

Received January 10, 2020, accepted February 21, 2020, date of publication February 27, 2020, date of current version March 9, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2976499

Nonlinearity Design With Power-Law Tails for Correlation Detection in Impulsive Noise

ZHONGTAO LUO¹, (Member, IEEE), AND EDMOND A. JONCKHEERE², (Life Fellow, IEEE)

¹School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

²Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA

Corresponding author: Zhongtao Luo (luozt@cqupt.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61701067, and in part by the U.S. National Science Foundation under Grant CCF-1423624.

ABSTRACT Impulsive noise plays an important role in power line communication among other applications. To improve the communication performance, this paper proposes a novel design of nonlinear processing which improves the fundamental performance of signal detection in impulsive noise. Power-law tails are firstly introduced in nonlinearity design to provide adjustable decay factors for different distributions. Four modes of nonlinearity functions are developed and analyzed. By taking the exponent and the threshold as two arguments, we formulate the nonlinearity design as an optimization problem of maximizing the efficacy function, which is the fundamental measurement for detecting a deterministic signal in impulsive noise. Given that the efficacy function is differentiable, unimodal but without closed-form derivatives, we propose to solve the optimization problem by derivative-free methods, e.g. the Nelder-Mead simplex method. As concept demonstration, our method is used for three commonly-used distribution examples. Results show that our nonlinearity design can achieve almost the same efficacy and detection performance as the locally optimal detector, with the advantage of easy-to-apply closed form expressions.

INDEX TERMS Impulsive noise, signal detection, nonlinearity, power-law tail, numerical optimization.

I. INTRODUCTION

While Gaussian noise is generally encountered in most systems, impulsive noise raises additional consideration for some scenarios, e.g. long-wave communications and ultra-wideband systems [1], [2]. Specifically, for power line communication (PLC) technology, which is very attractive as a smart grid application given such advantages as no additional installation costs [3], [4], the impulsive noise over the PLC channel may severely deteriorate the communication performance [5]. Although multi-carrier modulations, e.g. the orthogonal frequency division multiplexing (OFDM), are inherently more resistant to impulsive noise than single carrier modulations, the counter measures to the performance degradation caused by impulsive noise is still a challenging research area for communication engineers [6].

Impulsive noise possesses a unimodal probability density function (PDF) that is similar to the Gaussian PDF, but with significantly heavier tails. Up to now, various impulsive

noise models have been developed and used in research and applications, including the symmetric α -stable ($S\alpha S$) distribution [7], [8], the Middleton Class A/B distribution [9], [10], the Gaussian mixture model referring to as Bernoulli-Gaussian random process [11], [12], Poisson distribution of Class A noise and Nakagami-m noise [13]. Recently, a K -component Gaussian mixture model has been used to approximate the Class A noise and the $S\alpha S$ noise [6]. A hidden Markov Middleton model was adopted to characterize impulsive noise bursts [14].

For signal detection in impulsive noise, the maximum likelihood detector (MLD) is optimal; however, it has high computation complexity and needs prior knowledge of signal amplitude [15]. Therefore, investigators usually consider the locally optimal detector (LOD) for its simple structure, which contains two steps, i.e. nonlinear processing and linear correlation [16]. In low signal-to-noise ratio (SNR), the LOD can be almost optimal compared with the MLD. However, the LOD for impulsive noise has the drawback that its nonlinear function is possibly unavailable in closed-form. For instance, the $S\alpha S$ model and the Class A model do not admit closed-form PDFs.

The associate editor coordinating the review of this manuscript and approving it for publication was Qinghua Guo¹.

Methods for impulsive noise suppression can be classified into two categories. One kind is the zero-memory nonlinearity (ZMNL) transformation, which is developed to replace the nonlinearity of the LOD. Given a proper design, the ZMNL functions can be near optimal. Besides, the ZMNL functions are simple and easy to implement in traditional structures. The other kind is adaptive filtering, which is usually designed under a carefully chosen criterion, such as the least square criterion [14], the minimum dispersion criterion [17], the maximum correntropy criterion [18], the logarithmic least mean p th-power criterion [19].

This paper focuses on the ZMNL design. Generally, the ZMNL function works like a limiter for large-amplitude samples and consists of two basic regions, i.e. the linear (or near-linear) region and the nonlinear tail. Usually, traditional nonlinearity is designed based on a specific noise model, trying to optimize the breakpoint threshold under certain considerations, e.g. empirical generalization [20], output SNR maximization [5], [21] [22], Neyman-Pearson criteria [23], [24], the PDF analysis and approximation [25], [26], the signal envelop and peaks [27], [28] [29].

However, as traditional nonlinearity designs emphasize the evaluation of linear region thresholds, they ignore the design of tails but preset them instead. The most widely-used tails are the blanker and the clipper (or called soft limiter) [8], [25], while other tails are also used, such as joint blanking/clipping [5], multiple thresholds for blanking or clipping [6], deep clipping [30], Gaussian tail [20], algebraic tail [31], Cauchy tail [32]. These tails are proposed, inspired from the tails of the LOD. Certainly, the LOD function has the optimal “tail”, which varies for different noise models. However, up to now, researchers have not proposed an effective approach for approximating the tails of different noise models.

A fixed tail cannot be optimal for every distribution. The performances of existing tails vary greatly across different distributions. For example, the clipper is worse than the blanker in the Class A model, but better than the blanker in the $S\alpha S$ model. Moreover, the algebraic-tailed ZMNL (AZMNL) can approximate the LOD nonlinearity of the $S\alpha S$ distribution better than the blanker [26]. Our previous work has analyzed the ZMNL designs based on algebraic tail [33] and Gaussianization [34], both of which are only sub-optimal compared to the LOD. It is demonstrated that the match between tails and distributions is essential for nonlinearity in impulsive noise.

The tail design provides extra gain over the base of the threshold evaluation. In [35], we improved the Gaussian tail design and obtained nearly optimal performance for the $S\alpha S$ noise. Recently, two papers proposed new nonlinear tails. In [6], piecewise attenuation and clipping with multiple thresholds are developed for the K -component Gaussian mixture model. But for practical applications, the threshold number and the performance-complexity trade-off may be a problem. In [21], the nonlinear function for the $S\alpha S$ distribution is subject to the cumulative distribution

function (CDF) set. However, this design is only useful for known signal amplitude.

This paper proposes to optimize the tail by the power-law function in the nonlinearity design. The decay factor of tails is varied by employing the power-law function x^a for $a \leq 0$, so as to suit different distributions of impulsive noise. Compared with traditional designs that optimize the threshold T for fixed tails, our design optimizes an extra argument, i.e. the exponent a . New parametric nonlinear functions are developed in four modes with different properties.

To formulate the nonlinearity design, the objective function is defined by the efficacy function, which is closely related to the output SNR as well as the detection performance [21], [25]. Therefore, the design becomes an efficacy optimization problem with respect to two arguments (T, a). Analysis shows that the efficacy function is continuous, differentiable, and unimodal with respect to T and a , which makes the optimization problem convenient to solve by numerical derivative-free methods [36], such as the Powell’s method [37] and the Nelder-Mead simplex (NMS) method [38], [39]. The solution algorithm will be provided and simulated.

Our work in this paper is summarized as follows:

- First, we introduce the power-law tail for the nonlinearity design in impulsive noise. Unlike traditional clipping, blanking, or other fixed tails, the power-law function x^a can adapt a to vary its decay speed so that it can be suitable for various distributions of impulsive noise.
- Second, the nonlinearity design is converted into the problem of optimizing the power-law parameters for maximizing the efficacy. This problem is hard to solve theoretically, but can easily be solved by a derivative-free method, e.g. the NMS method.
- Third, the power-law nonlinearity is effective for the commonly-used models of impulsive noise, e.g. the $S\alpha S$ and the Class A models. Given in closed-form expressions, the power-law nonlinearity is nearly as optimal as the LOD. However, the LOD may bear high computational complexity without closed-form noise PDFs.

The remainder of this paper is organized as follows. Section II briefly describes the system model and nonlinear processing for impulsive noise. Section III develops the nonlinear functions in four modes. Section IV formulates the nonlinearity design as an optimization problem and analyzes its properties. Section V presents the solution algorithm. Then, Section VI simulates the proposed design in three models of impulsive noise and discusses the results. Then, Section VII presents the applications and advantages of the proposed method. Finally, conclusions are drawn in Section VIII.

II. SIGNAL MODEL AND NONLINEAR PROCESSING

Consider the detection of a deterministic signal in additive white noise. Given M samples available for correlation detection, under hypothesis H_i , the received signal model is

$$H_i : r[m] = \xi_i \cdot s_i[m] + w[m] \quad (1)$$

for $m = 1, 2, \dots, M$, where $s_i[m]$, ξ_i , and $w[m]$ denote the desired signal, the signal amplitude, and the noise, respectively. The model (1) can describe a wide range of signal detection scenarios, by adapting the definitions of $s_i[m]$ and ξ_i consistently with different hypotheses. For instance, in communication systems with H_i denoting bit i , $s_i[m]$ and ξ_i denote the corresponding waveform and amplitude, respectively.

For signal detection in impulsive noise, the correlation detector mostly uses a two-step structure which consists of nonlinear processing and linear correlation. By denoting the nonlinear function as $g(x)$, the correlation output is

$$T_i(g) = \sum_{m=1}^M g(r[m]) \cdot s_i[m]. \quad (2)$$

Since the noise is white, for low input SNR, by using the central-limit theorem, the test statistic $T_i(g)$ in (2) asymptotically follows a Gaussian distribution¹ as

$$T_i(g) \overset{a}{\sim} \mathcal{N}(\xi_i \mathcal{E}_{gl} E_i, \mathcal{E}_{gg} E_i) \quad (3)$$

for $M \rightarrow \infty$, $E_i = \sum_{m=1}^M (s_i[m])^2$, and

$$\mathcal{E}_{gl} = - \int_{-\infty}^{\infty} g(x) f'(x) dx, \quad \mathcal{E}_{gg} = \int_{-\infty}^{\infty} g^2(x) f(x) dx, \quad (4)$$

where $f(x)$ denotes the noise PDF [25], [31]. The output SNR can be derived from (3).

