)}(\boldsymbol{k}) \quad \boldsymbol{S}_{x\hat{x}}^{(2)}(\boldsymbol{k})\right]\left[\boldsymbol{S}_{x\hat{x}}^{(1)}(\boldsymbol{k}) \quad \boldsymbol{S}_{x\hat{x}}^{(2)}(\boldsymbol{k})\right]^T + \boldsymbol{Q} \tag{5}$$

where

$$\left\{\boldsymbol{S}_{x\hat{x}}^{(1)}(k)_{(i,j)}\right\} = \left\{\tfrac{1}{2h}\left(\boldsymbol{f}_i(\hat{x}_{k-1} + h\hat{s}_{x,j}) - \boldsymbol{f}_i(\hat{x}_{k-1} - h\hat{s}_{x,j})\right)\right\} \tag{6}$$

$\left\{\boldsymbol{S}_{x\hat{x}}^{(1)}(\boldsymbol{k})_{(i,j)}\right\}$ indicates the element (e_{ij}) of $\boldsymbol{S}_{x\hat{x}}^{(1)}(\boldsymbol{k}) \cdot \boldsymbol{S}_{x\hat{x}}^{(1)}(\boldsymbol{k})$ is first order approximation of the square root of a-priori error covariance based on interpolation formulae (Nørgaard, 2000). $\boldsymbol{S^{(1)}_{x\hat{x}}(k)}$ is to be computed using (6) for *i=1,…,n* and *j=1,…,n*.

$$\left\{\boldsymbol{S^{(2)}_{x\hat{x}}(k)}_{(i,j)}\right\} = \left\{\frac{\sqrt{h^2-1}}{2h^2}\left(\begin{array}{r}\left(f_i\left(\hat{x}_{k-1}+h\hat{s}_{x,j}\right)+f_i\left(\hat{x}_{k-1}-h\hat{s}_{x,j}\right)\right) \\ -2f_i\left(\hat{x}_{k-1}\right)\end{array}\right)\right\} \quad (7)$$

$\boldsymbol{S^{(2)}_{x\hat{x}}(k)}$ is the second order approximation to be computed in a similar way using (7).

The predicted information matrix $\boldsymbol{\overline{Z}_k}$ and information vector $\boldsymbol{\bar{z}_k}$ are related with the predicted estimate and error covariance as :

$$\boldsymbol{\overline{Z}_k} = \boldsymbol{\overline{P}_k^{-1}} \quad (8)$$

$$\boldsymbol{\bar{z}_k} = \boldsymbol{\overline{Z}_k \bar{x}_k} \quad (9)$$

Compute $\boldsymbol{\overline{S}_x(k)}$ such that it is a Cholesky factor of $\boldsymbol{\overline{P}_k}$. This factor has been involved for measurement update steps.

$$\boldsymbol{\overline{S}_x(k)} = CholeskyFactor(\boldsymbol{\overline{P}_k}) \quad (10)$$

Measurement update steps:
For $\zeta = 1,\cdots,M$ the following steps are to be executed:
Propagation of a-priori estimate of measurement:

$$\boldsymbol{\bar{y}^{\zeta}_k} = \frac{h^2-n}{h^2}\boldsymbol{g^{\zeta}}(\boldsymbol{\bar{x}_k}) + \frac{1}{2h^2}\sum_{p=1}^{n}\left\{\boldsymbol{g^{\zeta}}\left(\boldsymbol{\bar{x}_k}+h\boldsymbol{\bar{s}_{x,p}}\right)+\boldsymbol{g^{\zeta}}\left(\boldsymbol{\bar{x}_k}-h\boldsymbol{\bar{s}_{x,p}}\right)\right\} \quad (11)$$

The cross covariance is

$$\boldsymbol{P^{xy}_k} = \left[\boldsymbol{\overline{S}_x(k)}\right]\left[\boldsymbol{S^{(1)}_{y\bar{x}}(k)}\right]^T \quad (12)$$

where

$$\left\{\boldsymbol{S^{(1)}_{y\bar{x}}(k)}_{(i,j)}\right\} = \left\{\frac{1}{2h}\left(\boldsymbol{g^{\zeta}_i}\left(\boldsymbol{\bar{x}_k}+h\boldsymbol{\bar{s}_{x,j}}\right)-\boldsymbol{g^{\zeta}_i}\left(\boldsymbol{\bar{x}_k}-h\boldsymbol{\bar{s}_{x,j}}\right)\right)\right\} \quad (13)$$

$\left\{\boldsymbol{S^{(1)}_{y\bar{x}}(k)}_{(i,j)}\right\}$ in the similar way of (6) for *i=1,…,m* and *j=1,…,n*
Computation of Pseudo Measurement Matrix:
Now, to make the information contribution equations compatible to those of the EIF, a pseudo-measurement matrix is defined by (14) following the approach of (Lee, 2008).

$$\boldsymbol{\Psi^{\zeta}_k} = \left(\boldsymbol{\overline{P}^{-1}_k P^{xy}_k}\right)^T \quad (14)$$

