Abstract—A new affine projection sign algorithm (APSA) is proposed, which is robust against non-Gaussian impulsive interferences and has fast convergence. The conventional affine projection algorithm (APA) converges fast at a high cost in terms of computational complexity and it also suffers performance degradation in the presence of impulsive interferences. The family of sign algorithms (SAs) stands out due to its low complexity and robustness against impulsive noise. The proposed APSA combines the benefits of the APA and SA by updating its weight vector according to the $L_1$-norm optimization criterion while using multiple projections. The features of the APA and the $L_1$-norm minimization guarantee the APSA an excellent candidate for combating impulsive interference and speeding up the convergence rate for colored inputs at a low computational complexity. Simulations in a system identification context show that the proposed APSA outperforms the normalized least-mean-square (NLMS) algorithm, APA, and normalized sign algorithm (NSA) in terms of convergence rate and steady-state error. The robustness of the APSA against impulsive interference is also demonstrated.

Index Terms—Adaptive filter, affine projection, sign algorithm.

I. INTRODUCTION

Adaptive filters have been commonly used in various applications of system identification, such as channel estimation, noise cancelation, echo cancelation, image restoration, and seismic system identification [1]. The most popular adaptive filters are the least-mean-square (LMS) and normalized LMS (NLMS) algorithms due to their simplicity. However, their major drawbacks are slow convergence and performance degradation with colored input signals or in the presence of heavy-tailed impulsive interferences [2].

To overcome the deterioration of convergence performance caused by colored input signals, an affine projection algorithm (APA), which is based on affine subspace projections, has been proposed in [3]. Many variants of the APA have been developed in recent years [4]. The family of APAs updates the weight coefficients by multiple, most recent input vectors instead of a single, current data vector used in the LMS and NLMS algorithms. As the projection order of the APA increases, the convergence rate increases and so does the computational complexity.

This is why computational efficient methods have also been developed to reduce the computational cost, such as the fast affine projection (FAP) algorithm [5]. In addition to the drawback of computational complexity, the APA also suffers performance degradation in non-Gaussian interference due to the nature of the $L_2$-norm optimization. Interfering signals with heavy-tailed distributions produce more outliers than Gaussian models and the $L_2$-norm minimization criterion is no longer a proper choice.

Many studies have shown that lower-order norms lead to robustness against impulsive and intensive interference. The least mean $p$-norm (LMP) algorithm based on the $L_p$-norm is proposed in [2]. Among all the lower-order algorithms, the family of sign algorithms based on the $L_1$-norm minimization has attracted more attention due to its considerably low computational cost and easy implementation. Only the sign of the error signal is involved in the updating process. Many variants of the sign algorithm have been developed, including the normalized sign algorithm (NSA) [6], dual sign algorithm (DSA) [7], and variable step-size sign algorithm [8], [9]. The mixed-norm algorithm based on the weighted combination of the $L_1$ and $L_2$ norms is proposed in [10]. The switched-norm algorithm is proposed in [11] which switches between $L_1$ and $L_2$ norms. Although the sign algorithms achieve good performance in many applications due to their low complexity and robustness against impulsive noise, they suffer from slow convergence rate, especially for highly correlated input signals.

We propose an affine projection sign algorithm (APSA) which updates the weight vector with the $L_1$-norm optimization criterion by using multiple input vectors. The combination of the benefit of affine projection and $L_1$-norm minimization improves performance on combatting impulsive interference, speeding up the convergence rate with colored input signals, and lowering the computational complexity. The weight adaptation of the proposed algorithm does not involve any matrix inversion but only uses the sign operation of the error vector. The increase of the computational burden caused by high projection orders is much lower than the conventional APA.

The performance of the proposed APSA is evaluated in the context of system identification and compared with the NSA, APA, and NLMS algorithm. Simulation results with Bernoulli-Gaussian (BG) interference and colored input signals demonstrate the robustness of the APSA against impulsive interference, outperforming the APA and NLMS algorithm. The APSA also converges much faster and reaches a smaller steady-state misalignment than the NSA algorithm.