Among the ZMNL functions for nonlinear processing, the LOD function

$$g_{lo}(x) = -f'(x)/f(x). \quad (5)$$

is theoretically optimal in low SNR. However, the LOD $g_{lo}(x)$ is hard to manipulate when the PDF $f(x)$ or its derivative $f'(x)$ is not in closed-form. For instance, the α -stable distribution cannot provide a closed-form PDF, while the Class A distribution gives its PDF as an infinite series.

Closed-form nonlinearity can be obtained by developing ZMNL functions that have similar shapes as the LOD function. Two kinds of limiters, i.e. the blanker and the clipper, are most popular, formulated as

$$g_{bl}(x) = \begin{cases} x, & |x| \leq T, \\ 0, & |x| > T, \end{cases} \quad (6)$$

$$g_{cl}(x) = \begin{cases} x, & |x| \leq T, \\ T \cdot \text{sgn}(x), & |x| > T, \end{cases} \quad (7)$$

where T is the threshold, and $\text{sgn}(\cdot)$ denotes the signum function. The effectiveness of ZMNL functions in impulsive noise has been demonstrated in the literature and in practical applications. However, compared to the LOD, most ZMNL functions bear some performance loss. Moreover, traditional ZMNLs are generally developed based on specific models and applicable to a limited class of impulsive noises.

¹Generally, an impulsive noise has zero-mean, is symmetric, and has unimodal PDF, so that $f(x)$ is even and $f'(x)$ is odd. Like the LOD nonlinearity $g_{lo}(x)$ in (5), $g(x)$ is designed as an odd function. To suppress the impulsive nature of the noise, $g(x)$, as a random variable, has finite variance, which satisfies the condition for validity of the central-limit theorem.

III. NONLINEARITY DESIGN WITH POWER-LAW TAILS

This section introduces the power function for tail design and proposes new nonlinear functions with power-law tails.

A. ADVANTAGES OF A POWER-LAW TAIL

We aim to develop a nonlinear processor that can approach the optimal detection performance for a wide range of heavy-tailed distributions instead of a specific model. Thus, our design should work for different distributions and approximate the LOD better than the traditional ZMNL designs.

Traditional ZMNL designs focus on placing the threshold of the linear region for preset tails. However, the designed tail should vary according to the noise distribution, like the LOD tail which always depends on the noise PDF. The difference between the designed tail and the LOD tail can lead to significant loss in the optimality performance.

This paper proposes to optimize the tail function in the nonlinearity design and at the same time add an extra degree-of-freedom in approximating the LOD function and improving the detection performance. Since a tail function is generally odd, we will shape $g(x)$ for $x \geq 0$, as the nonlinear function over $x \leq 0$ can be obtained as $g(x) = g(|x|) \cdot \text{sgn}(x)$.

The power function $y = x^a$ has two significant advantages. (i) $y = x^a$ is simple and analytical. It is much easier to use than the LOD. (ii) $y = x^a$ can achieve various levels of decay rate by varying the exponent $a \leq 0$. We use it in the nonlinear region and construct a new nonlinearity with a power-law tail that can generate different levels of decay rate in the nonlinear region so as to match various distributions.

In addition, traditional ZMNL tails can be viewed as special cases of power-law tails. For example:

- (i) $a = 1$ coincides with the linear processing $g(x) = x$;
- (ii) $a = 0$ obtains the clipper by multiplying by a constant T ;
- (iii) $a \rightarrow -\infty$ approximates the blanker² by $(x/T)^a$.

Therefore, the power-law tail suite includes the traditional tails. By employing the exponent a as an additional degree-of-freedom, the new nonlinearity design can achieve significant gain, no less than the traditional designs of blanker and clipper.

B. PARAMETRIC NONLINEAR FUNCTIONS IN FOUR MODES

After defining the linear region as directly proportional to x and the nonlinear region as power-law x^a , the next question is how to combine them as a continuous function. Herein, the continuity is emphasized, not only for the similarity to the LOD $g_{lo}(x)$ which is continuous in common distributions, but also for the analysis on nonlinearity optimization.

As for combining the linear and nonlinear regions, four modes can be considered for the nonlinear functions.

²As shown in (6), the blanker has a sudden drop when x exceeds the threshold T . Consider a continuous limiter with a power-law tail, as will be introduced in (8), $g_{lm}(x) = x$ for $0 \leq x < T$ and $g_{lm}(x) = T(x/T)^a$ for $x \geq T$. When x increases from T to $T + \epsilon$, $g_{lm}(x)$ decreases at a rate $g'_{lm}(T + \epsilon) \approx a < 0$, where ϵ denotes a small positive number. In this context, the limiter $g_{lm}(x)$ approximates the blanker $g_{bl}(x)$ for $a \rightarrow -\infty$.

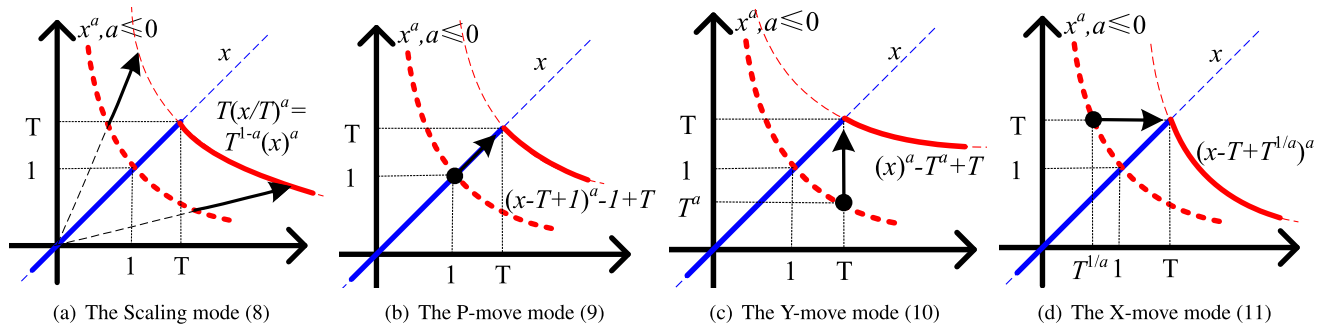


FIGURE 1. Nonlinear functions in four modes, with continuity across linear regions and power-law tails for $T > 1$.

Fig. 1 depicts the four modes, where the blue dotted line denotes the linear function x and the blue solid line denotes the linear region within the threshold $x \leq T$, the red dotted line denotes the power function x^a and the red solid line denotes the power-law tail beyond the threshold $x > T$.

In all four modes, the linear region always meets the power-law tails at the connecting point or breakpoint (T, T) . With this common goal, the power function x^a is transformed in four different ways to pass through (T, T) . The first way is by scaling, and the others are by moving axes.

The four modes are listed in detail as follows.

1) SCALE TRANSFORM (SCALING)

The scaling of x^a makes the dilated or contracted function graph pass through the breakpoint (T, T) . Three kinds of scaling, including along the X-axis, the Y-axis, or both axes, are equivalent, because of the properties of the power function. Fig. 1(a) draws the diagram of the scaling x^a to pass through (T, T) . Thus, the Scaling mode has a ZMNL function formulated as

$$g_{sc}(x, T, a) = \begin{cases} x, & |x| \leq T, \\ T|x/T|^a \cdot \text{sgn}(x), & |x| > T. \end{cases} \quad (8)$$

2) MOVE A FIXED POINT (P-MOVE)

The power function x^a always passes through the fixed point $(1, 1)$. Thus, x^a can be moved along the 45-deg. axis, hence moving the original point $(1, 1)$ to reach the breakpoint (T, T) , as shown in Fig. 1(b) for $T > 1$. Noting that, when $a < 0$ and $T < 1$, we get $T^a > T$, so that the scaling function x^a is translated parallel to the 225-deg. axis to meet the (T, T) point. Fig.1(b) now has to be interpreted with the red solid curve below the red dotted line. Therefore, the scaling function g_{pm} will generate negative values for $x \rightarrow \infty$. By setting such negative values to zero, as shown in Fig. 7(a), the P-move mode formulates the ZMNL function as

$$g_{pm}(x, T, a) = \begin{cases} x, & |x| \leq T, \\ \max [(|x| - T + 1)^a - 1 + T, 0] \cdot \text{sgn}(x), & |x| > T. \end{cases} \quad (9)$$

TABLE 1. Properties of nonlinear functions $g(x, T, a)$.

| $a \leq 1, x > 0$ | continuous | differentiable | breakpoint $X(T, a)$ except T |
|-------------------|----------------|----------------|---------------------------------|
| | $T > 0$ | $T > 0$ | $0 < T < 1, a < 0$ |
| $g_{sc}(x, T, a)$ | Yes | piecewise | none |
| $g_{pm}(x, T, a)$ | Yes | piecewise | $\sqrt[a]{1 - T} - 1 + T$ |
| $g_{ym}(x, T, a)$ | Yes | piecewise | $\sqrt[a]{T^a} - T$ |
| $g_{xm}(x, T, a)$ | Not at $a = 0$ | No | - |

3) MOVE ALONG Y-AXIS (Y-MOVE)

The breakpoint (T, T) can also be reached by moving the graph of x^a parallel to the Y-axis. This is equivalent to moving (T, T^a) to (T, T) , as shown in Fig. 1(c). Like the P-move mode, for $T < 1$ and $a < 0$, x^a after a downward translation parallel to the Y-axis will have negative values for x sufficiently large. Those negative values are set as zero, as shown in Fig. 7(b). Finally, the ZMNL function in the Y-move mode is

$$g_{ym}(x, T, a) = \begin{cases} x, & |x| \leq T, \\ \max (|x|^a - T^a + T, 0) \cdot \text{sgn}(x), & |x| > T. \end{cases} \quad (10)$$

4) MOVE ALONG X-AXIS (X-MOVE)

The breakpoint (T, T) can be reached by moving the graph of x^a parallel to the X-axis. Actually, this is obtained by moving point $(T^{1/a}, T)$ to (T, T) , as shown in Fig. 1(d). Then, the ZMNL function in the X-move mode is formulated as

$$g_{xm}(x, T, a) = \begin{cases} x, & |x| \leq T, \\ (|x| - T + T^{1/a})^a \cdot \text{sgn}(x), & |x| > T, a \neq 0, \\ T \cdot \text{sgn}(x), & |x| > T, a = 0. \end{cases} \quad (11)$$

C. PROPERTIES OF ZMNL FUNCTIONS

The properties of the four ZMNL functions, including continuity, differentiability, and breakpoints along x are listed in TABLE 1, based on the analysis presented in Appendix A. The P-move and the Y-move modes have breakpoints besides the common breakpoint at T , due to the $\max(\cdot)$ functions, as analyzed in Appendix A. Note that $X(T, a)$ is differentiable relative to T and a .

