An Optimisation Approach to Robust Estimation of Multicomponent Polynomial Phase Signals in Non-Gaussian Noise

Duc Son Pham, Yee Hong Leung, Kok Lay Teo
Curtin University of Technology
Perth, Western Australia
Email: dspham@ieee.org

Abdelhak Zoubir
Darmstadt University of Technology
Germany
Email: zoubir@ieee.org

***Abstract*— In this paper [1] we address the problem of estimating the parameters of multicomponent polynomial phase signals in impulsive noise which arises in many practical situations. In the presence of this non-standard noise, existing techniques perform can poorly. We propose a nonlinear M-estimation approach to improve the existing techniques. The phase parameters are obtained by solving a nonlinear optimisation problem. A procedure is proposed to find the global minimum at low computational cost. Simulation examples show the proposed method performs better than existing methods.**

I. Introduction

In many signal processing applications such as those in synthetic aperture radar (SAR) or in radio communications where the phase is continuously modulated, the signal of interest is *nonstationary*. The most commonly used model in parametric analysis of analytic and nonstationary signals is the polynomial phase signal (PPS) model which is motivated by Weierstrass' theorem. This theorem implies that for a finite duration of observations, an arbitrary time varying phase can be well approximated by a polynomial of sufficiently higher order.

Most signal processing techniques for extracting useful information from received signals are usually developed under the assumption that the noise is additive Gaussian [1], [8], [10]. However, practical measurements reveal that the noise behaviour is of a more impulsive nature and the probability density function does not follow the Gaussian distribution. Non-Gaussian noise has been observed over a wide range of important practical applications such as those in wireless communication channels [5], switching transients in power lines [6], automobile ignition noise [7], [12], and synthetic aperture radar [3]. Under these circumstances conventional signal processing techniques suffer from a considerable performance loss. Thus there is a need for alternative methods. With this motivation we propose an effective method in this paper.

The paper is organised as follows. In Section 2 we explain the signal model and derive the Cramér-Rao bound. In Section 3 we propose a nonlinear M-estimation approach to obtain robust estimates of the parameters, which can be formulated as a nonlinear optimisation problem. Then we propose a computationally efficient procedure for solving this optimisation problem. Section 4 studies the performance of the proposed approach via a simulation example. Section 5 concludes the paper.

[1]This research was in part supported by the Australian Research Council (ARC) under grant DP0211849.

II. Signal Model and the Cramér-Rao Bound

Consider a complex-valued K-component polynomial phase signal embedded in complex circular white non-Gaussian noise with total variance σ^2:

$$y(t) = \sum_{k=1}^{K} \alpha_k s_k(t) + x(t), \tag{1}$$

where α_k is the amplitude of the kth component,

$$s_k(t) = \exp\left\{ j \sum_{m=0}^{M} \omega_m^k t^m \right\}, \tag{2}$$

and $x(t)$ is the noise whose real and imaginary parts are uncorrelated and each follows the ε-contaminated model [13]

$$f(x) = (1-\varepsilon) f_G(x;\nu^2) + \varepsilon f_G(x;\kappa\nu^2), \tag{3}$$

where $f_G(x;\nu^2)$ denotes the zero-mean Gaussian distribution with variance ν^2 and $(1+(\kappa-1)\varepsilon)\nu^2 = \sigma^2/2$. The parameter $\kappa > 1$ represents the impulsive strength of the noise. This model has been used to characterise impulsive noise in practical situations [13].

The problem is as follows: given N samples of the received signal $y(t)$ at time instances $t_0, \ldots, t_{N-1}$, find estimates of the amplitudes α_k, the polynomial phase coefficients ω_m^k for $k = 1, \ldots, K$, $m = 0, 1, \ldots, M$, and the noise variance σ^2. We have assumed, without loss of generality, that all components have the same order M. Denote $\boldsymbol{\omega}_k = [\omega_0^k, \ldots, \omega_M^k]^T$, $\mathbf{t} = [t_0, \ldots, t_{N-1}]^T$, $\mathbf{y} = [y(t_0), \ldots, y(t_{N-1})]^T$, $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_K]^T$, $\mathbf{x} = [x(t_0), \ldots, x(t_{N-1})]^T$, $s_k(t, \boldsymbol{\omega}_k) = \exp\left\{ j \sum_{m=0}^{M} \omega_m^k t^m \right\}$, $\mathbf{s}_k(\mathbf{t}, \boldsymbol{\omega}_k) = [s_k(t_0, \boldsymbol{\omega}_k), \ldots, s_k(t_{N-1}, \boldsymbol{\omega}_k)]^T = \mathbf{s}_k(\boldsymbol{\omega}_k)$[2],

[2]The parameter $\mathbf{t}$ is omitted for notational simplification.

