

MDPI

# Article $\ell_p$ -STFT: A Robust Parameter Estimator of a Frequency Hopping Signal for Impulsive Noise

# Yang Su, Lina Wang \*, Yuan Chen and Xiaolong Yang

School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China

\* Correspondence: wln\_ustb@126.com

Received: 7 August 2019; Accepted: 6 September 2019; Published: 11 September 2019



**Abstract:** Impulsive noise is commonly present in many applications of actual communication networks, leading to algorithms based on the Gaussian model no longer being applicable. A robust parameter estimator of frequency-hopping (FH) signals suitable for various impulsive noise environments, referred to as  $\ell_p$ -STFT, is proposed. The  $\ell_p$ -STFT estimator replaces the  $\ell_2$ -norm by using the generalized version  $\ell_p$ -norm where  $1 for the derivation of the short-time Fourier transform (STFT) as an objective function. It combines impulsive noise processing with any time-frequency analysis algorithm based on STFT. Considering the accuracy of parameter estimation, the double-window spectrogram difference (DWSD) algorithm is used to illustrate the suitability of <math>\ell_p$ -STFT with STFT and fractional low-order STFT (FLOSTFT), Cauchy noise, and Gaussian mixture noise as supplements of different background noises to better demonstrate the robustness and accuracy of  $\ell_p$ -STFT.

**Keywords:** frequency-hopping signal; short-time Fourier transform;  $\ell_p$ -norm; blindly estimating

# 1. Introduction

The outstanding advantages of strong anti-jamming, a low interception rate, and a high-frequency spectrum utilization rate [1–3] have led to the rapid development of frequency hopping communication technology in military radio, military wireless communication, and civil wireless local area networks. Under the condition of less prior information, the blind parameter estimation of captured frequency-hopping (FH) signals is the core problem of FH communication. Accurate estimation parameters are helpful to monitor and interfere with enemy communications [4] and then steal enemy signals precisely and effectively, which is a crucial part of communication countermeasures.

The popular parameter estimation methods for FH signals mostly utilize a Gaussian model as background noise. However, a large amount of impulsive noise often occurs in the practical intricate noise environment, such as low-frequency atmospheric noise, low-voltage power communication, seismic signals, and radar clutter [5–7], which are characterized by remarkable impulses [8] and heavier tails. Since they have no second-order or higher moments, the performance of the existing time-frequency analysis methods based on the Gaussian noise model is usually unacceptable or even invalid: for example, the commonly used short-time Fourier transform (STFT), which is easy to calculate for simple principles [9]; Wigner–Ville distribution (WVD), with high time-frequency resolution; and pseudo-Wigner–Ville distribution (PWVD), with fewer cross-terms. Even improved algorithms that have been proposed in recent years will exhibit undesirable performance degradation, such as the spectrogram method, double-window spectrogram difference (DWSD) [10] method, and STFT-WVD joint algorithm.

noise. A parameter estimation method based on fractional low-order (FLO) spectrum is proposed in reference [12], which can achieve accurate parameter estimation even at a low signal-to-noise ratio (SNR), but it is commonly directed to obtain linear frequency modulation signals. A parameter estimation algorithm is suggested in reference [13] based on second-order cyclic spectrum, which analyzes the relationship between the frequency and the corresponding second-order cyclic spectrum parameters. The estimation performance is good under high SNR, but poor in the case of low SNR. The median filtering method [14,15] has low computational complexity, but it blurs the characteristics of the signal while filtering out noise. Since the time-frequency distribution of the FH signal satisfies the sparse characteristic, some parameters estimation methods [16,17] are also proposed which are based on sparse time-frequency representation (STFD) of the signal obtained by establishing an optimization problem model. The signal parameter estimation technique under impulsive noise satisfies the requirements of security and interference in the complex communication environment, whereas many methods are not suitable for blind estimation of the parameters of FH signals.

Based on the above issues, we first focus on the impulsive noise processing problem by considering noise as the error [18] in the estimation. In many practical applications, squared error functions are insufficient in minimizing the magnitude of the estimation error. Therefore, under the premise that the least squares (LS) method makes less sense in impulsive noise, a more generalized method, the  $\ell_p$ -norm [19], which has high tolerance to noise, will be a better choice to capture noise outliers. The  $\ell_p$ -norm utilizes the minimum dispersion (MD) criterion instead of the minimum mean squared error (MMSE) criterion [20] to minimize the error between the estimated value and the true value. In other words, the  $\ell_p$ -norm means minimizing the *p*th power of the error so that the fitted object can be infinitely close to the target.

Inspired by the fact that STFT is the mathematical basis of commonly used time-frequency analysis algorithms, it can be considered as the objective function of  $\ell_p$ -norm estimation. Since STFT describes the energy change of the signal from two dimensions, time and frequency, the ensuing problem is that the objective function will change from a one-dimensional (1D) vector to a two-dimensional (2D) matrix. To resolve this problem, the matrix  $\ell_p$ -norm estimation can be converted into a column vector  $\ell_p$ -norm estimation. Then STFT will be used as a bridge to combine the  $\ell_p$ -norm algorithm with the time-frequency analysis algorithm to accomplish parameter estimation of the FH signal while processing impulsive noise. An attractive and reliable method of accurate parameter estimation named  $\ell_p$ -STFT is proposed in this paper.