Computation of information state contribution and its associated matrix:
Each sensor presents local information state contribution and its associated information matrix as

$$\boldsymbol{\varphi^{\zeta}_k} = \left(\boldsymbol{\Psi^{\zeta}_k}\right)^T\left(\boldsymbol{R^{\zeta}_k}\right)^{-1}\left(\vartheta^{\zeta}_k + \boldsymbol{\Psi^{\zeta}_k \bar{x}_k}\right) \quad (15)$$

$$\boldsymbol{\Phi^{\zeta}_k} = \left(\boldsymbol{\Psi^{\zeta}_k}\right)^T\left(\boldsymbol{R^{\zeta}_k}\right)^{-1}\boldsymbol{\Psi^{\zeta}_k} \quad (16)$$

Multi Sensor Estimation:
For reliable estimation the information regarding the measurements obtained from all the sensors are combined using the Divided Difference information filter. The decentralized approach has been followed for multiple sensor estimation to economize computational effort.
As described in the problem statement, measurements are available from ζ^{th} sensor where $\zeta = 1,\cdots,M$. The local information state contribution and its associated information matrix from each sensor can be obtained by (15) and (16).
The measurement update for the information vector and information matrix after fusion is simply expressed as a linear combination of these local information contribution terms by:

$$\boldsymbol{\hat{z}_k} = \boldsymbol{\bar{z}_k} + \sum_{\zeta=1}^{M}\boldsymbol{\varphi^{\zeta}_k} \quad (17)$$

$$\boldsymbol{\hat{Z}_k} = \boldsymbol{\overline{Z}_k} + \sum_{\zeta=1}^{M}\boldsymbol{\Phi^{\zeta}_k} \quad (18)$$

The a posteriori estimates of systems state and error covariance matrix are extracted using the formula by:

$$\boldsymbol{\hat{x}_k} = \boldsymbol{\hat{Z}^{-1}_k \hat{z}_k} \quad (19)$$

$$\boldsymbol{\hat{P}_k} = \boldsymbol{\hat{Z}^{-1}_k} \quad (20)$$

Computation of state residual:
The state residual is defined as

$$\boldsymbol{\rho_k} = \boldsymbol{\hat{x}_k} - \boldsymbol{\bar{x}_k} \quad (21)$$

Adaptation of process noise covariance:
Using the estimated residual covariance from a sliding window (size *N*) the adapted $\boldsymbol{Q}$ can be expressed as

$$\boldsymbol{\hat{Q}_k} = \frac{1}{N}\sum_{j=k-N+1}^{k}\boldsymbol{\rho_k \rho^T_k} \quad (22)$$

The adapted $\boldsymbol{\hat{Q}_k}$ of current instant can be used in (5) to refine $\boldsymbol{\overline{P}_k}$ so that the measurement update can be further refined by re-computation of (11) to (20). Adaptation step is mathematically derived and provided in the subsection 2.3.

2.3 Q –Adaptation Steps

The $\boldsymbol{Q}$ adaptation formula used in the proposed algorithm is derived using MLE technique. The steps followed for derivation of adapted $\boldsymbol{Q}$ are inspired from the work of (Maybeck, 1982), (Mohamed, 1999) for linear signal models. The probability density function of the measurements conditioned on adaptive parameter, α at specific epoch k is chosen based on innovation sequence. It should be borne in mind that the algorithm presented in the paper is meant for multiple sensor fusion problems. Hence, for the simplicity of the mathematical derivation we augment all the available measurements to get a single measurement vector as $\vartheta_k = \begin{bmatrix} \vartheta_k^1 & \vartheta_k^2 & \dots & \vartheta_k^M \end{bmatrix}^T$ with order mM. Therefore, the corresponding measurement noise covariance becomes $\boldsymbol{R}_k = diag\left(\boldsymbol{R}_k^1, \boldsymbol{R}_k^2, \dots, \boldsymbol{R}_k^M\right)$ and the pseudo measurement matrix can be expressed as $\boldsymbol{\Psi}_k = diag\left(\boldsymbol{\Psi}_k^1, \boldsymbol{\Psi}_k^2, \dots, \boldsymbol{\Psi}_k^M\right)$

The likelihood function is chosen following (Mohamed, 1999) as

$$\boldsymbol{P}_{(y|\alpha)_k} = \frac{1}{\sqrt{(2\pi)^{mM}\left|\boldsymbol{C}_{\vartheta_k}\right|}} \exp\left(-\tfrac{1}{2}\vartheta_k^T \boldsymbol{C}_{\vartheta_k}^{-1} \vartheta_k\right) \tag{23}$$

or,

$$\ln\left(\boldsymbol{P}_{(y|\alpha)_k}\right) = -\frac{1}{2}\left\{mM\ln(2\pi) + \ln\left|\boldsymbol{C}_{\vartheta_k}\right| + \vartheta_k^T \boldsymbol{C}_{\vartheta_k}^{-1} \vartheta_k\right\} \tag{24}$$

Multiplying both sides with −2 and neglecting the constant term we get

$$E = \ln\left|\boldsymbol{C}_{\vartheta_k}\right| + \vartheta_k^T \boldsymbol{C}_{\vartheta_k}^{-1} \vartheta_k \tag{25}$$