$\boldsymbol{\omega} = [\boldsymbol{\omega}_1^T, \ldots, \boldsymbol{\omega}_K^T]^T$, $\mathbf{S}(\boldsymbol{\omega}) = [\mathbf{s}_1(\boldsymbol{\omega}_1), \ldots, \mathbf{s}_K(\boldsymbol{\omega}_K)]$, and $\boldsymbol{\vartheta} = \{\boldsymbol{\alpha}, \boldsymbol{\omega}, \sigma^2\}$. We can rewrite (1) as:

$$\begin{aligned} \mathbf{y} &= \mathbf{S}(\boldsymbol{\omega})\boldsymbol{\alpha} + \mathbf{x} \\ &= \boldsymbol{\eta}(\boldsymbol{\vartheta}) + \mathbf{x} \end{aligned} \tag{4}$$

When the noise is purely Gaussian, it was shown in [11, ch. 5] that the Fisher information matrix is

$$\mathbf{J}_G = \begin{bmatrix} \mathbf{J}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} & \mathbf{J}_{\boldsymbol{\alpha}\boldsymbol{\omega}} & \mathbf{0} \\ \mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\alpha}} & \mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\omega}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{J}_{\sigma^2\sigma^2} \end{bmatrix}, \tag{5}$$

where

$$\begin{aligned} \mathbf{J}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} &= \frac{2}{\sigma^2}\Re\left\{\mathbf{S}(\boldsymbol{\omega})^H\mathbf{S}(\boldsymbol{\omega})\right\}, \\ J_{\boldsymbol{\omega}\boldsymbol{\omega}}(\omega_m^k, \omega_{m'}^{k'}) &= \frac{2}{\sigma^2}\Re\left\{\alpha_k\alpha_{k'}\mathbf{s}_k(\boldsymbol{\omega}_k)^H\mathbf{T}^{m+m'}\mathbf{s}_{k'}(\boldsymbol{\omega}_{k'})\right\}, \\ J_{\boldsymbol{\omega}\boldsymbol{\alpha}}(\omega_m^k, \alpha_{k'}) &= \frac{2}{\sigma^2}\Im\left\{\alpha_k\mathbf{s}_k(\boldsymbol{\omega}_k)\mathbf{T}^m\mathbf{s}_{k'}(\boldsymbol{\omega}_{k'})\right\}, \\ \mathbf{J}_{\sigma^2\sigma^2} &= N/\sigma^4, \end{aligned} \tag{6}$$

$\mathbf{T} = \text{diag}(t_0, \ldots, t_{N-1})$, and $\Re$ and $\Im$ denote the real and imaginary operators. The Cramér-Rao bound is given by

$$\mathbf{C}_G = \mathbf{J}_G^{-1}. \tag{7}$$

Denote the maximum likelihood score function

$$\psi_{ML}(x) = -\frac{\partial \ln f(x)}{\partial x}. \tag{8}$$

It has been shown [11, ch. 5] that the Cramér-Rao bound in the case of non-Gaussian noise is

$$\mathbf{C} = \frac{\mathsf{E}\{\psi_{ML}^2(x)\}}{(\mathsf{E}\{\psi_{ML}(x)\})^2}\frac{2}{\sigma^2}\mathbf{C}_G. \tag{9}$$

III. Nonlinear M-Estimation

Since the noise parameters in (3) are unknown, the exact maximum likelihood (ML) solution is not available. We propose to estimate the parameters of the polynomial phase signals via the M-estimation approach

$$\begin{aligned} \hat{\boldsymbol{\vartheta}} &= \arg\min_{\boldsymbol{\vartheta}} \mathcal{C}(\boldsymbol{\vartheta}) \\ &= \arg\min_{\boldsymbol{\vartheta}} \sum_{n=0}^{N-1} \rho(\Re\{z_n\}) + \rho(\Im\{z_n\}) \end{aligned} \tag{10}$$

where $z_n = y_n - \eta_n(\boldsymbol{\vartheta})$. The penalty function is designed to suppress the outliers in the residuals z_n. The least-squares (LS) solution is equivalent to choosing $\rho(x) = x^2$. For robust estimation, the penalty function is required to be less rapidly increasing as the aforementioned quadratic function. In this work we shall use Huber's minimax solution given by

$$\rho(x) = \begin{cases} \frac{x^2}{2\nu^2} & \text{for } |x| \le k\nu^2 \\ k|x| - \frac{k^2\nu^2}{2} & \text{for } |x| > k\nu^2 \end{cases} \tag{11}$$