The rest of the paper is organized as follows: In Section 2, we briefly review the impulsive noise distribution. The  $\ell_p$ -STFT method is presented in detail in Section 3, where the steps of parameter estimation of FH signals are introduced. In Section 4, the computer simulation analysis demonstrates that the  $\ell_p$ -STFT algorithm is superior to the STFT and fractional low-order STFT (FLOSTFT) algorithms in terms of accuracy and robustness. Some conclusions are drawn in Section 5.

### 2. Impulsive Noise Distribution

The  $\alpha$ -stable distribution and Gaussian mixture distribution (GMD) are two basic and widely used empirical models for non-Gaussian noise.

If a random variable satisfies the  $\alpha$ -stable distribution, its characteristic function has the following form:

$$\varphi(t) = \left\{ \begin{array}{l} \exp\{jut - \gamma|t|^{\alpha}[1 + j\beta sign(t)\tan(\alpha\pi/2)]\}\\ \exp\{jut - \gamma|t|^{\alpha}[1 + j\beta sign(t)(\pi/2)\log|t|]\} \end{array} \right\} \begin{array}{l} \alpha = 1\\ \alpha \neq 1 \end{array}$$
(1)

where

$$sign(t) = \left\{ \begin{array}{c} 1, t > 0\\ 0, t = 0\\ -1, t < 0 \end{array} \right\}$$
(2)

with  $-\infty < \mu < \infty, \gamma > 0, 0 < \alpha \le 2, -1 \le \beta \le 1$ .

The  $\alpha$ -stable distribution is completely determined by four unique parameters:  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\mu$ . The characteristic exponent  $\alpha$  reflects the degree of the impulsive characteristics, which is embodied as the smaller the value of  $\alpha$ , the stronger the impulse characteristics. The symmetry  $\beta$  determines the symmetry of the distribution. The dispersion  $\gamma$  is used to measure the extent to which the sample deviates from the mean, and finally,  $\mu$  is the location parameter. If  $\beta = 0$ ,  $\gamma = 1$ , and  $\mu = 0$ , the distribution corresponds to the standard symmetric  $\alpha$ -stable distribution (S $\alpha$ S). For  $\alpha = 1$  with  $\beta = 0$  or  $\alpha = 2$ , they correspond to two special cases of Cauchy distribution and Gaussian distribution [21], which is consistent with the implication of  $\alpha$ .

The second-order GMD can be expressed as a weighted summation of two independent and identically distributed (IID) Gaussian distributions [22]; that is:

$$p(x_1, x_2) = \frac{1 - \varepsilon_1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x_1 - \mu_1)^2}{2\sigma_1^2}\right) + \frac{\varepsilon_1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(x_2 - \mu_2)^2}{2\sigma_2^2}\right)$$
(3)

where  $0 < \varepsilon_1 < 1$  is the weighting coefficient, which is used to control the ratio of impulsive noise to Gaussian noise; and  $\mu_1$ ,  $\mu_2$  and  $\sigma_1^2, \sigma_2^2$  are the mean and variance of the Gaussian distribution, and the impulse characteristics are represented by the large difference between the two variances; that is,  $\sigma_2^2/\sigma_1^2$ . Figure 1 shows the magnitude of the Cauchy noise, the standard S $\alpha$ S noise, and mixed Gaussian noise in the time domain.



**Figure 1.** Time-domain waveform diagrams of impulsive noise under different conditions: (**a**) Cauchy noise with  $\alpha$  = 1; (**b**) standard symmetric  $\alpha$ -stable (S $\alpha$ S) noise with  $\alpha$  = 1.2; (**c**) mixed Gaussian noise.

#### 3. Proposed Algorithm

The received FH signal is formulated as:

$$y(t) = A \sum_{k} rect_{T_{H}}(t - kT_{H} - \alpha T_{H}) * e^{j2\pi f_{k}(t - kT_{H} - \alpha T_{H})} + n(t)$$
(4)

with  $0 < t \le T$ ; *A* is a numerical constant denoting the amplitude of the FH signal;  $rect_{T_H} = \begin{cases} 1 & t \in (-T_H/2, T_H/2) \\ 0 & others \end{cases}$  denotes a rectangular window;  $T_H$  and  $f_k$  are the hopping period and hopping frequency to be estimated separately; and n(t) represents the impulsive noise.

#### 3.1. Implementation of $\ell_p$ -STFT

The definition of STFT indicates that a long non-stationary signal is transformed into a short stationary signal matrix by adding a window function in the signal time domain, then perform DFT (Discrete Fourier Transform) on the windowed signal matrix to obtain the time-frequency spectrum matrix. Therefore, the 1D signal model is innovatively represented as a 2D signal matrix:

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{Z} \tag{5}$$

where  $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]$  denotes a  $K \times N$  reconstructed data matrix and has the form

$$\mathbf{Y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \\ y_2 & y_3 & \cdots & y_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_m & y_{m+1} & \cdots & y_{n+m} \\ \vdots & \vdots & \ddots & \vdots \\ y_k & y_{k+1} & \cdots & y_{n+k} \end{bmatrix}$$
(6)

