

Détection sous optimale en contexte α -stable : performances et algorithmes

Steeve ZOZOR, Jean-Marc BROSSIER, Pierre-Olivier AMBLARD et Cédric DUCHÊNE

Laboratoire des Images et des signaux, CNRS UMR 5083
961 Rue de la Houille Blanche, B.P. 46, 38402 Grenoble Cedex, France

(steeve.zozor), (jean-marc.brossier), (bidou.amblard), (cedric.duchene)@lis.inpg.fr

Résumé – Nous nous intéressons à la détection d’un signal constant dans un bruit α -stable symétrique à échantillons indépendants et identiquement distribué. Pour remédier au fait que le récepteur optimal ne peut pas être mis en œuvre du fait de l’absence de forme explicite de la densité de probabilité, nous proposons une forme sous optimale paramétrée qui inclut le détecteur optimal Gaussien et le détecteur localement optimal adapté à un bruit de Cauchy. Après avoir optimisé les paramètres de ce détecteur, optimisation facilitée par le choix du récepteur, nous comparons ses performances à celles des détecteurs plus classiques que sont le filtre adapté, le détecteur de Cauchy et le détecteur proposé par Brown : le récepteur proposé montre une plus grande robustesse à l’indice de stabilité du bruit, ainsi qu’à l’amplitude du signal.

Abstract – This paper focus on the detection problem of a constant signal corrupted by additive symmetrical α -stable noise of samples independent identically distributed. The optimal detector cannot be implemented due to the non existence of a closed form for the probability density function of the noise. Hence, we propose the use of a suboptimal parametric detector which includes the matched filter, optimal in the Gaussian case, and the locally optimal detector of a Cauchy noise. We first optimize the parameters of the proposed detector. This optimization is easy due to the particular choice of the detector. Then, we compare the performance of the detector with classical detectors such as the matched filter, the Cauchy receiver and the receiver proposed by Brown: we show that the detector exhibits a higher robustness to the stability index and to the signal amplitude.

1 Introduction

Les lois α -stables constituent un modèle classique de bruits impulsifs tels qu’il en existe dans beaucoup de phénomènes naturels (par exemple en turbulence) et artificiels (par exemple dans des données financières, trafic internet). Une variable α -stable symétrique (notée $S\alpha S$) se définit via sa fonction caractéristique : $\Phi(u) = e^{-|\sigma u|^\alpha}$, où $\alpha \in]0; 2]$ est l’indice de stabilité de la loi et $\sigma > 0$ son paramètre d’échelle¹. La densité de probabilité (d.d.p.) – transformée de Fourier de Φ – de ces lois n’admet de forme explicite que dans les cas $\alpha = 2$ (Gauss) et $\alpha = 1$ (Cauchy)².

Le problème de détection considéré ici consiste à choisir entre l’hypothèse bruit seul H_0 et l’hypothèse H_1 correspondant à l’observation d’un signal constant c bruité. L’observation se compose de T échantillons x_t :

$$\begin{cases} H_0 & : & x_t & = & n_t \\ H_1 & : & x_t & = & c + n_t \end{cases} \quad t = 0, \dots, T-1 \quad (1)$$

Les n_t sont des variables $S\alpha S$ indépendantes, identiquement distribuées (i.i.d.) de d.d.p. (resp. fonction caractéristique) f_n (resp. Φ_n) et c est le signal connu dont on cherche à détecter la présence. σ est également supposé connu et nous choisissons $\sigma = 1$ sans perte de généralité. Les deux hypothèses sont supposées équiprobables.

Sous ces hypothèses, la solution optimale (probabilité d’erreur minimale (Pe min), Neyman-Pearson (probabilité de

¹La fonction caractéristique dépend en général d’un paramètre de centralité et d’un facteur de dissymétrie [1], tout deux nuls ici

²Pour les lois asymétriques, il existe aussi la loi de Lévy (ou Pearson) obtenue pour $\alpha = 0.5$ [1]

détection maximale à probabilité de fausse alarme fixée), approche Bayésienne ou maximum *a posteriori*) s’écrit à l’aide du test de rapport de vraisemblance [2] :

$$\Lambda_{\text{opt}} = \sum_{t=0}^{T-1} (\log(f_n(x_t - c)) - \log(f_n(x_t))) \geq \eta$$