For details on the minimax score function and its approximation, see [13]. Denote the minimax score function $\psi(x) = \partial\rho(x)/\partial x$, it follows directly from (9) that the theoretical bound on the performance of the proposed robust estimator is given by

$$\mathbf{C} = \frac{\mathsf{E}\{\psi^2(x)\}}{(\mathsf{E}\{\psi(x)\})^2}\frac{2}{\sigma^2}\mathbf{C}_G. \tag{12}$$

The formulation (10) is a difficult optimisation problem since for a general PPS problem it contains numerous local minima. Several sub-optimal techniques are available in the literature such as the product high-order ambiguity function (PHAF) [4] or nonlinear instantaneous least-squares (NILS) [2]. However, they are not robust against impulsive noise and they are not optimal for the multicomponent case.

To solve the nonlinear optimisation problem (10) we propose the following computational procedure

- Step 1: Use the PHAF technique to initialise the estimates (see [4] for details).
- Step 2: Move the parameter estimates to the vicinity of the global minimum via a global optimisation search. In this work, we shall use a recently proposed filled function approach [14].

The motivation of the proposed method is as follows. Step 1 makes use of a computationally attractive technique available in the literature to avoid the exhaustive grid search. However, in the presence of outliers and multicomponents the PHAF estimates can be far away from the global minimum. In Step 2, the filled function is used to search for a better local minimum within the current neighbourhood. Two key features of the technique are the construction of the filled function and the directions for which the search should be carried out.

A. Initialisation

The PHAF method [4] is a generalised version of the HAF method originally introduced by Peleg and Porat [9]. We define the multi-lag high-order instantaneous moment (ml-HIM) of $y(t)$ as follows:

$$\begin{aligned} y_1(t) &= y(t), \\ &\cdots \\ y_M(t; \boldsymbol{\tau}_{M-1}) &= y_{M-1}(t + \tau_{M-1}; \boldsymbol{\tau}_{M-1}) \\ &\quad \times y_{M-1}(t - \tau_{M-1}; \boldsymbol{\tau}_{M-1})^*, \end{aligned} \tag{13}$$

where $\boldsymbol{\tau}_i = [\tau_1, \tau_2, \ldots, \tau_{i-1}]$. The multi-lag HAF is defined as the finite Fourier transform of the ml-HIM

$$Y_m(f; \boldsymbol{\tau}_{M-1}) = \sum_{t=t_0}^{t=t_{N-1}} y_M(t; \boldsymbol{\tau}_{M-1})e^{-j2\pi ft}. \tag{14}$$

Introduce L sets of lags $\mathbf{T}_{M-1}^L = [\boldsymbol{\tau}_{M-1}^{(1)}, \boldsymbol{\tau}_{M-1}^{(2)}, \ldots, \boldsymbol{\tau}_{M-1}^{(L)}]$, where $\boldsymbol{\tau}_{M-1}^{(l)} = [\tau_1^{(l)}, \tau_2^{(l)}, \ldots, \tau_{M-1}^{(l)}]^T$. The PHAF is then defined as:

$$Y_M^L(f; \mathbf{T}_{M-1}^L) = \prod_{l=1}^{L} Y_M(\beta^{(l)}f; \boldsymbol{\tau}_{M-1}^{(l)}), \tag{15}$$

where $\beta^{(l)}$ is the scale factor for the lth set of lags

$$\beta^{(l)} = \frac{\left(\prod_{k=1}^{M-1} \tau_k^{(l)}\right)}{\left(\prod_{k=1}^{M-1} \tau_k^{(1)}\right)}. \tag{16}$$

By searching for the minima of $Y_M^L(f; \mathbf{T}_{M-1}^L)$ in the frequency domain, the highest-order phase parameters $(\omega_M^1, \ldots, \omega_M^K)$ can be found. The parameters of lower orders can be found via demodulation and local search. The details can be found in [11].