 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N] \in \mathbb{C}^{K \times N}$  denotes the time-frequency spectrum matrix that needs to be estimated. The coefficient matrix  $\mathbf{A}_K = [\mathbf{a}_K(w_0), \mathbf{a}_K(w_1), ..., \mathbf{a}_K(w_{K-1})]$  is a complex component matrix, where  $\mathbf{a}_K(w_k) = [1, e^{jw_k}, ..., e^{j(K-1)w_k}]^T$  denotes a steering vector with  $w_k = 2\pi k/K$ , k = 0, 1, ..., K-1.  $\mathbf{y}_N = [y_1, y_2, ..., y_m, ..., y_k]$  is the short discrete stationary signal that is achieved by moving the position of the window function on the observed signal at a different time. Since the window function length *M* is less than the estimated spectrum length *K*, that is M < K, the zero-padding on signal matrix is required.  $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_N] \in \mathbb{C}^{K \times N}$  is the impulsive noise. The estimation task is to construct the above STFT cost function to obtain time-frequency spectrum **X** from signal **Y** mixed with impulsive noise.

In conventional estimators, the cost function of spectrum estimation by DFT is in form of  $\ell_2$ -norm, which is exploited in reference [18]. Considering the relationship between STFT and DFT, we know that the cost function of STFT has the same form of  $\ell_2$ -norm, that means the estimation of **X**, denoted by  $\hat{\mathbf{X}}$ , can be obtained by

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2 \tag{7}$$

which corresponds to  $\ell_2$ -norm minimization, and  $\|\cdot\|_F$  represents the  $\ell_2$ -norm. Note that **X** is corresponding to the spectrum obtained by directly performing STFT on the FH signal. To illustrate this point visually, we verify this result in the first part of simulation.

It is well known that the LS solution is equivalent to the maximum likelihood (ML) estimate when noise is a zero-mean white Gaussian process. In fact, it was also verified in reference [23] for LS-based parameter estimation in the presence of white Gaussian noise via several representative signal processing examples. Nevertheless, the validity of the Gaussian assumption is approximate at best, while the occurrence of non-Gaussian noise is common in many fields. In particular, if the noise is impulsive noise, unreliable parameter estimation will result, since the performance of the  $\ell_2$ -norm minimizer is very sensitive to outliers. To achieve robust estimation,  $\ell_p$ -norm minimization with  $1 is widely used since it is less sensitive to outliers than the square function. In this work, we focus on the cost function of <math>\ell_p$ -norm  $\|\cdot\|_p$ .

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{Y}} \left\| \mathbf{Y} - \mathbf{A} \mathbf{X} \right\|_{p}^{p} \tag{8}$$

where the matrix  $\ell_p$ -norm commonly uses an easier form defined as

$$\left\|\mathbf{E}\right\|_{p}^{p} = \sum_{n} \left\|\mathbf{e}_{n}\right\|_{p}^{p} \tag{9}$$

In light of the  $\ell_p$ -STFT model, it first calculates the initial spectrum of **X** without filtering the outliers. Then, the powerful iterative reweighted least squares (IRLS) algorithm [24] is employed to eliminate the noise interference. One thing to point out is that the cost function of  $\ell_p$ -STFT is a matrix that needs to be converted to column vectors before processing noise by IRLS [25]. IRLS will continuously iterate the weight matrix to obtain a satisfactory solution that converges to optimal accuracy. Finally, a time-frequency analysis algorithm via  $\ell_p$ -STFT is constructed.

Since each column in matrix  $\mathbf{Y}$  is independent, we can convert a big equation of the cost function matrix into small questions of N vectors [26] to tackle. That is, the matrix objective function of Equation (9) can be further expressed as

$$J(\mathbf{x}) = \sum_{n=0}^{N-1} \left\| \mathbf{y}_n - \mathbf{A} \mathbf{x}_n \right\|_p^p$$
(10)

where  $\mathbf{y}_n$  and  $\mathbf{x}_n$  are the *n*th column of  $\mathbf{Y}$  and  $\mathbf{X}$ , respectively. From there we can separately process the  $\ell_p$ -norm of each column vector and then arrange the processed N sub-problems into a matrix in order, which is the time-frequency spectrum we ultimately need. Now the vector data model can be written as:

$$\mathbf{y}_K = \mathbf{A}\mathbf{x}_K + \mathbf{z}_K \tag{11}$$

where  $\mathbf{y}_K = [y_1, y_2, ..., y_K]$  and  $\mathbf{x}_K = [x_1, x_2, ..., x_K]$ . Equation (11) is solved by minimizing the vector objective function as follows:

$$\hat{x}_{k} = \arg\min_{x_{k}} \left\| \mathbf{y}_{K} - \mathbf{a}_{K}(w_{k}) x_{k} \right\|_{p}^{p}, k = 1, 2, \dots, K$$
(12)

with  $\hat{\mathbf{x}}_K = [\hat{x}_1, \hat{x}_2, \dots, \hat{x}_K]$ . Employing (12) yields:

$$J(x) = \sum_{k=1}^{K} |y_k - e^{jkw_k} x_k|^p, k = 1, 2, \dots, K$$
(13)

To solve Equation (13), IRLS is applicable to settle such minimizing cost function problem of  $\ell_p$ -norm. It starts with a rough estimate of vector  $\mathbf{x}_K$  and then adjusts the weight coefficient to gradually approximate the optimal value. The objective function f(x) can be reformulated as

$$J(x) = \sum_{k=1}^{K} \left[ \left| y_k - e^{jkw_k} x_k \right|^{\frac{p-2}{2}} \right]^2 \left| y_k - e^{jkw_k} x_k \right|^2$$
(14)