Innovation sequence has been considered inside a window size N as the filter uses a fixed length memory. The innovation inside the window will be summed. Therefore, the Maximum Likelihood condition becomes:

$$\min\left\{\sum_{j=j_0}^{k}\left(\ln\left|\boldsymbol{C}_{\vartheta_j}\right| + \vartheta_j^T \boldsymbol{C}_{\vartheta_j}^{-1} \vartheta_j\right)\right\} \tag{26}$$

which results in

$$\sum_{j=j_0}^{k}\left[tr\left\{\boldsymbol{C}_{\vartheta_j}^{-1}\left(\frac{\partial \boldsymbol{C}_{\vartheta_j}}{\partial \boldsymbol{\alpha}_k}\right)\right\} - \vartheta_j^T \boldsymbol{C}_{\vartheta_j}^{-1}\left(\frac{\partial \boldsymbol{C}_{\vartheta_j}}{\partial \boldsymbol{\alpha}_k}\right)\boldsymbol{C}_{\vartheta_j}^{-1}\vartheta_j\right] = 0 \tag{27}$$

$$\sum_{j=j_0}^{k}\left[tr\left\{\boldsymbol{C}_{\vartheta_j}^{-1}\left(\frac{\partial \boldsymbol{C}_{\vartheta_j}}{\partial \boldsymbol{\alpha}_k}\right)\right\} - tr\left\{\boldsymbol{C}_{\vartheta_j}^{-1}\vartheta_j\vartheta_j^T \boldsymbol{C}_{\vartheta_j}^{-1}\left(\frac{\partial \boldsymbol{C}_{\vartheta_j}}{\partial \boldsymbol{\alpha}_k}\right)\right\}\right] = 0 \tag{28}$$

$$\sum_{j=j_0}^{k}\left[tr\left\{\left[\boldsymbol{C}_{\vartheta_j}^{-1} - \boldsymbol{C}_{\vartheta_j}^{-1}\vartheta_j\vartheta_j^T \boldsymbol{C}_{\vartheta_j}^{-1}\right]\left(\frac{\partial \boldsymbol{C}_{\vartheta_j}}{\partial \boldsymbol{\alpha}_k}\right)\right\}\right] = 0 \tag{29}$$

The formulae for matrix operation are given in(Mohamed, 1999). Here, tr indicate trace of matrix and $j_0=k-N+1$. The deduction of the relation between innovation covariance, $\boldsymbol{C}_{\vartheta_k}$ and the measurement noise covariance, $\boldsymbol{R}_k$ necessitates the augmented pseudo measurement matrix of the nonlinear measurement equation. The use of the pseudo measurement matrix is justified as reported in (Lee, 2008), (Soken, 2015).

Using the pseudo measurement matrix the innovation covariance can be represented as:

$$\boldsymbol{C}_{\vartheta_k} = \boldsymbol{R}_k + \boldsymbol{\Psi}_k \boldsymbol{P}_k^- \boldsymbol{\Psi}_k^T \tag{30}$$

For adaptation of $\boldsymbol{Q}$, the adaptive parameter $\boldsymbol{\alpha}$ is chosen as $\boldsymbol{\alpha}_i = \boldsymbol{Q}_{ii}$

$$\frac{\partial \boldsymbol{C}_{\vartheta_k}}{\partial \boldsymbol{Q}_{kk}} = \boldsymbol{\Psi}_k \frac{\partial}{\partial \boldsymbol{Q}_{kk}}\left(\overline{\boldsymbol{P}}_k\right)\boldsymbol{\Psi}_k^T \tag{31}$$

$$\frac{\partial \boldsymbol{C}_{\vartheta_j}}{\partial \boldsymbol{Q}_{kk}} = \boldsymbol{\Psi}_k \frac{\partial}{\partial \boldsymbol{Q}_{kk}}\left(\begin{matrix}\left[\boldsymbol{S}_{x\hat{x}}^{(1)}(\boldsymbol{k}) & \boldsymbol{S}_{x\hat{x}}^{(2)}(\boldsymbol{k})\right]\left[\boldsymbol{S}_{x\hat{x}}^{(1)}(\boldsymbol{k}) & \boldsymbol{S}_{x\hat{x}}^{(2)}(\boldsymbol{k})\right]^T \\ + \boldsymbol{Q}_k\end{matrix}\right)\boldsymbol{\Psi}_k^T \tag{32}$$

The term $\left[\boldsymbol{S}_{x\hat{x}}^{(1)}(\boldsymbol{k}) \quad \boldsymbol{S}_{x\hat{x}}^{(2)}(\boldsymbol{k})\right]\left[\boldsymbol{S}_{x\hat{x}}^{(1)}(\boldsymbol{k}) \quad \boldsymbol{S}_{x\hat{x}}^{(2)}(\boldsymbol{k})\right]^T$ is analogous to the a priori error covariance when process noise covariance is absent. It is assumed following the work of (Mohamed, 1999) that the within the estimation window the a priori error covariance is in steady state. Hence the derivative of this term may be ignored.

