

# 大气低频噪声混合模型的 MCMC 参数估计

应文威, 蒋宇中, 刘月亮

(海军工程大学电子工程学院, 湖北 武汉 430033)

**摘要:** 大气噪声是低频通信中的主要干扰, 且具有严重非高斯分布特性, 对非高斯噪声模型的参数估计对于提高低频接收机的性能具有重要意义。设计了估计非高斯混合模型参数的马尔可夫链蒙特卡罗 (Markov chain Monte Carlo, MCMC) 算法, 该算法通过构建贝叶斯层次模型, 利用 Gibbs 抽样和 M-H 抽样更新迭代参数。利用乘积特性, 将稳定分布作为等价的高斯分布来处理, 并在层次模型中设置多个额外参数, 以增强其灵活性。仿真实验与实测数据表明, 该算法迭代收敛快、精度高, 有很高的实用价值。

**关键词:** 混合模型; 马尔可夫链蒙特卡罗; 非高斯噪声;  $\alpha$  稳定分布

**中图分类号:** TN 911.23

**文献标志码:** A

**DOI:** 10.3969/j.issn.1001-506X.2012.06.29

## Parameter estimation for mixture model of atmospheric noise through MCMC method

YING Wen-wei, JIANG Yu-zhong, LIU Yue-liang

(College of Electronic Engineering, Naval University of Engineering, Wuhan 430033, China)

**Abstract:** Atmospheric noise is the main interference in a low-frequency communication system, which is highly impulsive. So the work for estimating the parameters of model of non-Gaussian noises is of great significance to improve the performance of the low-frequency receiver. This paper proposes a Markov chain Monte Carlo (MCMC) method to estimate the parameters of a mixture model. The method updates the parameters through a Gibbs sampler and M-H algorithm, which are based on the Bayesian hierarchical model. The  $\alpha$  stable distribution in the mixture model is equivalent to the normal distribution by using the product properties. An extra layer is added to the hierarchy for full flexibility. The result shows that the new method has a good performance, high precision and can be excellently applied in practice.

**Keywords:** mixture model; Markov chain Monte Carlo (MCMC); non-Gaussian noise;  $\alpha$  stable distribution

## 0 引言

低频通信系统中的最优接收机一直以来都是基于高斯噪声的假设, 然而这类接收机在实际应用, 特别是在低频段通信系统中, 往往达不到预期的性能。原因在于低频段受雷电等影响, 其噪声特性往往具有非高斯特征<sup>[1-3]</sup>, 概率密度函数在尾部较高斯分布收敛更慢, 具有重尾的特点。为了使接收机能够进行最优接收, 则必须对噪声模型进行参数估计。非高斯分布和高斯分布的混合模型对自然界的非高斯噪声有很强的代表性, 有极为广泛的应用<sup>[4-7]</sup>。本文研究的是对称  $\alpha$  稳定 (symmetric  $\alpha$  stable, SaS) 分布和高斯分布的混合模型<sup>[8]</sup>。该模型涵盖了在非高斯模型应用广泛的  $\alpha$  稳定分布和混合高斯分布<sup>[9]</sup>, 因而受到学者的广泛关

注。对该混合模型的参数估计, 传统上采用最大期望 (expectation-maximization, EM) 算法或最大似然估计<sup>[9]</sup>。但由于该混合模型参数较多, 且对称  $\alpha$  稳定分布的概率密度没有闭合表达式, 计算难度大, 精度也不够理想。

马尔可夫链蒙特卡罗 (Markov chain Monte Carlo, MCMC) 算法, 是一类在统计计算中有优异性能的方法。本文通过设计 MCMC 多层次算法来估计混合模型参数, 迭代收敛快、精度高、具有明显优势。

## 1 $\alpha$ 稳定分布

### 1.1 定义

如果一个随机变量服从稳定分布, 则其特征函数满足:

$$\varphi(t) = \exp \{j\delta t - | \gamma t |^\alpha [1 + j\beta \text{sgn}(t) \omega(t, \alpha)]\} \quad (1)$$

$$w(t, \alpha) = \begin{cases} -\tan(\frac{\alpha\pi}{2}), & \alpha \neq 1 \\ \frac{2}{\pi} \ln |t|, & \alpha = 1 \end{cases} \quad (2)$$

