Article:

Yu, Yi, Huang, Zongxin, He, Hongsen et al. (2022) Sparsity-aware robust normalized subband adaptive filtering algorithms with alternating optimization of parameters. IEEE Transactions on Circuits and Systems II: Express Briefs. pp. 3934-3938. ISSN: 1549-7747

https://doi.org/10.1109/TCSII.2022.3171672

Sparsity-Aware Robust Normalized Subband Adaptive Filtering algorithms with Alternating Optimization of Parameters

Yi Yu, *Member, IEEE*, Zongxin Huang, Hongsen He, *Member, IEEE*, Yuriy Zakharov, *Senior Member, IEEE*, and Rodrigo C. de Lamare, *Senior Member, IEEE*

***Abstract*—This paper proposes a unified sparsity-aware robust normalized subband adaptive filtering (SA-RNSAF) algorithm for identification of sparse systems under impulsive noise. The proposed SA-RNSAF algorithm generalizes different algorithms by defining the robust criterion and sparsity-aware penalty. Furthermore, by alternating optimization of the parameters (AOP) of the algorithm, including the step-size and the sparsity penalty weight, we develop the AOP-SA-RNSAF algorithm, which not only exhibits fast convergence but also obtains low steady-state misadjustment for sparse systems. Simulations in various noise scenarios have verified that the proposed AOP-SA-RNSAF algorithm outperforms existing techniques.**

***Index Terms*—Impulsive noises, subband adaptive filters, sparse systems, time-varying parameters.**

I. INTRODUCTION

FOR highly correlated input signals, the normalized subband adaptive filtering (NSAF) [1] algorithm provides faster convergence than the normalized least mean square (NLMS) algorithm and retains comparable complexity. The NSAF algorithm was proposed based on the multiband structure of subband filters [2], which adjusts the fullband filter's coefficients to remove the aliasing and band edge effects of the conventional subband structure [2]. However, in practice the non-Gaussian noise with impulsive samples could commonly happen such as in echo cancellation, underwater acoustics, audio processing, and communications [3], [4], and in this scenario, the NSAF performance degrades. To deal with impulsive noises, several robust subband algorithms based on different robust criteria were proposed, see [5]–[10] and references therein, and most of them can be unified as the NSAF update with a specific scaling factor.

Furthermore, it is interesting to improve the adaptive filter performance by exploiting the system sparsity. For example, the impulse responses of propagation channels in underwater acoustic and radio communications are usually sparse [11], [12], only a few coefficients of which are non-zero. Aiming at sparse systems, existing examples are classified into the proportionate type and sparsity-aware type. The family of proportionate NSAF (PNSAF) algorithms [13] assigns an individual gain to each filter coefficient, which has faster convergence than the NSAF algorithm. Later, robust PNSAF algorithms were also presented [10], [14] to deal with impulsive noises. On the other hand, the family of sparsity-aware algorithms incorporates the sparsity-aware penalty into the original NSAF's and PNSAF's cost functions; as a result, sparsity-aware NSAF (SA-NSAF) [15], [16] and sparsity-aware PNSAF [17] algorithms were developed. In sparse system identification, these sparsity-aware algorithms can obtain better convergence and steady-state performance than their original counterparts.

However, the superiority of sparsity-aware algorithms depends mainly on the sparsity-penalty parameter, which is often chosen in an exploratory way thus reducing the practicality of the algorithms. Besides, they encounter the problem of choosing the step-size, which controls the tradeoff between convergence rate and steady-state misadjustment. Hence, adaptation techniques for the sparsity-penalty and the step-size parameters are necessary. In the literature, they are rarely discussed simultaneously regardless of the Gaussian noise or impulsive noise scenarios. In [18], the variable parameter SA-NSAF (VP-SA-NSAF) algorithm was proposed for the Gaussian noise, in which these two parameters are jointly adapted based on a model-driven method, but it requires knowledge of variances of the subband noises. In [19], by optimizing the parameters in the sparsity-aware individual-weighting-factors-based sign subband adaptive filter (S-IWF-SSAF) algorithm with the robustness in the impulsive noise, the variable parameters S-IWF-SSAF (VP-S-IWF-SSAF) algorithm was presented, while it lacks the generality in sparsity-aware subband algorithms.

In this paper, we propose a unified sparsity-aware robust NSAF (SA-RNSAF) framework to handle impulsive noises, which can result in different algorithms by only changing the robustness criterion and the sparsity-aware penalty. We then devise adaptive schemes for adjusting the step-size and the sparsity-aware penalty weight, and develop the alternating optimization of the parameters based SA-RNSAF (AOP-SA-RNSAF) algorithm, with fast convergence and low steady-state

This work was supported in part by National Science Foundation of P.R. China (Nos. 61901400, 62071399), Doctoral Research Fund of Southwest University of Science and Technology in China (No. 19zx7122), Sichuan Science and Technology Program (No. 2021YFG0253), and Natural Science Foundation of Sichuan in 2022, China. The work of Y. Zakharov was supported in part by the U.K. EPSRC through Grants EP/R003297/1 and EP/V009591/1. Corresponding author: Hongsen He.