It is not difficult to obtain the matrix form of Equation (14), and it can be expressed as

$$J(x) = (\mathbf{y}_K - \mathbf{a}_K(w_k)x_k)^{\mathsf{H}}\mathbf{W}_K(\mathbf{y}_K - \mathbf{a}_K(w_k)x_k)$$
(15)

where

$$\mathbf{W}_{K} = \operatorname{diag}\left\{ \left( |r_{1}|^{p-2/2} \right)^{2}, \left( |r_{2}|^{p-2/2} \right)^{2}, \dots, \left( |r_{K}|^{p-2/2} \right)^{2} \right\}$$
(16)

is a diagonal weighting matrix represented by a residual vector; that is,  $\mathbf{r}_K = \mathbf{y}_K - \mathbf{a}_K(w_k)x_k$ . We further analyze Equation (16) to get the estimate of  $\hat{x}_k$  as

...

$$\hat{x}_{k} = \left(\mathbf{a}_{K}^{H}(w_{k})\mathbf{W}_{K}\mathbf{a}_{K}(w_{k})\right)^{-1}\mathbf{a}_{K}^{H}(w_{k})\mathbf{W}_{K}\mathbf{y}_{K}$$
(17)

However, here  $x_k$  is taken as a function of  $W_K$ , and  $W_K$  is also taken as a residual function of  $x_k$ , and we cannot get the optimal estimate of  $x_k$  in one step. Usually, all elements in the weighted matrix are initially set to  $\mathbf{W}_{K}^{(0)} = 1$ , and the initial value of vector  $\mathbf{x}_{K}$  can be given by

$$\mathbf{x}_{K} = \sum_{k=1}^{K} \left( \mathbf{a}_{K}^{H}(w_{k}) \mathbf{a}_{K}(w_{k}) \right)^{-1} \mathbf{a}_{K}^{H}(w_{k}) \mathbf{y}_{K} = \left( \mathbf{A}^{H} \mathbf{A} \right)^{-1} \mathbf{A}^{H} \mathbf{y}_{K}.$$
 (18)

The core of the IRLS method lies in continuously updating the estimated vector by updating weighted coefficients that utilize the previous residual vector. After each iteration of vector  $\mathbf{x}_{K_r}$  IRLS will automatically judge whether it converges. If the accuracy requirement does not meet the tolerance  $\varepsilon$ , which is an extremely small positive number, it will continue iterating to guarantee convergence.

Afterwards, following the order of the previous matrix vectorization, we rearrange the series of vectors processed by  $\ell_p$ -STFT into a time-frequency matrix in order:

$$\mathbf{X}_{STFT} = [\mathbf{x}_1, \dots, \mathbf{x}_N]. \tag{19}$$

At this point, we choose an appropriate time-frequency analysis algorithm and construct it via the previously generated  $\ell_p$ -STFT. The complete  $\ell_p$ -STFT algorithm is summarized in Algorithm 1.

## Algorithm 1: For $\ell_p$ -norm short-time Fourier transform ( $\ell_p$ -STFT)

1. Reconstruct the original signal into a signal matrix and build the STFT cost function.

2. Simplify the matrix  $\ell_p$ -norm optimization into sequential vector  $\ell_p$ -norm optimization.

- 3. Initialize the weighted matrix  $\mathbf{W}_K$  with  $\mathbf{W}_K(0) = \mathbf{I}$  and the estimated  $\mathbf{x}_K$  with  $\mathbf{x}_K^{(0)} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{y}_K$ .
- 4. Compute the residual vector  $\mathbf{r}_{K}(k) = \mathbf{y}_{K} \mathbf{a}_{K}(w_{k})x_{k}$  and use  $\mathbf{r}_{K}(k)$  to reconstruct the weighting matrix

$$\mathbf{W}_{K}(k) = diag\{(|r_{1}|^{p-2/2})^{2}, (|r_{2}|^{p-2/2})^{2}, \dots, (|r_{K}|^{p-2/2})^{2}\}.$$

5. Update 
$$\hat{x}_k(k+1) = \left(\mathbf{a}_K^H(w_k)\mathbf{W}_K\mathbf{a}_K(w_k)\right)^{-1}\mathbf{a}_K^H(w_k)\mathbf{W}_K\mathbf{y}_K$$

5. Update  $x_k(k+1) = (\mathbf{a}_{K^{(n)}}(w_k)\mathbf{v}_{K}\mathbf{a}_{K}(w_k)) \quad \mathbf{a}_{K^{(n)}}(w_k)\mathbf{v}_{K}\mathbf{y}_{K}$ . 6. Repeat steps 4 and 5 until k = K to get the updated  $\hat{\mathbf{x}}_{k}^{(n)}$ . If  $\frac{\|\hat{\mathbf{x}}_{K}^{(n+1)} - \hat{\mathbf{x}}_{K}^{(n)}\|_{p}}{\|\hat{\mathbf{x}}_{K}^{(n)}\|_{p}} < \varepsilon$ , stop the iteration; otherwise,

go to step 4.

