Digital Signal Processing ••• (••••) •••–•••

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Improved proportionate-type sparse adaptive filtering under maximum correntropy criterion in impulsive noise environments

Vinay Chakravarthi Gogineni *, Subrahmanyam Mula

Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, India

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Impulsive interference
Maximum correntropy criterion
Sparse adaptive filters
Proportionate adaptive filters

ABSTRACT

An improved proportionate adaptive filter based on the maximum correntropy criterion (IP-MCC) is proposed for identifying the system with variable sparsity in an impulsive noise environment. Utilization of MCC mitigates the effect of impulse noise while the improved proportionate concepts exploit the underlying system sparsity to improve the convergence rate. The performance analysis of the proposed IP-MCC reveals that the steady-state excess mean square error (EMSE) of the proposed IP-MCC filter is similar to the MCC filter. Extensive simulations demonstrate that the proposed IP-MCC outperforms the state-of-the-art in terms of convergence rate, and the detailed complexity analysis reveals that IP-MCC requires much less computational effort.

1. Introduction

The popular least mean square (LMS) family of algorithms that developed under minimum mean square error (MMSE) criterion (i.e., minimizing the ℓ_2-norm of the error) are effective in Gaussian noise environment. However, they perform poorly for the non-Gaussian impulsive interference such as low frequency atmospheric noise, many types of man-made noise and underwater acoustic noise [1]. Impulsive interference is generally characterized by a heavy-tail distribution, and research shows that a lower-order statistical measure of the error in cost function offers more robustness against impulsive noise [2]. The sign algorithm [3] which was later extended to the normalized sign algorithm (NSA) [4], both employ the mean absolute value of the error (i.e., ℓ_1-norm of the error) as the cost function, exhibit robustness against the impulsive interference. To attain the improved convergence rate in colored input conditions, the affine projection sign algorithm (APSA) [2] is proposed by minimizing ℓ_1-norm of the *a posteriori* error vector.

For applications like network echo cancellation (NEC), the system (network echo path) to be estimated is sparse in nature [5]. Since the APSA is sparsity agnostic, it is not a best choice for sparse system identification. Inspired from the proportionate adaptation [6], two proportionate affine projection sign algorithms, namely real-coefficient proportionate APSA (RP-APSA) [7] and real-coefficient improved proportionate APSA (RIP-APSA) [7] are proposed by employing the proportionate concepts [6] and improved proportionate concept [8] to the APSA. Both the RP-APSA and RIP-APSA yield faster convergence rate and lower misadjustment over APSA at the cost of increased complexity. To reduce the computational complexity, the RIP-APSA was further modified to the memory improved proportionate APSA (MIP-APSA) [9]. Recently, by employing the Lorentzian norm, a Lorentzian based adaptive filter (LAF) is proposed in [10], which was later extended to the normalized Lorentzian based hard thresholding adaptive filter (normalized LHTAF) and normalized Lorentzian based variable hard thresholding adaptive filter (normalized LVHTAF) by incorporating the iterative hard thresholding concepts [11] into the normalized LAF. Since both the normalized LVHTAF and LHTAF use the hard thresholding operator H_K (that sets all coefficients of weight vector to zero, except K largest (in magnitude) coefficients), their performance is hugely dependent on the assumed sparsity value K. In general, the sparsity of the network echo path may vary with time and context and hence these algorithms suffer while tracking these variations. Moreover, as we will show subsequently, these algorithms involve huge computational complexity, hence not suitable for real-time applications.