Si Λ_{opt} dépasse le seuil η , on décide H_1 sinon H_0 . Pour mettre en pratique ce test, il faut connaître c ainsi que la d.d.p. f_n . Dans le cas α -stable, seule une méthode d’approximation numérique coûteuse peut être envisagée. La question pratique est alors celle de trouver des récepteurs sous-optimaux de mise en œuvre économique aux performances aussi proches que possible de l’*optimum*. Une solution couramment envisagée est celle du récepteur localement optimal (LO) [2] performant pour un signal “faible”, ce qui donne dans le cadre de (1) :

$$\Lambda_{\text{lo}} = \sum_{t=0}^{T-1} \frac{-f'_n}{f_n}(x_t) \geq \eta$$

Ce détecteur fait apparaître la fonction score $-f'_n/f_n$ (f'_n est la dérivée de f_n). Dans le cas α -stable, le récepteur LO est aussi complexe que l’*optimum* lui-même. Une solution est d’explorer d’autres formes de récepteur sous optimaux.

2 Famille de récepteurs sous-optimaux

La solution que nous proposons utilise un récepteur sous optimal de la forme :

$$\Lambda_g = \sum_{t=0}^{T-1} g(x_t) \quad \text{avec} \quad g(x) = \frac{\lambda x}{1 + \beta x^2} \quad (2)$$

g est une fonction dérivable, paramétrée, proche du récepteur LO, qui capture deux caractéristiques essentielles d'une loi α -stable : $g_{\text{lo}}(x) = \frac{-f'_n}{f_n}(x)$ est linéaire autour de 0 et $g_{\text{lo}}(x) \sim \frac{\alpha+1}{x}$ en $x \rightarrow \pm \infty$.

La structure du détecteur étant fixée, l'objectif est alors d'optimiser les paramètres λ et β afin de maximiser les performances pour ensuite rapporter les résultats à ceux que produit le récepteur optimal. Contrairement aux récepteurs proposés dans [3, 4, 5, 6], la forme (2) englobe le filtre adapté ($\beta = 0$) et le récepteur LO de Cauchy ($\beta = 1$).

3 Optimisation

Le type de critère envisagé est soit géométrique (ajustement du maximum de g sur celui de g_{LO} , distance euclidienne minimale entre g et g_{LO}), soit statistique (Erreur Quadratique Moyenne (EQM) $E[(g_{\text{LO}}(X) - g(X))^2]$ minimale), soit un critère de détection (efficacité, probabilité d'erreur minimum (Pe min)).

Nous optimisons maintenant les paramètres pour ces critères pour ensuite analyser l'évolution des paramètres optimaux et les performances atteintes en fonction de α et de l'amplitude du signal à détecter.

Position du max. : Notons $g_{x_{\text{lo}}}$ le maximum de g_{lo} , atteint pour $x = x_{\text{lo}}$. Ce critère, inspiré de [4, 5], conduit à

$$\begin{cases} \lambda_{x_{\text{lo}}} &= 2\beta_{x_{\text{lo}}}^{\frac{1}{2}} g_{x_{\text{lo}}} \\ \beta_{x_{\text{lo}}} &= \frac{1}{x_{\text{lo}}^2} \end{cases} \quad (3)$$

Il est donc nécessaire de déterminer numériquement la position du maximum de la fonction score (la valeur du maximum est inutile, voir plus loin). Néanmoins, dans la mesure où $\beta_{x_{\text{lo}}}$ ne dépend que de α , ce paramètre peut être déterminé une fois pour toute et être tabulé en fonction de α . Notons également que dans [4, 5] Brown propose l'approximation $x_{\text{lo}} = 2.73\alpha - 1.75$ précise pour $\alpha \in [1, 1.8]$. Le β optimal est représenté en fonction de α en figure 1.

Distance L_2 : On cherche à minimiser $L = \int_{\mathbb{R}} (g_{\text{lo}}(x) - g(x))^2 dx$. Résoudre $\frac{\partial L}{\partial \lambda} = 0$ permet de déterminer le λ optimal, fonction explicite de β . Puis on cherche à résoudre $\frac{\partial L}{\partial \beta} = 0$ pour obtenir le β optimal noté β_{L_2} . On remplace λ par son expression optimale dans la seconde équation pour obtenir alors β_{L_2} , implicitement donné par

$$F_{L_2}(\beta) = \int_0^{+\infty} \frac{(\beta x^3 - 3x) f'_\alpha(x)}{(1 + \beta x^2)^2 f_\alpha(x)} dx = 0 \quad (4)$$

Ce β optimal est issu d'une optimisation de type gradient, néanmoins il est nécessaire de connaître la fonction score pour cette optimisation. Toutefois, comme β_{L_2} ne dépend pas du signal, ce paramètre pourra être tabulé suivant α . La figure 1 représente une estimation numérique de β_{L_2} optimal en fonction de α .