II. CONVENTIONAL AFFINE PROJECTION ALGORITHM

Consider a system identification problem where all signals are real. The output signal from an unknown system with a weight coefficients vector $w$ is $y(k) = w^T x(k) + v(k)$, where $x(k) = [x(k), x(k-1), \ldots, x(k-L+1)]^T$ is the input signal vector of length $L$. The variable $v(k)$ represents the background noise plus interference signal. The superscript
\((\cdot)^T\) denotes vector transpose operation. Let \(\hat{w}(k)\) be an estimate of \(w\) at iteration \(k\). The a priori error is defined as 
\[e(k) = y(k) - \hat{w}(k)^T x(k),\]
while the a posteriori error is defined as 
\[e(k) = y(k) - \hat{w}(k + 1)^T x(k).\]
Grouping the \(M\) recent input vectors \(x(k)\) together gives the input signal matrix: 
\[X(k) = [x(k), x(k-1), \ldots, x(k-M+1)].\]
We define the a priori and a posteriori error vectors as 
\[e_p(k) = [e(k), e(k-1), \ldots, e(k-M+1)]^T,\]
\[e_p(k) = [e(k), e(k-1), \ldots, e(k-M+1)]^T,\]
and they can be computed as 
\[e(k) = y(k) - X^T(k)\hat{w}(k),\]
\[e_p(k) = y(k) - X^T(k)\hat{w}(k+1).\]
where \(y(k)\) is the output vector defined as 
\[y(k) = [y(k), y(k-1), \ldots, y(k-M+1)]^T.\]
The classical APA [3] is obtained by 
\[
\begin{align*}
\min_{\hat{w}(k+1)} & \quad \|y(k) - X^T(k)\hat{w}(k+1)\|_2 \\
\text{subject to} & \quad \|\hat{w}(k+1) - \hat{w}(k)\|_2 \leq \delta^2
\end{align*}
\]
III. AFFINE PROJECTION SIGN ALGORITHM

The proposed affine projection sign algorithm is obtained by minimizing the \(L_1\)-norm of the a posteriori error vector with a constraint on the filter coefficients, 
\[
\begin{align*}
\min_{\hat{w}(k+1)} & \quad \|y(k) - X^T(k)\hat{w}(k+1)\|_1 \\
\text{subject to} & \quad \|\hat{w}(k+1) - \hat{w}(k)\|_2 \leq \delta^2
\end{align*}
\]
where \(\delta^2\) is a parameter ensuring that the updating weight coefficients vector does not change dramatically [11]. We can also view \((8)\) as the minimum disturbance constraint. The minimum disturbance \(\delta\) controls the convergence level of the algorithm and it shall be as small as possible. Using the method of Lagrange multipliers, the unconstrained cost function can be obtained by combining \((7)\) and \((8)\), 
\[
J(\hat{w}(k+1)) = \|e_p(k)\|_1 + \beta \|\hat{w}(k+1) - \hat{w}(k)\|_2^2 - \delta^2
\]
where \(\beta\) is a Lagrange multiplier. The derivative of the cost function \((9)\) with respect to the weight vector \(\hat{w}(k+1)\) is 
\[
\frac{\partial J(\hat{w}(k+1))}{\partial \hat{w}(k+1)} = -\sum_{m=0}^{M-1} \text{sgn}(\varepsilon(k-m))x(k-m) \\
+ 2\beta [\hat{w}(k+1) - \hat{w}(k)] \\
= -X(k)\text{sgn}(e_p(k)) \\
+ 2\beta [\hat{w}(k+1) - \hat{w}(k)],
\]
where \(\text{sgn}(\cdot)\) denotes the sign function and \(\text{sgn}(e_p(k)) = [\text{sgn}(\varepsilon(k)), \ldots, \text{sgn}(\varepsilon(k-M+1))]^T.\)
Setting the derivative of \(J(\hat{w}(k+1))\) equal to zero, we get 
\[
\hat{w}(k+1) = \hat{w}(k) + \frac{1}{2\beta} X(k) \text{sgn}(e_p(k)),
\]
Substituting \((11)\) into the constraint \((8)\), we obtain 
\[
1 = \frac{\delta}{\sqrt{\text{sgn}(e_p(k))X(k) X^T(k) \text{sgn}(e_p(k))}},
\]
Substituting \((12)\) into \((11)\), the update equation for the weight vector is then: 
\[
\hat{w}(k+1) = \hat{w}(k) + \frac{\delta X(k) \cdot \text{sgn}(e_p(k))}{\sqrt{\text{sgn}(e_p(k))X(k) X^T(k) \text{sgn}(e_p(k))}}.
\]
Since the a posteriori error vector \(e_p(k)\) depends on \(\hat{w}(k+1)\) which is not accessible before the current update, it is reasonable to approximate it with the a priori error vector \(e(k)\). The minimum disturbance \(\delta\) controls the convergence level of the algorithm and it should be much smaller than one to guarantee convergence. It serves the similar purpose as the step-size parameter in conventional adaptive algorithms. Following the conventions, we replace \(\delta\) by the step-size parameter \(\mu\). Defining \(x_{\text{ns}}(k) = X(k) \text{sgn}(e(k))\) with \(\text{sgn}(e(k)) = [\text{sgn}(\varepsilon(k)), \ldots, \text{sgn}(\varepsilon(k-M+1))]^T,\)
we obtain the APA:
\[
\hat{w}(k+1) = \hat{w}(k) + \mu \frac{x_{\text{ns}}(k)}{\sqrt{x_{\text{ns}}(k) x_{\text{ns}}(k) + \epsilon}},
\]
where \(\epsilon\) represents the regularization parameter which should be a positive number. Since \(\mu\) comes from the minimum disturbance constraint \(\delta,\) we should choose \(0 < \mu \ll 1\) to ensure the stability of the algorithm and to achieve a small steady-state misalignment.