Y. Yu, Z. Huang and H. He are with School of Information Engineering, Robot Technology Used for Special Environment Key Laboratory of Sichuan Province, Southwest University of Science and Technology, Mianyang, 621010, China (e-mail: yuyi_xyuan@163.com).

Y. Zakharov is with the Department of Electronics, University of York, York YO10 5DD, U.K. (e-mail: yury.zakharov@york.ac.uk).

R. C. de Lamare is with CETUC, PUC-Rio, Rio de Janeiro 22451-900, Brazil. (e-mail: delamare@cetuc.puc-rio.br).

misadjustment for sparse systems.

II. Statement of Problem and SA-RNSAF Algorithm

Consider a system identification problem. The relationship between the input signal $u(n)$ and desired output signal $d(n)$ at time n is given by

$$d(n) = \boldsymbol{u}^{\mathrm{T}}(n)\boldsymbol{w}^o + \nu(n), \tag{1}$$

where the $M \times 1$ vector $\boldsymbol{w}^o$ is the impulse response of the sparse system that we want to identify, $\boldsymbol{u}(n) = [u(n), u(n-1), ..., u(n-M+1)]^{\mathrm{T}}$ is the $M \times 1$ input vector, and $\nu(n)$ is the additive noise independent of $u(n)$. For estimating $\boldsymbol{w}^o$, the SAF with a coefficient vector $\boldsymbol{w}(k)$ is used, shown in Fig. 1 with N subbands, where k denotes the sample index in the decimated domain. The input signal $u(n)$ and the desired output signal $d(n)$ are decomposed into multiple subband signals $u_i(n)$ and $d_i(n)$ via the analysis filters $\{\boldsymbol{h}_i\}_{i=1}^N$, respectively. For each subband input signal $u_i(n)$, the corresponding output of the fullband filter $\boldsymbol{w}(k)$ is $y_i(n)$. Then, both $d_i(n)$ and $y_i(n)$ are critically decimated to yield signals $d_{i,D}(k)$ and $y_{i,D}(k)$, respectively, with lower sampling rate, namely, $d_{i,D}(k) = d_i(kN)$ and $y_{i,D}(k) = \boldsymbol{u}_i^{\mathrm{T}}(k)\boldsymbol{w}(k)$, where $\boldsymbol{u}_i(k) = [u_i(kN), u(kN-1), ..., u(kN-M+1)]^{\mathrm{T}}$. By subtracting $y_{i,D}(k)$ from $d_{i,D}(k)$, the decimated subband error signals are obtained:

$$e_{i,D}(k) = d_{i,D}(k) - \boldsymbol{u}_i^{\mathrm{T}}(k)\boldsymbol{w}(k),\ i = 1, 2, ..., N, \tag{2}$$

which are used to adjust the coefficient vector $\boldsymbol{w}(k)$ [1].

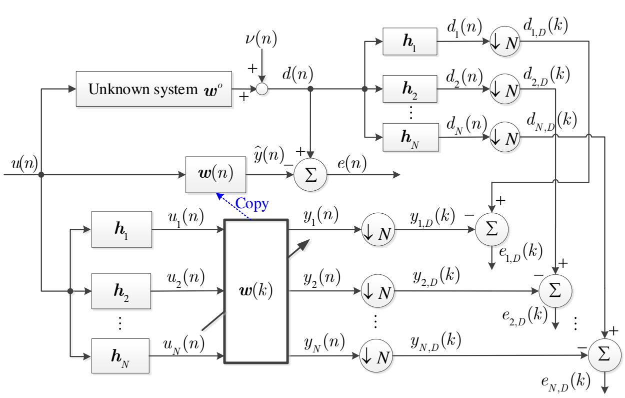

Fig. 1. Multiband structure of subband adaptive filter.

In practice, the additive noise $\nu(n)$ can be non-Gaussian consisting of Gaussian and impulsive components. Hence, for the identification of a sparse vector $\boldsymbol{w}^o$ in the presence of impulsive noise, we define the following minimization problem:

$$\arg \min_{\boldsymbol{w}(k+1)} \left[||\boldsymbol{w}(k+1) - \boldsymbol{w}(k)||_2^2 + \rho f(\boldsymbol{w}(k+1)) \right], \tag{3}$$

[1]In some applications, we could also eventually need the output error $e(n)$ in the original time domain. To this end, we obtain $\boldsymbol{w}(n)$ by copying $\boldsymbol{w}(k)$ for every N input samples, and then compute the output error by $e(n) = d(n) - \boldsymbol{u}^{\mathrm{T}}(n)\boldsymbol{w}(n)$.

subject to

$$e_{p,i}(k) = [1 - \mu\phi_i(k)]\, e_{i,D}(k), \tag{4a}$$

$$\phi_i(k) = \frac{\varphi'(e_{i,D}(k))}{e_{i,D}(k)}, \tag{4b}$$

for subbands $i = 1, ..., N$, where $e_{p,i}(k) = d_{i,D}(k) - \boldsymbol{u}_i^{\mathrm{T}}(k)\boldsymbol{w}(k+1)$ denotes the *a posteriori* decimated subband error, $\mu > 0$ will be called the step-size in the sequel, and $\phi_i(k)$ is called the scaling factor of the i-th subband. In (3), $f(\boldsymbol{w})$ is a sparsity-aware penalty function and $\rho > 0$ is the weight of this penalty term. In (4b), $\varphi'(e) \triangleq \frac{\partial\varphi(e)}{\partial e}$, where $\varphi(e) \geq 0$ is an even function of variable e, defining the robustness to impulsive noise.