Substituting this value in the ML equation we get

$$\sum_{j=j_0}^{k}\left[tr\left\{\left[\boldsymbol{C}_{\vartheta_j}^{-1} - \boldsymbol{C}_{\vartheta_j}^{-1}\vartheta_j\vartheta_j^T \boldsymbol{C}_{\vartheta_j}^{-1}\right]\left(\boldsymbol{\Psi}_j \boldsymbol{I} \boldsymbol{\Psi}_j^T\right)\right\}\right] = 0 \tag{33}$$

Alternatively,

$$\sum_{j=j_0}^{k}\left[tr\left\{\boldsymbol{\Psi}_j^T\left[\boldsymbol{C}_{\vartheta_j}^{-1} - \boldsymbol{C}_{\vartheta_j}^{-1}\vartheta_j\vartheta_j^T \boldsymbol{C}_{\vartheta_j}^{-1}\right]\boldsymbol{\Psi}_j\right\}\right] = 0 \tag{34}$$

$$\Rightarrow \sum_{j=j_0}^{k}\left[tr\left\{\left[\boldsymbol{\Psi}_j^T \boldsymbol{C}_{\vartheta_j}^{-1}\boldsymbol{\Psi}_j - \boldsymbol{\Psi}_j^T \boldsymbol{C}_{\vartheta_j}^{-1}\vartheta_j\vartheta_j^T \boldsymbol{C}_{\vartheta_j}^{-1}\boldsymbol{\Psi}_j\right]\right\}\right] = 0 \tag{35}$$

$$\Rightarrow \sum_{j=j_0}^{k}\left[tr\left\{\overline{\boldsymbol{P}}_j^{-1}\boldsymbol{K}_j\boldsymbol{\Psi}_j - \overline{\boldsymbol{P}}_j^{-1}\boldsymbol{K}_j\vartheta_j\vartheta_j^T \boldsymbol{K}_j^T \overline{\boldsymbol{P}}_j^{-1}\right\}\right] = 0 \tag{36}$$

$$\Rightarrow \sum_{j=j_0}^{k}\left[tr\left\{\overline{\boldsymbol{P}}_j^{-1}\left(\boldsymbol{K}_j\boldsymbol{\Psi}_j\overline{\boldsymbol{P}}_j - \boldsymbol{K}_j\vartheta_j\vartheta_j^T \boldsymbol{K}_j^T\right)\overline{\boldsymbol{P}}_j^{-1}\right\}\right] = 0 \tag{37}$$

The term $\boldsymbol{K}_k\vartheta_k$ can also be represented as $\boldsymbol{K}_k\vartheta_k = \hat{\boldsymbol{x}}_k - \overline{\boldsymbol{x}}_k$, which the state residual.

$$\boldsymbol{\rho}_k = \hat{\boldsymbol{x}}_k - \overline{\boldsymbol{x}}_k \tag{38}$$

Equation (35) can, therefore, be expressed as

$$\sum_{j=j_0}^{k}\left[tr\left\{\overline{\boldsymbol{P}}_j^{-1}\left(\boldsymbol{K}_j\boldsymbol{\Psi}_j\overline{\boldsymbol{P}}_j - \boldsymbol{\rho}_j\boldsymbol{\rho}_j^T\right)\overline{\boldsymbol{P}}_j^{-1}\right\}\right] = 0 \tag{39}$$

The expression of $\overline{\boldsymbol{P}}_k$ ensures the positive definiteness of $\overline{\boldsymbol{P}}_k$. Therefore, above expression vanishes only when

$$\sum_{j=j_0}^{k}\left[tr\left\{\left(\boldsymbol{K}_j\boldsymbol{\Psi}_j\overline{\boldsymbol{P}}_j-\boldsymbol{\rho}_j\boldsymbol{\rho}_j^T\right)\right\}\right]=0 \tag{40}$$

$$\Rightarrow\sum_{j=j_0}^{k}\left[tr\left\{\left(\overline{\boldsymbol{P}}_j-\hat{\boldsymbol{P}}_j-\boldsymbol{\rho}_j\boldsymbol{\rho}_j^T\right)\right\}\right]=0 \tag{41}$$

$$\Rightarrow\overline{\boldsymbol{P}}_k-\hat{\boldsymbol{P}}_k=\frac{1}{N}\sum_{j=j_0}^{k}\left[\boldsymbol{\rho}_j\boldsymbol{\rho}_j^T\right] \tag{42}$$

Using (5) $\overline{\boldsymbol{P}}_k$ can be replace as

$$\begin{aligned}&\left[\boldsymbol{S}^{(1)}_{x\hat{x}}(\boldsymbol{k})\quad \boldsymbol{S}^{(2)}_{x\hat{x}}(\boldsymbol{k})\right]\left[\boldsymbol{S}^{(1)}_{x\hat{x}}(\boldsymbol{k})\quad \boldsymbol{S}^{(2)}_{x\hat{x}}(\boldsymbol{k})\right]^T\\&+\hat{\boldsymbol{Q}}_k-\hat{\boldsymbol{P}}_k=\frac{1}{N}\sum_{j=j_0}^{k}\left[\boldsymbol{\rho}_j\boldsymbol{\rho}_j^T\right]\end{aligned} \tag{43}$$

$$\begin{aligned}\hat{\boldsymbol{Q}}_k=&\frac{1}{N}\sum_{j=j_0}^{k}\left[\boldsymbol{\rho}_j\boldsymbol{\rho}_j^T\right]\\&+\left(\hat{\boldsymbol{P}}_k-\left[\boldsymbol{S}^{(1)}_{x\hat{x}}(\boldsymbol{k})\quad \boldsymbol{S}^{(2)}_{x\hat{x}}(\boldsymbol{k})\right]\left[\boldsymbol{S}^{(1)}_{x\hat{x}}(\boldsymbol{k})\quad \boldsymbol{S}^{(2)}_{x\hat{x}}(\boldsymbol{k})\right]^T\right)\end{aligned} \tag{44}$$