易看出,  $\alpha$  稳定分布由 4 个参量决定<sup>[6]</sup>:  $\alpha$  是特征指数,  $\alpha$  的值越小, 远离中心值的出现概率越大;  $\beta$  是对称参数, 当  $\beta=0$  变成对称  $\alpha$  稳定分布;  $\gamma$  是尺度参数, 代表了杂散程度, 同能量密切相关;  $\delta$  是位移参数。对于 SaS 来说, 当  $1 < \alpha \leq 2$  时, 代表均值。通过傅里叶逆变换有

$$f(x; \alpha, \beta, \delta, \gamma) = F^{-1}(\phi(t)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-j\omega x} \phi(t) dt \quad (3)$$

然而对于式(3), 除了当  $\alpha$  取少数特定值的时候,  $f(x)$  没有闭合形式的表达式。这也造成了  $\alpha$  稳定分布参数估计上的难度。

### 1.2 乘法性质

$\alpha$  稳定分布有一个很有意思的性质, 下面不加证明引入该性质<sup>[10]</sup>:

**定理 1** 若有两个独立随机变量  $X, Y > 0$ , 若  $X \sim f(x; \alpha_1, 0, 0, \gamma), Y \sim f(x; \alpha_2, 1, 0, (\cos(\pi\alpha_2/2))^{\frac{2}{\alpha_2}})$ 。

则  $XY^{\frac{1}{\alpha_2}}$  仍服从  $\alpha$  稳定分布, 且满足  $XY^{\frac{1}{\alpha_2}} \sim f(x; \alpha_1\alpha_2, 0, 0, \gamma)$ 。

由于  $\alpha$  稳定分布具有 SMiN(scale mixture of normal) 的性质<sup>[11-12]</sup>, 则  $\alpha$  稳定分布可以表示为  $X = Y\sigma, Y$  为正态分布,  $\sigma$  为正的独立随机变量。又因为当  $\alpha=2$  时,  $\alpha$  稳定分布即为高斯分布。因此综合上述有: 如果  $x_k$  是对称  $\alpha$  稳定分布的独立同分布抽样, 满足  $x_k \sim f(\alpha, 0, 0, \gamma)$ , 则可得下列的等价式子:

$$x_k \sim N(0, \lambda_k \gamma^2) \quad (4)$$

$$\lambda_k \sim f(x; \alpha_2/2, 1, 0, 2(\cos(\pi\alpha_2/4))^{\frac{2}{\alpha_2}}) \quad (5)$$

## 2 混合模型

### 2.1 混合模型

对称  $\alpha$  稳定分布和高斯分布的混合模型如下:

$$p(x) = w_1 f(x; \alpha, 0, 0, \gamma) + w_2 g(x; 0, \sigma^2) \quad (6)$$

式中,  $f(x; \alpha, 0, 0, \gamma)$  为非高斯 SaS 部分;  $g(x; 0, \sigma^2)$  为高斯部分;  $w_1$  和  $w_2$  为权重因子。显然地, 满足  $w_1 + w_2 = 1$ 。因此, 需要估计的参数集为  $\{w, \alpha, \gamma, \sigma^2\}$ 。

假设观测到的数据集为  $\mathbf{X} = \{x_k | 1 \leq k \leq N\}$ , 引入标签变量  $\mathbf{Z} = \{z_k | 1 \leq k \leq N\}$ , 其中  $z_k$  的取值为:

(1) 当  $x_k$  属于对称  $\alpha$  稳定分布部分的时候,  $z_k$  的取值为 1;