By using the Lagrange multiplier method, we obtain the solution of (3) subject to (4a) as

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \mu \sum_{i=1}^N \phi_i(k) \frac{e_{i,D}(k)\boldsymbol{u}_i(k)}{||\boldsymbol{u}_i(k)||_2^2} - \rho \left[f'(\boldsymbol{w}(k+1)) - \sum_{i=1}^N \frac{\boldsymbol{u}_i(k)\boldsymbol{u}_i^{\mathrm{T}}(k)}{||\boldsymbol{u}_i(k)||_2^2} f'(\boldsymbol{w}(k+1)) \right]. \tag{5}$$

Note that the derivation of (5) also uses an approximation in the SAF, that is $\boldsymbol{u}_i^{\mathrm{T}}(k)\boldsymbol{u}_j(k) \approx 0$ for $i \neq j$ [1]. Then, by introducing an intermediate estimate $\boldsymbol{\psi}(k)$, we propose to implement (5) in two steps:

$$\boldsymbol{\psi}(k) = \boldsymbol{w}(k) + \mu \sum_{i=1}^N \phi_i(k) \frac{e_{i,D}(k)\boldsymbol{u}_i(k)}{||\boldsymbol{u}_i(k)||_2^2}, \tag{6a}$$

$$\boldsymbol{w}(k+1) = \boldsymbol{\psi}(k) - \rho \boldsymbol{P}(k), \tag{6b}$$

where

$$\boldsymbol{P}(k) = f'(\boldsymbol{\psi}(k)) - \sum_{i=1}^N \frac{\boldsymbol{u}_i(k)\boldsymbol{u}_i^{\mathrm{T}}(k)}{||\boldsymbol{u}_i(k)||_2^2} f'(\boldsymbol{\psi}(k)). \tag{7}$$

This completes the derivation of the update for the SA-RNSAF algorithm. In this algorithm, the steps (6a) and (6b) have their own roles. The former behaves like the RNSAF algorithm to obtain a coarse estimate $\boldsymbol{\psi}(k)$ of the sparse vector $\boldsymbol{w}^o$ in impulsive noise. Subsequently, the step (6b) forces the inactive coefficients in $\boldsymbol{\psi}(k)$ to zero, thus obtaining a more accurate sparse estimate $\boldsymbol{w}(k+1)$.

It is noteworthy that the parameters μ and ρ control the SA-RNSAF's performance. Specifically, the step-size μ controls the convergence rate and steady-state misadjustment of the algorithm. Moreover, the SA-RNSAF algorithm can be superior to the RNSAF algorithm when dealing with sparse systems, but ρ must be chosen within a theoretically existed range while this range is unpredictable actually (see Remark 1 below). As such, we will derive adaptive recursions for adjusting μ and ρ. However, it is challenging to solve the global optimization problem on μ and ρ, as (6a) and (6b) depend on each other. Interestingly, μ and ρ mainly affect the steps (6a) and (6b), respectively, thus we can use alternating optimization [20] to solve this global optimization problem. Accordingly, the adaptations of μ and ρ will be designed independently according to (6a) and (6b), respectively.

III. Proposed AOP-SA-RNSAF Algorithm

By using the band-dependent variable step-size (VSS) $\mu_i(k)$ and $\rho(k)$ instead of some fixed values, we rearrange (6a) and (6b) as follows:

$$\boldsymbol{\psi}(k) = \boldsymbol{w}(k) + \sum_{i=1}^{N} \mu_i(k)\phi_i(k)\frac{e_{i,D}(k)\boldsymbol{u}_i(k)}{||\boldsymbol{u}_i(k)||_2^2}, \tag{8a}$$

$$\boldsymbol{w}(k+1) = \boldsymbol{\psi}(k) - \rho(k)\boldsymbol{P}(k). \tag{8b}$$

