Robust Bounding Ellipsoidal Adaptive Constrained Least-squares Algorithm and its Performance Analysis

Sheng Zhang and Jiashu Zhang

Abstract—In this paper, the robust Bounding Ellipsoidal Adaptive Constrained least-squares (BEACON) algorithm based on new set-membership error bounds is proposed to improve robustness and steady-state misalignment in impulsive noise environments. The expressions for the steady-state excess mean square error (EMSE) of the proposed robust BEACON are obtained. Simulations on system identification and double-talk scenarios show that the robust BEACON algorithm considerably outperforms the traditional BEACON algorithm in the presence of impulsive noise and the theoretical predictions are in good agreement with simulation results.

Index Terms—Set-membership adaptive filtering, robust algorithm, performance analysis.

1. INTRODUCTION

Adaptive filters are frequently adopted in many applications such as system identification, adaptive control and adaptive echo cancellation (AEC), interference suppression in industrial and biomedical engineering [1, 2]. Whenever a large number of coefficients are required, especially for echo cancellation, it diminishes the usefulness of the adaptive algorithm owing to increased complexity. Several schemes for reducing the computational complexity of adaptive filters have emerged. Partial-coefficient-update techniques update a subset of the adaptive filter coefficients [3]. “A potential drawback of partial updates is the reduction in convergence speed” [4]. By maintaining a bound on the output error magnitude of an adaptive filter, set-membership (SM) adaptive filter coefficients are only updated when the a priori error magnitude exceeds a predetermined threshold. It reduces the power consumption and computational complexity. The objective of a SM filter is to estimate set membership function, which defines the SM filter’s performance specification [5, 6, 7].

The optimal bounding ellipsoids (OBE) algorithms approximate the membership function by tightly outer-bounding it with ellipsoids in the associated parameter space [7-9] that are well-established SM adaptive filter algorithms. A thorough review of the OBE family can be found in [10]. Among all the OBE algorithms, the BEACON algorithm [7, 11, 12] is particularly interesting since it is computationally more efficient than other OBE algorithms and because of its simple but efficient innovation check and the optimal weight calculation processes. However, as for the recursive least-squares (RLS) algorithm, the performance of the traditional BEACON algorithm is affected by outliers in the error signal samples that can be brought about by impulsive-noise

This work was partially supported by National Science foundation of China (Grant: 61271341) and by the Doctoral Innovation Fund of Southwest Jiaotong University.

The authors are with the Sichuan province Key Lab of Signal and Information Processing at Southwest Jiaotong University, Chengdu, 610031, China.(E-mail:zhangsheng@my.swjtu.edu.cn, jszhang@home.swjtu.edu.cn).

interference [13]. Recently, several algorithms of the normalized least mean squares (NLMS), affine projection algorithm (APA) and Quasi-Newton (QN) in the SM filter framework that are robust with respect to outliers can be found in [14-17]. Some recursive least M-estimate algorithms that are robust with respect to outliers can be found in [18].

In this paper, we propose the robust BEACON algorithm in impulsive noise environments. Suggested by the basic idea of [16, 17], the robust BEACON algorithm is presented to enhance the robustness against impulsive noise, which contains two bounds. One bound is active during the transient state, and the other is used in steady-state and in the presence of impulsive noise. Thus, it is robust with respect to outliers brought about by impulsive noise and it also fast tracks sudden system disturbances. Compared to the BEACON, a reduced steady-state misalignment in the robust BEACON is achieved. The second contribution of this work is the analysis of steady-state performance of the proposed robust BEACON with fixed set-membership low bound. Using the energy conservation approach [19-22], the resulting model accurately predicts the proposed algorithm steady-state behavior and reveals the effects of the outlier detection threshold (ODT) parameter. In addition, the probability of the update (PU) is considered, which suggests that the proposed algorithm possesses higher PU than the BEACON. Finally, experiments in system identification and double-talk scenarios are conducted to evaluate the performance of proposed algorithm.

The rest of the paper is organized as follows. In section 2, a brief review of the SM adaptive filtering and BEACON is presented. Section 3 describes the proposed robust BEACON algorithm. Performance analysis of the proposed algorithm is carried out in Section 4. Section 5 contains simulation results and conclusions are drawn in Section 6.

2. SET-MEMBERSHIP FILTER

Consider the case of system identification problem. The unknown coefficients and input signal at time instant n are denoted by $\mathbf{w}_{opt} = [w_0, w_1, \dots, w_{L-1}]^T$ and $\mathbf{x}_n = [x_n, x_{n-1}, \dots, x_{n-L+1}]^T$, respectively, where L is the filter length. The observed desired signal is assumed to be corrupted by an additive noise:

$$y_n = \mathbf{w}_{opt}^T \mathbf{x}_n + \nu_n,$$

where ν_n is the additive background noise. The estimated error between the output of unknown system and of the adaptive filter is

$$e_n = y_n - \mathbf{x}_n^T \hat{\mathbf{w}}_{n-1},$$

where $\hat{\mathbf{w}}_n = [w_{n,0}, w_{n,1}, \dots, w_{n,L-1}]^T$ denotes the adaptive filter tap-weights at time instant n . The SM criterion corresponds to finding $\hat{\mathbf{w}}$ that satisfies the *feasibility set*

$$\Theta = \bigcap_{any\ n} \left\{ \hat{\mathbf{w}} \in R^L : \left| y_n - \mathbf{x}_n^T \hat{\mathbf{w}} \right| \leq \gamma \right\}$$

where γ is a pre-specified bound. The *constraint set* that contains all vectors $\hat{\mathbf{w}}$ satisfying the error bound at time instant n , is defined as

$$H_n = \left\{ \hat{\mathbf{w}} \in R^L : \left| y_n - \mathbf{x}_n^T \hat{\mathbf{w}} \right| \leq \gamma \right\}$$

Hence, the SM adaptive filtering algorithm corresponds to finding a *membership set* ψ_n at time instant n that satisfies

$$\psi_n = \bigcap_{i=1}^n H_i$$

where the feasibility set Θ is a subset of ψ_n .