7. Arrange the vectors into a time-frequency matrix in order.

8. Construct time-frequency analysis algorithm using a  $\ell_p$ -STFT spectrum matrix.

#### 3.2. Parameter Estimation Based on $\ell_v$ -STFT

As usual, whether it is directly acquiring the characteristics of the FH signal for tracking interference or further constructing the local oscillator signal to steal enemy information, the blind estimation of parameters is a necessary condition for the next step. The double-window spectrogram difference (DWSD) method has a superior time-frequency representation that can accurately estimate important parameters such as hopping period, hopping frequency, hopping time set, and hopping rate. More specifically, the spectrogram (SPEC) method should be performed using a long window function  $h_l(t)$  and short window function  $h_s(t)$ , respectively on signal to obtain the time-frequency transform results. Combining the two sets of results with better time and frequency resolution will provide the time-frequency analysis of DWSD. The definition of DWSD is given as

$$DWSD(t, w) = SPEC_l(t, w) \cdot SPEC_s(t, w)$$
(20)

where

$$SPEC(t,w) = \left|STFT(t,w)\right|^2 = \left|\int_{-\infty}^{+\infty} x(\tau)h(t-\tau)e^{-jw\tau}d\tau\right|^2$$
(21)

using the  $\ell_p$ -STFT method to eliminate the presence of outliers and acquire the  $\ell_p$  –  $STFT_l(t, w)$  and  $\ell_p$  –  $STFT_s(t, w)$ . The normalized Stankovic entropy measure is an effective criterion used for evaluating the time-frequency concentration in the spectrogram [10], which has the expression

$$R_{3} = -\frac{1}{2} \log_{2} \left( \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} SP^{3}(n,m)}{\sum_{i=1}^{N} \sum_{j=1}^{M} |SP(n,m)|} \right)$$
(22)

where SP(n, m) denotes the signal entry of time-frequency distribution that is affected by the length of the window function. The smaller the value of  $R_3$ , the better the time-frequency performance and the more optimal the length of the window function.

Suppose N equals the length of signal, Figure 2 shows the curve of entropy measure  $R_3$  obtained by simultaneously varying the length of long window and the short window between  $0 \sim N/2$  and  $N/8 \sim 5N/8$  in a same sampling step. When the short window ranges from N/20 to 9N/40 and the long window ranges from N/4 to 9N/20, we can obtain a better time-frequency performance.



Figure 2. The Stankovic entropy measure versus length of window function.

Afterwards, we can apply the treated  $\ell_p$ -STFT spectrum to construct the crucial time-frequency diagram even in an impulsive noise environment. The  $\ell_p$ -DWSD can be expressed as

$$\ell_p \text{-}DWSD(t,w) = \left|\ell_p - STFT_l(t,w)\right|^2 \cdot \left|\ell_p - STFT_s(t,w)\right|^2$$
(23)

To more conveniently and clearly estimate parameters, we can take the maximum value of the time-frequency diagram to obtain the time-frequency ridge line, as shown in Figure 5b.

$$f(n) = \operatorname*{argmax}_{m} \{DWSD'(n,m)\} * f_{s}/(2N)$$
(24)

Performing first-order difference on the time-frequency ridge line will produce a frequency difference sequence  $d_1$ .

$$d_1 = \left| \Delta(f(n)) \right| > 0 \tag{25}$$

The time domain position of the value in  $d_1$  is an equally spaced difference sequence  $d_2$ , which corresponds to the hopping time set. The estimated hopping period is the mean value of a set of the hopping period approximations, which are acquired by performing first-order difference on  $d_2$ .

$$T_H = mean(\Delta d_2) / f_s \tag{26}$$

Then we can easily estimate the hopping rate, which is the reciprocal of the hopping period and the hopping time set, because of the changeless hopping period.

$$R_H = 1/T_H \tag{27}$$

$$t_k = T_0 + kT_H \tag{28}$$

The estimation of hopping frequency is also based on the time-frequency ridge line, and can be acquired by averaging the instantaneous frequency in each hopping period.

$$f_k = mean[f(n)], (k-1)T_H * f_s \le n \le kT_H \cdot f_s$$
<sup>(29)</sup>

After obtaining the above parameters, it is not difficult to determine that the hopping period and hopping frequency are two basic parameters. The estimation accuracy of these two parameters is an important criterion to evaluate the performance of the  $\ell_p$ -STFT algorithm.

## 4. Simulation Result and Analysis

Computer simulations were conducted to investigate the performance of the derived estimator. The signal is generated by Equation (4), where the parameters in the FH signal are set as follows: eight hopping periods with  $T_H$ = 0.1s, the sampling rate  $f_s$  = 10 kHz, the sampling point N = 800, the hopping frequencies  $f_k$  = {0.5, 2.5, 1, 3.5, 2, 3, 1.5, 4} kHz, and the noise n(t) being standard S $\alpha$ S noise with  $\alpha$  = 1.5,  $\beta$  = 0,  $\gamma$  = 1. Since impulsive noise has no finite second-order moment, the generalized signal-to-noise ratio (GSNR) [27] will be more applicable than the traditional signal-to-noise ratio, and is set to 3 dB. The order of  $\ell_p$ -STFT satisfies 1 < p < 2, p = 1.2, the explanation for p is given in Appendix A, and the tolerance is taken as  $\varepsilon$  = 1 × 10<sup>-4</sup>. All results were simulated using MATLAB running on Intel(R) Core(TM) i7-4790 CPU@3.60 GHz and Windows 10 for observations.