The X-move mode has a discontinuous nonlinearity, whereas the other modes are piecewise continuously differentiable. The piecewise differentiability of $g(x, T, a)$ is important, for it yields the differentiability of related integrals, as will be shown in the next section.

The four functions $g_{sc}(x, T, a)$, $g_{xm}(x, T, a)$, $g_{ym}(x, T, a)$, and $g_{pm}(x, T, a)$ share the same arguments T, a to be optimized, where the differentiability described by Theorem 1 will contribute to the optimization.

As the optimizations of the modes are similar, this paper provides a generic presentation of the optimization and its computational solution. The performances of four modes will be compared in three commonly-used noise models.

IV. OPTIMIZATION OF THE EFFICACY FUNCTION

This section introduces the efficacy function in the nonlinearity design and analyzes its properties.

A. OPTIMIZATION PROBLEM FORMULATION

Basically, nonlinearity design is to improve the detection performance; for example, increasing the detection probability in radar systems or decreasing the bit error ratio (BER) in communication systems. Instead of direct analysis of the detection performances or via the distance to the LOD nonlinearity, the efficacy function has been proposed by researchers to test the optimality of nonlinear processors [7].

For the nonlinearity with power-law tails, the efficacy function can be calculated by

$$\mathcal{E}(g, T, a) = \frac{\mathcal{E}_{gl}^2}{\mathcal{E}_{gg}} = \frac{[\int_{-\infty}^{\infty} g(x, T, a)f'(x)dx]^2}{\int_{-\infty}^{\infty} g^2(x, T, a)f(x)dx}. \quad (12)$$

As can be seen from (3), the efficacy function represents a measurement of the asymptotic output SNR for detecting a deterministic signal in white noise [7], [21]. The signal amplitude ξ_i , or its distribution characterized by the channel gain, does not change the optimization of the output SNR.

The efficacy function is closely related to the fundamental performance for signal detection. As the test statistics are approximately Gaussian distributed, the BER performance can be calculated in a similar way as for the Gaussian noise channels.³

For instance, in the minimum shift keying (MSK) system, assuming the flat-fading channel $\xi_0 = \xi_1 = \xi$, a bit-by-bit detector can achieve the minimum BER as

$$P_e = Q \left[\xi \sqrt{E_s \cdot \mathcal{E}(g, T, a)/2} \right], \quad (13)$$

where E_s denotes the bit energy, $Q(\cdot)$ is the tail distribution function of the standard normal distribution $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp(-t^2/2) dt$. Obviously, the detection performance is improved when we increase $\mathcal{E}(g, T, a)$.

³Actually, for radar systems, the efficacy function is also closely related to the receiver operation characteristics. In [34], the relationship between the false alarm probability and the detection probability is analyzed in Appendix B and developed as formula (25).

Given the concordant relation between the efficacy and detection performance, by taking the efficacy as the objective function, the nonlinearity design is formulated as the optimization problem

$$\max_{T,a} \mathcal{E}(g, T, a) \text{ s.t.}, T > 0, a \leq 1. \quad (14)$$

For each mode, $g(x, T, a)$ is substituted into (14) and the arguments (T, a) are to be optimized for maximum efficacy.

Unfortunately, it is hard to develop a theoretical maximizer or analytical solution to the problem (14), since the efficacy $\mathcal{E}(g, T, a)$ has a complex expression and $f(x)$ does not denote a specific PDF. This also causes great difficulty in the strict mathematical analysis of the objective function $\mathcal{E}(g, T, a)$.

However, the numerical solution is a good choice for solving (14). To seek such a solution, we need to investigate the properties of $\mathcal{E}(g, T, a)$ with respect to (T, a) , e.g., continuity, differentiability, and monotonicity.

B. DIFFERENTIABILITY OF THE EFFICACY

First of all, $\mathcal{E}(g, T, a)$ is *continuous* with respect to T and a , for the nonlinear functions in the Scaling, the Y-move, and the P-move modes. This is obtained directly based on two points: $g(x, T, a)$ is continuous; the calculation in (12) does not change the continuity.

Then, $\mathcal{E}(g, T, a)$ is *differentiable* with respect to T and a . This can be proved based on the continuity and the piecewise differentiability of $g(x, T, a)$.

Given that $g(x, T, a)$ is odd and $f(x)$ is even, we can rewrite the efficacy function in (12) as

$$\mathcal{E}(g, T, a) = \frac{2 [\int_0^{\infty} g(x, T, a)f'(x)dx]^2}{\int_0^{\infty} g^2(x, T, a)f(x)dx}. \quad (15)$$

Given that the denominator never vanishes, as long as the integrals are differentiable with respect to (T, a) , $\mathcal{E}(g, T, a)$ is also differentiable. Moreover, the integral can be simplified by substituting the sub-functions of $g(x, T, a)$. Then, all the related integrals can be proved to be differentiable, based on applications of Theorem 1.

For instance, the Scaling mode $g_{sc}(x, T, a)$ has two subdomains. In the numerator of (15), we have

$$\int_0^{\infty} g(x, T, a)f'(x)dx = \int_0^T xf'(x)dx + \int_T^{\infty} g(x, T, a)f'(x)dx.$$

Obviously, the first term in the right-hand side is differentiable relative to T and trivially relative to a . The culprit is hence the second term of the right-hand side, for which the following holds:

Theorem 1: Given a differentiable function $h(x) > 0$, for $g(x, T, a)$ as in (8), (9), or (10), the integral

$$E(T, a) = \int_T^{\infty} g(x, T, a)h(x)dx \quad (16)$$

is differentiable with respect to T and a .

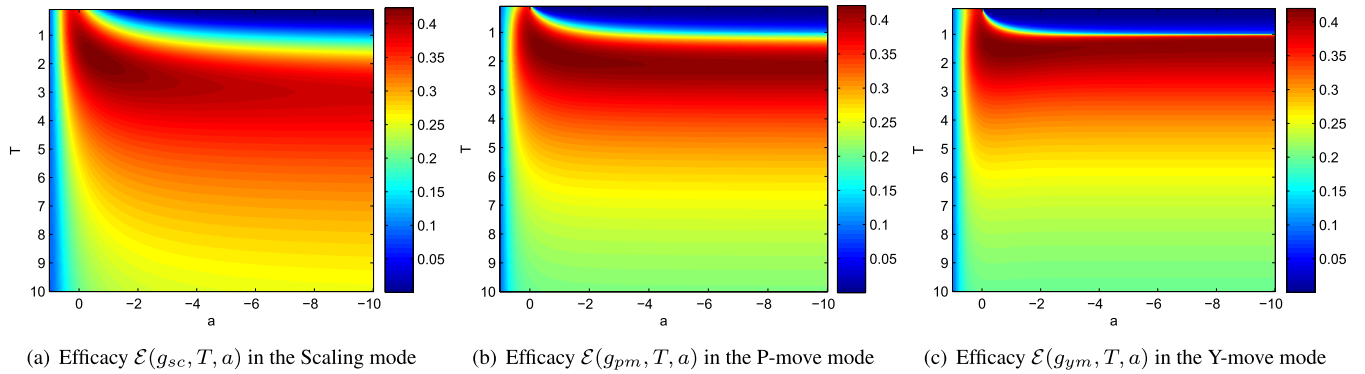


FIGURE 2. The efficacy $\mathcal{E}(g, T, a)$ is unimodal in both T and a , in the Scaling, the P-move, and the Y-move mode for the $S\alpha S$ noise, $\alpha = 1.5$ and $\gamma = 1$.

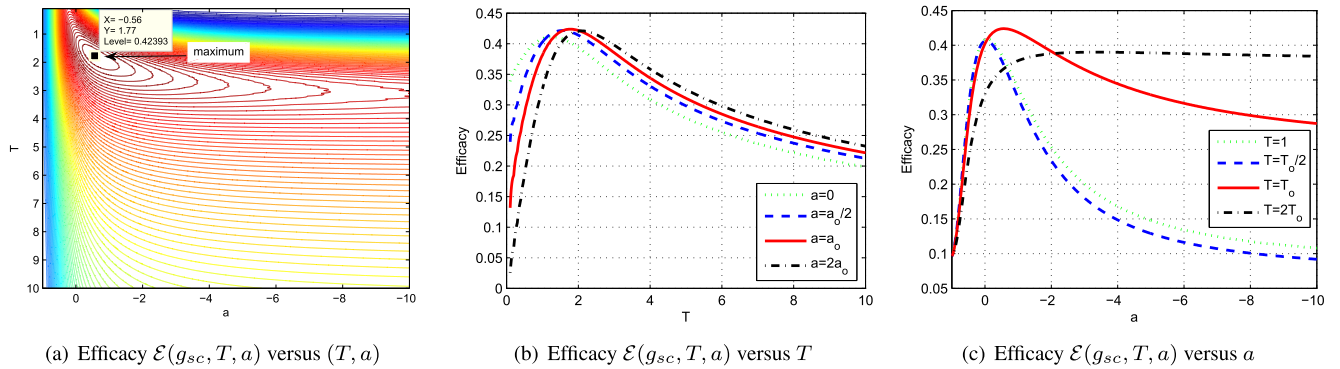


FIGURE 3. The efficacy $\mathcal{E}(g, T, a)$ is unimodal in both T and a , in the Scaling mode for the $S\alpha S$ noise $\alpha = 1.5$ and $\gamma = 1$. The maximum efficacy is achieved by the optimal threshold and exponent $(T_o, a_o) = (1.77, -0.56)$.