B. Global Search

Consider the problem (10) with some initial estimate $\hat{\boldsymbol{\vartheta}}_0$ which is supposed to be a local minimiser. We define the basin $\mathcal{B}_0$ about $\hat{\boldsymbol{\vartheta}}_0$ as the neighbourhood such that the steepest descent trajectory of $\mathcal{C}(\boldsymbol{\vartheta})$ converges to $\hat{\boldsymbol{\vartheta}}_0$ from any point within $\mathcal{B}_0$. To move the current point to a better position, a filled function $p(\boldsymbol{\vartheta}; \hat{\boldsymbol{\vartheta}}_0)$ is to be constructed with the following requirements [14]:

- $\hat{\boldsymbol{\vartheta}}_0$ is a maximiser of $p(\boldsymbol{\vartheta}; \hat{\boldsymbol{\vartheta}}_0)$ and the whole basin $\mathcal{B}_0$ is covered by $p(\boldsymbol{\vartheta}; \hat{\boldsymbol{\vartheta}}_0)$.
- $p(\boldsymbol{\vartheta}; \hat{\boldsymbol{\vartheta}}_0)$ does not have minimisers or saddle points in any basin of $\mathcal{C}(\boldsymbol{\vartheta})$ higher than $\mathcal{B}_0$.
- Suppose that there is a basin $\mathcal{B}_1$ lower than $\mathcal{B}_0$, then there exists a minimiser $\boldsymbol{\vartheta}_1 \in \mathcal{B}_1$ of $\mathcal{C}(\boldsymbol{\vartheta})$ along the line through $\boldsymbol{\vartheta}_0$ to $\boldsymbol{\vartheta}_1$.

The following filled function was proposed in [14]:

$$\begin{aligned} p(\boldsymbol{\vartheta}; \hat{\boldsymbol{\vartheta}}_0) &= \min\left\{\mathcal{C}(\boldsymbol{\vartheta}), \mathcal{C}(\hat{\boldsymbol{\vartheta}}_0)\right\} - \varrho \parallel \boldsymbol{\vartheta} - \hat{\boldsymbol{\vartheta}}_0 \parallel^2 \\ &\quad + \mu\left(\max\left\{0, \mathcal{C}(\hat{\boldsymbol{\vartheta}}_0 - \mathcal{C}(\boldsymbol{\vartheta})\right\}\right)^2 \end{aligned} \tag{17}$$

The parameters ϱ and μ are chosen such that $\varrho > 0$ and $0 \leq \mu < \varrho/\mathcal{L}^2$ where $\mathcal{L}$ is the Lipschitz constant of $\mathcal{C}(\boldsymbol{\vartheta})$. Depending on the conditions of the current point, one carries a search along one of the following directions

$$\mathbf{D}_1 = -\nabla p(\boldsymbol{\vartheta}; \hat{\boldsymbol{\vartheta}}_0) \tag{18}$$

$$\mathbf{D}_2 = -\frac{\nabla \mathcal{C}(\boldsymbol{\vartheta})}{\parallel \nabla \mathcal{C}(\boldsymbol{\vartheta}) \parallel} - \frac{\nabla p(\boldsymbol{\vartheta}; \hat{\boldsymbol{\vartheta}}_0)}{\parallel p(\boldsymbol{\vartheta}; \hat{\boldsymbol{\vartheta}}_0) \parallel}. \tag{19}$$

The algorithm iteratively adjusts ϱ and μ to find a better local minimiser until ρ reaches its minimum value. The algorithm then accepts the current minimiser as the global minimiser. For a more complete description of the steps, see [14, p.30].

To demonstrate the filled function technique, we consider a second-order monocomponent PPS in the absence of noise with the following parameters: $\alpha = 1$, $(\omega_0, \omega_1, \omega_2) = (0.942, 2, -0.025)$. In a nonlinear least-squares (NLS) formulation, minimising the NLS cost function which involves $(\alpha, \omega_0, \omega_1, \omega_2)$ is equivalent to minimising an augmented cost function of only (ω_1, ω_2) [11]. Fig. 1 shows the contour plot of the augmented NLS cost function whose global minimum is at (2, -0.0125) on the (ω_1, ω_2) plane. We arbitrarily initialise the estimates at (1.00,0.08) and perform the global search. As can be seen, the cost function contains numerous local minima and the radius of convergence at the global minimum is very small. The simulation example indicates that the method first found a local minimum, then in the second move it reached the boundary. Finally, the region containing the global minimum was successfully identified.