First, we verified the equivalence between the spectrogram obtained by STFT and the spectrogram estimated from the cost function of STFT based on  $\ell_2$ -norm. We select the length of the window function is 100 and carry out the experiment in non-noise environment to removal interference. Figure 3 shows that the time-frequency spectrograms produced by these two ways are almost consistent. It serves well to prove that the cost function of STFT is effective and the spectrogram obtained by STFT is equivalent to the spectrogram estimated by the cost function base on  $\ell_2$ -norm.



**Figure 3.** Time-frequency spectrograms of short-time Fourier transform (STFT) obtained by (**a**) the STFT; (**b**) the cost function of STFT based on  $\ell_2$ -norm.

Second, we studied the performance of the proposed method applied to the DWSD algorithm in  $\alpha$ -stable noise. The lengths of the long and short windows are 251 and 51, respectively. This means they are within the range we have illustrated and the selection of window function length is suitable

for the DWSD method. The time-frequency diagram of STFT with a narrow window function is shown in Figure 4a, which is also the initial spectrogram of  $\ell_p$ -STFT. We can see that there are several impulse interferences that are maintained for a short time but with a high amplitude, as well as many small noise points, both of which seriously disturb the signal estimation. In contrast, the spectrogram in Figure 4b, which was obtained by  $\ell_p$ -STFT, effectively extracts the energy of signal submerged in impulsive noise while removing noise outliers. In the same way, the result comparison of time-frequency diagrams using a wide window function are presented in Figure 4c,d. The result is consistent with the effect of  $\ell_p$ -STFT adding the narrow window function. Therefore,  $\ell_p$ -STFT provides excellent robustness that appears to be hardly affected by the characteristics of noise outliers. Then, Figure 5a,b shows the time-frequency diagram and the time-frequency ridge line of  $\ell_p$ -DWSD, which combines the above two sets of optimal results processed by  $\ell_p$ -STFT.  $\ell_p$ -DWSD exhibits extraordinarily powerful time-frequency concentration and accurately describes the parameter information of FH signals, even in an impulsive noise environment.

Third, we selected the algorithms related to the STFT to compare the performance of parameter estimation: STFT, FLOSTFT, and  $\ell_p$ -STFT. The FLOSTFT method represents a class of algorithms based on fractional low-order statistics, and the fractional lower moments assumes p = 0.4. After repeating 100 experiments on these three methods in  $\alpha$ -stable noise with  $\alpha = 1.2$ ,  $\beta = 0$ ,  $\gamma = 1$ , we compared the relative error of hopping period and the mean squared error (MSE) of hopping frequency to evaluate their parameter estimation performance.



**Figure 4.** Time-frequency spectrograms of STFT with different window functions in impulsive noise. (a) Initial spectrogram with narrow window function; (b) spectrogram of  $\ell_p$ -STFT with narrow window function; (c) initial spectrogram with wide window function; (d) spectrogram of  $\ell_p$ -STFT with wide window function.



**Figure 5.** (a) Time-frequency spectrogram and (b) time-frequency ridge line of  $\ell_p$ -double-window spectrogram difference (DWSD).

On the one hand, in terms of the estimated performance of the hopping period that is shown in Figure 6a, the accuracy of STFT without any processing is the worst among the three methods but is also the most primitive. STFT can be used as an intuitive indicator of performance comparison. From the holistic perspective of the curve, the larger the impulse amplitude, the lower the fault tolerance of the algorithm and the worse the performance becomes. When GSNR > 3 dB, both FLOSTFT and  $\ell_p$ -STFT methods almost accurately estimate the hopping period and maintain stable estimation under high GSNR, and for STFT, the same result requires a higher GSNR to be completed, while the robustness of the algorithm under low GSNR can better prove the superiority of the algorithm. From Figure 4a we can find that, first, at low GSNR, the overall performance of the  $\ell_p$ -STFT algorithm is superior to FLOSTFT and FLOSTFT is better than STFT. Second, the periodic relative error curve of the  $\ell_p$ -STFT algorithm has the fastest rate of decline compared with the other two methods. Furthermore, at GSNR = 1 dB, the relative error value of  $\ell_p$ -STFT is almost equal to zero. This effectively demonstrates that the  $\ell_p$ -STFT algorithm not only maintains stable and accurate estimation at high GSNR, but also quickly achieves relatively high parameter estimation accuracy at low GSNR.

On the other hand, from the perspective of estimating the performance of the hopping frequency, in addition to  $\alpha$ -stable noise, Cauchy noise with  $\alpha = 1$  and GWD noise with  $\varepsilon = 0.03$ ,  $\mu = 0$ ,  $\sigma_1 = 2$ ,  $\sigma_2^2/\sigma_1^2 = 3000$  are also used as impulsive noise environments to better illustrate the robustness of the  $\ell_p$ -STFT algorithm. Comparing the MSE curve of the three methods presented in Figure 6b–d in any kind of noise environment,  $\ell_p$ -STFT has higher accuracy for frequency estimation than the other two methods. For Figure 6b, with smaller impulsive amplitudes, when GSNR < 0 dB, the estimated performance of FLOSTFT is seriously degraded, whereas  $\ell_v$ -STFT can still estimate the hopping frequency comparatively accurately, and this point is more evident in the Cauchy noise environment of Figure 6c. The strong impulsive noise features further highlight the advantages of  $\ell_p$ -STFT, and the estimation precision of  $\ell_p$ -STFT is nearly 10 dB higher than FLOSTFT when GSNR < 0 dB. Figure 6d showing GWD noise represents the case where the impulse appears more frequently, and the noise removal effect of the  $\ell_v$ -STFT algorithm is still the best under this scenario, at about 8 dB higher than FLOSTFT in low GSNR. Even when GSNR > 0 dB, although the estimated performance of both algorithms is improved,  $\ell_v$ -STFT is also maintained at nearly 3 dB higher than FLOSTFT for all impulsive noise environments. This result convincingly proves the robustness of  $\ell_p$ -STFT to outliers and the accuracy of parameter estimation in the face of strong impulsive noise.