(2) 当  $x_k$  属于高斯分布部分的时候,  $z_k$  的取值为 2。

因此,  $z_k$  可以看成是独立变量, 满足:

$$P(z_k = j) = w_j, \quad j = 1, 2 \quad (7)$$

### 2.2 贝叶斯层次模型和先验

贝叶斯推断通过先验分布来推断后验分布

$$P(A | B) \propto P(B | A)P(A) \quad (8)$$

根据贝叶斯层次理论, 参变量集  $\{w, \alpha, \gamma, \sigma^2\}$  的联合概率密度为

$$p(\alpha, \gamma, \sigma^2, w, z, x) = p(x | \alpha, \gamma, \sigma^2, w, z) \times p(\alpha, \gamma, \sigma^2, w, z) \quad (9)$$

选取合理的共轭先验能够充分利用先验信息, 提高贝叶斯模型的性能。选取以下先验:

由于高斯分布的参数具有共轭先验, 因此对于式(4)中高斯部分的参数  $\sigma^2$  有

$$\sigma^{-2} \sim \Gamma(\alpha_0, \beta) \quad (10)$$

式中,  $\Gamma(\cdot)$  为 Gamma 分布;  $\alpha_0, \beta$  分别为形状参数和反尺度参数。

利用第 1.2 节中的结论, 则对称  $\alpha$  稳定分布参数  $\gamma^2$  有

$$\gamma^{-2} \sim \Gamma(\epsilon, \kappa) \quad (11)$$

对于指数参数  $\alpha$ , 选取先验使得  $\alpha$  的值在  $(0, 2]$  上均匀分布, 为

$$P(\alpha | a) = \frac{1}{a} = \frac{1}{2}, \quad \alpha \in (0, 2] \quad (12)$$

参量  $w$  的共轭先验选取为 Dirichlet 分布:

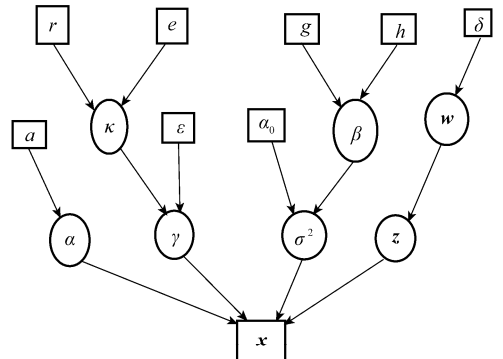
$$w \sim D(\delta, \delta) \quad (13)$$

使超参数  $\beta, \kappa$  具有更强的灵活性, 多增加一个额外层。它们的共轭先验为

$$\beta \sim \Gamma(g, h), \kappa \sim \Gamma(r, e) \quad (14)$$

混合模型的非循环图 (directed acyclic graph, DAG) 如图 1 所示。为表述简单, 设  $\theta = \{\alpha, \gamma, \sigma^2\}$  和  $\theta$  的超参数  $\eta = \{a, \kappa, \epsilon, \alpha_0, \beta, r, e, g, h\}$ , 则式(8)可扩展为

$$P(w, \theta, \eta, z, \delta, x) = P(\eta)P(\delta)P(x | \theta, z) \times P(z | w)P(w | \delta)P(\theta | \eta) \quad (15)$$



○: 参量; □: 定值。

图 1 混合模型的 DAG

### 3 MCMC 算法

在本节将设计具体的 MCMC 算法来估计混合模型的参数。在 MCMC 算法中, Gibbs 抽样和 Metropolis-Hastings(M-H)算法是两种常用的计算方法。一般来说,若是后验概率较易计算则常采用 Gibbs,不然采用 M-H 较多。对于本文,采用 Gibbs 和 M-H 的混合算法,算法流程如下:

- 步骤 1 通过 Gibbs 抽样更新权重系数  $w$ ;
- 步骤 2 通过 Gibbs 抽样更新参数  $\gamma$ ;
- 步骤 3 通过 Gibbs 抽样更新参数  $\sigma^2$ ;
- 步骤 4 通过 M-H 算法更新  $\alpha$ ;
- 步骤 5 通过 Gibbs 抽样更新参数  $(\kappa, \beta)$ ;
- 步骤 6 更新标签变量  $z$ ;
- 步骤 7 通过 M-H 算法更新变量  $\lambda$ ;