Proof: See Appendix B.

For the P-move or Y-move mode, there may be two or three subdomains, according to TABLE 1. The integrals involved in efficacy calculation are differentiable, invoking Theorem 1.

In all, the efficacy $\mathcal{E}(g, T, a)$ in (12) is differentiable with respect to (T, a) , for $T > 0$ and $a \leq 1$, for the nonlinearity in the Scaling, the P-move and the Y-move modes.

C. UNIMODALITY OF SIMULATED EFFICACY

However, as for monotonicity, it is hard to analyze what interval of T or a can support an efficacy $\mathcal{E}(g, T, a)$ that is monotonically decreasing or increasing. This difficulty is partly due to our analysis focusing on a generic case of “impulsive noise” instead of a specified model or PDF $f(x)$. A “heavy-tailed” PDF is not enough to provide a theoretical proof of the monotonicity of $\mathcal{E}(g, T, a)$.

Alternatively, we employ numerical simulations as an effective way for shedding light on the monotonicity of the objective function. As shown by massive simulations, the various models of impulsive noise share the common property of unimodality of the efficacy.

A typical simulation is illustrated in Fig. 2, which shows the $\mathcal{E}(g, T, a)$ of the Scaling, the P-move, and the Y-move modes for the $S\alpha S$ noise, $\alpha = 1.5$, $\gamma = 1$. As can be seen, the surface plotting $\mathcal{E}(g, T, a)$ has only one local maximum,

which is also the global maximum. The efficacy surfaces are smooth, which demonstrates that $\mathcal{E}(g, T, a)$ is differentiable.

Fig. 3(a) depicts the the level sets of the efficacy $\mathcal{E}(g_{sc}, T, a)$ of the Scaling mode. We can see that the level sets are not convex except in the neighborhood of maximum point. Figs. 3(b) and 3(c) plot the curves of $\mathcal{E}(g_{sc}, T, a)$ versus T or a when the other one is fixed. Obviously, $\mathcal{E}(g_{sc}, T, a)$ is unimodal for either argument.⁴

Based on the above analysis, we reach the conclusion that the objective function $\mathcal{E}(g, T, a)$ is *unimodal* for either argument T or a , for $T > 0$ and $a \leq 1$, in the Scaling, the P-move and the Y-move modes, though this is hard to prove since $f(x)$ is not specified.

As for the X-move mode, since $g_{xm}(x, T, a)$ is discontinuous at $a = 0$, its efficacy $\mathcal{E}(g_{xm}, T, a)$ behaves differently. Discussion of the X-move mode is deferred to Appendix C.

V. SOLUTION TO THE OPTIMIZATION

Based on relevant properties of the efficacy function, we provide solution guidelines as well as a practical algorithm for the optimization problem (14).

⁴The definition of a “unimodal” real function is as follows: $y(x)$ is a unimodal function in an interval if for some value x_0 , it is monotonically increasing for $x \leq x_0$ and monotonically decreasing for $x \geq x_0$. In that case, the maximum value of $y(x)$ is $y(x_0)$ and there is not another local maximum.

A. SOLVABLE BY DIRECT SEARCH METHODS

To summarize the property analysis, $\mathcal{E}(g, T, a)$ has two important properties for $T > 0$ and $a \leq 1$.

- $\mathcal{E}(g, T, a)$ is continuous and differentiable for (T, a) ;
- $\mathcal{E}(g, T, a)$ is unimodal relative to both T and a .

This reveals that it is promising to develop a numerical method to maximize $\mathcal{E}(g, T, a)$.

Herein, due to the fact that “*partial derivatives are unavailable in closed-form*”, derivative-free methods are of interest for their advantage of no requirements on partial derivatives. Especially, direct search methods are reliable and applicable for using function values instead of calculating or approximating any gradients [36].

To solve problem (14) as a two-dimensional optimization, two notable direct search methods can be considered, i.e. Powell’s method and Nelder-Mead simplex (NMS) method. Though both methods are proposed for minimization, they are equivalent to maximizing the negative of the objective functions. Brief introductions to the methods are given as follows.

Powell’s method adjusts one argument at a time and finds the minimum in a finite number of steps [37]. The minimization might be accomplished by the application of the golden section or Fibonacci search as long as the function is unimodal in this search direction [40]. Powell’s method is workable, since $\mathcal{E}(g, T, a)$ is unimodal for both T and a .

The Nelder-Mead method is inspired by simplex-based direct search methods in [41] and deliberately modified to avoid assuming knowledge of relative steps [38]. The NMS method has four coefficients, including α_{NM} for reflection, γ_{NM} for expansion, β_{NM} for contraction, and σ_{NM} for shrinkage. By constant coefficients, the NMS method “adapts itself to the local landscape and contracts on to the final minimum” [38].

The optimization problem (14) is solvable by the Nelder-Mead method or Powell’s method. To use the methods, some details need to be addressed accordingly.

B. SOLUTION BY THE NMS METHOD

This section illustrates the Nelder-Mead method for designing the algorithmic solution to the optimization of efficacy $\mathcal{E}(g, T, a)$. For the NMS procedure, some algorithm details are summarized as follows.

1) OBJECTIVE FUNCTION

Considering the constraints on the domain of arguments to be searched, the efficacy is set as zero for $T < 0$ or $a > 0$. The objective function is defined as

$$\mathcal{E}_{\text{NM}}(g, T, a) = \begin{cases} -\mathcal{E}(g, T, a), & T > 0 \text{ and } a \leq 1 \\ 0, & T \leq 0 \text{ or } a > 1 \end{cases} \quad (17)$$

Then, the optimization of the efficacy function becomes the problem of the unconstrained minimization of the objective function $\mathcal{E}_{\text{NM}}(g, T, a)$ with respect to (T, a) .

2) COEFFICIENTS

Four coefficients are set as standard values, e.g.

$$\alpha_{\text{NM}} = 1, \quad \beta_{\text{NM}} = 0.5, \quad \gamma_{\text{NM}} = 2, \quad \text{and } \sigma_{\text{NM}} = 0.5. \quad (18)$$

This is a nearly universal choice in the NMS method. It also reflects the fact that the efficacy optimization is a regular problem for the NMS method.

3) STARTING POINT

Without prior knowledge about the noise distribution, it is recommended to set the starting point at $(T_0, a_0) = (1, 0)$. The NMS method is also convergent even if the starting point is not within the neighborhood of the maximum point. Besides, any information about noise covariance can be used to define the starting point and help improve the convergence speed.

4) STOPPING CRITERIA

A final point concerns the criterion used for halting the procedure. The criterion adopted in this paper is to take the preset value 10^{-4} as final increment on either the objective function or its arguments.

Based on the above issues, the optimal values of T and a in each mode can be achieved by substituting the corresponding nonlinear function into the objective function $\mathcal{E}_{\text{NM}}(g, T, a)$ and then minimizing it via the NMS iteration procedure [39].

VI. PERFORMANCE ANALYSIS IN DISTRIBUTIONS

This section deals with simulations of the proposed nonlinearity design in three commonly-used models of impulsive noise, including the S α S noise, the Class A noise, and the Gaussian mixture noise. Note that these noise models include the background Gaussian noise.

In each noise, the four modes (8), (9), (10), and (11) are considered, and the NMS method is implemented to find the optimal thresholds and exponents. For comparison, traditional Blanking and Clipper are optimized by substituting (6) and (7) for $g(\cdot, \cdot, \cdot)$ into the efficacy optimization problem (14) which is also solved by the NMS method.

A. DESIGN IN THE S α S NOISE

The α -stable distribution is considered to be symmetric relative to zero, the so-called S α S distribution. The S α S probability density does not exist in closed-form, except for the Gaussian and the Cauchy distributions. As a result, the S α S PDF is calculated, numerically, by

$$f_{\alpha, \gamma}(x) = \mathcal{LFT} \left[\exp(-\gamma |w|^\alpha) \right] \quad (19)$$

where $0 < \alpha \leq 2$ is the characteristic exponent, γ is the dispersion, and $\mathcal{LFT}(\cdot)$ denotes the inverse Fourier transformation. A smaller α means a heavier tail.

The nonlinearity functions of the four modes are designed in the S α S noise for α running from 1.0 to 1.9, and for $\gamma = 1$. The optimal efficacy of the four modes are shown in Fig. 4(a). Clearly, the optimal efficacy is achieved by the LOD.

