USC–SIPI REPORT #274

General $\mathcal{L}^{(p,q)}$-Metric Estimator of Arbitrary Complex Impulsive Interference in Linear Systems

by

Jun Shen and Chrysostomos L. Nikias

November 1994

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Room 404
Los Angeles, CA 90089-2564 U.S.A.

General $\mathcal{L}^{(p,q)}$-Metric Estimator of Arbitrary Complex Impulsive Interference in Linear Systems*

Jun Shen, Student Member, IEEE, and Chrysostomos L. Nikias, Fellow, IEEE

Signal & Image Processing Institute
Department of Electrical Engineering-Systems
University of Southern California
Los Angeles, CA 90089-2564

Tel: (213)740-2218 Fax:(213)740-8931

(09/15/94)

Abstract

A general $\mathcal{L}^{(p,q)}$-metric, $p, q > 0$, on the probability space is defined and the corresponding optimality criterion derived. This criterion is adopted to the estimation problem of complex impulsive interference in linear systems presented by state-space equations. The closed-form a posteriori density of the state (interference) is computed recursively for both arbitrary i.i.d. state noise and any discrete-type measurement noise (multi-level complex signal), and the optimal $\mathcal{L}^{(p,q)}$-metric interference estimators based on different values of p and q are developed. As a test, the proposed algorithms are effectively applied to estimate highly impulsive state processes driven by noise with symmetric α-stable distribution.

*This work was supported by the Office of Naval Research under contract N00014-92-J-1034.

I. Introduction

The detection and estimation of impulsive interference have been the subject of research by many investigators. We mention, for example, the work by Poor *et al.* [1, 2], in which the approximate conditional mean(ACM) filter [3] was developed for estimating the state of a linear system with Gaussian state noise and non-Gaussian measurement noise. For the early works in this area, it is worthwhile to mention Middleton *et al.* [4, 5] who proposed the Class A impulsive interference model, for which the p.d.f. of the interference was given by a weighted sum of Gaussian distributions with increasing variance, and the corresponding coherent binary detection scheme [5] and a recursive decision-directed algorithm (MBDD) for on-line identification the class parameters [6] were developed. However, all previous work on the interference mitigation problem was mainly done under the assumption that the noise sequence was Gaussian or Gaussian mixtures, so that the *a posteriori* density function could be approximated by Gaussian density function(s) [3, 7]. Since the interference sequence is usually very impulsive or obeying a heavily tailed distribution, the family of exponential distributions (e.g. Gaussian) are decaying too fast to adequately describe the impulsiveness of the interference. So, it is desirable to choose more appropriate models for the impulsive interference other than the Gaussian or Gaussian mixtures. Another disadvantage for these approaches is that, the *a posteriori* density functions could only be approximated so that the interference estimators are not optimal. It is also need to point out that these previous schemes are interference-model dependent and their applications in the real world are limited. To overcome the shortcomings of the above algorithms, Shen and Nikias [8, 9] proposed the general conditional mean (GCM) and absolute value criterion (AVC) for the development of impulsive interference mitigation algorithms. However, these new algorithms were developed only for the case of real time series.

In this paper, we first define a general metric on the probability space and derive the corresponding optimality criterion which involves n-dimensional random vectors. Then, we apply this general metric optimality criterion to the complex dynamic systems described by the state space equations and the optimal estimators for complex interference sequence are obtained. Hence, we extend the results of [8, 9] to the complex interference sequence. And for this interference model, it is assumed that the complex noise sequence has arbitrary density function and the signal can be any discrete-type complex sequence so that the impulsive interference can be modeled properly according to the interference observed in real world applications. In the simulation, we choose the symmetric α-stable (SαS) process [10, 11, 12, 13] as the impulsive interference model. Then the proposed algorithms are applied to estimate and cancel the interference from the measurement data.

This paper has the following structure. Section II describes the setting of the problem to be solved. In Section III, a new general metric optimality criterion is introduced to accommodate the n-dimensional random vectors and the objective function is derived for the optimal estimators. Section IV is devoted to the design of interference estimator based on new general metric criterion. Some simulation results are presented in Section V. Finally, the proposed new algorithms are summarized in Section VI.

Some notations used in this paper are defined as follows.