#### 3.1 通过 Gibbs 抽样更新权重系数 $w$

直接计算  $w$  后验概率,易得  $w$  的全条件后验分布仍服从 Dirichlet 分布,即

$$w | \dots \sim D(\delta + n_1, \delta + n_2) \quad (16)$$

式中,  $w | \dots$  表示在除  $w$  外的所有变量的条件下;  $n_1$  代表属于对称  $\alpha$  稳定分布部分的观测数据数目;  $n_2$  代表属于高斯部分的观测数据数目。通过生成 Dirichlet 分布随机数更新  $w$  的值。

#### 3.2 通过 Gibbs 抽样更新参数 $\gamma$

通过参数  $\gamma^{-2}$  的全条件分布更新参数  $\gamma$  的值,易算得  $\gamma^{-2}$  的全条件分布为

$$\gamma^{-2} | \dots \sim \Gamma\left(\epsilon + 0.5n_1, 0.5 \sum_{i=1; z_i=1}^N \frac{x_i^2}{\lambda_i} + \kappa\right) \quad (17)$$

#### 3.3 通过 Gibbs 抽样更新参数 $\sigma^2$

通过参数  $\sigma^{-2}$  的全条件分布更新参数  $\sigma^2$  的值,  $\sigma^{-2}$  的全条件分布为

$$\sigma^{-2} | \dots \sim \Gamma\left(\alpha_0 + 0.5n_2, 0.5 \sum_{i=1; z_i=2}^N x_i^2 + \beta\right) \quad (18)$$

#### 3.4 通过 M-H 算法更新 $\alpha$

对于参数  $\alpha$ , 由于其全条件概率分布不是常用的概率分布,因此采用 M-H 算法更新  $\alpha$  的值。

(1) 从第  $t$  次迭代时,从建议分布  $q(\alpha^{t+1} | \alpha^t)$  产生一个新值  $\alpha^{t+1}$ 。在本文中,选用正态分布。即  $\alpha^{t+1} \sim N(\alpha^t, 0.3)$ 。

(2) 以接受概率  $\min(1, A)$  接收新值  $\alpha^{t+1}$ , 其中  $A$  为

$$A = \frac{\prod_{i=1; z_i=1}^N P(x_i | w_{z_i}, \gamma, \sigma, \alpha^{t+1}) \times P(w_{z_i}, \gamma, \sigma, \alpha^{t+1}) q(\alpha^t | \alpha^{t+1})}{\prod_{i=1; z_i=1}^N P(x_i | w_{z_i}, \gamma, \sigma, \alpha^t) \times P(w_{z_i}, \gamma, \sigma, \alpha^t) q(\alpha^{t+1} | \alpha^t)} \quad (19)$$

由于  $\alpha$  是独立的且先验服从均匀分布,建议分布是对称的,于是式(19)可化为

$$A = \frac{\prod_{i=1; z_i=1}^N P(x_i | w_{z_i}, \gamma, \sigma, \alpha^{t+1})}{\prod_{i=1; z_i=1}^N P(x_i | w_{z_i}, \gamma, \sigma, \alpha^t)} \quad (20)$$

若不接收新值,则  $\alpha^{t+1} = \alpha^t$ 。

#### 3.5 通过 Gibbs 抽样更新参数 $\kappa, \beta$

分别对  $\kappa, \beta$  的全条件分布进行 Gibbs 采样来更新参数的值,它们的全条件分布分别为

$$\kappa | \dots \sim \Gamma(r + \alpha, e + \gamma^{-2}) \quad (21)$$

$$\beta | \dots \sim \Gamma(g + \alpha, h + \sigma^{-2}) \quad (22)$$

#### 3.6 更新标签变量 $z$

变量  $z_i$  指示了观测数据  $x_i$  所属的概率分布,对于其全条件分布满足

$$P(z_i = 1 | \dots) = \frac{w_1 f(x_i | \alpha, \gamma)}{w_1 f(x_i | \alpha, \gamma) + w_2 g(x_i | \sigma)} \quad (23)$$

$$P(z_i = 2 | \dots) = \frac{w_2 f(x_i | \sigma)}{w_1 f(x_i | \alpha, \gamma) + w_2 g(x_i | \sigma)} \quad (24)$$