EQM : Comme dans le cas précédent, la minimisation de $M = E[(g_{\text{lo}}(X) - g(X))^2]$ conduit à un λ optimal fonction explicite de β . β optimal, noté β_{eqm} , est solution d'une équation implicite. Mais contrairement au critère L_2 , l'équation donnant β_{eqm} fait apparaître des intégrales en forme de produit scalaire $\int_{\mathbb{R}} h(x) f_n(x) dx$. Ces produits scalaires peuvent s'exprimer dans l'espace de Fourier inverse,

faisant apparaître la fonction caractéristique du bruit au lieu de la d.d.p., le choix du récepteur permettant pour sa part d'écrire facilement ces produits scalaires en Fourier inverse. Quelques calculs fastidieux aboutissent finalement à l'expression

$$\begin{aligned} F_{\text{mse}}(\beta) &= \int_0^{+\infty} u e^{-u - \beta^{\frac{\alpha}{2}} u^\alpha} du \\ &\times \int_0^{+\infty} (u^2 - 5u + 3) e^{-u - \beta^{\frac{\alpha}{2}} u^\alpha} du \\ &- 2 \int_0^{+\infty} (u - 1) e^{-u - \beta^{\frac{\alpha}{2}} u^\alpha} du \\ &\times \int_0^{+\infty} (u^2 - 2u) e^{-u - \beta^{\frac{\alpha}{2}} u^\alpha} du \\ &= 0 \end{aligned} \quad (5)$$

Ce critère ne requiert que l'évaluation d'intégrales simples utilisant uniquement des fonctions connues explicitement. La figure 1 représente β_{eqm} (indépendant du signal) en fonction de α .

Pe min : Pour le problème de détection présenté ici, la Pe min s'écrit en utilisant le fait que $g(x_t)$ a une variance finie, et en supposant que le nombre d'observations T est assez grand pour considérer que Λ_g (cf. (2)) est Gaussien sous H_0 (resp. H_1), de moyenne E_0 (resp. E_1) et de variance σ_0^2 (resp. σ_1^2), en conséquence du théorème de la limite centrale. Son expression est alors donnée par

$$\begin{aligned} P_{e,\text{min}} &= \frac{1}{4} \operatorname{erfc} \left(\frac{E_1}{\sqrt{2}(\sigma_0 + \sigma_1)} \left(1 + \frac{\sigma_1(\delta - 1)}{\sigma_1 - \sigma_0} \right) \right) \\ &+ \frac{1}{4} \operatorname{erfc} \left(\frac{E_1}{\sqrt{2}(\sigma_0 + \sigma_1)} \left(1 - \frac{\sigma_0(\delta - 1)}{\sigma_1 - \sigma_0} \right) \right) \end{aligned} \quad (6)$$

avec

$$\delta = \left(1 + 2 \frac{\sigma_1^2 - \sigma_0^2}{E_1^2} \log \left(\frac{\sigma_1}{\sigma_0} \right) \right)^{\frac{1}{2}} \quad (7)$$

La fonction erfc est la fonction d'erreur complémentaire et les moyennes et variances sous chaque hypothèse sont données par $E_1 = T\lambda \int_{\mathbb{R}} \frac{(x+c)f_n(x) dx}{1+\beta(x+c)^2}$ et $\sigma_k^2 = T\lambda^2 \int_{\mathbb{R}} \frac{(x+k.c)^2 f_n(x) dx}{(1+\beta(x+k.c)^2)^2} - E_k^2/T$, $k = 0, 1$ ($E_0 = 0$). Lorsque l'amplitude du signal est assez faible, on peut facilement montrer que cette Pe min est une fonction décroissante (i.e. erfc) de l'efficacité, qui est une approximation au premier ordre de la déflection $D = \frac{E_1^2}{\sigma_0^2} : E = \frac{(\int_{\mathbb{R}} g'(x) f_n(x) dx)^2}{\int_{\mathbb{R}} g^2(x) f_n(x) dx}$. Si dans l'expression de l'EQM on remplace λ par son expression optimale, l'efficacité et l'EQM vérifient $M = J - E$ avec $J = \int_{\mathbb{R}} \left(\frac{f'_n(x)}{f_n(x)} \right)^2 f_n(x) dx$ l'information de Fisher du bruit lié au centrage (moyenne si $\alpha > 1$). J ne dépend pas des paramètres, minimiser la Pe min équivaut à minimiser l'EQM (sous l'hypothèse de signal faible). Puisque λ disparaît de l'expression à optimiser (voir expression de la Pe min et de E_k et σ_k^2), seul le β optimal doit être recherché pour la mise en œuvre³.