A. Adaptation of the step-size

By subtracting (8a) from $\boldsymbol{w}^o$, we obtain

$$\widetilde{\boldsymbol{\psi}}(k) = \widetilde{\boldsymbol{w}}(k) - \sum_{i=1}^{N} \mu_i(k)\phi_i(k)\frac{e_{i,D}(k)\boldsymbol{u}_i(k)}{||\boldsymbol{u}_i(k)||_2^2}, \tag{9}$$

where $\widetilde{\boldsymbol{w}}(k) = \boldsymbol{w}^o - \boldsymbol{w}(k)$ and $\widetilde{\boldsymbol{\psi}}(k) = \boldsymbol{w}^o - \boldsymbol{\psi}(k)$ define the deviation vectors for the final estimate $\boldsymbol{w}(k)$ and the intermediate estimate $\boldsymbol{\psi}(k)$ with respect to the true value. By pre-multiplying $\boldsymbol{u}_i^{\mathrm{T}}(k)$ on both sides of (9) and applying the approximation $\boldsymbol{u}_i^{\mathrm{T}}(k)\boldsymbol{u}_j(k) \approx 0$ for $i \neq j$ again, it is established that

$$e_{\varepsilon,i}(k) = [1 - \mu_i(k)\phi_i(k)]\, e_{i,D}(k) \tag{10}$$

for $i = 1, 2, ..., N$, where $e_{\varepsilon,i}(k) = d_{i,D}(k) - \boldsymbol{u}_i^{\mathrm{T}}(k)\boldsymbol{\psi}(k)$ defines the intermediate *a posteriori* error at the i-th subband resulting from the step (6a). By squaring both sides of (10) and taking the expectations over all the terms, the following relation is obtained:

$$\mathrm{E}\{e_{\varepsilon,i}^2(k)\} = [1 - \mu_i(k)\phi_i(k)]^2\, \mathrm{E}\{e_{i,D}^2(k)\}, \tag{11}$$

where $\mathrm{E}\{\cdot\}$ denotes the mathematical expectation. In (11), a common assumption is used that the step-size $\mu_i(k)$ and the scaling factors $\{\phi_i(k)\}_{i=1}^N$ are deterministic at iteration k in contrast with the randomness of $e_{i,D}(k)$ [7], [21].

Motivated by [21], we wish to compute the subband step-sizes in such a way that $\mathrm{E}\{e_{\varepsilon,i}^2(k)\} = \sigma_{\nu,i}^2$, $i = 1, 2, ..., N$, which means that the powers of the intermediate *a posteriori* subband errors always equal those of the subband noises, where $\sigma_{\nu,i}^2 \triangleq \mathrm{E}\{\nu_{i,D}^2(k)\}$ denotes the power of the i-th subband noise excluding impulsive interferences. on this requirement, then from (11) we can obtain the following equation:

$$\mu_i(k)\phi_i(k) = 1 - \sqrt{\frac{\sigma_{\nu,i}^2}{\sigma_{e_{i,D}}^2(k)}}, \tag{12}$$

where $\sigma_{e_{i,D}}^2(k) \triangleq \mathrm{E}\{e_{i,D}^2(k)\}$ indicates the power of $e_{i,D}(k)$ without impulsive noises. For robust adaptive algorithms with the scaling factors, there is a common property [6], [7] that when impulsive noises happen, the scaling factors $\phi_i(k)$ will become very small (close to zero), thereby preventing the adaptation (8a) from the interference caused by impulsive noises. If the impulsive noise is absent, $\phi_i(k)$ will approximately equal one to ensure fast convergence. As such, we can change (12) to

$$\mu_i(k) = 1 - \sqrt{\frac{\sigma_{\nu,i}^2}{\sigma_{e_{i,D}}^2(k)}}. \tag{13}$$

To implement (13), the statistical quantities $\sigma_{e_{i,D}}^2(k)$ and $\sigma_{\nu,i}^2$ are replaced with their estimates $\hat{\sigma}_{e_{i,D}}^2(k)$ and $\hat{\sigma}_{\nu,i}^2(k)$, respectively. Specifically, $\hat{\sigma}_{e_{i,D}}^2(k)$ is calculated in an exponential window way as

$$\hat{\sigma}_{e_{i,D}}^2(k) = \zeta\hat{\sigma}_{e_{i,D}}^2(k-1) + (1-\zeta)\phi_i^2(k)e_{i,D}^2(k), \tag{14}$$

where ζ is a weighting factor often chosen as $\zeta = 1 - 1/(\kappa M)$ with $\kappa \geq 1$. Similar to [21], $\hat{\sigma}_{\nu,i}^2(k)$ is calculated by the following equations:

$$\hat{\sigma}_{u_i}^2(k) = \zeta\hat{\sigma}_{u_i}^2(k-1) + (1-\zeta)u_i^2(kN), \tag{15a}$$

$$\hat{\boldsymbol{r}}_{ue_i}(k) = \zeta\hat{\boldsymbol{r}}_{ue_i}(k-1) + (1-\zeta)\phi_i(k)e_{i,D}(k)\boldsymbol{u}_i(k), \tag{15b}$$

$$\hat{\sigma}_{\nu,i}^2(k) = \hat{\sigma}_{e_{i,D}}^2(k) - \frac{||\hat{\boldsymbol{r}}_{ue_i}(k)||_2^2}{\hat{\sigma}_{u_i}^2(k) + \epsilon_1}, \tag{15c}$$

where ϵ_1 is a small positive number (e.g., 10^{-5}). Note that, we introduce the scaling factor $\phi_i(k)$ in (14) and (15b) for each subband to suppress impulsive noises.