通过生成  $U(0, 1)$  的均匀随机数,确定其所在区间来判断  $z_i$  的取值。

#### 3.7 M-H 算法更新变量 $\lambda$

由于变量  $\lambda_i$  的引入,使得对称  $\alpha$  稳定分布可以写成  $\lambda_i$  条件下的高斯分布。由于观测数据只有一部分属于对称  $\alpha$  分布,因此仅更新那一部分对应的值。易得变量  $\lambda_i$  的全条件分布为

$$P(\lambda_i | \alpha, x_i, \gamma, z_i = 1) \propto N(x_i | 0, \lambda_i \gamma^2) \times f(\lambda_i; \alpha/2, 1, 0, 2(\cos(\pi\alpha/4))^{2/\alpha}) \quad (25)$$

采用 M-H 算法更新  $\lambda_i$  的值,选用建议分布为

$$\lambda_i^{t+1} \sim f(\lambda; \alpha/2, 1, 0, 2(\cos(\pi\alpha/4))^{2/\alpha}) \quad (26)$$

则易得接受概率为

$$R = \min \left[ 1, \frac{N(x_i | 0, \lambda_i^{t+1} \gamma^2)}{N(x_i | 0, \lambda_i^t \gamma^2)} \right] \quad (27)$$

生成  $U(0, 1)$  的均匀随机数,若小于接受概率则接受新值,否则不变。

### 4 仿真及实测结果

在仿真实验中,按照预设的权重生成混合模型。其中  $\alpha$  稳定分布随机值生成采用常用的 Chambers-Mallows-Stuck 算法。对于预设参数,设置为  $r=0.2, e=1, g=0.2, h=1, \epsilon=2, \alpha_0=2$ 。设置预烧期为  $N_0=200$ ,总的迭代次数为  $N=1000$ 。在实验中,将混合模型参数设为设置为  $\alpha=1.5, w=(0.7, 0.3), \gamma=7, \sigma^2=4$ 。仿真数据数目设为 2000。仿真结果如图 2~图 5 所示。其中,  $w_2$  的迭代收敛情况由于限于篇幅并没有给出。但因为  $w_2 = 1 - w_1$ , 因此不影响读者理解。将各个参量在预烧期之后的迭代值作平均,得到结果为  $\alpha=1.52, w_1=0.71, \gamma=7.1, \sigma^2=3.9$ 。