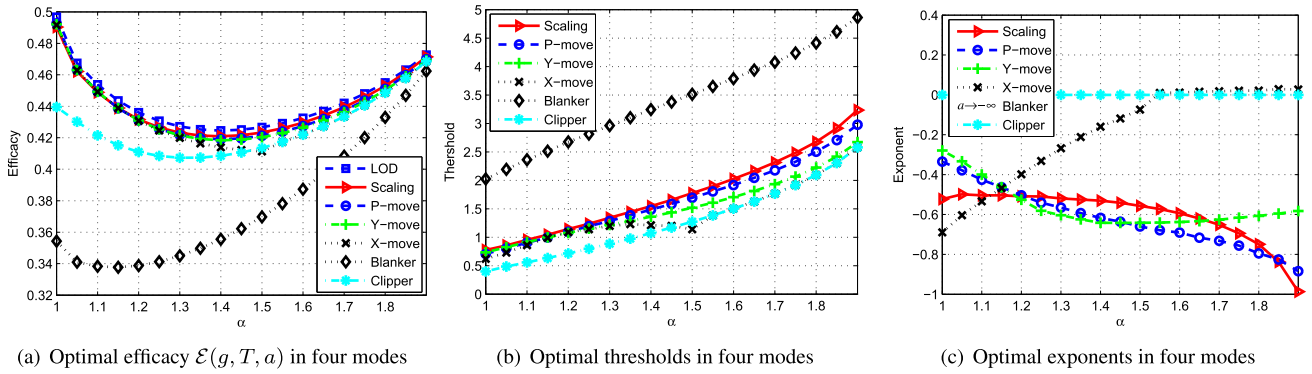


FIGURE 4. The efficacy, thresholds, and exponents of optimal design of four modes in the $S\alpha S$ noise, $\gamma = 1$.

TABLE 2. Optimal efficacy of nonlinearity designs, in the Class A model with parameters (A, Γ) for $\sigma = 1$.

| (A, Γ) | (0.1, 0.001) | (0.3, 0.001) | (0.5, 0.001) | (0.1, 0.01) | (0.3, 0.01) | (0.5, 0.01) | (0.1, 0.1) | (0.3, 0.1) | (0.5, 0.1) |
|---------------|--------------|--------------|--------------|-------------|-------------|-------------|------------|------------|------------|
| LOD | 891.3 | 697.2 | 542.5 | 88.16 | 65.97 | 49.05 | 9.189 | 6.409 | 4.608 |
| Scaling | 890.7 | 696.2 | 540.5 | 88.03 | 65.58 | 48.50 | 9.151 | 6.318 | 4.504 |
| P-move | 890.8 | 697.2 | 541.9 | 88.09 | 65.76 | 48.72 | 9.159 | 6.312 | 4.443 |
| Y-move | 890.2 | 693.9 | 532.3 | 87.64 | 64.40 | 46.87 | 9.161 | 6.311 | 4.443 |
| X-move | 890.7 | 696.5 | 541.0 | 88.04 | 65.63 | 48.56 | 9.152 | 6.317 | 4.491 |
| Blanker | 888.8 | 691.0 | 532.3 | 87.52 | 64.07 | 46.52 | 9.015 | 5.972 | 4.039 |
| Clipper | 671.6 | 600.7 | 257.3 | 69.54 | 42.79 | 28.47 | 7.786 | 5.176 | 3.805 |

The designs in all four modes achieve nearly optimal efficacy, above 99% of the LOD efficacy. More precisely, the Scaling mode achieves more than 99.5% of the LOD efficacy, while the Y-move mode and the P-move mode achieve slightly less. The Clipper is sub-optimal, while the Blanker is significantly worse than the others. This reveals that the Blanker method is not suitable for nonlinear processing in the $S\alpha S$ noise. Analysis on the X-move mode is relegated to Appendix C.

Optimal thresholds and exponents of the four modes are shown in Fig. 4(b) and Fig. 4(c) respectively. The Blanker has the maximum threshold and the most rapid decay $g(x) = 0$ for $|x| > T$, while the Clipper has the minimum threshold and the slowest decay $g(x) = \text{sgn}(x)T$ for $|x| > T$. Compared with the Blanker and the Clipper, the proposed designs in four modes have moderate thresholds and decay exponents. They keep changing for varying α . It demonstrates that different distributions require various decay factors in the tails.

B. DESIGN IN THE CLASS A MODEL

The Middleton Class A distribution has PDF given by

$$f_{A,\Gamma}(x) = \sum_{k=0}^{\infty} \frac{\exp(-A)A^k}{k!} \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{x^2}{2\sigma_k^2}\right) \quad (20)$$

where $\sigma_k^2 = \sigma^2(k/A + \Gamma)/(1 + \Gamma)$, σ^2 is the average power, A is the impulsiveness index, and Γ is the power ratio of the Gaussian component to the non-Gaussian component.

The nonlinearity of the four modes, as well as the Blanker and the Clipper, is designed for various cases of the Class A model. The optimization results are listed in TABLE 2. Compared with the optimal efficacy resulting from the LOD, three modes including the Scaling, the X-move, and the P-move obtain 99.3% average for all the cases. The Y-move mode is slightly worse, at 98.3%. The Blanker is sub-optimal, at 96.3%. The Clipper is the worst, at about 70%. It indicates that the Clipper is unsuitable for suppressing the Class A noise.

C. DESIGN IN THE GAUSSIAN MIXTURE MODEL

Then, we consider a two-component Gaussian mixture distribution. Its PDF is given by

$$f_{\varepsilon,\sigma_1,\sigma_2}(x) = \frac{1-\varepsilon}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{x^2}{2\sigma_1^2}\right) + \frac{\varepsilon}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{x^2}{2\sigma_2^2}\right) \quad (21)$$

where ε denotes the probability of occurrence of the impulsive noise, $\sigma_2^2 \gg \sigma_1^2$ are the variances of the impulsive and the Gaussian components, respectively.

In simulations, the nonlinear functions of the four modes are designed by the NMS method, for various cases of Gaussian mixture. The efficacy of the designed nonlinearities are listed in TABLE 3. Among the nonlinearity designs, the Scaling mode is the best, with 99.5% of the LOD efficacy

TABLE 3. Optimal efficacy of nonlinearity designs, in the Gaussian mixture model with parameters (ϵ, σ_2^2) for $\sigma_1 = 1$.

| (ϵ, σ_2^2) | (0.3, 10) | (0.3, 100) | (0.3, 1000) | (0.1, 10) | (0.1, 100) | (0.1, 1000) | (0.01, 10) | (0.01, 100) | (0.01, 1000) |
|--------------------------|-----------|------------|-------------|-----------|------------|-------------|------------|-------------|--------------|
| LOD | 0.5198 | 0.5709 | 0.6338 | 0.7935 | 0.8316 | 0.8678 | 0.9695 | 0.9776 | 0.9846 |
| Scaling | 0.5106 | 0.5644 | 0.6308 | 0.7884 | 0.8281 | 0.8663 | 0.9686 | 0.9771 | 0.9844 |
| P-move | 0.5043 | 0.5081 | 0.5358 | 0.7734 | 0.7705 | 0.7818 | 0.9628 | 0.9595 | 0.9604 |
| Y-move | 0.4946 | 0.4962 | 0.5278 | 0.7609 | 0.7513 | 0.7639 | 0.9592 | 0.9528 | 0.9533 |
| X-move | 0.5110 | 0.5633 | 0.6298 | 0.7877 | 0.8273 | 0.8658 | 0.9568 | 0.9453 | 0.9409 |
| Blanker | 0.4637 | 0.5421 | 0.6196 | 0.7624 | 0.8160 | 0.8611 | 0.9647 | 0.9752 | 0.9837 |
| Clipper | 0.4592 | 0.3906 | 0.3662 | 0.7407 | 0.6958 | 0.6793 | 0.9568 | 0.9453 | 0.9409 |

in average. The other three modes obtain over 92.5%, the Blanker obtains 97.2%, and the Clipper obtains 84.5%.

Next we investigate the robustness of each design. Considering the worst cases of all designs, the smallest efficacy percentages of the Scaling, the X/Y/P-move, the Blanker, and the Clipper are 98.2%, 95.6%, 83.3%, 84.6%, 89.2%, 57.8% respectively. Thus, we can see that the Scaling mode is very robust, while the Clipper is not robust.

D. SUMMARY

From the above analysis of the efficacy performance in the $S\alpha S$ noise, the Class A noise, and the Gaussian mixture noise, we can conclude that the proposed nonlinearity design is more robust and effective than traditional limiters. Since the exponent in power-law tails is optimized to provide an adjustable decay factor, the proposed design almost achieves the optimal efficacy in various noise models. Among the four modes of power-law tails, the Scaling mode in (8) works best and so is recommended above all other methods.

VII. DISCUSSION AND SIMULATION OF APPLICATIONS

This section discusses the applications of the proposed method and compares it with other detectors, such as the LOD. Detection performance for various nonlinearity is simulated.

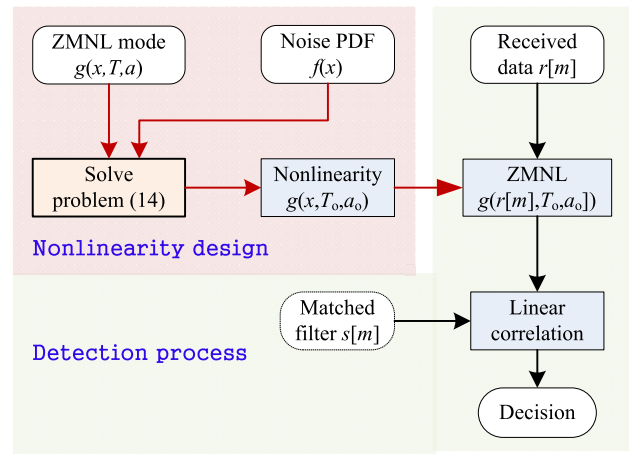
A. TWO STAGES IN PRACTICAL APPLICATION

When the proposed design is used for signal detection in impulsive noise, any practical application would contain two stages.

Stage 1: Design — Obtain the optimal parameters (T_o, a_o) by solving the optimization problem (14). For example, solve (17) by the NMS method in Section V-B.

Stage 2: Process — Use the nonlinearity $g(x, T_o, a_o)$ to transform the received data $r(m)$. Then the output $g(r[m], T_o, a_o)$ will be used for the signal detection.