The weights of the BEACON algorithm [7, 11] are updated as follows

$$V_n = V_{n-1} - \frac{\lambda V_{n-1} \mathbf{x}_n \mathbf{x}_n^T V_{n-1}}{1 + \lambda \mathbf{x}_n^T V_{n-1} \mathbf{x}_n} \quad (1)$$

$$\hat{\mathbf{w}}_n = \hat{\mathbf{w}}_{n-1} + \lambda V_n e_n \mathbf{x}_n \quad (2)$$

where V_n is the inverse input autocorrelation matrix, the forgetting factor is

$$\lambda = \begin{cases} \frac{1}{G_n} \left(\frac{|e_n|}{\gamma} - 1 \right) & \text{if } |e_n| > \gamma \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

and

$$G_n = \mathbf{x}_n^T V_{n-1} \mathbf{x}_n.$$

However, the performance of the BEACON algorithm deteriorates in the presence of outliers in the error signal samples that can be brought about by impulsive noise interference.

3. PROPOSED ROBUST BEACON ALGORITHM

In order to mitigate the effect of impulsive noise, two versions of robust BEACON algorithm are presented in this section, one with fixed set-membership low bound and the other with time-varying set-membership low bound.

3.1. Proposed robust BEACON with fixed set-membership low bound (robust BEACON-FB)

In order to achieve the robust performance against impulsive noise, the set-membership error bound (SMEB) γ in the BEACON is chosen as

$$\gamma_n = \begin{cases} \frac{e_n^2}{\nu \theta_n + |e_n|} & \text{if } |e_n| > \xi \theta_n \\ \gamma_c & \text{otherwise} \end{cases} \quad (4)$$

where we usually choose $0 < \nu < 1$ to obtain low steady-state misalignments, $\gamma_c = \sqrt{5\sigma_v^2}$ (named set-membership low bound) [5], ξ (named outlier detection threshold, ODT) is chosen to be less than $\sqrt{5}$ in order to ensure that $\xi \theta_n \leq \gamma_c$ at steady-state (a similar condition is used in [16]), θ_n is chosen to be $\theta_n = \hat{\sigma}_n$, which is the impulsive-free estimation of $E[e_n^2]$ and is estimated as

$$\hat{\sigma}_n = \beta_1 \hat{\sigma}_{n-1} + (1 - \beta_1) \text{median}(\psi_n) \quad (5)$$

where β_1 is a forgetting factor, $\text{median}(\cdot)$ is the median operator and

$$\psi_n = [|e_n|, |e_{n-1}|, \dots, |e_{n-N+1}|]^T$$

where N is the length of the estimation window.

Remark 1: Note that the γ_n in (4) is different from that used in the robust QN and SMAPA in [16, 17]. The SMEB in [17] was chosen as

$$\gamma_n = \begin{cases} |e_n| - \nu \theta_n & \text{if } |e_n| > \theta_n \\ \gamma_c & \text{otherwise} \end{cases} \quad (6)$$

where $\theta_n = Q\sigma_{1,k}$ with $1.86 < Q < 1.98$, $\sigma_{1,k}^2$ is the impulsive-free estimation of $E[e_n^2]$. When the impulsive noise occurs at time instant n , using (4) and (6) in (3) yields $\lambda G_n = \nu \theta_n / |e_n|$, $\nu \theta_n / (|e_n| - \nu \theta_n)$, respectively. For $\nu \theta_n < |e_n|$, $\nu \theta_n / |e_n| \approx \nu \theta_n / (|e_n| - \nu \theta_n)$. However, for multiple outliers the proposed SMEB (4) would achieve improved robustness, compared with (6). For multiple outliers, the

inaccurate estimator θ_n may cause $|e_n| \approx v\theta_n$. In this case, $v\theta_n/|e_n| < v\theta_n/(|e_n| - v\theta_n)$. Therefore, (4) achieves improved robustness with respect to multiple outliers. The improved robustness had been verified in the simulations.

3.2. Robust BEACON with time-varying set-membership low bound (Robust BEACON-VB)

Here, we use an error bound function that can automatically adjust the set-membership low bound with the update of the filter coefficients [9]. The time-varying set-membership low bound is $\gamma_c = \eta_n$, which can be expressed as

$$\eta_n = \beta_2 \eta_{n-1} + (1 - \beta_2) \sqrt{\tau \|\hat{\mathbf{w}}_{n-1}\|^2 \times \sigma_v^2} \quad (7)$$

where $0 < \beta_2 < 1$ is a forgetting factor that should be adjusted to ensure an appropriate time-averaged estimate of the evolution of the power of the adaptive filter tap-weights $\hat{\mathbf{w}}_n$, τ is the tuning parameter, and σ_v^2 is the variance of the impulsive-free noise. The tuning parameter in (7) should be set by experimental results and too high or low values will lead to inappropriate time-varying low bound. Table I shows a summary of the robust BEACON algorithm with time-varying set-membership low bound, which will be used for the simulations.

TABLE I
SUMMARY OF THE ROBUST BEACON-VB ALGORITHM

3.3. Discussion of computational complexity

The computational complexity of the robust BEACON-VB is compared with that of conventional RLS and BEACON-VB algorithms in terms of the total number of additions, multiplications, comparisons, square-roots, and media operation (MO). With filter length L and the estimation window N , the computational requirements for these adaptive filters with corresponding adaptive algorithms are summarized in Table II. With only slightly more computations in the computation of γ_n , the robust BEACON-VB behaves much better than the conventional BEACON algorithm using (7) (referred to as BEACON-VB) in the presence of impulsive noise, shown in Section 5.

TABLE II
COMPUTATIONAL COMPLEXITY OF ALGORITHMS' COEFFICIENTS UPDATE PER ITERATION
WHERE $P_1 = P(|e_n| > \xi \hat{\sigma}_n)$, $P_2 = P(|e_n| > \gamma_n)$

4. STEADY-STATE PERFORMANCE ANALYSIS

In this section, we study the steady-state properties and provide insights about how the proposed algorithms operate in steady-state. Compared with the robust BEACON-FB algorithm, the steady-state performance of the robust BEACON-VB algorithm only depends on the equation (7). The similar steady-state performance can be derived by approximating the expectations of the equation (7). Thus, we will focus on the steady-state analysis of the proposed robust BEACON-FB algorithm. To simplify the performance analysis, the following assumptions are adopted.