On the other hand, a new robust optimal criterion, namely the maximum correntropy criterion (MCC) has been successfully applied in adaptive filtering [12–17]. As the correntropy is resilient to the outliers for appropriate kernel width, MCC cost became a good choice in impulsive interference environment. In recent past, the MCC cost has been widely used to develop a series of robust adaptive filters in various non-sparse applications. MCC based quantized kernel adaptive filters have been developed in [18]. In [19,20], MCC based schemes are applied for direction of arrival

* Corresponding author.
E-mail addresses: vinaychakravarthi@ece.iitkgp.ernet.in (V.C. Gogineni), svmula@iitkgp.ac.in (S. Mula).

https://doi.org/10.1016/j.dsp.2018.04.011

(DOA) estimation in impulsive noise environments. A distributed implementation of MCC algorithm is presented in [21] to improve the performance of the estimation over network in impulsive noise environments. To extend the MCC methodology to the sparse system identification, in [22], proportionate MCC (PMCC) algorithm is proposed. However, the performance of PMCC algorithm degrades with the time-varying system sparsity. Moreover, after initial phase of rapid convergence, the convergence rate of PMCC slows down dramatically due to the stalling of inactive coefficients.

To address the aforesaid issues of Lorentzian based algorithms (i.e., robustness against the time-varying system sparsity and computational complexity) and PMCC (i.e., robustness against the time-varying system sparsity and stalling of inactive coefficients), in this paper, we propose an improved proportionate MCC (IP-MCC) algorithm by combining the MCC based adaptive filter with the improved proportionate concept [8]. Usage of MCC makes the algorithm robust against the impulsive interference and the improved proportionate adaptation ensures high convergence rate and robustness against the time varying system sparsity. Our key contributions include:

1. We designed a novel IP-MCC algorithm with modified gain factors which ensures the stability of the algorithm.
2. Without any white assumption on input, we carried out the performance analysis of the proposed IP-MCC and the analysis shows that the steady-state excess mean square error (EMSE) of the proposed IP-MCC is same as that of MCC.
3. Through extensive simulations, we demonstrate the superiority of the proposed IP-MCC in terms of convergence rate, steady-state mean square deviation (MSD) and tracking capability over the state-of-the-art.
4. We carried out a detailed complexity analysis of the IP-MCC and compared it with the existing algorithms to show that this improvement in performance is achieved at much less computational complexity.

2. Algorithm design

We consider here the problem of identifying a system that takes an input signal $u(n)$ and produces the observable output $d(n) = \mathbf{u}^T(n)\mathbf{w}_{\text{opt}} + \vartheta(n)$, where $\mathbf{u}(n) = [u(n), u(n-1), \cdots, u(n-L+1)]^T$ is the input data vector at time index n, $\mathbf{w}_{\text{opt}}$ is the $L \times 1$ system impulse response vector (to be identified) which is known *a priori* to be sparse with variable sparsity and $\vartheta(n)$ constitutes the observation noise plus impulsive interference with mean zero and variance σ_ϑ^2 which is taken to be i.i.d. and independent of input $u(m)$ for all n, m.

The MCC based stochastic gradient adaptive filter in [13] is derived by maximizing the cost function $E\big[\exp\big(-\frac{e^2(n)}{2\sigma^2}\big)\big]$ and the corresponding update equation is given by

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \exp\Big(-\frac{e^2(n)}{2\sigma^2}\Big)\, e(n)\, \mathbf{u}(n), \tag{1}$$

where μ is the adaptation step size, $e(n) = d(n) - \mathbf{w}^T(n)\mathbf{u}(n)$ is the estimation error and σ is the kernel width. For a constant step size μ in (1), the MCC filter with a small kernel width leads to a low steady-state misalignment but a slow convergence rate, and with a large kernel width it provides a fast convergence rate but a high steady-state misalignment [23]. Also note that the MCC algorithm is equivalent to LMS algorithm with a variable step size $\mu(n) = \mu \exp\big(-\frac{e^2(n)}{2\sigma^2}\big)$, with $0 \le \mu(n) \le \mu$ [23]. It can be seen that when impulsive noise occurs $e(n)$ value becomes large, then $\exp\big(-\frac{e^2(n)}{2\sigma^2}\big) \to 0$, thereby stopping the adaptation process. On the other hand, when $e(n)$ value is small, $\exp\big(-\frac{e^2(n)}{2\sigma^2}\big) \to 1$, implying MCC algorithm reduces to the LMS algorithm.