Fig. 5 depicts the block diagram of designing the nonlinearity based on the noise PDF and applying it for the signal detection from the received data. In practical applications, the optimal threshold and exponent can be calculated once and off-line. The noise parameters that are unknown can be estimated based on noise samples before the design stage.

**FIGURE 5. Block diagram of nonlinearity design and application.**

The merits of this method can be mainly summarized as three points:

- The nonlinearity design with power-law tails is almost as optimal as the LOD, as demonstrated in Section VI.
- The nonlinearity design with power-law tails is effective for various distributions of impulsive noise.
- The nonlinearity $g(x, T_o, a_o)$ has closed-form expression, so that its calculation is accurate and efficient.

The first and second points show that our method does not bear any degradation when it replaces the LOD. The last point reveals that our method can outperform the LOD on computational efficiency in the Process Stage, when the PDF is unavailable in closed-form.

B. ADVANTAGES OVER THE LOD IN PROCESSING

As discussed before, the distribution models of impulsive noises may not provide closed-form PDFs, e.g. the $S\alpha S$ and the Class A models. Thus, we cannot process the received data $r[m]$ easily or directly by the LOD.

To use the LOD, a workable approach is by interpolation. It may consist of two steps. The first step is to generate discrete samples $g_{lo}(kx_\Delta)$, $k = -K, -K + 1, \dots, -1, 0, +1, \dots, K$, where x_Δ is a uniform interval and K controls the range. The second step is the nonlinear transformation of $r[m]$

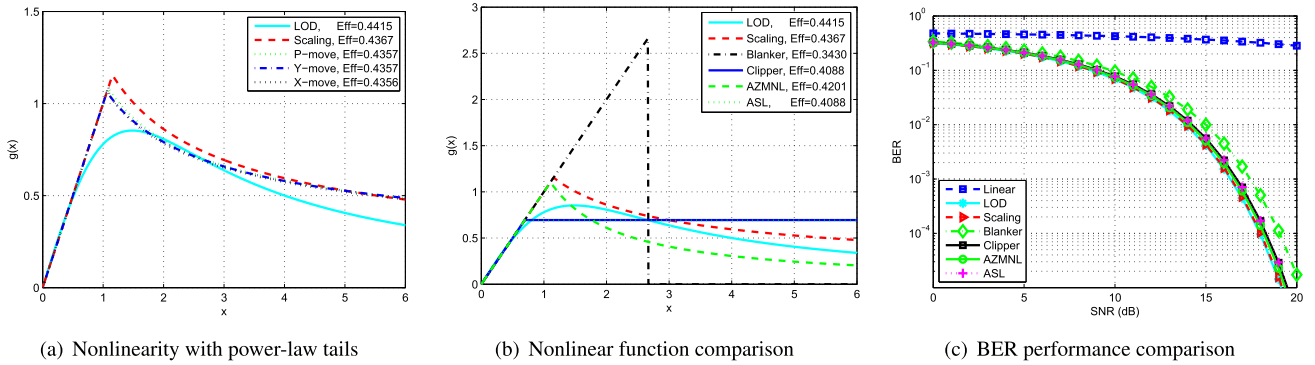


FIGURE 6. Nonlinear functions and detection performances in the $S\alpha S$ noise, for $\alpha = 1.2$ and $\gamma = 1$.

by interpolation. For each output $g_{lo}(r[m])$, the LOD must firstly find the location of $r[m]$ and then use linear or nonlinear interpolation.

Contrarily, our method applies more easily to signal processing. The optimal parameters (T_o, a_o) are designed only once in the Design Stage. Then, in the Process Stage, $g(x, T_o, a_o)$ is directly used for calculating $g(r[m], T_o, a_o)$ for each received data $r[m]$. It has less computational complexity and better accuracy than the LOD. It is also more suitable for practical applications.

C. PERFORMANCE SIMULATION AND COMPARISON

The nonlinearity design and application is simulated for communication in the $S\alpha S$ noise for $\alpha = 1.2, \gamma = 1$. Besides the proposed nonlinearity in four modes, the Blanker, the Clipper and other nonlinearity preprocessors are also presented for comparison. The AZMNL sets the tail as $1/x$ and the threshold as $T_{\alpha z} = \alpha^2 \Gamma(1/\alpha) / \Gamma(3/\alpha)$ [31]. The adaptive soft limiter (ASL) method [21] uses a clipper threshold which is obtained by solving the probability equation $Pr(|x| \leq T_{asl}) = 1 + 0.7756(\alpha - 2)$.

Firstly, the nonlinearity in four modes are designed and depicted in Fig. 6. The LOD function, which corresponds to the optimal nonlinearity, is continuous and differentiable everywhere. The four modes have less decay rates than the LOD, with similar thresholds for the linear regions. From the comparison of these functions, we can see that the four modes do not strictly follow the LOD’s shape, but maximize the efficacy in their own function sets.

Secondly, other nonlinear functions are compared in Fig. 6(b). The ASL method performs the same way as the Clipper, which demonstrates that [21] achieves optimal design of the clipper. From the efficacy values denoted as “Eff” in the legends, we can see that the AZMNL and the ASL achieve less efficacy than our design.

Finally, the BER is simulated for the MSK modulation in the flat-fading channel. The SNR is defined as $\xi^2 E_s / \gamma$, for $M = 1024$ samples in one bit. The output of LOD is calculated by linear interpolation, as discussed in Section VII-B. The BER results of 10^7 Monte Carlo simulations are plotted in Fig. 6(c).

As can be seen, the Scaling mode obtains the same BER as the LOD. The Clipper is sub-optimal, while the Blanker is worse. The AZMNL is near-optimal, better than the Clipper but clearly worse than the Scaling mode. Note that the simulated BERs are consistent with the theoretical BERs (13) which are not drawn in the figure to avoid the clutter.

To quantify the performance gain of the proposed scheme and other nonlinearities, we can use the efficacy of the nonlinearity to measure its SNR loss with relative to the LOD

$$SNR_{loss}(g) = -10 \times \log_{10} \left(\frac{\text{Efficacy of } g(x)}{\text{Efficacy of } g_{lo}(x)} \right). \quad (22)$$

Based on the efficacy values denoted as “Eff” in the legends of Fig. 5(b), we can see that the Scaling mode bears a loss of 0.05dB, the Blanking scheme a loss of 1.06dB, the Clipping method a loss of 0.33dB, and the AZMNL method a loss of 0.22dB. Hence, the nonlinearity with power-law tails bear little loss compared with the LOD and outperforms all the optimal designs of traditional tails.

VIII. CONCLUSION

This paper has introduced power-law tails in nonlinearity design and provided a nearly optimal solution for correlation detection in various models of impulsive noise. The nonlinear function is defined as a piecewise-constructed function consisting of a linear function and a power-law function where the threshold T and the exponent a are two design arguments. Then, nonlinearity design is formulated as the problem of the efficacy optimization with respect to T and a . As the efficacy is differentiable and unimodal, the optimization problem can be efficiently solved by derivative-free methods, such as the Nelder-Mead simplex method.

The proposed design has been tested in the $S\alpha S$ noise, the Class A noise, and the Gaussian mixture noise. Simulation results have shown that our proposed design is significantly more efficient and robust than the traditional blanker and clipper, since the power-law tail provides a suitable decay factor to match the tail of the noise model. Compared with the locally optimal detector, the designed power-law tail achieves almost the same optimality and has the advantage of a simple closed-form formula for nonlinear computation.

This paper has analyzed the fundamental performance of signal detection by the power-law tail design. Future work will discuss the applications in wireless communication systems under specific modulation and unknown noise distribution. For instance, when the PLC system employs the OFDM, the OFDM waveform envelope and the average peak-to-average ratio have impacts on the nonlinearity and the BER [5]. Besides, for real-time processing when the noise model is unknown or time-varying, it is meaningful to develop efficient approaches for designing the power-law tails based on the received data instead of prior known noise distributions.

APPENDIXES
APPENDIX A
PROPERTIES OF NONLINEARITY
IN FOUR MODES

The nonlinearity $g(x, T, a)$ in the four modes, i.e. $g_{sc}(x, T, a)$, $g_{xm}(x, T, a)$, $g_{ym}(x, T, a)$, and $g_{pm}(x, T, a)$, possesses similar but different properties, which can affect the optimization. The following analyzes the properties of $g(x, T, a)$ for $x > 0$, $T > 0$, and $a \leq 1$.

(1). *Continuity.* The Scaling mode, the P-move mode and the Y-move mode have continuous nonlinearity functions. However, in the X-move mode, $g_{xm}(x, T, a)$ is discontinuous at $a = 0$.

The proof of discontinuity is simple. Let $a \rightarrow 0^+$ or $a \rightarrow 0^-$ denote that a may approach 0 from above (right) or below (left). For $x > T$, we can calculate the limits:

$$\lim_{a \rightarrow 0^+} g_{xm}(x, T, a) = \begin{cases} T, & 1 \leq T, \\ 1, & 0 < T < 1, \end{cases} \quad (23)$$

$$\lim_{a \rightarrow 0^-} g_{xm}(x, T, a) = \begin{cases} 1, & 1 \leq T, \\ T, & 0 < T < 1, \end{cases} \quad (24)$$

which yields

$$\lim_{a \rightarrow 0^+} g_{xm}(x, T, a) \neq \lim_{a \rightarrow 0^-} g_{sc}(x, T, a), \text{ for } T \neq 1. \quad (25)$$

Therefore, $g_{xm}(x, T, a)$ is not continuous at $a = 0$.

(2). *Differentiability.* All the nonlinearity functions $g_{sc}(x, T, a)$, $g_{pm}(x, T, a)$ and $g_{ym}(x, T, a)$ are made up with three types of functions, i.e. $y = x$, x^a , and 0, all of which are differentiable. Furthermore, all such pieces are connected in a continuous fashion. Therefore, the nonlinearity functions are piecewise continuous and continuously differentiable except at the breakpoints.