Assumption (1): The input signal $\mathbf{x}_n$ is independent and identically distributed (i.i.d.) with zero-mean and covariance matrix $E(\mathbf{x}_n \mathbf{x}_n^T) = \mathbf{R}$. This so-called independence assumption is commonly used in [18-23]. This assumption also allows us to deal with $\{\mathbf{x}_k\}$ independently from $\mathbf{v}_n$ for $n \leq k$ in the following analysis, where $\mathbf{v}_n = \mathbf{w}_{opt} - \hat{\mathbf{w}}_n$ denotes the weight deviation vector.

Assumption (2): The noise $\{v_n\}$ is i.i.d. with zero-mean and variance σ_v^2 , and is independent of the input $\mathbf{x}_j$ for all n, j .

Assumption (3): In steady-state, $\|V_{n-1} \mathbf{x}_n / G_n\|^2$ is approximately independent of the priori excess errors $e_{a,n} = \mathbf{v}_{n-1}^T \mathbf{x}_n$ and posteriori

excess errors $e_{p,n} = \mathbf{v}_n^T \mathbf{x}_n$. This is the so-called separation principle invoked in [11].

Assumption (4): In steady-state, the priori excess errors, $e_{a,n}$, is white Gaussian with zero mean and variance of $\sigma_{a,n}^2 = E(e_{a,n}^2)$. This assumption is reasonable for longer filters by central limit arguments. It is commonly used in the literature [11, 19].

Applying the weight deviation vector, the estimation error e_n is rewritten as $e_n = \mathbf{v}_{n-1}^T \mathbf{x}_n + v_n$. The mean square error (MSE) at time instant n is expressed as [1]: $E[e_n^2] = \sigma_v^2 + E[e_{a,n}^2]$, where the assumption (1) is used and $E[e_{a,n}^2] = \text{Tr}(RE[\mathbf{v}_{n-1} \mathbf{v}_{n-1}^T])$ is the EMSE. The mean-square derivation (MSD) at time instant n is $MSD(n) = \text{Tr}(E[\mathbf{v}_n \mathbf{v}_n^T])$.

4.1. Steady-state analysis

According to [11], the weight-error vector update equation can be written as

$$\mathbf{v}_n = \mathbf{v}_{n-1} - \frac{V_{n-1} \mathbf{x}_n f(e_n)}{G_n} \quad (8)$$

where

$$f(e_n) = e_n \left(1 - \frac{\gamma_n}{|e_{0,n}|} \right) \quad (9)$$

$$e_{0,n} = \begin{cases} |e_n| & \text{if } |e_n| > \gamma_n \\ \gamma_n & \text{otherwise} \end{cases} \quad (10)$$

and

$$\gamma_n = \begin{cases} \frac{e_n^2}{\nu \theta_n + |e_n|} & \text{if } |e_n| > \xi \theta_n \\ \gamma_c & \text{otherwise} \end{cases} \quad (11)$$

By pre-multiplying both sides of (8) by $\mathbf{x}_n^T$, we get the following relation

$$e_{p,n} = e_{a,n} - f(e_n) \quad (12)$$

By substituting $f(e_n) = e_{a,n} - e_{p,n}$ into (8), we can get

$$\mathbf{v}_n + \frac{V_{n-1} \mathbf{x}_n}{G_n} e_{a,n} = \mathbf{v}_{n-1} + \frac{V_{n-1} \mathbf{x}_n}{G_n} e_{p,n} \quad (13)$$

Taking the Euclidean norms of (13) and the expectation, we obtain the following energy conservation relation

$$\begin{aligned} E[\|\mathbf{v}_n\|^2] + E\left[\left\|\frac{V_{n-1} \mathbf{x}_n}{G_n}\right\|^2 e_{a,n}^2\right] \\ = E[\|\mathbf{v}_{n-1}\|^2] + E\left[\left\|\frac{V_{n-1} \mathbf{x}_n}{G_n}\right\|^2 e_{p,n}^2\right] \end{aligned} \quad (14)$$

Assuming (14) is operating in steady-state when $n \rightarrow \infty$, we observe from (14) that

$$E\left[\left\|\frac{V_{n-1} \mathbf{x}_n}{G_n}\right\|^2 e_{a,n}^2\right] = E\left[\left\|\frac{V_{n-1} \mathbf{x}_n}{G_n}\right\|^2 e_{p,n}^2\right] \quad (15)$$

Then adopting the assumption (3) and substituting (12) into (15), it simplifies to

$$E[f^2(e_n)] - 2E[e_{a,n}f(e_n)] = 0 \quad (16)$$

Using the property of conditional expectation, (16) are calculated as follows.

$$\begin{aligned} & P(|e_n| > \xi\theta_n) \left(E \left[f^2(e_n) \middle| \gamma_n = \frac{e_n^2}{\nu\theta_n + |e_n|}, |e_n| > \xi\theta_n \right] \right. \\ & \quad \left. - 2E \left[e_{a,n}f(e_n) \middle| \gamma_n = \frac{e_n^2}{\nu\theta_n + |e_n|}, |e_n| > \xi\theta_n \right] \right) \\ & + P(|e_n| \leq \xi\theta_n) \left(E \left[f^2(e_n) \middle| \gamma_n = \gamma_c, |e_n| \leq \xi\theta_n \right] \right. \\ & \quad \left. - 2E \left[e_{a,n}f(e_n) \middle| \gamma_n = \gamma_c, |e_n| \leq \xi\theta_n \right] \right) = 0 \end{aligned} \quad (17)$$

In general, the behavior of (17) can become very complex to be studied analytically. Equation (17) should be differentiated into three categories and handled accordingly.

Case 1): $P(|e_n| > \xi\theta_n) = 0$, i.e., $\xi \rightarrow \infty$

The equation (17) simplifies to

$$E[f^2(e_n) | \gamma_n = \gamma_c] - 2E[e_{a,n}f(e_n) | \gamma_n = \gamma_c] = 0 \quad (18)$$

where the expectations on the left hand side (LHS) of (18) are given in Appendix A. When the noise is a Gaussian random process, equation (18) is further rewritten as (28) in [11].