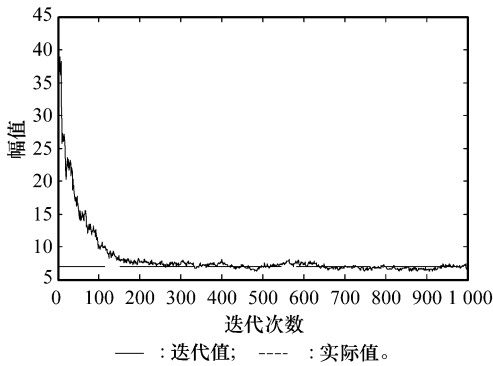


图 2  $\gamma$  的迭代收敛情况

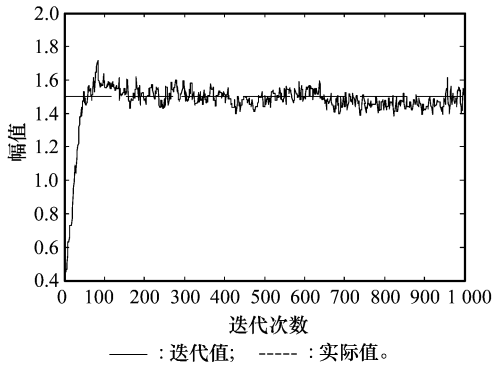


图 3  $\alpha$  的迭代收敛情况

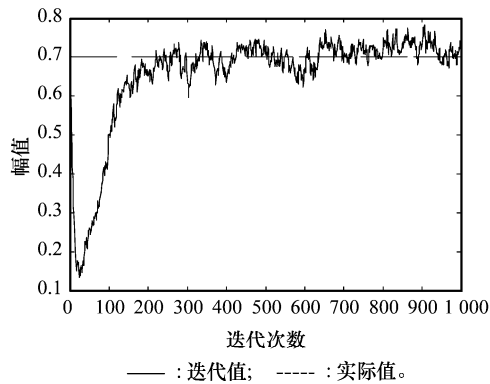


图 4  $w_1$  的迭代收敛情况

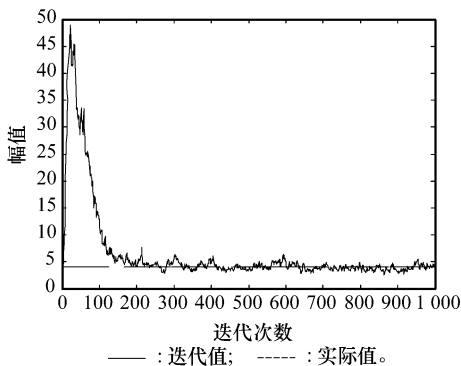


图 5  $\sigma^2$  的迭代收敛情况

在图 2~图 5 中,容易发现在迭代 200 次后,各个参量都能趋于收敛,总体上收敛速度快,结果精度较高。在实验中, $\alpha$  的更新过程需要计算  $\alpha$  稳定分布的概率密度值,以往一般采用 Nolan 提出的数值积分进行计算,但耗时长,计算过大。在本文中,通过对特征函数做快速傅里叶变换,求特定序列的概率密度值,再插值求各个点的概率密度值,这样可以大幅降低时间消耗<sup>[13]</sup>。

图 6 为实测大气噪声数据,地点为河南某地,采样频率为 40 960 Hz,带宽 80 Hz,测量仪器为 GSM-06,该仪器为地球物理电磁探测领域事实上的标准。选择该地的主要因为该地区属前寒武纪玄武岩地区,电导率低,绝缘性好,对大地电磁脉冲的影响较弱,易获得质量较好的信号。特别地,选择这样的场地,对保留脉冲的细节部分特别有利。从图 6 为可以看到明显的脉冲干扰。图 7 为实际值及估计的幅度概率分布(amplitude probability distribution, APD)图(X 轴为  $-\lg(-\lg P(|x| > x_0))$ ),Y 轴单位为 dB,其中,圆圈代表实际数据,曲线代表通过本文方法估计出的结果。预测的结果为  $\alpha=1.5, w=(0.36, 0.64), \gamma=8.49, \sigma^2=30.95$ 。从图 7 中可以看出,MCMC 方法确实能够有效估计模型参数,具有较高精度,同时也再次验证了对称  $\alpha$  稳定分布和高斯分布混合模型对大气低频噪声的适用性。