As recommended in (Mohamed, 1999) the term $\left(\hat{\boldsymbol{P}}_k-\left[\boldsymbol{S}^{(1)}_{x\hat{x}}(\boldsymbol{k})\quad \boldsymbol{S}^{(2)}_{x\hat{x}}(\boldsymbol{k})\right]\left[\boldsymbol{S}^{(1)}_{x\hat{x}}(\boldsymbol{k})\quad \boldsymbol{S}^{(2)}_{x\hat{x}}(\boldsymbol{k})\right]^T\right)$ becomes often low and may be negligible during steady state. Hence, adapted $\boldsymbol{Q}$ is approximately represented by

$$\hat{\boldsymbol{Q}}_k=\frac{1}{N}\sum_{j=j_0}^{k}\left[\boldsymbol{\rho}_j\boldsymbol{\rho}_j^T\right] \tag{45}$$

2.3.1 Notes on Adaptation

The expression (45) given above has been presented in a simplified approach. Mathematically derived expression of adapted $\boldsymbol{Q}$ is given by (44). The term $\left[\boldsymbol{S}^{(1)}_{x\hat{x}}(\boldsymbol{k})\quad \boldsymbol{S}^{(2)}_{x\hat{x}}(\boldsymbol{k})\right]\left[\boldsymbol{S}^{(1)}_{x\hat{x}}(\boldsymbol{k})\quad \boldsymbol{S}^{(2)}_{x\hat{x}}(\boldsymbol{k})\right]$ of a posteriori error covariance $\hat{\boldsymbol{P}}_k$ is implicitly dependent on $\hat{\boldsymbol{P}}_k$ and acquires a steady value (often low) as the filter approaches steady state. Therefore, ignoring their effect from the adapted $\boldsymbol{Q}$ is justified and does not induce large error in the adapted $\boldsymbol{Q}$. This approximation as mentioned in (Mohamed, 1999) for linear systems is followed here so that the symmetry and the positive definiteness of adapted $\boldsymbol{Q}$ can be ensured. Otherwise singularity cannot be over ruled. Because of this assumption the adaptation becomes more accurate as the filter reaches steady state.

It is to be noted that the expression (45) which same as (22) is appropriate only when the step index k is greater than or equal to the window length N. When the step index k is less than N, adaptation begins with available state residual. The window length is gradually increased till it reaches the desired window length N. Afterward the sliding window concept becomes appropriate.

The window size N, should be appropriately chosen considering several factors. A smaller window size generally ensures lower computational burden but reportedly may be prone to divergence. A larger window size ensures unbiased estimates. However, it is not suitable for short term variation in process noise covariance.

3 CASE STUDY

The performance of adaptive DDIF is demonstrated using a problem of multi sensor tracking of an aircraft which is executing a maneuvering turn. The dynamic model of the system is presented in two dimensional spaces as given in (Jia, 2013). The turn rate of the aircraft is considered to be unknown and time varying which makes the tracking problem significantly nonlinear. Therefore, this problem may be an appropriate one to validate the performance of the proposed $\boldsymbol{Q}$ adaptive information filter.

3.1 System Dynamics

The dynamic equation of the above mentioned tracking problem is presented below. The turn rate of the aircraft being unknown it is modelled as a state and augmented with the state vector of the system model. As knowledge about the nature of variation of the unknown turn is unavailable the augmented parameter is considered to follow a simple random walk model. The dynamic model is taken from the work of (Jia,2013).

$$\xi_k=\begin{bmatrix}1 & \frac{\sin(\omega_{k-1}\tau)}{\omega_{k-1}} & 0 & \frac{\cos(\omega_{k-1}\tau)-1}{\omega_{k-1}} & 0\\ 0 & \cos(\omega_{k-1}\tau) & 0 & -\sin(\omega_{k-1}\tau) & 0\\ 0 & \frac{1-\cos(\omega_{k-1}\tau)}{\omega_{k-1}} & 1 & \frac{\sin(\omega_{k-1}\tau)}{\omega_{k-1}} & 0\\ 0 & \sin(\omega_{k-1}\tau) & 0 & \cos(\omega_{k-1}\tau) & 0\\ 0 & 0 & 0 & 0 & 1\end{bmatrix}\xi_{k-1}+w_k \tag{46}$$

where the state vector $\xi_k=\left[p_{x_k}\quad v_{x_k}\quad p_{y_k}\quad v_{y_k}\quad \omega_k\right]^T$; p_{x_k} and p_{y_k} are the position in x and y coordinate; v_{x_k} and v_{y_k} are the corresponding velocity components at the instant k. τ indicates the time interval between two consecutive measurements. w_k is zero mean Gaussian noise (white) which indicates the modeling error. The process noise for this noise sequence is considered as

$$Q_k = \begin{bmatrix} \frac{0.1\tau^3}{3} & \frac{0.1\tau^2}{2} & 0 & 0 & 0 \\ \frac{0.1\tau^2}{2} & 0.1\tau & 0 & 0 & 0 \\ 0 & 0 & \frac{0.1\tau^3}{3} & \frac{0.1\tau^2}{2} & 0 \\ 0 & 0 & \frac{0.1\tau^2}{2} & 0.1\tau & 0 \\ 0 & 0 & 0 & 0 & q\tau \end{bmatrix} \quad (47)$$

Note, that the element $Q_k(5,5)$ is the noise covariance of corresponding augmented parameter, i.e., turn rate. As the turn rate is unknown and time varying the accurate knowledge of q is unavailable and has to be assumed for traditional non adaptive information filter. However, for simulation study we have considered $q = \left(1.323\times10^{-2} \text{ rad s}^{-1}\right)^2$ to generate the true state trajectories.