However, for the X-move mode, $g_{xm}(x, T, a)$ is not continuous at $a = 0$ and thus not piecewise differentiable.

(3). *Breakpoints.* Breakpoints along the x -axis affect the efficacy, which involves an integral along the x axis. All the modes share the same breakpoint at $x = T$. However, the P-move and the Y-move modes may have additional breakpoints.

In the Scaling mode, $g_{sc}(x, T, a)$ always has only one breakpoint at $x = T$.

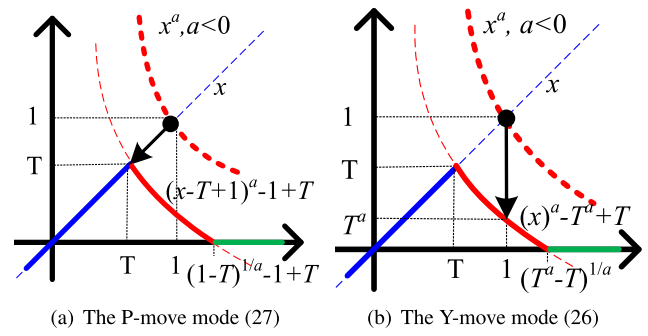


FIGURE 7. In the P-move and the Y-move modes, breakpoints $X_p(T, a) = \sqrt[a]{1-T} - 1 + T$ and $X_y(T, a) = \sqrt[a]{T^a - T}$, resp., occur when x^a is moved down for $0 < T < 1$ and $a < 0$.

In the Y-move mode, $g_{ym}(x, T, a)$ can be rewritten as

$$g_{ym}(x, T, a) = \begin{cases} x, & 0 < x \leq T, \\ 0, & x > X_y(T, a), \\ x^a - T^a + T, & \text{otherwise,} \end{cases} \quad 0 < T < 1, \quad a < 0 \quad (26)$$

where $X_y(T, a) = \sqrt[a]{T^a - T}$ is the breakpoint in addition to T .

Similarly, in the P-move mode, $g_{pm}(x, T, a)$ is rewritten as

$$g_{pm}(x, T, a) = \begin{cases} x, & 0 < x \leq T, \\ 0, & x > X_p(T, a), \\ (x - T + 1)^a - 1 + T, & \text{otherwise,} \end{cases} \quad 0 < T < 1, \quad a < 0, \quad (27)$$

where $X_p(T, a) = \sqrt[a]{1-T} - 1 + T$ is the breakpoint in addition to T .

Both $X_y(T, a)$ and $X_p(T, a)$ occur only for $0 < T < 1$ and $a < 0$ when we move x^a down for the continuity of $g(x, T, a)$, as depicted in Fig. 7.

Finally, notice that all the breakpoints, i.e. T , $X_y(T, a)$ and $X_p(T, a)$, are differentiable with respect to T and a .

APPENDIX B
PROOF OF THEOREM "DIFFERENTIABILITY OF INTEGRAL"

To prove differentiability, we use the related condition that the partial derivatives $\partial E(T, a)/\partial T$ and $\partial E(T, a)/\partial a$ exist and are continuous. As the proofs of the two partial derivatives are similar, in the following we prove continuity of $\partial E(T, a)/\partial T$.

First of all, consider the P/Y-move mode for $T \geq 1$ or $a \geq 0$, as well as the Scaling mode. In such a case, as analyzed in Appendix A, $g(x, T, a)$ is the power-law function for $x \in [T, \infty)$. Since $g(x, T, a)$ is differentiable in $[T, \infty)$, by the Leibniz integral rule on (16), we have

$$\frac{\partial E(T, a)}{\partial T} = \int_T^\infty \frac{\partial g(x, T, a)}{\partial T} h(x) dx - g(T, T, a)h(T). \quad (28)$$

Obviously, (28) is continuous.

Then, in the P/Y-move mode for $0 < T < 1$ and $a < 0$, $g(x, T, a)$ equals zero for $x \in [X(T, a), \infty)$. So we can rewrite $E(T, a)$ as

$$E(T, a) = \int_T^{X(T,a)} g(x, T, a)h(x)dx, \quad (29)$$

where $g(x, T, a)$ is differentiable for $x \in [T, X(T, a)]$. Thus, by the Leibniz integral rule, we have

$$\begin{aligned} \frac{\partial E(T, a)}{\partial T} &= \int_T^{X(T,a)} \frac{\partial g(x, T, a)}{\partial T} h(x)dx - g(T, T, a)h(T) \\ &\quad + g[X(T, a), T, a]h[X(T, a)]\frac{\partial X(T, a)}{\partial T} \\ &= \int_T^{X(T,a)} \frac{\partial g(x, T, a)}{\partial T} h(x)dx - g(T, T, a)h(T), \quad (30) \end{aligned}$$

where $g[X(T, a), T, a] = 0$ is used. Obviously, (30) is also continuous.

At last, we need to prove that the derivatives in the P/Y-move modes are continuous at $T = 1$ and $a = 0$, i.e.

$$\lim_{T \rightarrow 1^-} \frac{\partial E(T, a)}{\partial T} = \lim_{T \rightarrow 1^+} \frac{\partial E(T, a)}{\partial T}, \quad (31)$$

$$\lim_{a \rightarrow 0^-} \frac{\partial E(T, a)}{\partial T} = \lim_{a \rightarrow 0^+} \frac{\partial E(T, a)}{\partial T}. \quad (32)$$

Actually, this is easy to obtain, since both $\lim_{T \rightarrow 1^-} X(T, a) = \infty$ and $\lim_{a \rightarrow 0^-} X(T, a) = \infty$ make the integrations limits in (30) and (28) the same.

Therefore, $\partial E(T, a)/\partial T$ is continuous. Similarly, $\partial E(T, a)/\partial a$ is also continuous. Hence, $E(T, a)$ is differentiable.

APPENDIX C PROPERTIES AND PERFORMANCE OF THE X-MOVE MODE

Unlike the other three modes, the X-move mode has discontinuity with respect to a , not only for the nonlinear function $g_{xm}(x, T, a)$ in (11), but also for its efficacy function $\mathcal{E}(g_{xm}, T, a)$ in (12). This makes the X-move mode behave differently from the other three modes.

(1) *Discontinuity.* The efficacy $\mathcal{E}(g_{xm}, T, a)$ is not continuous at $a = 0$. This is due to the discontinuity of $g_{xm}(x, T, a)$ at $a = 0$.

Fig. 8 shows the efficacy $\mathcal{E}(g_{xm}, T, a)$ for the S α S noise $\alpha = 1.5$ and $\gamma = 1$. As can be seen, the efficacy surface is divided into two regions by the vertical break line at $a = 0$. Contrarily, the break line does not exist in any efficacy surface of the Scaling, P-move or Y-move mode.

(2) *Differentiability.* The efficacy $\mathcal{E}(g_{xm}, T, a)$ is differentiable in two subdomains, i.e. $a < 0$ and $a > 0$, respectively. In every subdomain, $g_{xm}(x, T, a)$ is differentiable and thus $\mathcal{E}(g_{xm}, T, a)$ is differentiable.

As can be seen in Fig. 8, $\mathcal{E}(g_{xm}, T, a)$ is smooth in each region of the surface, though it is discontinuous at $a = 0$.

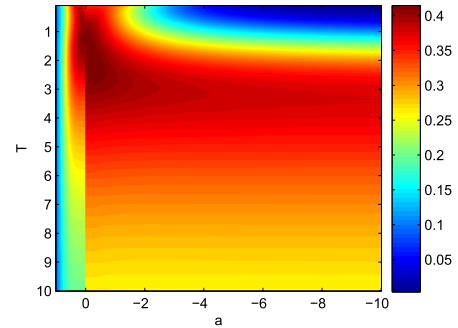


FIGURE 8. The efficacy $\mathcal{E}(g_{xm}, T, a)$ in the S α S noise, for $\alpha = 1.5$ and $\gamma = 1$. The surface is divided into two regions by the line $a = 0$.

(3) *Unimodality.* Simulations show that the efficacy function $\mathcal{E}(g_{xm}, T, a)$ is unimodal with respect to a in each subdomain for $a > 0$ or $a < 0$, respectively.

(4) *Optimization results.* The NMS method still works for the optimization of $\mathcal{E}(g_{xm}, T, a)$. As shown in Fig. 4(a), TABLE 2, and TABLE 3, the optimal efficacy in the X-move mode is close to the efficacy of the other modes. Actually, the NMS method succeeds in finding the bigger one of two local maximums for $a > 0$ and $a < 0$.

Figs. 4(b) and 4(c) show that the threshold T and the exponent a for the X-move mode change dramatically around $\alpha = 1.55$. In Fig. 4(c) for the X-move mode in $\alpha > 1.55$, the optimal exponent a_o is a small positive number. The corresponding nonlinear function and efficacy are close to those of the optimal clipper.

Finally, due to the discontinuity, the X-move mode is less recommended than the other modes.