As with many global optimisation techniques, the method does not always guarantee the true global minimum to be found. This depends on a number of factors such as the radius of convergence about the global minimum, its distance to the initialisation point, and the preset threshold values of the algorithm's parameters. Besides, there may be computational time contraint whereby one must terminate the search when a maximum number of iterations has been reached. However, this method at least improves over conventional gradient search techniques in that the solution found is always at an equal or lower basin. Extensive numerical studies in the context of the polynomial phase signal problem also confirm this.

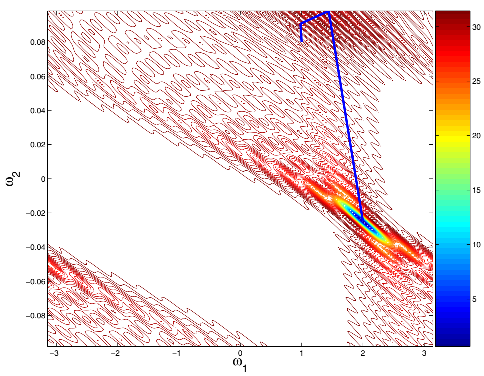

Fig. 1. Illustration of the filled function method.

IV. Simulation Example

In this section we illustrate the proposed nonlinear M-estimation approach with a simulation example. Consider a two-component second-order PPS with amplitudes $\alpha_1 = \alpha_2 = 1$, and frequency parameters $(\omega_0^1, \omega_1^1, \omega_2^1) = (0, 0.20\pi, 0.22\pi)$ and $(\omega_0^2, \omega_1^2, \omega_2^2) = (0, 0.80\pi, -0.31\pi)$. The sampling period is $\Delta = 1$ and the total number of observations is $N = 64$. This choice of parameters satisfies the Nyquist-like criteria for polynomial phase signals [11]. The Wigner-Ville distribution plot which indicates the time-frequency behaviour of individual components is given in Fig. 2. To model impulsive noise, we select $\varepsilon = 0.1$ and $\kappa = 100$. For comparison, we also include the following methods:

- The NLS method which was originally formulated for Gaussian noise. It consists of PHAF initialisation and a local search via the simplex method over the NLS cost function. It is known to improve over the original PHAF method, especially for the multicomponent case [11] .

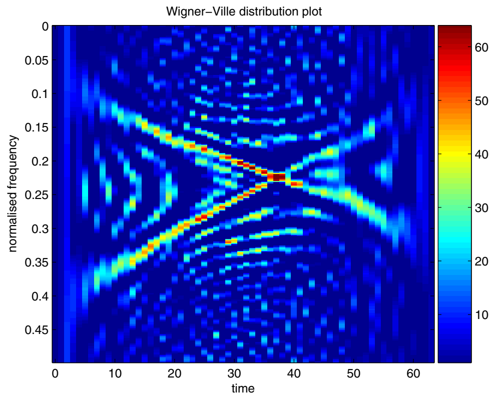

Fig. 2. Wigner-Ville distribution plot of the studied mc-PPS signal.

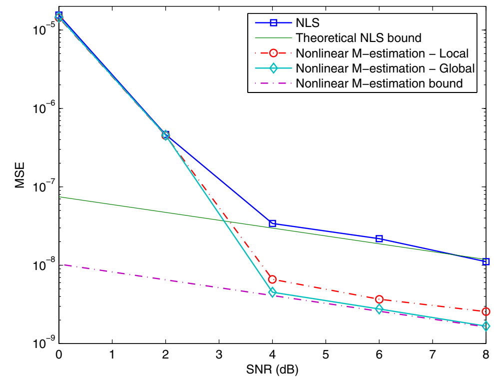

Fig. 3. Example of a 2nd-order mc-PPS in impulsive noise with $\varepsilon = 0.1$, $\kappa = 100$.

- The nonlinear M-estimator with PHAF initialisation and a local search over the Huber's cost function (11) via the simplex method.