The trajectories of the aircraft are tracked by the fusion of the bearing angle signal from two tracking radars located in two different places. The measurement equation can be represented as

$$\theta_k^{\zeta} = \tan^{-1}\left(\frac{p_{y_k} - p_{y_{ref}}^{\zeta}}{p_{x_k} - p_{x_{ref}}^{\zeta}}\right) + v_k^{\zeta} \qquad \zeta = 1,2 \quad (48)$$

ζ indicates position of the ζ^{th} radar. $p_{y_{ref}}^1 = -10^4 m;\ p_{x_{ref}}^1 = -10^4 m;\ p_{y_{ref}}^2 = 10^4 m;\ p_{x_{ref}}^2 = 10^4 m$. The zero mean measurement noise (Gaussian) has covariances $R_1 = \left(\sqrt{30} mrad\right)^2$ and $R_2 = \left(\sqrt{40} mrad\right)^2$. The interval between two successive measurements, $\tau = 1\sec$.

3.2 Simulation Procedure

The proposed filter for sensor fusion has been evaluated with help of an extensive Monte Carlo simulation with 10000 runs. True state trajectories are generated with an initial state $x_0 = \left[1000\text{m} \quad 300\text{ms}^{-1} \quad 1000\text{m} \quad 0\text{ms}^{-1} \quad -0.05235 \text{ rad s}^{-1}\right]^T$ and the unknown element q_{true} .The filter is initialized with a Gaussian prior with mean x_0 and $\hat{P}_0$, where $\hat{P}_0 = diag(\left[100 \quad 10 \quad 100 \quad 10 \quad 10^{-4}\right])$. As the element q_{true} is practically unknown, we assume q_{filter} to be 20 times higher q_{true} during initialization of the filter. The choice of a high q_{filter} is justified as it indicates high degree of uncertainty about the nature of variation of unknown time varying turn rate.

Note that only the element of $\boldsymbol{Q}$ which is associated with the turn rate is unknown while the other elements are known. Therefore, we need to adapt the element of $\boldsymbol{Q}$ related to turn rate leaving the other known element frozen at the truth value. It can be verified from the derivation of adapted $\boldsymbol{Q}$ that partial differential is taken with respect to each diagonal element of $\boldsymbol{Q}$. Therefore, the adaptation formula can easily be reformulated only for the unknown elements. More details are provided in (Dey, 2014).

Root means square error (RMSE) for position, velocity and turn rate are computed for performance analysis of the proposed filter. The RMSE for position and velocity are computed using the formula given in (Jia, 2013).

$$RMSE = \sqrt{\frac{1}{N_{mc}}\sum_{i=1}^{N_{mc}}\left(\left(\xi_{k,i}\boldsymbol{e_j} - \hat{\xi}_{k,i}\boldsymbol{e_j}\right)^2 + \left(\xi_{k,i}\boldsymbol{e_l} - \hat{\xi}_{k,i}\boldsymbol{e_l}\right)^2\right)} \quad (49)$$

where *j=1* and *l=3* for RMSE of position estimation. For RMSE of velocity estimation *j=2* and *l=4*. RMSE for turn rate estimation is obtained with *j=5* and replacing the unit vector $\boldsymbol{e_l}$ by a zero vector.

Further investigation with this tracking problem revealed that this particular bearing only tracking problem is susceptible to track losses because of its measurement equation. It has been considered that the turn rate is unknown and time varying. As a consequence the trajectory of the aircraft become such that the difference between the bearing angle from two different radars may either be negligibly small or become closer to π. Practically the line of sight of two radars does not intersect each other to find the object in some of such situations. Consequently the measurement loses its uniqueness of information as the aircraft tracked by the radar cannot be specifically located in the atmosphere with the measured bearing angles. It is to be noted that in the work of (Jia, 2013) which has considered the same tracking problem track loss phenomenon has not been discussed. The authors have presented a representative run for illustration where track loss occurs for the non adaptive filter in ideal case when knowledge of process noise covariance is available. The performance of the proposed filter has also been compared to its non adaptive counter part in context of its susceptibility to track losses. To detect the occurrence of track loss a condition has been considered as given below.

$$\left\|\sqrt{\left(x_k - \hat{x}_k\right)^2 + \left(y_k - \hat{y}_k\right)^2}\right\|_{\infty} \geq \quad 800m \quad (50)$$

If the condition given by (50) is satisfied it is understood that the estimated trajectory fails to track the true trajectory of the aircraft. Such a situation for a representative run is illustrated by Figure 1 where in the ideal case the non adaptive DDIF with known $\boldsymbol{Q}$ fails to track the trajectory.

In case of Monte Carlo simulation the track loss cases are detected and those error sequences are omitted while calculating the RMS errors. The track loss count from 10000 Monte Carlo run are also presented for each filter in percentage.