As shown in \((14)\), no matrix inversion is needed for the proposed APSA and it only requires \(L\) multiplications at each iteration for the normalization. In comparison, the computational complexities of the APA and FAP algorithm are \(2LM + K_{\text{inv}}M^2\) and \(2L + 20M\) [5] multiplications respectively, where \(K_{\text{inv}}\) is the factor associated with the complexity required in matrix inversion. The proposed APSA is much simpler in implementation than the APA and even the FAP. Besides, it does not have the numerical problems that the FAP exhibits. It is also worth mentioning that the APSA with \(M = 1\) reduces to a new kind of normalized sign algorithm, whose normalization is based on the Euclidean norm of the input vector. This is different from the normalized least-mean-square deviation (NLMAD) algorithm in [6] which is normalized by the \(L_1\) norm of the input vector.

ALGORITHM PERFORMANCE

The proposed APSA is compared to the NLMS, APA, and NSA via system identification applications. The adaptive filter has a length \(L = 256\) taps. The input signal is chosen to be a colored Gaussian process. This input is generated by filtering a white Gaussian noise through a first order system with a pole at 0.8 or 0.95. An independent white Gaussian noise is added to the system background with a 30 dB signal-to-noise ratio (SNR). In addition, a strong interference signal is also added to the system output \(y(k)\) with a signal-to-interference ratio (SIR) of \(-30\) to 10 dB. The Bernoulli-Gaussian (BG) distribution [11] is used for modeling the interference signal, which is generated as the product of a Bernoulli process and a Gaussian process, i.e., \(z(k) = \omega(k)n(k),\) where \(n(k)\) is a white Gaussian random sequence with zero mean and variance \(\sigma_n^2,\) and \(\omega(k)\) is

Authorized licensed use limited to: Inst Natl de la Recherche Scientific EMT. Downloaded on February 8, 2010 at 14:57 from IEEE Xplore. Restrictions apply.
Fig. 1. Misalignment of the APSA with varying values of $M = 1, 2, 5, 10, 20$ and same step size of $\mu = 0.01$. The input is an AR(1) with a pole at 0.8. The background noise is Gaussian with SNR = 30 dB. The interference is a BG with SIR=-30 dB and $P_F = 0.001$.

Fig. 2. Misalignment of the APSA with varying step sizes of $\mu = 0.1, 0.01, 0.0025, 0.001$ and same projection order of $M = 2$. Other parameters are the same as those in Fig. 1.

Fig. 3. Misalignment comparison of the APSA, NLMS, APA, and NSA. Other parameters are the same as those in Figs. 1 and 2.

Fig. 4. Misalignment comparison of the APSA, NLMS, APA, and NSA. SIR = -10 dB. Other parameters are the same as those in Fig. 3.

Fig. 5. Misalignment comparison of the APSA, NLMS, APA, and NSA. SIR = -10 dB. Other parameters are the same as those in Fig. 3.

A Bernoulli process with the probability mass function given as $P(\omega) = 1 - P_F$ for $\omega = 0$, and $P(\omega) = P_F$ for $\omega = 1$. The average power of a BG process is $P_F \cdot \sigma_n^2$. Keeping the average power constant, a BG process is spikier when $P_F$ is smaller. It reduces to a Gaussian process when $P_F = 1$.

The convergence is evaluated by the normalized misalignment [1] defined as $M(k) = 20 \log_{10} \left( \frac{||\hat{w}(k) - w||_2}{||w||_2} \right)$. The ensemble average of 20 trails is used for $M(k)$. The regularization parameter $\epsilon$ is set to 0.0001 for the APA and 0 for the APSA.

This work first examines the performance of the APSA with different projection orders $M$, as shown in Fig. 1, where $M = 1, 2, 5, 10, 20$, and the interference is a BG with SIR = -30 dB and $P_F = 0.001$ is used. The APSA with higher projection order achieves both faster convergence and lower misalignment for $M = 1$ to 10. When $M$ is larger than a certain value (in this case is 10), the convergence is faster with a larger $M$ but the
IV. CONCLUSIONS

This paper has proposed an affine projection sign algorithm (APSA) that updates its weight vector according to the sign of the a priori error vector based on the $L_1$-norm optimization. A constraint is applied on variations of the weight vector, leading to normalization based on the correlation matrix of the input signal. The proposed APSA combines the benefits of the APA and sign algorithm. The affine projection makes the APSA converge fast with colored input signals while the $L_1$ optimization guarantees its robustness against impulsive interference. In addition, the APSA has much lower computational complexity than the conventional APA because its adaptation only involves the sign operation. As a result, a large projection order can be selected to achieve faster convergence rate with affordable computational cost. Simulations have also confirmed that the proposed APSA algorithm improves the capability of combating impulsive interferences and accelerating the convergence rate with colored input signals. Its performance in Gaussian noise is also better than that of the conventional sign algorithm.

REFERENCES