Case 2): $P(|e_n| > \xi\theta_n) = 1$, i.e., $\xi \rightarrow 0$

The equation (17) simplifies to

$$E \left[f^2(e_n) \middle| \gamma_n = \frac{e_n^2}{\nu\theta_n + |e_n|} \right] - 2E \left[e_{a,n}f(e_n) \middle| \gamma_n = \frac{e_n^2}{\nu\theta_n + |e_n|} \right] = 0 \quad (19)$$

In this case, the robust BEACON-FB is in all update stages. The steady-state EMSE is

$$E[e_{a,n}^2] = \nu\sigma_v^2 / (\nu + 2) \quad (20)$$

where the detailed derivation of the equation (20) is presented in Appendix A.

Case 3): $P(|e_n| > \xi\theta_n) \neq 0, 1$

Let $\alpha = P(|e_n| \leq \xi\theta_n)$ and $0 < \alpha < 1$, (17) is

$$\begin{aligned} & (1 - \alpha) \left(E \left[\left(\frac{\nu\theta_n e_n}{\nu\theta_n + |e_n|} \right)^2 \middle| |e_n| > \xi\theta_n \right] \right. \\ & \quad \left. - 2E \left[\frac{\nu\theta_n e_{a,n} e_n}{\nu\theta_n + |e_n|} \middle| |e_n| > \xi\theta_n \right] \right) \\ & + \alpha \left(E \left[f^2(e_n) \middle| \gamma_n = \gamma_c, |e_n| \leq \xi\theta_n \right] \right. \\ & \quad \left. - 2E \left[e_{a,n}f(e_n) \middle| \gamma_n = \gamma_c, |e_n| \leq \xi\theta_n \right] \right) = 0 \end{aligned} \quad (21)$$

We know that if $\theta_n \approx \sigma_v$ at steady-state and $\xi\theta_n \leq \gamma_c$ due to choosing $\xi < \sqrt{5}$, the following equations in (21) is zero and

$$E[f^2(e_n) | \gamma_n = \gamma_c, |e_n| \leq \xi\theta_n] = 0$$

$$E[e_{a,n}f(e_n)|\gamma_n = \gamma_c, |e_n| \leq \xi\theta_n] = 0$$

Therefore, equation (21) simplifies to:

$$E\left[\left(\frac{\nu\theta_n e_n}{\nu\theta_n + |e_n|}\right)^2 \middle| |e_n| > \xi\theta_n\right] - 2E\left[\frac{\nu\theta_n e_n e_n}{\nu\theta_n + |e_n|} \middle| |e_n| > \xi\theta_n\right] = 0 \quad (22)$$

Using $|e_n|/\theta_n \approx 1$ in (22) at steady-state, we obtain that

$$\left(\frac{\nu}{\nu+1}\right)^2 E[e_n^2 | |e_n| > \xi\theta_n] - 2\frac{\nu}{\nu+1} E[e_{a,n}e_n | |e_n| > \xi\theta_n] = 0 \quad (23)$$

The second expectation of the LHS of the equation (23) is:

$$\begin{aligned} & E\left[(e_{a,n}(e_{a,n} + \nu_n)) \middle| |e_n| > \xi\theta_n\right] \\ & \approx E[e_{a,n}^2 | |\nu_n| > \xi\theta_n] + E[e_{a,n}\nu_n | |\nu_n| > \xi\theta_n] \\ & = E[e_{a,n}^2] \end{aligned} \quad (24)$$

where the statistical dependence of $e_{a,n}(e_{a,n} + \nu_n)$ and $|e_n| > \xi\theta_n$ is neglected due to $e_n \approx \nu_n$ at steady-state.

Combining (24) into (23), we obtain the steady-state EMSE as

$$E[e_{a,n}^2] = \frac{\nu}{2(\nu+1)} E[e_n^2 | |e_n| > \xi\theta_n] \quad (25)$$

Remark 2: Though we limit that $\xi < \sqrt{5}$ in equation (21), it is shown from (21) that with a large ξ , i.e., $\alpha \rightarrow 1$, the robust BEACON-FB algorithm yields a steady-state EMSE, which is the solution of (18), and with a small ξ , i.e., $\alpha \rightarrow 0$, it offers another steady-state EMSE, which is the solution of (19), i.e., (20).

4.2. The probability of the update at steady-state

Due to the use of the SMEB (4), the probability of the update can be computed as

$$\begin{aligned} & P(|e_n| > \gamma_n) \\ & = P(|e_n| > \gamma_n | \gamma_n = \gamma_c, |e_n| \leq \xi\theta_n) P(|e_n| \leq \xi\theta_n) \\ & + P\left(|e_n| > \gamma_n \middle| \gamma_n = \frac{e_n^2}{\nu\theta_n + |e_n|}, |e_n| > \xi\theta_n\right) P(|e_n| > \xi\theta_n) \\ & = P(|e_n| > \gamma_n | \gamma_n = \gamma_c, |e_n| \leq \xi\theta_n) P(|e_n| \leq \xi\theta_n) + P(|e_n| > \xi\theta_n) \end{aligned} \quad (26)$$

which contains two parts, i.e., $\gamma_n = \gamma_c$ and $\gamma_n = e_n^2/(\nu\theta_n + |e_n|)$. To find a close form solution is rather difficult since the variance of e_n for the robust BEACON-FB algorithms are unknown at arbitrary time instant n . When the algorithm works at steady-state, $\xi\theta_n \leq \gamma_c$ and

$$P(|e_n| > \gamma_n | \gamma_n = \gamma_c, |e_n| \leq \xi\theta_n) P(|e_n| \leq \xi\theta_n) = 0$$

Then, equation (26) is

$$P(|e_n| > \gamma_n) = P(|e_n| > \xi\theta_n)$$

Compared with the BEACON (B) algorithm, the extra PU in the robust BEACON-FB (RB) is:

$$\begin{aligned}
& P(|e_n| > \gamma_n)_{RB} - P(|e_n| > \gamma_c)_B \\
&= P(|e_n| > \gamma_c | |e_n| \leq \xi\theta_n) P(|e_n| \leq \xi\theta_n)_{RB} \\
&\quad + P(|e_n| > \xi\theta_n)_{RB} - P(|e_n| > \gamma_c | |e_n| \leq \xi\theta_n) P(|e_n| \leq \xi\theta_n)_B \\
&\quad - P(|e_n| > \gamma_c | |e_n| > \xi\theta_n) P(|e_n| > \xi\theta_n)_B
\end{aligned} \tag{27}$$