REFERENCES

- [1] Y. Yabuuchi, D. Umehara, M. Morikura, T. Morita, S. Ishiko, and S. Horihata, "Low rate and high reliable modulation schemes for in-vehicle power line communications," in *Proc. IEEE ISPLC*, Udine, Italy, Apr. 2011, pp. 249–254.
- [2] N. C. Beaulieu, H. Shao, and J. Fiorina, "P-order metric UWB receiver structures with superior performance," *IEEE Trans. Commun.*, vol. 56, no. 10, pp. 1666–1676, Oct. 2008.
- [3] S. Galli, A. Scaglione, and Z. Wang, "For the grid and through the grid: The role of power line communications in the smart grid," *Proc. IEEE*, vol. 99, no. 6, pp. 998–1027, Jun. 2011.
- [4] W. Liu, M. Sigle, and K. Dostert, "Channel characterization and system verification for narrowband power line communication in smart grid applications," *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 28–35, Dec. 2011.
- [5] K. M. Rabie and E. Alsusae, "On improving communication robustness in PLC systems for more reliable smart grid applications," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2746–2756, Nov. 2015.
- [6] N. Rozic, P. Banelli, D. Begusic, and J. Radic, "Multiple-threshold estimators for impulsive noise suppression in multicarrier communications," *IEEE Trans. Signal Process.*, vol. 66, no. 6, pp. 1619–1633, Mar. 2018.
- [7] K. Vastola, "Threshold detection in narrow-band non-Gaussian noise," *IEEE Trans. Commun.*, vol. COM-32, no. 2, pp. 134–139, Feb. 1984.
- [8] C. L. Nikias and M. Shao, *Signal Processing With Alpha-Stable Distributions and Applications*, New York, NY, USA: Wiley, 1995.
- [9] D. Middleton, "Statistical-physical models of electromagnetic interference," *IEEE Trans. Electromagn. Compat.*, vol. EMC-19, no. 3, pp. 106–127, Aug. 1977.
- [10] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: New methods and results for class A and class B noise models," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1129–1149, May 1999.

- [11] K. M. Rabie, B. Adebisi, A. M. Tonello, and G. Nauryzbayev, "For more energy-efficient dual-hop DF relaying power-line communication systems," *IEEE Syst. J.*, vol. 12, no. 2, pp. 2005–2016, Jun. 2018.
- [12] S. V. Zhidkov, "Performance analysis and optimization of OFDM receiver with blanking nonlinearity in impulsive noise environment," *IEEE Trans. Veh. Technol.*, vol. 55, no. 1, pp. 234–242, Jan. 2006.
- [13] A. Mathur, M. R. Bhatnagar, and B. K. Panigrahi, "Performance evaluation of PLC under the combined effect of background and impulsive noises," *IEEE Commun. Lett.*, vol. 19, no. 7, pp. 1117–1120, Jul. 2015.
- [14] L. Bai, M. Tucci, and M. Raugi, "Impulsive noise mitigation with interleaving based on MUSIC in power line communication," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3575–3584, Jul. 2019.
- [15] A. Spaulding and D. Middleton, "Optimum reception in an impulsive interference environment—Parts I: Coherent detection," *IEEE Trans. Commun.*, vol. COM-25, no. 9, pp. 910–923, Sep. 1977.
- [16] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. 2. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.
- [17] B. Weng and K. E. Barner, "Nonlinear system identification in impulsive environments," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2588–2594, Jul. 2005.
- [18] S. Peng, B. Chen, L. Sun, W. Ser, and Z. Lin, "Constrained maximum correntropy adaptive filtering," *Signal Process.*, vol. 140, pp. 116–126, Nov. 2017.
- [19] L. Lu, H. Zhao, and B. Champagne, "Distributed nonlinear system identification in α -stable noise," *IEEE Signal Process. Lett.*, vol. 25, no. 7, pp. 979–983, Jul. 2018.
- [20] A. Swami and B. M. Sadler, "On some detection and estimation problems in heavy-tailed noise," *Signal Process.*, vol. 82, no. 12, pp. 1829–1846, Dec. 2002.
- [21] G. Zhang, J. Wang, G. Yang, Q. Shao, and S. Li, "Nonlinear processing for correlation detection in symmetric alpha-stable noise," *IEEE Signal Process. Lett.*, vol. 25, no. 1, pp. 120–124, Jan. 2018.
- [22] H. Oh, H. Nam, and S. Park, "Adaptive threshold blanker in an impulsive noise environment," *IEEE Trans. Electromagn. Compat.*, vol. 56, no. 5, pp. 1045–1052, Oct. 2014.
- [23] K. A. Saaifan and W. Henkel, "Decision boundary evaluation of optimum and suboptimum detectors in Class-A interference," *IEEE Trans. Commun.*, vol. 61, no. 1, pp. 197–205, Jan. 2013.
- [24] G. Ndo, P. Siohan, and M.-H. Hamon, "Adaptive noise mitigation in impulsive environment: Application to power-line communications," *IEEE Trans. Power Del.*, vol. 25, no. 2, pp. 647–656, Apr. 2010.
- [25] H. Oh and H. Nam, "Design and performance analysis of non-linearity preprocessors in an impulsive noise environment," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 364–376, Mar. 2017.
- [26] X. Li, Y. Jiang, and M. Liu, "A near optimum detection in alpha-stable impulsive noise," in *Proc. IEEE ICASSP*, Taipei, Taiwan, Apr. 2009, pp. 3305–3308.
- [27] E. Alsusa and K. M. Rabie, "Dynamic peak-based threshold estimation method for mitigating impulsive noise in power-line communication systems," *IEEE Trans. Power Del.*, vol. 28, no. 4, pp. 2201–2208, Oct. 2013.
- [28] K. M. Rabie, E. Alsusa, A. D. Familua, and L. Cheng, "Constant envelope OFDM transmission over impulsive noise power-line communication channels," in *Proc. IEEE Int. Symp. Power Line Commun. Its Appl. (ISPLC)*, Austin, TX, USA, Mar. 2015, pp. 13–18.
- [29] A. Ikpehai, B. Adebisi, K. M. Rabie, M. Fernando, and A. Wells, "Energy-efficient vector OFDM PLC systems with dynamic peak-based threshold estimation," *IEEE Access*, vol. 5, pp. 10723–10733, 2017.
- [30] F. H. Juwono, Q. Guo, D. Huang, and K. P. Wong, "Deep clipping for impulsive noise mitigation in OFDM-based power-line communications," *IEEE Trans. Power Del.*, vol. 29, no. 3, pp. 1335–1343, Jun. 2014.
- [31] X. Li, J. Sun, S. Wang, L. Fan, and L. Chen, "Near-optimal detection with constant false alarm ratio in varying impulsive interference," *IET Signal Process.*, vol. 7, no. 9, pp. 824–832, Dec. 2013.
- [32] S. Zozor, J.-M. Brossier, and P.-O. Amblard, "A parametric approach to suboptimal signal detection in α -stable noise," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4497–4509, Dec. 2006.
- [33] Z. Luo, P. Lu, and G. Zhang, "Locally optimal detector design in impulsive noise with unknown distribution," *EURASIP J. Adv. Signal Process.*, vol. 2018, no. 34, pp. 1–10, Jun. 2018.
- [34] Z. Luo and Y. Zhang, "Novel nonlinearity based on Gaussianization and generalized matching for impulsive noise suppression," *IEEE Access*, vol. 7, pp. 65163–65173, 2019.
- [35] Z. Luo, R. Guo, X. Zhang, and Y. Nie, "Optimal and efficient designs of Gaussian-tailed non-linearity in symmetric α -stable noise," *Electron. Lett.*, vol. 55, no. 6, pp. 353–355, Mar. 2019.
- [36] A. R. Conn, K. Scheinberg, and L. N. Vicente, *Introduction to Derivative-Free Optimization*. Philadelphia, PA, USA: SIAM, 2009.
- [37] M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *Comput. J.*, vol. 7, no. 2, pp. 155–162, Feb. 1964.
- [38] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, no. 4, pp. 308–313, Jan. 1965.
- [39] M. H. Wright, "Direct search methods: Once scorned, now respectable," in *Proc. Numer. Anal.*, Harlow, U.K., 1996, pp. 191–208.
- [40] J. H. Mathews. (Jun. 2017). Module for Powell search method for a minimum. California State University, Fullerton. [Online]. Available: <http://mathfaculty.fullerton.edu/mathews/n2003/PowellMethodMod.html>
- [41] W. Spendley, G. R. Hext, and F. R. Himsforth, "Sequential application of simplex designs in optimisation and evolutionary operation," *Technometrics*, vol. 4, no. 4, pp. 441–461, Nov. 1962.



ZHONGTAO LUO (Member, IEEE) received the B.S. degree in electronic engineering and the Ph.D. degree in signal and information processing from the University of Electronic Science and Technology of China, China, in 2007 and 2015, respectively. From 2007 to 2008, he was an Electronics Engineer with Haier Group, China. From 2018 to 2019, he was a Visiting Postdoctoral Scholar with the University of Southern California. Since July 2015, he has been with the Chongqing University of Posts and Communications, China. His research interests include signal processing in wireless communication and radar systems, statistical signal processing, and digital image processing.



EDMOND A. JONCKHEERE (Life Fellow, IEEE) received the degree in electrical engineer from the University of Louvain, Belgium, in 1973, the Dr.-Eng. degree in aerospace engineering from Universite Paul Sabatier, Toulouse, France, in 1975, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1978. From 1973 to 1975, he was a Research Fellow of the European Space Agency. From 1975 to 1978, he was a Teaching Assistant, a Research Assistant, and subsequently a Research Associate with the Department of Electrical Engineering-Systems, University of Southern California. In 1979, he was with the Philips Research Laboratory, Brussels, Belgium. In 1980, he returned to the University of Southern California, where he is currently a Full Professor of electrical engineering and mathematics, a member of the Center for Applied Mathematical Sciences (CAMS), and a member of the Center for Quantum Information Science and Technology (CQUIST). He has had short-term visiting appointments with the Max-Planck-Institute, Göttingen, Germany, The Australian National University, Canberra, Australia, Cardiff University, Wales, U.K., and Swansea University, Wales. He has had consulting affiliations with the Memorial Medical Center of Long Beach, Lockheed-Martin, The Aerospace Corporation, and Honeywell. He is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE) for Contribution to the spectral theory of the linear-quadratic and H-infinity problems. His current research interests include conventional versus quantum networks, adiabatic quantum computations, and the power grid.

● ● ●