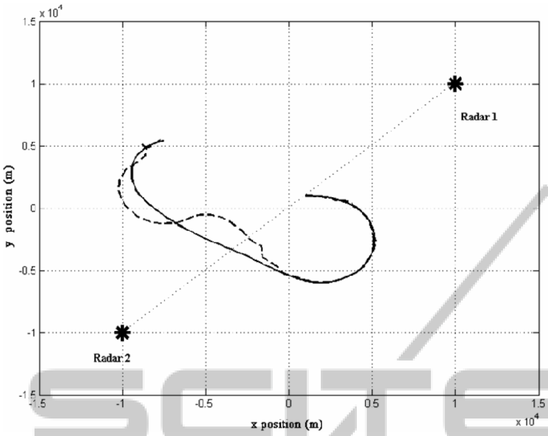

Figure 1: A representative run to illustrate track loss for the ideal case when non adaptive DDIF has known ***Q***.

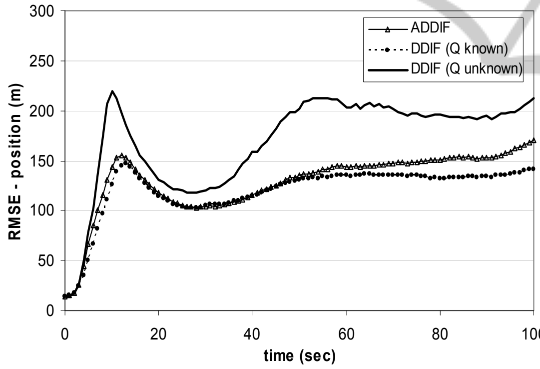

Figure 2: Comparison of RMS error (position estimation) of ADDIF & DDIF for 10000 MC runs.

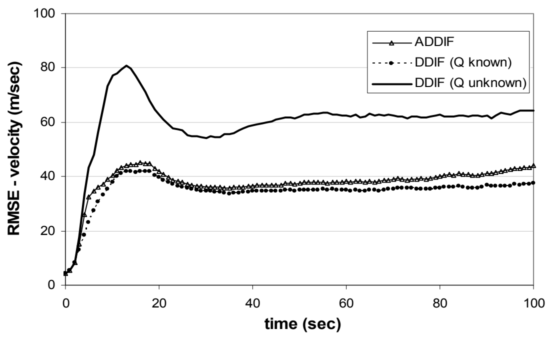

Figure 3: Comparison of RMS error (velocity estimation) of ADDIF & DDIF for 10000 MC runs.

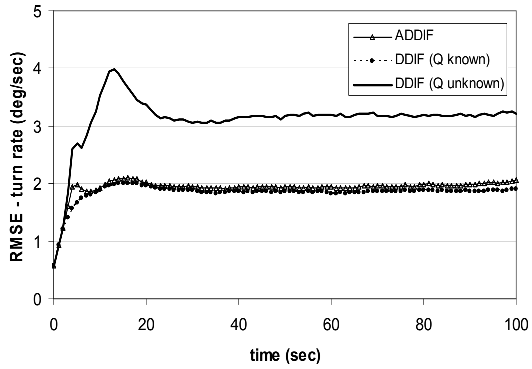

Figure 4: Comparison of RMS error (turn rate estimation) of ADDIF & DDIF for 10000 MC runs.

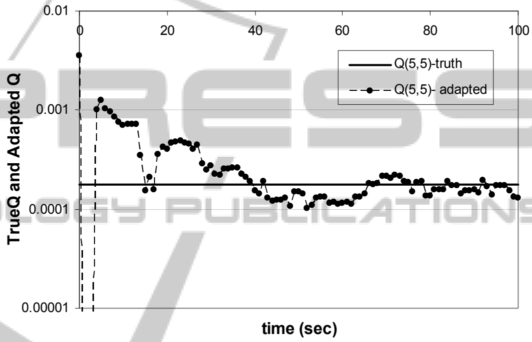

Figure 5: Plot of estimated process noise covariance ($\boldsymbol{Q}_{5,5}$) for a representative run.

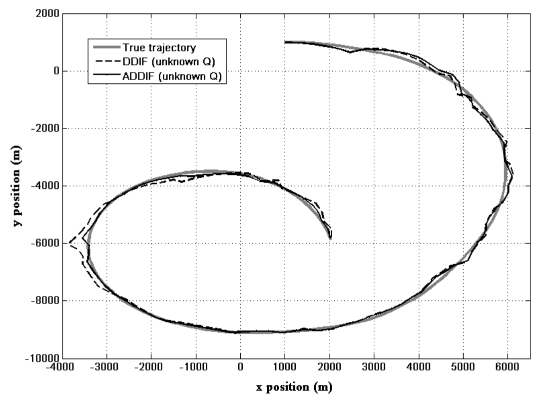

Figure 6: Tracking performance of ADDIF and non adaptive DDIF with unknown ***Q*** for a run.

3.3 Simulation Results

From the results of Monte Carlo simulation, performance of proposed ADDIF is compared with that of non adaptive DDIF in the situation when the turn rate of the aircraft is unknown and time varying.