- $\mathcal{R}$: the real line;
- $\mathcal{R}^+$: the set of non-negative real numbers, i.e., $\mathcal{R}^+ = \{x : x \in \mathcal{R}, x \geq 0\}$;
- Ω: a sample space;
- $\mathcal{F}_\Omega$: a σ-field of subsets of Ω;
- $\mathcal{R}^n$: the n-dimensional real space;
- $\mathcal{B}_{\mathcal{R}^n}$: the Borel σ-field of the subsets of $\mathcal{R}^n$;
- $\mathcal{C}$: the complex plane, i.e., $\mathcal{C} = \{x_1 + jx_2 : x_1, x_2 \in \mathcal{R}\}$;
- $\mathcal{B}_\mathcal{C}$: the smallest σ-algebra containing the sets

  $\{y : y = x_1 + jx_2, x_1 \in (a_1, b_1], x_2 \in (a_2, b_2], a_1, a_2, b_1, b_2 \in \mathcal{R}\} \subseteq \mathcal{C}$;
- Operation $\wedge$: $(a \wedge b) = \min\{a, b\}, \forall a, b \in \mathcal{R}$;
- A *complex random variable* Y is a measurable function on the space $(\Omega, \mathcal{F}_\Omega)$ to the space $(\mathcal{C}, \mathcal{B}_\mathcal{C})$ and is denoted by $Y(\omega) = Y_1(\omega) + jY_2(\omega), \forall \omega \in \Omega$, where $Y_1(\omega)$ and $Y_2(\omega)$ are two real ordinary random variables.

II. Problem Statement

We consider the following scalar linear system

$$X_{k+1} = A_k X_k + W_k, \qquad k = 1, 2, \cdots \tag{1}$$

$$Z_k = H_k X_k + V_k, \qquad k = 1, 2, \cdots \tag{2}$$

where the following assumptions are made.

[A1]: $\{W_k,\ k = 1, 2, \cdots\}$ is the i.i.d. complex noise sequence, i.e., at time k, $W_k = W_k^R + jW_k^I$ is a complex random variable, where W_k^R and W_k^I are two real random variables with arbitrarily chosen joint density function $f_{W^R W^I}(w_k^R, w_k^I)$. For simplicity, we adopt the notation [14] $f_W(w_k^R, w_k^I) = f_W(w_k^R + jw_k^I)$.

[A2]: $\{V_k,\ k = 1, 2, \cdots\}$ is the complex signal sequence with any discrete-type distribution, i.e., $V_k = V_k^R + jV_k^I$ is a complex random variable with V_k^R and V_k^I being two real discrete-type random variables, $\forall k$. At each time k, V_k^R and V_k^I take the values from the two real sets $\{c_i :\ i = 1, 2, ..., I\}$ and $\{d_j :\ j = 1, 2, ..., J\}$, respectively. The joint distribution is $Pr(V_k^R = c_i, V_k^I = d_j) = q_{ij}$, $i = 1, 2, ..., I$, $j = 1, 2, ..., J$, $\sum_{i=i}^{I} \sum_{j=1}^{J} q_{ij} = 1, \forall k$.

[A3]: The two sequences $\{W_k, k = 1, 2, ...\}$ and $\{V_k, k = 1, 2, ...\}$ are assumed to be mutually independent at any time k.

[A4]: For the system parameters, it is assumed that $A_k,\ H_k \in \mathcal{R},\ \forall k$.

[A5]: The complex interference sequence (state) is, defined in (1), $\{X_k = X_k^R + jX_k^I,\ k = 1, 2, ...\}$ with X_k^R and X_k^I, $\forall k$, being two real ordinary random variables.

[A6]: Let $Z^k = \{Z_1, Z_2, ..., Z_k\}$ be the set of possible observations (complex random variables) from past up to time k. At any fixed time k, the set of past complex measurement (data) $z^k = \{z_l :\ Z_l = z_l, z_l = z_l^R + jz_l^I,\ l = 1, 2, ..., k\}$ are known to the observer.

At time k, for the true complex random variable X_k, define $\hat{X}_k$ as the filtered estimator for the observation Z^k. Our objective is to find $\hat{X}_k$ such that the error between $\hat{X}_k$ and X_k is minimized under some criterion.

III. General $\mathcal{L}^{(p,q)}$-Metric and Optimality Criteria

In this section, we will first introduce the general $\mathcal{L}^{(p,q)}$-metric (GLM), which is the generalization of the classic $\mathcal{L}^{p}$-norm to incorporate the case for $p \in (0,1)$ and can also be applied to random vectors. Then we formulate the optimal interference estimation problem under the GLM criterion.