**Figure 6.** Performance comparison of STFT, fractional low-order STFT (FLOSTFT), and  $\ell_p$ -STFT methods versus generalized signal-to-noise ratio (GSNR). (a) Periodic relative error of hopping period; (b) mean squared error (MSE) of hopping frequency in  $\alpha$ -stable noise; (c) MSE of hopping frequency in Cauchy noise; (d) MSE of hopping frequency in Gaussian mixture noise.

# 5. Conclusions

Almost all algorithms approximately applicable to the Gaussian distribution model have varying degrees of performance degradation in an impulsive noise environment. To solve such a knotty problem, a new algorithm called  $\ell_p$ -STFT is proposed for robust and accurate parameter estimation of FH signals. It takes STFT as an associated bond to combine the  $\ell_p$ -norm algorithm with the time-frequency analysis algorithm. The simulation results demonstrate that the  $\ell_p$ -STFT method has strong robustness and parameter estimation accuracy, especially in low-GSNR environments. However, the computational complexity caused by iteration requires further research in the future.

**Author Contributions:** Conceptualization, Y.S., L.W., Y.C. and X.Y.; methodology, Y.S. and L.W.; software, Y.S. and L.W.; validation, Y.S., L.W. and Y.C.; formal analysis, Y.C. and X.Y.; writing—original draft preparation, Y.S., L.W., Y.C. and X.Y.; writing—review and editing, Y.S., L.W., Y.C. and X.Y.; visualization, Y.S. and L.W.; supervision, Y.C. and X.Y.; funding acquisition, L.W. and Y.C.

**Funding:** This research was funded by the National Natural Science Found of China under grant number 61701021 and 61701020.

Conflicts of Interest: The authors declare no conflict of interest.

# Appendix A

We explore the choice of order p, which is important in the proposed method. In 50 repeated experiments in  $\alpha$ -stable noise, seven values of p from large to small were selected to compare the

impact on  $\ell_p$ -STFT. As shown by the MSE of hopping frequency in Figure A1a, it is not difficult to find that  $\ell_p$ -STFT with a smaller order p has better estimation performance. It is difficult to find the optimal p precisely, whereas the performance of  $\ell_p$ -STFT is generally consistent and satisfactory when the order p within the range  $1.1 \le p \le 1.4$ . Hence the p we need is not necessarily the best but a suitable one, and p = 1.2 is preferred in  $\ell_p$ -STFT.

# Appendix **B**

The convergence rate of  $\ell_p$ -STFT method in  $\alpha$ -stable noise, Cauchy noise and GWD noise are also compared. In fact, we primarily care about the number of iterations of  $\ell_p$ -STFT method under different tolerances. Seven tolerances are taken for verification, that are  $\varepsilon = 1 \times 10^{-6}$ ,  $1 \times 10^{-5}$ ,  $1 \times 10^{-4}$ ,  $1 \times 10^{-3}$ ,  $1 \times 10^{-2}$ ,  $1 \times 10^{-1}$  and  $1 \times 10^{-0}$ . From the Figure A1b we observed that the proposed algorithm has similar convergence rate under different impulsive noises, which has a linear approximation relationship to the tolerance accuracy. In addition, it can be seen that for the tolerance  $\varepsilon = 1 \times 10^{-4}$  we provided, the relative error is less than the tolerance after 9 iterations.



**Figure A1.** (a) MSE of hopping frequency with *p* = 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7; (b) Number of iterations versus tolerance.

# References

- Jin, Y.; Liu, J. Parameter estimation of frequency hopping signals in alpha stable noise environment. In Proceedings of the International Conference on Signal Processing, Beijing, China, 21–25 October 2012; pp. 250–253.
- Fan, H.N.; Guo, Y.; Meng, Q.W. Blind parameter estimation of frequency-hopping signals based on atomic decomposition. In Proceedings of the 2009 First International Workshop on Education Technology and Computer Science, Wuhan, China, 7–8 March 2009.
- 3. Chen, T.C. Joint signal parameter estimation of frequency-hopping communications. *IET Commun.* 2012, *6*, 381–389. [CrossRef]
- 4. Liu, S.H.; Zhang, Y.M. Structure-aware Bayesian compressive sensing for frequency-hopping spectrum estimation with missing observations. *IEEE Trans. Signal Process.* **2018**, *66*, 2153–2166. [CrossRef]
- 5. Jin, Y.; Liu, J. Parameter estimation of frequency hopping signals based on the robust S-transform algorithms in alpha stable noise environment. *AEU-Int. J. Electron. Commun.* **2016**, *70*, 611–616. [CrossRef]
- Yue, B.B.; Peng, Z.M.; He, Y.M. Impulsive noise suppression using fast Myriad filter in seismic signal processing. In Proceedings of the International Conference on Computational and Information Sciences, Shiyang, China, 21–23 June 2013.
- Amirshahi, P.; Navidpour, S.M.; Kavehrad, M. Performance analysis of uncoded and coded OFDM broadband transmission over low voltage power-line channels with impulsive noise. *IEEE Trans. Power Deliv.* 2006, 21, 1927–1934. [CrossRef]