Accordingly, (13) can be rewritten as

$$\mu_i(k) = 1 - \sqrt{\frac{\hat{\sigma}_{\nu,i}^2(k)}{\hat{\sigma}_{e_{i,D}}^2(k) + \epsilon_2}}, \tag{16}$$

where ϵ_2 is also a small positive number. It is stressed that the estimated values of multiple statistical quantities are used in (15c), and thus $\hat{\sigma}_{\nu,i}^2(k)$ could be negative at some iterations. To avoid this, we add the step $\hat{\sigma}_{\nu,i}^2(k) \leftarrow \hat{\sigma}_{\nu,i}^2(k-1)$ after (15c).

B. Adaptation of the sparsity penalty weight

By subtracting (8b) from $\boldsymbol{w}^o$, we obtain

$$\widetilde{\boldsymbol{w}}(k+1) = \widetilde{\boldsymbol{\psi}}(k) + \rho(k)\boldsymbol{P}(k). \tag{17}$$

By pre-multiplying both sides of (17) by their transpose, we obtain

$$||\widetilde{\boldsymbol{w}}(k+1)||_2^2 = ||\widetilde{\boldsymbol{\psi}}(k)||_2^2 + \triangle(k), \tag{18}$$

where

$$\triangle(k) = 2\rho(k)\widetilde{\boldsymbol{\psi}}^{\mathrm{T}}(k)\boldsymbol{P}(k) + \rho^2(k)||\boldsymbol{P}(k)||_2^2. \tag{19}$$

Remark 1: (18) clearly reveals that the proposed SA-RNSAF algorithm will outperform the RNSAF algorithm for identifying sparse systems, if and only if $\triangle(k) < 0$ holds[2]. It follows that $\rho(k)$ should satisfy the inequality

$$0 < \rho(k) < 2\frac{[\boldsymbol{\psi}(k) - \boldsymbol{w}^o]^{\mathrm{T}}\boldsymbol{P}(k)}{||\boldsymbol{P}(k)||_2^2}. \tag{20}$$

Moreover, since $\triangle(k)$ is the quadratic function of $\rho(k)$, there is an optimal $\rho(k)$ such that $\triangle(k)$ arrives at the negative maximum value. Consequently, the optimal $\rho(k)$ is given as

$$\rho_{\mathrm{opt}}(k) = \frac{[\boldsymbol{\psi}(k) - \boldsymbol{w}^o]^{\mathrm{T}}\boldsymbol{P}(k)}{||\boldsymbol{P}(k)||_2^2}. \tag{21}$$

Although Remark 1 states that the relations (20) and (21) are existing in sparse systems, they are incalculable due to the fact that the sparse vector $\boldsymbol{w}^o$ is unknown. To solve this

[2]Following a derivation similar to that in Appendix D in [19], $\triangle(k) < 0$ is likely to be true as long as $\boldsymbol{w}^o$ is sparse.

problem, we use the previous estimate $\boldsymbol{w}(k)$ to approximate $\boldsymbol{w}^o$, then (21) can be reformulated as

$$\hat{\rho}_{\text{opt}}(k) = \max\left\{\frac{[\boldsymbol{\psi}(k) - \boldsymbol{w}(k)]^{\mathrm{T}}\boldsymbol{P}(k)}{||\boldsymbol{P}(k)||_2^2}, 0\right\}, \tag{22}$$

where $\hat{\rho}_{\text{opt}}(k)$ is set to zero at $k = 0$.

The recursion (8) equipped with $\mu_i(k)$ in (16) and $\hat{\rho}_{\text{opt}}(k)$ in (22) constitutes the proposed AOP-SA-RNSAF algorithm.

Remark 2: The proposed AOP-SA-RNSAF update generalizes different algorithms, depending on the choice of $\varphi(e)$ in (4b) and $f(\boldsymbol{w})$ in (3). In the literature, several robust criteria against impulsive noises [6], [7], [9], [10], [14] defined by $\varphi(e)$ and sparsity-aware penalties [15], [16], [18], [19], [22] defined by $f(\boldsymbol{w})$ have been studied, which can be applied in the AOP-SA-RNSAF. Nevertheless, this paper does not consider the effect of different choices of $\varphi(e)$ and/or $f(\boldsymbol{w})$, which is worth studying in future work. Note that, when setting $\varphi(e) = \frac{1}{2}e^2$, the proposed algorithm is called the alternating optimization of parameters based SA-NSAF (AOP-SA-NSAF) suited for Gaussian noise environments, which is a sparsity-aware variant of the VSS-NSAF algorithm presented in [21].

Remark 3: By firstly computing the inner product $\boldsymbol{u}_i^{\mathrm{T}}(k)f'(\boldsymbol{\psi}(k))$ in $\boldsymbol{P}(k)$, and then calculating $\boldsymbol{P}(k)$ only requires $2M$ multiplications, $2M - M/N$ additions, and 1 division. Therefore, the complexity of the proposed AOP-SRNSAF algorithm is still low with $\mathcal{O}(M)$ arithmetic operations per input sample.