A. General $\mathcal{L}^{(p,q)}$-metric

Let $(\Omega, \mathcal{F}_\Omega)$ and $(\mathcal{R}^n, \mathcal{B}_{\mathcal{R}^n})$ be measurable spaces. We say a function $\mathbf{X} : \Omega \to \mathcal{R}^n$ is a *n-dimensional random vector* if

$$\{\omega : \omega \in \Omega, \mathbf{X}(\omega) = (X_1(\omega), X_2(\omega), ..., X_n(\omega)) \in B\} \in \mathcal{F}_\Omega, \ \ \forall B \in \mathcal{B}_{\mathcal{R}^n},$$

where $X_k(\omega)$ is an ordinary random variable which is the projection of $\mathbf{X}(\omega)$ on the k^{th} coordinate axis of $\mathcal{R}^n$. Notice that, if $n = 2$, a random vector is equivalent to a complex random variable. Next, let us introduce two operators.

Definition 3.1 Let $\mathbf{x} = (x_1, x_2, ..., x_n)$ be a fixed point in $\mathcal{R}^n$. Define the operator $\|\cdot\|_p$ on $\mathcal{R}^n$ as

$$\|\mathbf{x}\|_p = \left[\sum_{i=1}^{n} |x_i|^p\right]^{(1 \wedge 1/p)}, \quad p > 0. \tag{3}$$

□

By the Minkowski inequality and the following elementary inequality ([15], pp.155)

$$|a+b|^p \le c_p(|a|^p + |b|^p), \quad \forall a, b \in \mathcal{R}, \ \ p > 0, \tag{4}$$

where $c_p = 1$ if $0 < p \le 1$ and $c_p = 2^{p-1}$ if $p \ge 1$, we can easily show that

$$\|\mathbf{x}+\mathbf{y}\|_p \le \|\mathbf{x}\|_p + \|\mathbf{y}\|_p, \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{R}^n, \ \ \text{for } p > 0, \tag{5}$$

and $\tilde{d}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_p$ defines a metric on $\mathcal{R}^n$.

Definition 3.2 Let the n-dimensional random vector be $\mathbf{X} = (X_1, X_2, ..., X_n)$ with the distribution function $F_{\mathbf{X}}(x_1, x_2, ..., x_n)$ and density $f_{\mathbf{X}}(x_1, x_2, ..., x_n)$. For the fixed real numbers $p, q > 0$, define the operator $\langle\!\langle \cdot \rangle\!\rangle_{(p,q)}$ on the n-dimensional Borel space $(\mathcal{R}^n, \mathcal{B}_{\mathcal{R}^n})$ as

$$\langle\!\langle \mathbf{X} \rangle\!\rangle_{(p,q)} = \left\{\int_{\mathcal{R}^n} \left[\|\mathbf{x}\|_p\right]^q dF_{\mathbf{X}}\right\}^{(1 \wedge 1/q)} \tag{6}$$

$$\overset{\text{or}}{=} \quad \left\{ \int_{\mathcal{R}^n} \left[\left(\sum_{i=1}^{n} |x_i|^p \right)^{(1 \wedge 1/p)} \right]^q f_{\mathbf{X}}(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n \right\}^{(1 \wedge 1/q)} \tag{7}$$

with the assumption that $\mathbf{X}$ is equivalent to the n-dimensional zero vector $\mathbf{0}$ if $\mathbf{X} = \mathbf{0}$ almost surely. □

With the operator $\langle\!| \cdot |\!\rangle_{(p,q)}$, a induced metric space is defined and stated in the following theorem.

Theorem 3.1 *For the fixed real numbers $p, q > 0$, let $\mathcal{L}^{(p,q)} = \mathcal{L}^{(p,q)}(\mathcal{R}^n, \mathcal{B}_{\mathcal{R}^n})$ be the space of the n-dimensional random vectors $\mathbf{X}$ with $\langle\!|\mathbf{X}|\!\rangle_{(p,q)} < \infty$. Then $\mathcal{L}^{(p,q)}$ is a metric space and for any two random vectors $\mathbf{X}_1$, $\mathbf{X}_2 \in \mathcal{L}^{(p,q)}$, the corresponding $\mathcal{L}^{(p,q)}$-metric $d_{(p,q)}$ is defined as*

$$d_{(p,q)}(\mathbf{X}_1, \mathbf{X}_2) = \langle\!|\mathbf{X}_1 - \mathbf{X}_2|\!\rangle_{(p,q)}. \tag{8}$$

Proof: For any given random vectors $\mathbf{X}_1, \mathbf{X}_2 \in \mathcal{L}^{(p,q)}$, it is easy to verify that, i) $d_{(p,q)}(\mathbf{X}_1, \mathbf{X}_2) \in \mathcal{R}^+$; ii) $d_{(p,q)}(\mathbf{X}_1, \mathbf{X}_2) = d_{(p,q)}(\mathbf{X}_2, \mathbf{X}_1)$; and iii) $d_{(p,q)}(\mathbf{X}_1, \mathbf{X}_2) = 0 \Longleftrightarrow \mathbf{X}_1 = \mathbf{X}_2$ $(a.s.)$. Now, we want to show that the triangle inequality also holds.