If we assume that $|e_{a,n}| \ll |\nu_n|$ for both BEACON and robust BEACON-FB at steady-state, we have that from (27)

$$\begin{aligned}
& P(|e_n| > \gamma_n)_{RB} - P(|e_n| > \gamma_c)_B \\
&\approx P(|\nu_n| > \xi\theta_n) - P(|\nu_n| > \gamma_c | |\nu_n| > \xi\theta_n) P(|\nu_n| > \xi\theta_n) \\
&= \left(1 - P(|\nu_n| > \gamma_c | |\nu_n| > \xi\theta_n)\right) P(|\nu_n| > \xi\theta_n) \\
&= P(|\nu_n| \leq \gamma_c | |\nu_n| > \xi\theta_n) P(|\nu_n| > \xi\theta_n)
\end{aligned} \tag{28}$$

Due to $\xi\theta_n \leq \gamma_c$ at steady-state, the above equation is

$$P(\xi\theta_n < |\nu_n| \leq \gamma_c) > 0 \tag{29}$$

Remark 4: As can be seen from (29), the proposed robust BEACON-FB algorithm offers improved robustness at the expense of the PU compared with the BEACON algorithm. In order to achieve low PU, a large ξ should be chosen.

4.3. Discussion for Gaussian noise

In order to obtain close form solution that allow some understanding of the steady-state EMSEs in Case 2 and Case 3, we study here the particular case of noise $\{\nu_n\}$ drawn from a Gaussian distribution. The variance of e_n at the steady-state is computed as $\sigma_n^2 = \sigma_{a,n}^2 + \sigma_\nu^2$. Using assumption (4), equation (25) is calculated as:

$$\begin{aligned}
E[e_{a,n}^2] &= \frac{\nu}{\nu+1} \frac{1}{\sqrt{2\pi\sigma_n^2}} \int_{\xi\theta_n}^{\infty} x^2 \exp\left(-\frac{x^2}{2\sigma_n^2}\right) dx \\
&= \frac{\nu}{\nu+1} \left(\frac{\sigma_n \xi\theta_n}{\sqrt{2\pi}} \exp\left(-\frac{\xi^2\theta_n^2}{2\sigma_n^2}\right) + \frac{\sigma_n^2}{2} \left(1 - \operatorname{erf}\left(\frac{\xi\theta_n}{\sqrt{2\sigma_n^2}}\right)\right) \right)
\end{aligned} \tag{30}$$

where the approximated EMSE (30) is given in Appendix A.

Remark 3: As can be seen from (30) that: (1) A small ν will leads to a low EMSE. (2) When $\xi = 0$, it is consistent to the result in Case 2 with Gaussian noises. (3) For sufficiently small value $0 < \nu \ll 1$, we can assume that $\sigma_n^2 \approx \sigma_\nu^2$ ($\sigma_{a,n}^2 \ll \sigma_\nu^2$) at steady-state. Then, the difference between (30) and (35) is

$$\begin{aligned}
& E[e_{a,n}^2]_{\text{case 2}} - E[e_{a,n}^2]_{\text{case 3}} \\
&\approx \frac{\nu}{(\nu+1)} \frac{1}{\sqrt{2\pi\sigma_\nu^2}} \left(\int_0^{\infty} x^2 \exp\left(-\frac{x^2}{2\sigma_\nu^2}\right) dx - \int_{\xi\theta_n}^{\infty} x^2 \exp\left(-\frac{x^2}{2\sigma_\nu^2}\right) dx \right) > 0
\end{aligned} \tag{31}$$

Thus, the steady-state EMSE in Case 3 is less than that in Case 2. This fact was also verified in the simulations. Under the condition $\xi\theta_n \leq \gamma_c$, it can be seen from (31) that for the robust BEACON-FB a larger ξ gives a lower steady-state EMSE. In [16], the steady-state analysis of the RSMAP was carried out without any consideration for this point.

Moreover, using assumption (4) and Gaussian noises assumption, we have

$$\begin{aligned} P(|e_n| > \xi\theta_n) &= 1 - \frac{1}{\sqrt{2\pi\sigma_n^2}} \int_{-\xi\theta_n}^{\xi\theta_n} \exp\left(-\frac{t^2}{2\sigma_n^2}\right) dt \\ &= 1 - \operatorname{erf}\left(\xi\theta_n / \sqrt{2\sigma_n^2}\right) \approx 1 - \operatorname{erf}\left(\frac{\xi}{\sqrt{2}}\right) \end{aligned} \quad (32)$$

where $\theta_n^2 \approx \sigma_n^2$ and the error function is defined as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt$$

In order to achieve the robustness and low PU simultaneously, we choose $P(|e_n| > \xi\theta_n) = 0.03$, i.e., $\xi = 2.17$, which has 97% confidence to down-weight the error in the interval $[-\xi\theta_n, \xi\theta_n]$.

5. SIMULATIONS

In this section, we evaluate the performance of the proposed two versions of robust BEACON (i.e., robust BEACON-FB and robust BEACON-VB). The unknown system to be estimated in all experiments is an order L finite impulse response (FIR) filter, and the weight vector of the unknown system is randomly generated and normalized to unit energy. The initial weight vector of the adaptive filter is an all zero vector. All the learning curves are obtained by averaging the results of 100 independent runs. The signal-to-noise ratio (SNR) at the system output is given by: $SNR = 10\log_{10}(\sigma_d^2 / \sigma_v^2)$, where σ_d^2 is the variance of the uncorrupted output of the unknown system. We set the estimation window length, $N=9$ used here for all experiments. For illustration purpose, the impulsive noise [24] is generated as

$$\varpi_n = b_n \nu_{b,n} \quad (33)$$

where b_n is an i.i.d. Bernoulli random sequence whose value at any time instant n is either zero or one, with occurrence probability $p_i = P(b_n=i)$ with $i=0, 1$, and $\nu_{b,n}$ is an i.i.d. zero mean Gaussian sequence with variance $\sigma_b^2 = G\sigma_d^2$ with $G \gg 1$.