IV. Simulation Results

To evaluate the proposed AOP-SA-RNSAF algorithm, simulations are conducted to identify the acoustic echo paths with $M = 512$ taps. The sparsenesses, defined as $\chi(\boldsymbol{w}^o) = \frac{M}{M-\sqrt{M}}\left(1 - \frac{||w^o||_1}{\sqrt{M}||w^o||_2}\right)$, of two echo paths are $\chi(\boldsymbol{w}_1^o) = 0.9357$ (sparse) [19] and $\chi(\boldsymbol{w}_2^o) = 0.3663$ (dispersive or non-sparse) [14], respectively. The length of the adaptive filters is the same as that of $\boldsymbol{w}^o$. The correlated input signal $u(n)$ is a first-order autoregressive (AR) process with the pole at 0.9, generated by filtering a white Gaussian noise with zero-mean and unit variance. The analysis filters $\{\boldsymbol{h}_i\}_{i=0}^{N-1}$ for decomposing signals $d(n)$ and $u(n)$ are obtained by cosine-modulated filter banks, where the length of the prototype filter for $N = 4$ subbands is 33 to obtain 60 dB stopband attenuation. The high stopband attenuation is to guarantee that adjacent analysis filters have almost no overlap and the cross-correlation of nonadjacent subbands is negligible [2]. The normalized mean square deviation (NMSD), defined as $\mathrm{E}\{||\boldsymbol{w}(n) - \boldsymbol{w}^o||_2^2/||\boldsymbol{w}^o||_2^2\}$, is the performance measure. All the results are the average over 50 independent runs.

For the AOP-SA-RNSAF algorithm, we use the modified Huber (MH) function for $\varphi(e)$ and the log-penalty for $f(\boldsymbol{\psi}(k))$. The MH function is formulated as $\varphi(e_{i,D}(k)) = e_{i,D}^2(k)/2$ if $|e_{i,D}(k)| < \xi_i$ and $\varphi(e_{i,D}(k)) = 0$ if $|e_{i,D}(k)| \geq \xi_i$ [10], where ξ_i is the threshold. Accordingly, when $|e_{i,D}(k)| \geq \xi_i$ (usually impulsive noises occur), then the scaling factor in (4b) is $\phi_i(k) = 0$, which makes the adaptation step (8a) freeze to suppress impulsive interferences; otherwise, $\phi_i(k) = 1$, which retains fast convergence. Note that, the threshold ξ_i for each subband i is set to $\xi_i = 2.576\hat{\sigma}_{\varepsilon,i}(k)$, where $\hat{\sigma}_{\varepsilon,i}^2(k)$ is the variance of $e_{i,D}(k)$ excluding impulsive samples. $\hat{\sigma}_{\varepsilon,i}^2(k)$ is computed by $\hat{\sigma}_{\varepsilon,i}^2(k) = \lambda\hat{\sigma}_{\varepsilon,i}^2(k-1) + c_\sigma(1-\lambda)\text{med}(\boldsymbol{a}_{\varepsilon,i})$, where $\lambda \in (0.9, 1)$ is the forgetting factor (but $\lambda = 0$ at $k = 0$), med$(\cdot)$ denotes the median operator to remove outliers in the data window $\boldsymbol{a}_{\varepsilon,i} = [e_{i,D}^2(k), e_{i,D}^2(k-1), ..., e_{i,D}^2(k-N_w+1)]$ with a length of N_w, and $c_\sigma = 1.483(1 + 5/(N_w - 1))$ is the correction factor. The log-penalty is given as $f(\boldsymbol{\psi}_k) = \sum_{m=1}^{M}\ln(1 + |\psi_{m,k}|/\theta)$ [19] which characterizes the sparsity of systems, where $\psi_{m,k}$ is the m-th element of $\boldsymbol{\psi}_k$, and the shrinkage factor $\theta > 0$ cuts apart inactive and active entries in $\boldsymbol{\psi}_k$. Thus, $f'(\boldsymbol{\psi}_{m,k})$ in (7) is computed element-wise as $f'(\psi_{m,k}) = \frac{\text{sgn}(\psi_{m,k})}{\theta + |\psi_{m,k}|}, m = 1, ..., M$. In our simulations, the additive noise $\nu(n)$ is described by the symmetric α-stable process, also called the α-stable noise, whose characteristic function is formulated as $\phi(t) = \exp(-\gamma|t|^\alpha)$ [3]. The parameter $\alpha \in (0, 2]$ represents the impulsiveness of the noise that for smaller α leads to stronger impulsive noises, and $\gamma > 0$ behaves like the variance of the Gaussian density. In particular, it reduces to the Gaussian noise for the case of $\alpha = 2$. In the following simulations, we set $\gamma = 0.02$.