Let $F_{\mathbf{X}_1\mathbf{X}_2}$ be the joint distribution function of $\mathbf{X}_1$ and $\mathbf{X}_2$. Then,

$$\begin{aligned} \langle\!|\mathbf{X}_1 - \mathbf{X}_2|\!\rangle_{(p,q)} &= \left\{ \int_{\mathcal{R}^{2n}} \|\mathbf{x}_1 - \mathbf{x}_2\|_p^q dF_{\mathbf{X}_1\mathbf{X}_2} \right\}^{(1 \wedge 1/q)} \quad \text{(by the definition)} \\ &\leq \left\{ \int_{\mathcal{R}^{2n}} [\|\mathbf{x}_1\|_p + \|\mathbf{x}_2\|_p]^q dF_{\mathbf{X}_1\mathbf{X}_2} \right\}^{(1 \wedge 1/q)} \quad \text{(due to (5))} \end{aligned}$$

Now, if $0 < q \leq 1$, then $(1 \wedge 1/q) = 1$ and by (4) we have

$$\begin{aligned} \int_{\mathcal{R}^{2n}} [\|\mathbf{x}_1\|_p + \|\mathbf{x}_2\|_p]^q dF_{\mathbf{X}_1\mathbf{X}_2} &\leq \int_{\mathcal{R}^{2n}} \|\mathbf{x}_1\|_p^q dF_{\mathbf{X}_1\mathbf{X}_2} + \int_{\mathcal{R}^{2n}} \|\mathbf{x}_2\|_p^q dF_{\mathbf{X}_1\mathbf{X}_2} \\ &= \langle\!|\mathbf{X}_1|\!\rangle_{(p,q)} + \langle\!|\mathbf{X}_2|\!\rangle_{(p,q)}, \end{aligned}$$

and for $q \geq 1$, so $(1 \wedge 1/q) = 1/q$ and by the Minkowski inequality ([15], pp.161) it follows that

$$\begin{aligned} \left\{ \int_{\mathcal{R}^{2n}} [\|\mathbf{x}_1\|_p + \|\mathbf{x}_2\|_p]^q dF_{\mathbf{X}_1\mathbf{X}_2} \right\}^{1/q} &\leq \left\{ \int_{\mathcal{R}^{2n}} \|\mathbf{x}_1\|_p^q dF_{\mathbf{X}_1\mathbf{X}_2} \right\}^{1/q} + \left\{ \int_{\mathcal{R}^{2n}} \|\mathbf{x}_2\|_p^q dF_{\mathbf{X}_1\mathbf{X}_2} \right\}^{1/q} \\ &= \langle\!|\mathbf{X}_1|\!\rangle_{(p,q)} + \langle\!|\mathbf{X}_2|\!\rangle_{(p,q)}. \end{aligned}$$

Finally, we have

$$\langle\!\langle \mathbf{X}_1 - \mathbf{X}_2 \rangle\!\rangle_{(p,q)} \leq \langle\!\langle \mathbf{X}_1 \rangle\!\rangle_{(p,q)} + \langle\!\langle \mathbf{X}_2 \rangle\!\rangle_{(p,q)}, \quad \text{for } p,\, q > 0.$$

Q.E.D.

Remarks:

1. For the definition of $\langle\!\langle \cdot \rangle\!\rangle_{(p,q)}$ in (6), the $\|\cdot\|_p$ operator may not necessarily be the form in (3) as long as (5) is valid.

2. It can be shown that the space $\mathcal{L}^{(p,q)}$ is also complete in the metric $d_{(p,q)}$, and for $p, q \geq 1$, the operator $\langle\!\langle \cdot \rangle\!\rangle_{(p,q)}$ defines a norm. Thus, $\mathcal{L}^{(p,q)}$, $p, q > 1$, is a Banach space.

3. For the one-dimensional random variable X with symmetric α-stable distribution [10, 12], the number $\langle\!\langle X \rangle\!\rangle_{(p,q)}$, for $p, q \in (0, 2]$, is similar to the so-called dispersion of X.