- 8. Yang, C.; Zhang, Y.; Jian, Y. Structure-adaptive fuzzy estimation for random-valued impulse noise suppression. *IEEE Trans. Circuits Syst. Video Technol.* **2016**, *28*, 414–427.
- 9. Tong, L.; Tang, Y.; Lv, J. Parameter estimation of FH signals based on STFT and music algorithm. In Proceedings of the International Conference on Computer Application and System Modeling, Taiyuan, China, 22–24 October 2010.
- He, Y.; Su, Y.; Chen, Y. Double window spectrogram difference method: A blind estimation of frequency-hopping signal for battlefield communication environment. In Proceedings of the Asia-Pacific Conference on Communications, Ningbo, China, 12–14 November 2018; pp. 439–443.
- 11. Tsihrintzis, G.A.; Nikias, C.L. Performance of optimum and suboptimum receivers in the presence of impulsive noise modeled as an alpha-stable process. *IEEE Trans. Commun.* **1995**, *43*, 904–914. [CrossRef]
- Gu, G.D.; Zhang, Y.S.; Tian, B. Estimation of LFM signal's time parameters under the alpha-stable distribution noise. In Proceedings of the IEEE Circuits and Systems International Conference on Testing and Diagnosis, Chengdu, China, 28–29 April 2009.
- 13. Zhao, C.H.; Yang, W.C.; Cheng, B.Z. Parameter estimation of MPSK signals in alpha stable distribution noise environment. *J. Shenyang Univ. Technol.* **2013**, *35*, 194–199.
- 14. Wierzchowski, W.; Pawełczyk, M. Median filtering approach for active control of impulsive noise. In Proceedings of the International Conference on Methods and Models in Automation and Robotics, Miedzyzdroje, Poland, 2–5 September 2014.
- Ismaeil, M.; Pritamdas, K.; Devi, K.J.K. Performance analysis of new adaptive decision based median filter on FPGA for impulsive noise filtering. In Proceedings of the International Conference on Electronics, Materials Engineering and Nano-Technology, Kolkata, India, 28–29 April 2017.
- 16. Kowalski, M. Sparse regression using mixed norms. *Appl. Comput. Harmon. Anal.* **2009**, *27*, 303–324. [CrossRef]
- 17. Gardner, T.J.; Magnasco, M.O. Sparse time-frequency representations. *Proc. Natl. Acad. Sci. USA* 2006, 103, 6094–6099. [CrossRef] [PubMed]
- 18. Ming, X.; Xu, L.; Jian, L. IAA spectral estimation: Fast implementation using the Gohberg–Semencul factorization. *IEEE Trans. Signal Process.* **2011**, *59*, 3251–3261. [CrossRef]
- 19. Li, T.H. A nonlinear method for robust spectral analysis. *IEEE Trans. Signal Process.* **2010**, *58*, 2466–2474. [CrossRef]
- 20. Chen, B.S.; Chen, J.M.; Shern, S.C. An ARMA robust system identification using a generalized *l*p-norm estimation algorithm. *IEEE Trans. Signal Process.* **1994**, *42*, 1063–1073. [CrossRef]
- 21. Kuruoglu, E.E.; Fitzgerald, W.J.; Rayner, P.J.W. Nonlinear autoregressive modeling of non-Gaussian signals using l/sub p/-norm techniques. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Munich, Germany, 21–24 April 1997.
- 22. Zhao, X.; Li, F. Sparse Bayesian compressed spectrum sensing under Gaussian mixture noise. *IEEE Trans. Veh. Technol.* **2018**, *67*, 6087–6097. [CrossRef]
- 23. So, H.C.; Chan, Y.T.; Chen, Y. Simple formulas for bias and mean square error computation. *IEEE Signal Process Mag.* **2013**, *30*, 162–165.
- 24. Lai, M.J.; Xu, Y.; Yin, W. Improved iteratively reweighted least squares for unconstrained smoothed *l*p minimization. *Siam J. Numer. Anal.* **2013**, *51*, 927–957. [CrossRef]
- Wen, F.; Liu, W. Iteratively reweighted optimum linear regression in the presence of generalized Gaussian noise. In Proceedings of the IEEE International Conference on Digital Signal Processing, Beijing, China, 16–18 October 2016.
- 26. Zeng, W.J.; So, H.C.; Huang, L. *l*p-MUSIC: Robust direction-of-arrival estimator for impulsive noise environments. *IEEE Trans. Signal Process.* **2013**, *61*, 4296–4308. [CrossRef]
- Li, L.; Nicolas, H.; Shi, X.F. Joint Estimation of Doppler Stretch and Time Delay of Wideband Echoes for LFM Pulse Radar Based on Sigmoid-FRFT Transform under the Impulsive Noise Environment. *Electronics* 2019, *8*, 121. [CrossRef]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).