³Dans tous les cas de figures, λ n'influence pas la Pe min, même s'il intervient dans l'optimisation de la distance L_2 et de l'EQM (voir (6)-(7) et les expressions des moyennes et variances). Ce paramètre peut être rejeté dans le seuil. Le problème sera de déterminer le seuil optimal.

En général, l'expression de P_e min est trop compliquée pour envisager une optimisation directe. On suppose alors que T est suffisamment grand pour simplifier son expression (en négligeant les infiniment petits) pour aboutir à

$$P_{e,m} \approx \frac{1}{2} \operatorname{erfc} \left(\frac{E_1}{\sqrt{2}(\sigma_0 + \sigma_1)} \right) \quad (8)$$

La minimisation de la P_e asymptotique minimale équivaut à la maximisation de l'argument de la fonction erfc . La fonction à annuler pour obtenir le paramètre optimal, $\beta_{P_{e,\min}}$, n'est pas donnée ici. Mais comme pour l'EQM, il s'agit de d'annuler la dérivée de la P_e min (8) en β , puis d'exprimer les E_k et σ_k^2 dans l'espace de Fourier inverse : le paramètre optimal, $\beta_{P_{e,\min}}$, est solution d'une équation implicite décrite via des intégrales de fonction connues explicitement. On peut montrer alors que $\beta_{P_{e,\min}}$ optimal est maintenant fonction de α et du signal c à détecter (voir expression de E_1 et σ_k). La figure 1 illustre le comportement de $\beta_{P_{e,\min}}$ en fonction de α pour plusieurs valeurs de c et la figure 2 donne $\beta_{P_{e,\min}}$ en fonction de α et c .

Dans les cas distance L_2 , EQM et P_e min, le paramètre optimum est solution d'une équation implicite : la solution est recherchée par une méthode de type gradient. Nous avons choisi l'algorithme de Newton-Raphson qui, expérimentalement, exhibe une convergence rapide. Enfin, dans le cas de l'EQM et P_e min, étant donné la forme des intégrales à évaluer, nous utilisons une méthode de type quadrature de Gauss (Gauss-Laguerre [7]) : expérimentalement cette méthode se montre très précise et à l'avantage d'être rapide : plutôt que de tabuler β , une optimisation en ligne peut être envisagée.

La section suivante analyse le comportement des β optimaux en fonction de α .

4 Paramètres optimaux et performances de Λ_g

Les figures 1 et 2 représentent les β optimaux pour les différents critères :

- Pour $\alpha = 2$, quelque soit l'amplitude du signal, les critères conduisent tous au filtre adapté ($\beta = 0$, solution optimale en Gaussien).
- Les critères position, L_2 et EQM donnent des β optimaux voisins pour $\alpha \leq 1$. Quand α tend vers 2, la distance L_2 tend à donner un β très différents des autres critères. Ceci semble lié au caractère spécial de f'_n/f_n en $\alpha = 2$ (transition d'un comportement asymptotique en $1/x$ à un comportement linéaire). La fonction score n'apparaît pas pour l'EQM ou la P_e minimale, les solutions correspondantes ne semblent pas présenter cette particularité.
- Quand c croît, le critère de la P_e minimale donne des β de plus en plus petits : P_e est de moins en moins liée à l'EQM et le LO est de plus en plus éloignée du récepteur optimal. En particulier, pour $\alpha = 1$, ce critère ne donne pas $\beta = 1$. Tout se passe comme si le récepteur tendait à réaliser le filtre adapté pour les observations dans une gamme d'échelle plus importante à mesure que c est grand (g est approximativement linéaire dans une gamme d'entrée de plus en plus importante à mesure que β décroît).

- Lorsque α tend vers 0, β diverge. Ce cas dégénéré correspond à un bruit déterministe, nul sous notre hypothèse de symétrie. On tend à comparer x_t à 0 ($g \sim \delta_{x,0}$) et sommer ces comparaisons.