- It has been observed from Figure 2, Figure 3 and Figure 4 that the performance of ADDIF is substantially superior to that of non adaptive DDIF as the RMSE for all three states converged to a lower steady state value within comparatively less time. RMSE of non adaptive DDIF deteriorates as the $\boldsymbol{Q}$ remains unknown due to unknown turn rate.
- It is to be also pointed out that though the elements of $\boldsymbol{Q}$ related to position and velocity are known RMSE of position and velocity for the non adaptive DDIF is degraded because of the implicit influence of poorly estimated turn rate.
- Figure 5 indicates that for ADDIF the unknown process noise element is converged to the truth value in about 30 sec.
- The RMSE results of ADDIF are also compared with non adaptive DDIF in the ideal situation when q is known only to the latter. Though this comparison may sound unusual, this comparison illumines on how far the performance ADDIF even with unknown $\boldsymbol{Q}$ is close to the performance of traditional filter in ideal situation with known $\boldsymbol{Q}$. It is demonstrated that the RMSE of ADDIF for all the states are very closed to that nature of RMSE of non adaptive filter in ideal condition. The initial mismatch in RMSE is because of the time taken for adapted $\boldsymbol{Q}$ to converge.
- It is also found from the Monte Carlo simulation that the track loss cases cannot be ruled out even for the ideal situation when the non adaptive DDIF has the knowledge of $\boldsymbol{Q}$. In the MC simulation 1.7% of track loss has been observed for the ideal case. When $\boldsymbol{Q}$ is unknown, the percentage of track loss for ADDIF is 2.2% and that for non adaptive DDIF is 15%. The track loss percentage for ADDIF is comparable with the ideal case and substantially low compared its non adaptive version which is prone to track loss cases.

These observations indicate the superiority of ADDIF over non adaptive DDIF when Q remains unknown for parametric uncertainties.

4 CONCLUSIONS

An Adaptive Divided Difference Information filter has been proposed for multiple sensor fusion in face of unknown parameter variation and exemplified with the help of an aircraft tracking problem. The proposed filter is found to carry out multiple sensor estimation successfully by online adaptation of process noise covariance ($\boldsymbol{Q}$) where the knowledge of $\boldsymbol{Q}$ remains unavailable due to parametric uncertainty. The adapted $\boldsymbol{Q}$ from the filter converges on the true value of $\boldsymbol{Q}$ and continues to track it for subsequent time. The results from Monte Carlo study indicate that the RMS error performance of the proposed filter, as expected, is significantly superior to the non adaptive Divided Difference Information filter in face of unknown $\boldsymbol{Q}$. Because of the capability of adaptation, flexibility for multiple sensor estimation and good error settling performance the proposed filter may be a recommended for multiple sensor fusion for the systems affected by unknown parameter variation.

ACKNOWLEDGEMENTS

The First author thanks Council of Scientific & Industrial Research (CSIR), New Delhi, India for financial support and expresses his gratitude to Centre for Knowledge Based System, Jadavpur University, Kolkata, India for infrastructural support.

REFERENCES

T. Vercauteren, and X. Wang., 2005. Decentralized sigma-point information filters for target tracking in collaborative sensor networks. In *IEEE Trans. Signal Processing*, 53 (8), 2997-3009.

D. J. Lee, 2008. Nonlinear estimation and multiple sensor fusion using unscented information filtering. In *IEEE Signal Processing Letters*, 15, 861-864.

B. Jia, M. Xin, K. Pham, E. Blasch, & G. Chen, 2013. Multiple sensor estimation using a high-degree cubature information filter. In conference *SPIE Defense, Security, and Sensing. International Society for Optics and Photonics*, 87390T-87390T.

G. Liu, F. Worgotter and I. Markelic, 2011. Nonlinear estimation using central difference information filter. In *IEEE Workshop Statistical Signal Processing*. 593-596.

Q. Ge, Daxing Xu, and Chenglin Wen, 2014. Cubature information filters with correlated noises and their applications in decentralized fusion. In *Journal of Signal Processing*, 94, 434-444.

B. D. O. Anderson and J.B. Moore, 1979. *Optimal Filtering*, Prentice Hall, Englewood Cliffs, 1st edition.

M. Nørgaard, N. K. Poulsen, and O. Ravn, 2000. New developments in state estimation for nonlinear systems. In *Automatica*, 36 (11), 1627-1638.

P. S. Maybeck, 1982. *Stochastic models, estimation, and control*, Academic Press, New York, 1st edition.

A. H. Mohamed and K. P. Schwarz, 1999.Adaptive Kalman filtering for INS/GPS. In *Journal of geodesy*, vol. 73 (4), 193-203.

M. Das, S. Sadhu, T. K. Ghoshal, 2013. An Adaptive Sigma Point Filter for Nonlinear Filtering Problems. In *International Journal of Electrical, Electronics and Computer Engineering*, 2 (2), 13-19.

C. Hajiyev, H. E. Soken, 2014. Robust adaptive unscented Kalman filter for attitude estimation of pico satellites. In *International Journal of Adaptive Control and Signal Processing*, 28 (2), 107-120.

D. J. Lee, 2005. Nonlinear Bayesian filtering with applications to estimation and navigation,. PhD thesis, Texas A&M University.

H. Soken, S. Sakai, 2015. Adaptive Tuning of the UKF for Satellite Attitude Estimation. In *Journal of Aerospace Engineering*, 28 (3), DOI: 10.1061/(ASCE)AS.1943-5525.0000412.

C. D. Karlgaard, H. Schaub, 2011. Adaptive nonlinear Huber-based navigation for rendezvous in elliptical orbit, In *Journal of Guidance, Control, and Dynamics*, 34 (2), pp. 388-402.

A, Dey, S. Sadhu, T.K. Ghoshal, 2014. Adaptive Gauss Hermite Filter for Parameter Varying Nonlinear Systems. In *International Conference on Signal Processing and Communication*, 1-5, DOI: 10.1109/SPCOM.2014.6983948.