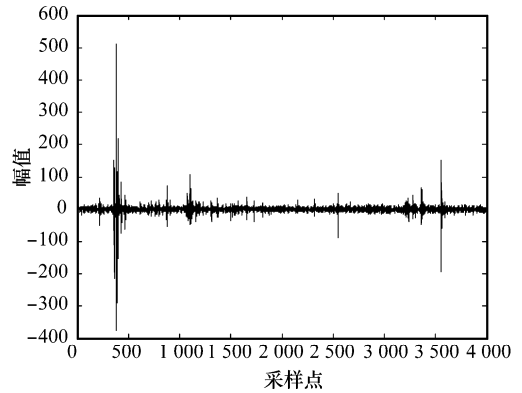


图 6 大气低频噪声

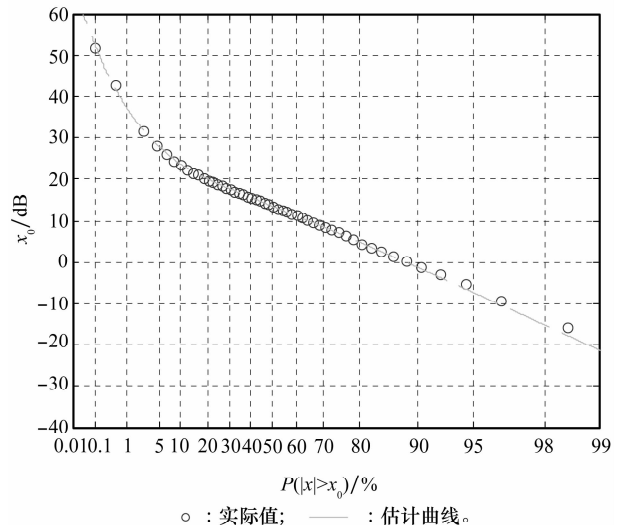


图 7 实际值及估计的 APD 图

## 5 结 论

本文设计了基于 MCMC 算法的大气低频噪声混合模型参数估计方法。通过利用  $\alpha$  稳定分布的乘积性质, 将混合模型中的  $\alpha$  稳定分布等价于高斯分布, 便于 MCMC 算法的更新迭代。设计了贝叶斯层次模型, 充分利用各个参数的先验, 以 Gibbs 和 M-H 混合抽样更新参数。同时在层次模型中引入了多个超参数, 增强了灵活性。在算法中, 利用 FFT 算法计算  $\alpha$  稳定分布的概率密度, 加速了算法运行的速度。仿真实验和实测数据表明, 该算法迭代收敛快, 精度高, 有很高的实用价值, 对日后设计高性能的非高斯低频接收机起了重要的铺垫作用。

## 参考文献:

- [1] Abraham D A. Detection-threshold approximation for non-Gaussian backgrounds [J]. *IEEE Journal of Oceanic Engineering*, 2010, 35(8): 335 - 341.
- [2] Goken C, Gezici S, Arıkan O. Optimal signaling and detector design for power-constrained binary communications systems over non-Gaussian channels [J]. *IEEE Trans. on Communications Letters*, 2010, 14(2): 100 - 111.
- [3] Middleton D. Non-Gaussian noise models in signal processing for telecommunications: new methods and results for class A and class B noise models [J]. *IEEE Trans. on Information Theory*, 1999, 45(4): 1129 - 1149.
- [4] Stople D, Zamir R. Capacity and error probability in single-tone and multitone multiple access over an impulsive channel [J]. *IEEE Trans. on Communication*, 2011, 49(3): 506 - 517.
- [5] Schulte S. A fuzzy impulse noise detection and reduction method [J]. *IEEE Trans. on Image Processing*, 2006, 15(5): 1153 - 1162.
- [6] Lind L, Mufti N. Efficient method for modeling impulse noise in a communication system [J]. *Electronics Letters*, 1996, 32(3): 1440 - 1441.
- [7] Ekrem E, Koca M. Robust ultra-wideband signal acquisition [J]. *IEEE Trans. on Wireless Communications*, 2008, 7(11): 4656 - 4669.
- [8] Nikias C L, Shao M. *Signal processing with alpha-stable distribution and applications* [M]. New York: Wiley, 1995.
- [9] Wang X D, Poor H V. Robust multi-user detection in non-Gaussian channels [J]. *IEEE Trans. on Signal Processing*, 1999, 47(2): 289 - 305.
- [10] Samorodnitsky G, Taqqu M. *Stable non-Gaussian random process: stochastic models with infinite variance* [M]. New York: Chapman & Hall, 1994.
- [11] Godsill S, Kuruoglu E E. Bayesian inference for time series with heavy-tailed symmetric alpha stable noise process [C] // *Proc. of the Applications of Heavy Tailed Distributions in Economics, Engineering and Statistics*, 1999: 101 - 128.
- [12] Salas-Gonzales D, Kuruoglu E E. Modeling with mixture of symmetric stable distributions using Gibbs sampling [J]. *Signal Processing*, 2010, 17(2): 774 - 783.
- [13] Menn C, Rachev S T. Calibrated FFT-based density approximations for stable distribution [J]. *Computational Statistics and Data Analysis*, 2006, 50(8): 1891 - 1904.