Les performances du détecteur proposé, du filtre adapté, du détecteur optimal de Cauchy, et du détecteur de Brown (linéaire sur $[-x_{10}, x_{10}]$ et égal à $\frac{\alpha+1}{x}$ sinon) sont illustrées sur les figures 3, 4, et 5, comparées avec les performances optimales. La P_e min est calculée en utilisant l'hypothèse gaussienne pour le récepteur proposé, le récepteur optimal, de Cauchy et de Brown (théorème de la limite centrale). Pour le filtre adapté, on utilise la propriété de stabilité des $S\alpha S$ pour en déduire que le filtre adapté est non plus asymptotiquement gaussien mais α -stable de même indice α que le bruit [1]. La probabilité d'erreur minimale est alors donnée par $P_{e,m} = F_n(-\frac{c}{2} T^{\frac{\alpha-1}{\alpha}})$ (où F_n est la fonction de répartition du bruit). La d.d.p. n'est pas requise pour évaluer la P_e minimale du récepteur Λ_g , elle est par contre nécessaire au calcul des performances optimales, du récepteur de Cauchy et de Brown tandis que la fonction de répartition du bruit intervient dans les performances du filtre adapté. La comparaison des performances aux performances optimales donne lieu aux conclusions suivantes :

- Pour $\alpha = 2$, quelque soit l'amplitude du signal, les critères conduisent tous au filtre adapté ($\beta = 0$, solution optimale en Gaussien).
- Pour un signal faible, la notion de "faible" dépendant clairement de α , les performances des récepteurs sont similaires, sauf celles du récepteur de Cauchy et du filtre adapté. Ces deux derniers récepteurs s'avèrent peu robustes quand α s'écarte respectivement de 1 et 2.
- Pour un signal fort et α loin de 2, le critère de P_e min est meilleur que les autres (dont les performances se dégradent lorsque l'amplitude du signal grandit, d'autant plus que α est petit). Le filtre adapté reste non robuste en α , tandis que le récepteur de Cauchy s'avère de plus en plus robuste à mesure que le signal devient grand.

5 Discussion

Le détecteur proposé montre une certaine robustesse vis-à-vis de l'indice de stabilité α et du signal c , en particulier pour les critères d'optimisation basés sur des considérations de détection (surtout P_e min). Pour l'optimisation du paramètre, les intégrations numériques entrant en jeu se font une fois pour toute et non pas à chaque échantillon du signal observé comme ce serait le cas pour la log-vraisemblance. On peut donc tabuler β optimal, en fonction de α (et c). Le choix de la forme du récepteur, essentiellement dictée par des considérations sur les lois α -stable, permet *a posteriori* de faciliter l'optimisation. Dans le cas où α serait inconnu, on peut envisager l'estimation de α et d'utiliser la table établie, mais aussi de faire l'optimisation en ligne avec l'estimation de α . De même, si l'on considère le problème de détection d'un signal non constant s_t , seul le critère P_e min est à revoir. Dans ce cas, on peut considérer le β optimal obtenu pour une constante, $\beta_{P_{e,\min}}(\alpha, c)$ et le remplacer par $\beta_{P_{e,\min}}(\alpha, s_t)$: dans ce cas encore on peut envisager une optimisation temps réel.

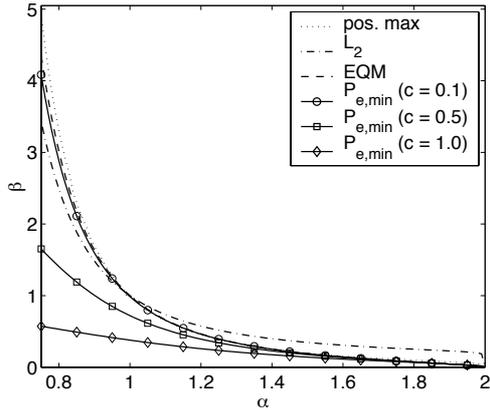


FIG. 1 – Comportement du β optimal pour le détecteur (2) en fonction de α et pour les différents critères.