5.1. Performance comparison between the proposed BEACON and other algorithms

In the first experiment, the convergence performance and robustness to impulsive noise of the robust BEACON-FB/VB are evaluated with different SNRs. The input is generated by filtering a white Gaussian noise through a first order system with a pole at 0.8 and the order of the unknown system was set at $L=32$. For illustration purposes, the impulsive noise is generated from (33) with an impulse occurrence probability of $p_1=0.1$ and $G=1,000$. The locations of the impulse noise are fixed at iterations 2,000, 8,000, 16,000. The initial parameters for all algorithms are set to $V_0=1000 \times I$, $\hat{\sigma}_0 = 10$, $\nu = 0.19$, $\beta_1 = \beta_2 = 1 - 1/L$. Figs. 1 and 2 depict the performance of the robust BEACON with $SNR=20$ dB and $SNR=40$ dB, respectively. It can be seen that the robust BEACON and robust QN-II [17] algorithms are robust to impulsive noise. In contrast, the performance of the BEACON algorithm is deteriorated in the presence of impulsive noise. Moreover, the robust BEACON-VB can converge at a faster speed than the robust BEACON-FB with the same steady-state MSD. In Table III the overall update ratio of the corresponding algorithms is exhibited. In addition, we have measured the average run time of the algorithms on a 2.9 GHz Intel Core i5 processor with 4 Gb of RAM, running Matlab R2009b on Windows 7. Results have been displayed in Table IV. The robust BEACON-FB/VB has better robustness and lower steady-state MSD than the BEACON [7] at the expense of marginally computational cost. But the computational cost of robust BEACON-FB/VB is lower than that of the robust QN-II.

Fig. 1. MSD comparison of the robust BEACON algorithms, and other algorithms, where $SNR=20\text{dB}$, $\tau=3$, $\eta_0=0.35$.

Fig. 2. MSD comparison of the robust BEACON algorithms, and other algorithms, where $SNR=40\text{dB}$, $\tau=3$, $\eta_0=0.35$.

TABLE III
UPDATE RATIO (20000 ITERATIONS)

TABLE IV
AVERAGE COMPUTATION TIME PER RUN (RESULTS IN SECONDS)

Fig. 3. MSD curves of the robust BEACON with an abrupt change in system coefficients, where $SNR=10\text{dB}$, $\tau=3$, $\eta_0=2$.

Fig. 4. MSD curves of the robust BEACON with an abrupt change in system coefficients, where $SNR=40\text{dB}$, $\tau=3$, $\eta_0=0.35$.

Fig. 5. MSD curves of the BEACON with different SMEBs in impulsive noise environments, where $SNR=0\text{dB}$, $p_1=0.1$, $Q=1.88$.

In the second experiment, the comparisons of the proposed robust BEACON-FB/VB with the BEACON are shown in the case of a system abrupt change environment. The unknown system coefficients were abruptly multiplied by -1 at iteration 10 000 in Figs. 3 and 4. The input signals, the impulses and the order of the unknown system were the same as those in the first experiment. The measurement noise added to the desired signal was a zero-mean white Gaussian noise signal to achieve SNRs of 10 dB and 40 dB in Figs. 3 and 4, respectively. It can be seen that the robust BEACON-FB/VB handles sudden system changes successfully and at the same time maintains its robustness with respect to impulsive noise. The robust BEACON-VB has a better tracking performance compared with the robust BEACON-FB.

In the third experiment, the robustness of the different SMEBs in the BEACON is examined. The input signals and the order of the unknown system were the same as those in the first experiment except that the position of the impulses is randomly selected. The measurement noise added to the desired signal was a zero-mean white Gaussian noise signal to achieve SNR of 0 dB in Fig. 5. It can be seen that the BEACON are more likely to diverge in more intensive interference. The proposed robust BEACON-FB using the SMEB (4) leads to lower steady-state MSD and better robustness than that using the SMEB (6).

5.2. Double-talk scenario

For the double-talk scenario, both near-end and far-end signals were speech segments. A classical Geigel double-talk detector (DTD) is used to suppress adaptation during periods of simultaneous far- and near-end signals [25]. The Geigel DTD declares double-talk if

$$|y_n| \geq T \max(x_n, \dots, x_{n-D+1})$$

where T is the detection threshold, x_n are the samples of the far-end signal and y_n are the samples of the far-end signal filtered by the acoustic impulse response and possibly contaminated by a background noise and a near-end signal.

Fig. 6. The impulse response of echo path.

Fig. 7. Acoustic echo cancellation. *Top*. Far-end speech signal. *Bottom*. Near-end speech signal.

Fig. 8. Performance of the BEACON family for speech signals during double-talk. $\tau=5$.

In Figs. 6 and 7, the impulse response of the acoustic echo path ($L=126$), the far-end and near-end signals are illustrated. The far-end and near-end signals are speech sampled at 8 kHz. The background noise was a zero-mean white Gaussian noise signal with variance 0.0001. The parameters of the DTD are $T=0.45$ and $D=L$. As shown in Fig. 8, the BEACON is highly affected by the double-talk situation. In comparison, the robust BEACON algorithms are more robust against double-talk, and the robust BEACON algorithms achieve lower steady-state misalignment.

Fig. 9. Learning curves of the robust BEACON-FB with varying ξ for Gaussian noises. $L=32$.

Fig. 10. Learning curves of the robust BEACON-FB with varying ν for Gaussian noises. $L=32$.

5.3. Verification of the analysis

In this experiment, this input was a colored Gaussian process with a pole at 0.8. The measurement noise added to the desired signal was a zero-mean white Gaussian signal with power 0.0001. The analytical steady-state EMSEs are calculated from the solution of (18) given by [11], (20) and (36), which are shown as the horizontal dash lines. It can be seen from Fig. 9 that the robust BEACON-FB outperforms BEACON. The robust BEACON-FB with $\xi=2.17$ achieves lower steady-state EMSE than that with $\xi=0$ at the expense of a poor convergence speed. In other words, the steady-state EMSE in Case 3 is lower than that in Case 2. As can be seen from Fig. 10, a small ν slows down the convergence rate but also lowers the steady-state EMSE. In contrast, a large ν speeds up the convergence rate but gives a higher steady-state EMSE. Figs. 9 and 10 show that simulation results are in good agreement with the proposed theoretical results.