Like conventional LMS and normalized LMS (NLMS) adaptive filters, MCC algorithm is also sparse agnostic and hence it can not exploit the underlying system sparsity. To achieve this, the proportionate adaptation concepts [6] can be extended to MCC algorithm by pre-multiplying the update vector with the proportionate gain matrix $\mathbf{G}(n)$. Among the proportionate adaptation algorithms, proportionate normalized LMS (PNLMS) [6] is the most popular one. However, in [24], proportionate LMS (PLMS) is proposed by omitting the normalization term (i.e., $\mathbf{u}^T(n)\mathbf{G}(n)\mathbf{u}(n)$) which is present in PNLMS algorithm. This saves significant computational cost and it is proved that the penalty for this omission is negligible. Motivated from this, MCC filter coefficients can also be proportionately adapted without the normalization. The resultant algorithm is termed as the proportionate MCC (PMCC) and its weight update is given as follows:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \exp\Big(-\frac{e^2(n)}{2\sigma^2}\Big)\, e(n)\, \mathbf{G}(n)\, \mathbf{u}(n), \tag{2}$$

where $\mathbf{G}(n) = diag\{g_0(n), g_1(n), \cdots, g_{L-1}(n)\}$, distributes the adaption energy among the filter taps in proportion to the individual filter tap magnitude i.e., $g_i(n) \propto |w_i(n)|$ [6]. Note that similar algorithm is reported in [22] which is developed in the unified analysis context. However, PMCC performance degrades when the sparsity of the system varies over time. For example, in echo cancellation application the sparsity of the echo path changes with time. For effective identification of these time varying sparse systems, improved proportionate concept [8] is preferred. These improved proportionate concept intrinsically combines the conventional adaptation with proportionate adaptation through a mixing parameter α and the corresponding gain factors are given by [8],

$$g_i(n) = \frac{1-\alpha}{2L} + (1+\alpha)\,\frac{|w_i(n)|}{2\,\|\mathbf{w}(n)\|_1 + \epsilon_p}, \tag{3}$$

where $-1 \le \alpha \le 1$ and ϵ_p is a small positive constant, employed to avoid division by zero. By ensuring minimum adaptation energy to all the taps, the improved proportionate concept addresses the slow convergence issue of PMCC and also makes it robust against the time varying system sparsity. In this work, we combine the improved proportionate concept with MCC and the resultant algorithm is termed as improved proportionate MCC (IP-MCC).

Since the normalization term is omitted in (2), the gain factors in (3) can not be incorporated directly into the proposed IP-MCC algorithm. The gain factors in their original form render the selection of step size challenging and also they can make the algorithm unstable intermittently, especially for correlated input. This issue is not discussed in both [24] and [22]. To address this issue, we reformulate the gain factors to the following:

$$g_i(n) = \frac{1-\alpha}{2} + (1+\alpha)\,\frac{L\,|w_i(n)|}{2\|\mathbf{w}(n)\|_1 + \epsilon_p}. \tag{4}$$

From (4), we observe that the proposed IP-MCC reduces to MCC for $\alpha = -1$, whereas for $\alpha = 1$, it behaves as PMCC [22]. With this reformulation, the bounds on μ of the proposed IP-MCC will be similar to MCC which will be proved subsequently in Section 3. The proposed IP-MCC is summarized in Algorithm 1.

3. Performance analysis

The proposed IP-MCC can be seen as the transform domain MCC with transform matrix $\mathbf{G}^{\frac{1}{2}}(n)$, transformed input $\mathbf{s}(n) = \mathbf{G}^{\frac{1}{2}}(n)\,\mathbf{u}(n)$ and transformed filter coefficient vector $\mathbf{w}_t(n) = \mathbf{G}^{\frac{-1}{2}}(n)\,\mathbf{w}(n)$, as suggested in [25] in the context of PNLMS analysis. The elements of transform matrix are given by $[\mathbf{G}^{\frac{1}{2}}(n)]_{i,i} =$