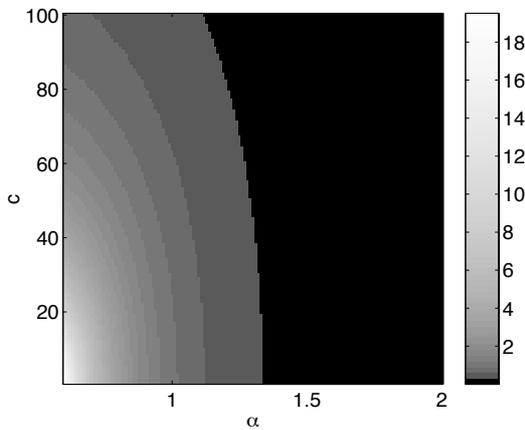


FIG. 2 – β optimal pour le critère P_e min, en fonction de α et du signal c à détecter.

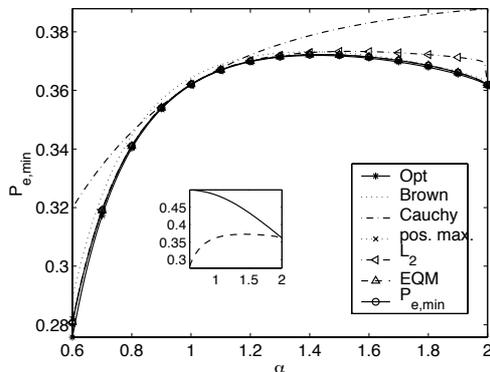


FIG. 3 – Probabilité d'erreur minimale pour le détecteur (2), fonction de α et pour les différents critères. Comparaison avec les performances des récepteurs optimal, de Cauchy et proposé par Brown. En insert : performances du filtre adapté (trait plein) et du critère EQM (tirets). $T = 100$ et $c = 0.1$.

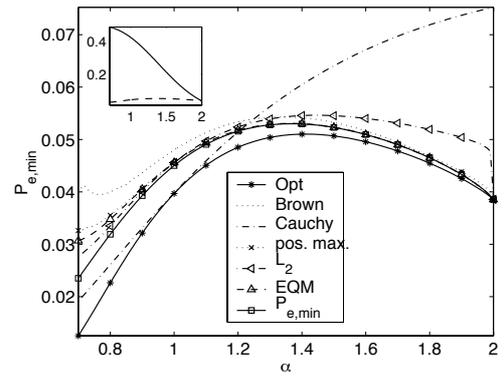


FIG. 4 – Comme la figure 3, pour $c = 0.5$.

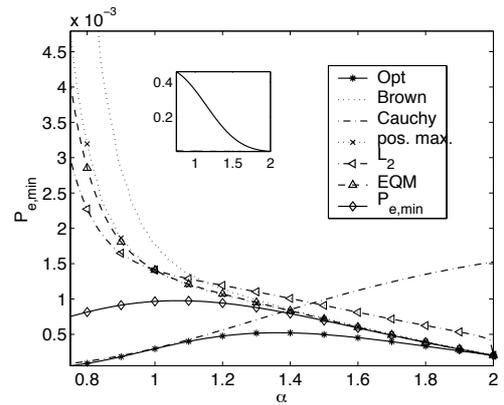


FIG. 5 – Comme la figure 3, pour $c = 1$.

Références

- [1] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes. Stochastic Models with infinite Variance*. New-York : Chapman & Hall, 1994.
- [2] S. M. Kay, *Fundamentals for Statistical Signal Processing : Detection Theory*. Englewood Cliffs, NJ : Prentice Hall, 1998.
- [3] G. A. Tsihrintzis and C. L. Nikias, "Performance of optimum and suboptimum receivers in the presence of impulsive noise modeled as an alpha-stable process," *IEEE Trans. on Com.*, vol. 43, pp. 904–914, February/March/April 1995.
- [4] C. L. Brown and A. M. Zoubir, "A nonparametric approach to signal detection in impulsive interference," *IEEE Trans. on Sig. Proc.*, vol. 48, pp. 2665–2669, September 2000.
- [5] C. L. Brown, "Score functions for locally suboptimum and locally suboptimum rank detection in alpha-stable interference," in *Proc. of the 11th Sig. Proc. workshop on Stat. Sig. Proc.*, (Singapore), pp. 58–61, August 2001.
- [6] A. Ambike, J. Ilow, and D. Hatzinakos, "Detection for binary transmission in a mixture of Gaussian noise and impulsive noise modeled as an alpha-stable process," *IEEE Sig. Proc. Let.*, vol. 1, pp. 55–57, March 1994.
- [7] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. 9th printing, New-York : Dover, 1970.