6. CONCLUSION

In this paper, we have proposed a robust BEACON adaptive-filtering algorithm. Robustness against impulsive noise is achieved by imposing a new set-membership error bound. We have analyzed the computational complexity, the probability of the update and devised analytical expressions to predict the excess MSE performance with a good accuracy. The two versions of the proposed algorithm were applied to system identification and double-talk echo cancellation applications, where the robust BEACON-VB makes it robust to changes in the environment. Simulations results have shown that the proposed robust BEACON-FB/VB algorithms can achieve better performance to the conventional BEACON.

APPENDIX A

Case 1:

For the general noise, under the assumption (4) the expectations on the LHS of (18) is calculated as

$$\begin{aligned} & E[f^2(e_n)] \\ &= \frac{1}{\sqrt{2\pi\sigma_{a,n}^2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^2(e_{a,n} + v_n) \exp\left(-\frac{e_{a,n}^2}{2\sigma_{a,n}^2}\right) \phi(v_n) de_{a,n} dv_n \end{aligned}$$

and

$$\begin{aligned}
& E[e_{a,n} f(e_n)] \\
&= \frac{1}{\sqrt{2\pi\sigma_{a,n}^2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e_{a,n} f(e_{a,n} + v_n) \exp\left(-\frac{e_{a,n}^2}{2\sigma_{a,n}^2}\right) \varphi(v_n) dv_n de_{a,n}
\end{aligned}$$

where $\varphi(v_n)$ is the p.d.f of the noise.

When the noise is a Gaussian random process, the expectations can be further calculated as shown in [11].

Case 2:

Substituting $\gamma_n = e_n^2 / (v\theta_n + |e_n|)$ into (9), we have

$$f(e_n) = \frac{v\theta_n e_n}{v\theta_n + |e_n|} \quad (34)$$

Using $|e_n| \approx \theta_n$ at steady-state and (34) in (19), we get

$$\left(\frac{v}{v+1}\right)^2 E[e_n^2] - 2\frac{v}{v+1} E[e_{a,n} e_n] = 0 \quad (35)$$

Then using assumption (2) leads to

$$E[e_{a,n}^2] = \frac{v\sigma_v^2}{v+2}$$

Case 3: for Gaussian noises

If we assume $\theta_n \approx \kappa\sigma_n$ at (30), one have

$$E[e_{a,n}^2] = \frac{v\varsigma}{v(1-\varsigma)+1} \sigma_v^2 \quad (36)$$

where

$$\varsigma = \left(\frac{\xi\kappa}{\sqrt{2\pi}} \exp\left(-\frac{\xi^2\kappa}{2}\right) + \frac{1}{2} \left(1 - \operatorname{erf}\left(\frac{\xi\kappa}{\sqrt{2}}\right)\right) \right)$$

REFERENCES

- [1] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ: Prentice-Hall, 2002.
- [2] A. H. Sayed, *Fundamentals of Adaptive Filtering*, New York: Wiley-Interscience, 2003.
- [3] S. Werner, M. L. R. de Campos, and P. S. R. Diniz, "Partial-update NLMS algorithms with data-selective updating," *IEEE Trans. on Signal Process.*, vol. 52, pp. 938–949, 2004.
- [4] K. Dogancay, *Partial-Update Adaptive Signal Processing: Design, Analysis and Implementation*. New York: Academic, 2008.
- [5] S. Gollamudi, S. Nagaraj, S. Kapoor, and Y.-F. Huang, "Set-membership filtering and a set-membership normalized LMS algorithm with an adaptive step size," *IEEE Signal Process. Lett.*, vol. 5, no. 5, pp. 111–114, 1998.
- [6] S. Werner and P. S. R. Diniz, "Set-membership affine projection algorithm," *IEEE Signal Process. Lett.*, vol. 8, pp. 231–235, Aug. 2001.
- [7] S. Nagaraj, S. Gollamudi, S. Kapoor, and Y. F. Huang, "BEACON: An adaptive set-membership filtering technique with sparse updates," *IEEE Trans. Signal Process.*, vol. 47, no. 11, pp. 2928–2941, Nov. 1999.
- [8] J. R. Deller Jr., S. Gollamudi, S. Nagaraj, D. Joachim, and Y. F. Huang, "Convergence analysis of the quasi-OBE algorithm and related performance issues," *International Journal of Adaptive Control and Signal Process.*, vol. 21, pp. 499–527, 2007.
- [9] T. Wang, R. C. De Lamare, and P. D. Mitchell, "Low-complexity set-membership channel estimation for cooperative wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2594–2607, 2011.
- [10] J. R. Deller Jr., Y. F. Huang, "Set-membership identification and filtering for signal processing applications," *Circuits Systems and Signal Process.*, vol. 21, no. 1, pp. 69–82, 2002.

- [11] R. Arablouei and K. Doğançay, “Steady-state mean squared error and tracking performance analysis of the quasi-OBE algorithm,” *Signal Process.*, vol. 93, pp.100–108, 2013.
- [12] R. Arablouei, K. Doğançay, “Modified quasi-OBE algorithm with improved numerical properties,” *Signal Process.*, vol. 93, pp. 797-803, Apr. 2013.
- [13] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*. New York: Wiley, 1987.
- [14] S. Zhang and J. Zhang, “Set-membership NLMS algorithm with robust error bound,” *IEEE Trans. Circuits Sys. II*, vol. 61, no. 7, pp. 536-540, July 2014.
- [15] M. Z. A. Bhotto and A. Antoniou, “A robust set-membership normalized least mean-square adaptive filter,” in *Proc. IEEE Can. Conf. on Elect. and Comp. Eng.*, May 2010.
- [16] M. Z. A. Bhotto and A. Antoniou, “Robust set-membership affine-projection adaptive-filtering algorithm,” *IEEE Trans. on Signal Process.*, vol. 60, no. 1, pp. 73–81, 2012.
- [17] M. Z. A. Bhotto and A. Antoniou, “Robust quasi-Newton adaptive filtering algorithms,” *IEEE Trans. Circuits Sys. II*, vol. 58, no. 8, pp.537–541, Aug. 2011.
- [18] Y. Zou, S. C. Chan, and T. S. Ng, “A recursive least M-estimate (RLM) adaptive filter for robust filtering in impulse noise,” *IEEE Signal Process. Lett.*, vol. 7, no. 11, pp. 324–326, Nov. 2000.
- [19] N. R. Yousef and A. H. Sayed, “A unified approach to the steady-state and tracking analyses of adaptive filters,” *IEEE Trans. Signal Process.*, vol. 49, no. 2, pp. 314–324, Feb. 2001.
- [20] T. Y. Al-Naffouri and A. H. Sayed, “Transient analysis of adaptive filters with error nonlinearities,” *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 653–663, Mar. 2003.
- [21] T. Y. Al-Naffouri and A. H. Sayed, “Transient analysis of data-normalized adaptive filters,” *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 639–652, Mar. 2003.
- [22] N. Takahashi and I. Yamada, “Steady-state mean-square performance analysis of a relaxed set-membership NLMS algorithm by the energy conservation argument,” *IEEE Trans. on Signal Process.*, vol. 57, no. 9, pp. 3361–3372, 2009.
- [23] S. Zhang and J. Zhang, “Transient analysis of zero attracting NLMS algorithms without Gaussian input Signal,” *Signal Process.*, vol. 97, pp. 100-109, 2014.
- [24] S. Zhang and J. Zhang, “Enhancing the tracking capability of recursive least p-norm algorithm via adaptive gain factor,” *Digital Signal Process.*, vol. 30, pp. 67–73, Jul. 2014.
- [25] T. Gansler, S. Gay, M. Sondhi, and J. Benesty, “Double-talk robust fast converging algorithms for network echo cancellation,” *IEEE Trans. Speech, Audio Process.*, vol. 8, no. 6, pp. 656 –663, Nov. 2000.