Example 1: the impulsive noise is absent, i.e., $\alpha = 2$. The proposed AOP-SA-NSAF algorithm in Remark 2 is compared with the NSAF, VP-SA-NSAF [18], VSS-NSAF, and VSS-PNSAF algorithms in Fig. 2, where both VSS-NSAF and VSS-PNSAF are obtained from [10] but in the Gaussian noise we reset $\varphi(e) = \frac{1}{2}e^2$ instead of using the MH function. For a fair evaluation, we select the log-penalty parameter $\theta = 0.005$ for all the sparsity-aware algorithms. As can be seen, the VSS-NSAF algorithm obtains fast convergence and low steady-state misadjustment, which overcomes the trade-off problem in the NSAF algorithm. By considering the sparsity of the underlying system, both VP-SA-NSAF and VSS-PNSAF algorithms further improve the convergence rate. As compared to the VSS-PNSAF algorithm, the proposed AOP-SNSAF algorithm shows slower initial convergence, but it achieves higher reduction in the steady-state misadjustment.

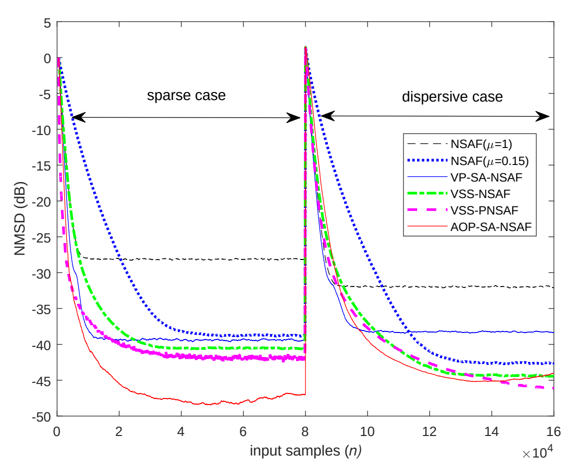

Fig. 2. NMSD performance of NSAF-type algorithms in the Gaussian noise. The parameters of algorithms are listed as follows: $\eta = 0.99$, $\lambda = 0.95$, and $\mu_{\max} = 1$ for VP-SA-NSAF; $\tau = 3$ for VSS-NSAF; $\tau = 5$ and $\zeta = 0$ for VSS-PNSAF; $\kappa = 6$ for AOP-SA-NSAF.

Example 2: $\alpha = 1.6$ displays the presence of impulsive noises. Fig. 3 depicts the NMSD performance of the NSAF,

M-NSAF [10], VSS-M-NSAF [10], VSS-M-PNSAF [10], VP-IWF-SSAF with RA [19], and the proposed AOP-SA-RNSAF algorithms[3]. For the M-estimate based algorithms, we choose the common M-estimate parameters $\lambda = 0.99$ and $N_w = 20$. It is seen that the NSAF algorithm shows poor misadjustment in the α-stable noise, and other algorithms exhibit robust convergence. Among these robust algorithms, the proposed AOP-SA-RNSAF algorithm is the best choice for identifying sparse systems, due to the fact that it has lower steady-state misadjustment than the VSS-M-PNSAF and VP-S-IWF-SSAF with RA algorithms, even if it has a slightly slower initial convergence than the VSS-M-PNSAF algorithm.

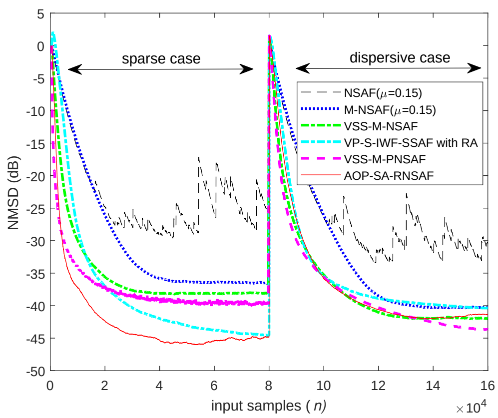

Fig. 3. NMSD performance of NSAF-type algorithms in the α-stable noise. The parameters of algorithms are listed as follows: $\mu_{\min} = 10^{-5}$, $\tau = 2$, and $\chi = 1$ for VP-S-IWF-SSAF; $\tau = 3$ for VSS-M-NSAF; $\tau = 5$ and $\zeta = 0$ for VSS-M-PNSAF; $\kappa = 6$ for AOP-SA-RNSAF.

It can be seen in Figs. 2 and 3 that, after $\boldsymbol{w}^o$ becomes dispersive at the middle of input samples, the proportionate-type (i.e., VSS-PNSAF, VSS-PNSAF) and sparsity-aware type (i.e., VP-SA-NSAF, AOP-SA-NSAF, AOP-SA-RNSAF) algorithms still show almost the same performance as the competing algorithms (i.e., VSS-NSAF and VSS-M-NSAF) in both Gaussian and α-stable noise scenarios. In addition, as α decreases from 2 to 1.6, the steady-state misadjustment of the proposed AOP-SA-RNSAF algorithm increases, but this algorithm is still convergent.

V. Conclusion

In this paper, a unified SA-RNSAF framework for algorithms was developed for identifying sparse systems in impulsive noise environments. By replacing directly the specified robustness criterion and sparsity-aware penalty, it can yield different SA-RNSAF algorithms. We then developed adaptive techniques for the step-size and the sparsity penalty weight in the SA-RNSAF algorithm, thus arriving at the AOP-SA-RNSAF algorithm with a further performance improvement in terms of the convergence rate and steady-state misadjustment. Simulations for the sparse system identification have demonstrated the effectiveness of the proposed algorithms.