The main purpose is to see the performance gain when using the global search technique. The mean-square error (MSE) performance is shown in Fig. 3. We made the following observations:

- All methods suffer from a threshold effect, i.e. when the signal-to-noise ratio (SNR) is low, the performance is poorer. This can be explained by the dependence of the methods on PHAF initialisation, which is poor at low SNRs. It appears that with the original settings the global optimisation method does not find the global minimum successfully at low SNRs.
- When the SNR is beyond the threshold, which is 3dB in this case, the proposed robust method clearly outperforms the NLS as expected. It also approaches the theoretical performance bound. Compared with the result on the local search, the global search helps further reduce the MSE. As the MSE of the proposed method approaches the theoretical bound, the figure indicates that the global minimum is more likely to be found with the fill function technique.

V. Conclusion

We have presented a nonlinear M-estimation approach to the robust estimation of multicomponent PPS's in the presence of impulsive noise modelled by a Gaussian mixture. The nonlinear optimisation problem is solved in two steps. Being different to other techniques, we employed a global optimisation search to find the estimates rather than relying only on gradient search. This makes the approach more robust against large deviation of the initial estimates about the global minimum. Simulation results show that our proposed approach offers considerable performance gain when compared to conventional techniques.

References

[1] T. J. Abatzoglou, "Fast Maximum Likelihood Joint Estimation of Frequency and Frequency Rate," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, no. 6, pp. 708–715, 1986.

[2] J. Angeby, "Estimating Signal Parameters Using the Nonlinear Instantaneous Least Squares Approach," *IEEE Transactions on Signal Processing*, vol. 48, no. 10, pp. 2721–2732, 2000.

[3] A. Banerjee, P. Burlina, and R. Chellappa, "Adaptive Target Detection in Foliage-Penetrating SAR Images Using Alpha-Stable Models," *IEEE Transactions on Image Processing*, vol. 8, no. 12, pp. 1823–31, Dec 1999.

[4] S. Barbarossa, A. Scaglione, and G. B. Giannakis, "Product High-Order Ambiguity Function for Multicomponent Polynomial Phase Signal Modeling," *IEEE Transactions on Signal Processing*, vol. 46, no. 3, pp. 691–708, Mar 1998.

[5] K. L. Blackard, T. S. Rappaport, and C. W. Bostian, "Measurements and Models of Radio Frequency Impulsive Noise for Indoor Wireless Communications," *IEEE Journal on Selected Areas in Communications*, vol. 11, no. 7, pp. 991–1001, 1993.

[6] V. Degardin, "Classification and Characterization of Impulsive Noise on Indoor Power Line Used for Data Communications," *IEEE Transactions on Consumer Electronics*, vol. 48, no. 4, pp. 913–918, Nov 2002.

[7] H. Hsu, R. Storwick, D. Schlick, and G. Maxam, "Measured Amplitude Distribution of Automotive Ignition Noise," *IEEE Transactions on Electromagnetic Compatibility*, vol. EMC-16, no. 2, pp. 57–63, May 1974.

[8] S. Peleg and B. Friedlander, "The Discrete Polynomial-Phase Transform," *IEEE Transactions on Signal Processing*, vol. 43, no. 8, pp. 1901–1914, Aug 1995.

[9] S. Peleg and B. Porat, "Estimation and Classification of Polynomial Phase Signals," *IEEE Transactions on Information Theory*, vol. 37, no. 2, pp. 422–430, Mar 1991.

[10] ——, "The Cramér-Rao Lower Bound for Signals with Constant Amplitude and Polynomial Phase," *IEEE Transactions on Signal Processing*, vol. 39, no. 3, pp. 749–752, Mar 1991.

[11] D. S. Pham, "Robust Estimation of Signals in Heavy Tailed Noise," Ph.D. dissertation, Curtin University of Technology, 2005.

[12] R. A. Shepherd, "Measurements of Amplitudes Probability Distributions and Power of Automobile Ignition Noise at HF," *IEEE Transactions on Vehicular Technology*, vol. 23, pp. 72–83, Aug 1974.

[13] X. Wang and H. V. Poor, "Robust Multiuser Detection in Non-Gaussian Channels," *IEEE Transactions on Signal Processing*, vol. 47, no. 2, pp. 289–305, Feb 1999.

[14] L.-S. Zhang, C.-K. Ng, D. Li, and W.-W. Tian, "A New Filled Function Method for Global Optimisation," *Journal of Global Optimisation*, vol. 28, pp. 17–43, 2004.