TABLE I
SUMMARY OF THE ROBUST BEACON-VB ALGORITHM

Initialization:

$$\hat{\mathbf{w}}_0 = 0, e_n = 0 \text{ for } n \leq 0$$

$$\mathbf{x}_n, N, \beta_1, \beta_2, \sigma_v^2, \hat{\sigma}_0 = 10, \eta_0 = 0, \xi, \nu,$$

For $n = 1, 2, 3 \dots$

$$e_n = y_n - \mathbf{x}_n^T \hat{\mathbf{w}}_{n-1}$$

$$\psi_n = [|e_n|, \dots, |e_{n-N+1}|]^T$$

$$\hat{\sigma}_n = \beta_1 \hat{\sigma}_{n-1} + (1 - \beta_1) \text{median}[\psi_n]$$

$$\eta_n = \beta_2 \eta_{n-1} + (1 - \beta_2) \sqrt{\tau \|\hat{\mathbf{w}}_{n-1}\|^2 \times \sigma_v^2}$$

$$\text{if } |e_n| > \xi \hat{\sigma}_n$$

$$\gamma_n = e_n^2 / (\nu \hat{\sigma}_n + |e_n|)$$

else

$$\gamma_n = \eta_n$$

end

$$\text{if } |e_n| > \gamma_n$$

$$G_n = \mathbf{x}_n^T V_{n-1} \mathbf{x}_n$$

$$\lambda = \frac{1}{G_n} \left(\frac{|e_n|}{\gamma_n} - 1 \right)$$

$$V_n = V_{n-1} - \frac{\lambda V_{n-1} \mathbf{x}_n \mathbf{x}_n^T V_{n-1}}{1 + \lambda G_n}$$

$$\hat{\mathbf{w}}_n = \hat{\mathbf{w}}_{n-1} + \lambda V_n e_n \mathbf{x}_n$$

else

$$\hat{\mathbf{w}}_n = \hat{\mathbf{w}}_{n-1}$$

$$V_n = V_{n-1}$$

end

TABLE II
 COMPUTATIONAL COMPLEXITY OF ALGORITHMS' COEFFICIENTS UPDATE PER ITERATION
 WHERE $P_1 = P(|e_n| > \xi \hat{\sigma}_n)$, $P_2 = P(|e_n| > \gamma_n)$

Algorithm	Multiplication	Addition	Square-root	Comparison	MO
RLS	$3L^2+4L+2$	$2L^2+2L$	0	0	0
BEACON-VB	$2L+2+P_2(2L^2+3L+6)$	$L+1+P_2(2L^2+L+1)$	1	1	0
Robust BEACON-VB	$2L+3P_1+5+P_2(2L^2+3L+6)$	$L+2+P_2(2L^2+L+1)$	1	2	$O(N \log N)$

TABLE III
UPDATE RATIO (20000 ITERATIONS)

Algorithm	Updates (ratio)	
	SNR=20dB	SNR=40dB
Robust QN-II[17]	21.5%	21.7%
BEACON [7]	12.3%	12.6%
Proposed robust BEACON-FB	13.7%	13.8%
Proposed robust BEACON-VB	14.1%	14.5%

TABLE IV
AVERAGE COMPUTATION TIME PER RUN (RESULTS IN SECONDS)

Algorithm	Computation time (seconds)
	SNR=20dB
Robust QN-II[17]	0.82
BEACON [7]	0.31
Proposed robust BEACON-FB	0.61
Proposed robust BEACON-VB	0.79

Fig. 1. MSD comparison of the robust BEACON algorithms, and other algorithms, where $SNR=20\text{dB}$, $\tau=3$, $\eta_0=0.35$

Fig. 2. MSD comparison of the robust BEACON algorithms, and other algorithms, where $SNR=40\text{dB}$, $\tau=3$, $\eta_0=0.35$.

Fig. 3. MSD curves of the robust BEACON with an abrupt change in system coefficients, where $SNR=10\text{dB}$, $\tau=3$, $\eta_0=2$.
 Fig. 10. Learning curves of the robust BEACON-FB with varying v for Gaussian noises. $L=32$.

Fig. 4. MSD curves of the robust BEACON with an abrupt change in system coefficients, where $SNR=40\text{dB}$, $\tau=3$, $\eta_0=0.35$.

Fig. 5. MSD curves of the BEACON with different SMEBs in impulsive noise environments, where $SNR=0dB$, $p_I=0.1$, $Q=1.88$.

Fig. 6. The impulse response of echo path.

Fig. 7. Acoustic echo cancellation. *Top*. Far-end speech signal. *Bottom*. Near-end speech signal.

Fig. 8. Performance of the BEACON family for speech signals during double-talk. $\tau=5$.

Fig. 9. Learning curves of the robust BEACON-FB with varying ζ for Gaussian noises. $L=32$.

Fig. 10. Learning curves of the robust BEACON-FB with varying v for Gaussian noises. $L=32$.