[3]Since the variance of the α-stable noise is nonexistent, here we do not show the performance of the VP-SA-NSAF algorithm.

References

[1] K.-A. Lee and W.-S. Gan, "Improving convergence of the NLMS algorithm using constrained subband updates," *IEEE Signal Processing Letters*, vol. 11, no. 9, pp. 736–739, 2004.

[2] K.-A. Lee, W.-S. Gan, and S. M. Kuo, *Subband adaptive filtering: theory and implementation*. John Wiley & Sons, 2009.

[3] C. L. Nikias and M. Shao, *Signal processing with Alpha-stable distributions and applications*. Wiley-Interscience, 1995.

[4] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications," *IEEE Transactions on Electromagnetic Compatibility*, vol. 44, no. 1, pp. 249–258, 2002.

[5] J.-H. Kim, J. Kim, J. H. Jeon, and S. W. Nam, "Delayless individual-weighting-factors sign subband adaptive filter with band-dependent variable step-sizes," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 7, pp. 1526–1534, 2017.

[6] F. Huang, J. Zhang, and S. Zhang, "Combined-step-size normalized subband adaptive filter with a variable-parametric step-size scaler against impulsive interferences," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 11, pp. 1803–1807, 2017.

[7] J. Hur, I. Song, and P. Park, "A variable step-size normalized subband adaptive filter with a step-size scaler against impulsive measurement noise," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 64, no. 7, pp. 842–846, 2016.

[8] Z. Zheng, Z. Liu, and X. Lu, "Robust normalized subband adaptive filter algorithm against impulsive noises and noisy inputs," *Journal of the Franklin Institute*, vol. 357, no. 5, pp. 3113–3134, 2020.

[9] Y. Yu, H. Zhao, B. Chen, and Z. He, "Two improved normalized subband adaptive filter algorithms with good robustness against impulsive interferences," *Circuits, Systems, and Signal Processing*, vol. 35, no. 12, pp. 4607–4619, 2016.

[10] Y. Yu, H. He, B. Chen, J. Li, Y. Zhang, and L. Lu, "M-estimate based normalized subband adaptive filter algorithm: Performance analysis and improvements," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 225–239, 2020.

[11] J. Radecki, Z. Zilic, and K. Radecka, "Echo cancellation in IP networks," in *The 45th Midwest Symposium on Circuits and Systems (MWSCAS)*, vol. 2, 2002, pp. II–II.

[12] W. F. Schreiber, "Advanced television systems for terrestrial broadcasting: Some problems and some proposed solutions," *Proceedings of the IEEE*, vol. 83, no. 6, pp. 958–981, 1995.

[13] M. S. E. Abadi and S. Kadkhodazadeh, "A family of proportionate normalized subband adaptive filter algorithms," *Journal of the Franklin Institute*, vol. 348, no. 2, pp. 212–238, 2011.

[14] Z. Zheng, Z. Liu, H. Zhao, Y. Yu, and L. Lu, "Robust set-membership normalized subband adaptive filtering algorithms and their application to acoustic echo cancellation," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 8, pp. 2098–2111, 2017.

[15] Y. Yu, H. Zhao, R. C. de Lamare, and L. Lu, "Sparsity-aware subband adaptive algorithms with adjustable penalties," *Digital Signal Processing*, vol. 84, pp. 93–106, 2019.

[16] E. Heydari, M. S. E. Abadi, and S. M. Khademiyan, "Improved multiband structured subband adaptive filter algorithm with l0-norm regularization for sparse system identification," *Digital Signal Processing*, p. 103348, 2021.

[17] N. Puhan and G. Panda, "Zero attracting proportionate normalized subband adaptive filtering technique for feedback cancellation in hearing aids," *Applied Acoustics*, vol. 149, pp. 39–45, 2019.

[18] L. Ji and J. Ni, "Sparsity-aware normalized subband adaptive filters with jointly optimized parameters," *Journal of the Franklin Institute*, vol. 357, no. 17, pp. 13 144–13 157, 2020.

[19] Y. Yu, T. Yang, H. Chen, R. C. de Lamare, and Y. Li, "Sparsity-aware ssaf algorithm with individual weighting factors: Performance analysis and improvements in acoustic echo cancellation," *Signal Processing*, vol. 178, p. 107806, 2021.

[20] M. Hong, Z.-Q. Luo, and M. Razaviyayn, "Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems," *SIAM Journal on Optimization*, vol. 26, no. 1, pp. 337–364, 2016.

[21] J. Ni and F. Li, "A variable step-size matrix normalized subband adaptive filter," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 6, pp. 1290–1299, 2010.

[22] R. C. de Lamare and R. Sampaio-Neto, "Sparsity-aware adaptive algorithms based on alternating optimization and shrinkage," *IEEE Signal Processing Letters*, vol. 21, no. 2, pp. 225–229, 2014.