B. GLM optimality criterion

Here, we start by considering GLM criterion for the design of optimal interference estimator. To be general, here it is assumed that X_k and Z_k in (1) and (2) are n-dimensional random vectors. In symbol, let $X_k = \mathbf{X}_k$ and Z_k the same. At time k, given the measurement set $Z^k = z^k$, let the random vector $\hat{\mathbf{X}}_k(Z^k)$, which is a function of Z^k, be the estimator of the true interference $\mathbf{X}_k$. The objective is to find the optimal estimator $\hat{\mathbf{X}}_k^{opt} \in \mathcal{L}^{(p,q)}$ such that

$$\langle\!\langle \mathbf{X}_k - \hat{\mathbf{X}}_k^{opt} \rangle\!\rangle_{(p,q)} = \min \left\{ \langle\!\langle \mathbf{X}_k - \hat{\mathbf{X}}_k \rangle\!\rangle_{(p,q)} : \hat{\mathbf{X}}_k \in \mathcal{L}^{(p,q)} \right\} \tag{9}$$

where $\langle\!\langle \cdot \rangle\!\rangle_{(p,q)}$ is the GLM operator, $p, q \in (0, \infty)$. By the definition of (6) and notice that $\hat{\mathbf{X}}_k$ is a function of Z^k, we have

$$\langle\!\langle \mathbf{X}_k - \hat{\mathbf{X}}_k \rangle\!\rangle_{(p,q)} = \left\{ \int\!\!\int \|\mathbf{x}_k - \hat{\mathbf{x}}_k\|_p^q f_{\mathbf{X}_k Z^k}(\mathbf{x}_k, z^k) d\mathbf{x}_k dz^k \right\}^{(1 \wedge 1/q)}.$$

Since

$$\left\{ \int\!\!\int \|\mathbf{x}_k - \hat{\mathbf{x}}_k\|_p^q f_{\mathbf{X}_k Z^k}(\mathbf{x}_k, z^k) d\mathbf{x}_k dz^k \right\}^{(1 \wedge 1/q)}$$

is the monotonically increasing function of

$$\int\!\!\int \|\mathbf{x}_k - \hat{\mathbf{x}}_k\|_p^q f_{\mathbf{X}_k Z^k}(\mathbf{x}_k, z^k) d\mathbf{x}_k dz^k,$$

to minimize $\langle\!| \mathbf{X}_k - \hat{\mathbf{X}}_k |\!\rangle_{(p,q)}$ is equivalent to minimize

$$\int\int \|\mathbf{x}_k - \hat{\mathbf{x}}_k\|_p^q f_{\mathbf{X}_k Z^k}(\mathbf{x}_k, z^k) d\mathbf{x}_k dz^k.$$

Now,

$$\int\int \|\mathbf{x}_k - \hat{\mathbf{x}}_k\|_p^q f_{\mathbf{X}_k Z^k}(\mathbf{x}_k, z^k) d\mathbf{x}_k dz^k = \int f_{Z^k}(z^k) dz^k \int \|\mathbf{x}_k - \hat{\mathbf{x}}_k\|_p^q f_{\mathbf{X}_k | Z^k}(\mathbf{x}_k | z^k) d\mathbf{x}_k.$$

From the above it is clear that we only need to minimize the functional

$$R(\hat{\mathbf{x}}_k(z^k)) = \int \|\mathbf{x}_k - \hat{\mathbf{x}}_k(z^k)\|_p^q f_{\mathbf{X}_k | Z^k}(\mathbf{x}_k | z^k) d\mathbf{x}_k. \tag{10}$$

So, the final GLM optimization problem of (9) can be stated as

$$\hat{\mathbf{x}}_k^{opt}(z^k) \in \mathcal{R}^n :$$

$$R(\hat{\mathbf{x}}_k^{opt}(z^k)) = \min\left\{ \int \|\mathbf{x}_k - \hat{\mathbf{x}}_k(z^k)\|_p^q f_{\mathbf{X}_k | Z^k}(\mathbf{x}_k | z^k) d\mathbf{x}_k : \ \hat{\mathbf{x}}_k(z^k) \in \mathcal{R}^n \right\}. \tag{11}$$

This objective functional $R(\mathbf{x})$, obtained from the GLM criterion, will be used in the sequel as the performance criterion for the impulsive interference estimators.

IV. Optimal Impulsive Interference Estimators

With the GLM criterion available, we can now attack the original problem stated in section II. In this section, the results for the general $\mathcal{L}^{(p,q)}$-metric will be applied to the complex random variables, i.e., the 2-dimensional random vectors. From (11), to estimate the state X_k, we have to find the *a posteriori* density function $f_{X_k|Z^k}(x_k|z^k)$ first. Then the corresponding estimates can be determined.

A. Computation of the *a Posteriori* Density Functions

Given the state space equations (1) and (2), according to the Bayesian law, the *a posteriori* density can be determined recursively by the following set of equations [7].

$$f_{X_k|Z^k}(x_k|z^k) = \frac{f_{X_k|Z^{k-1}}(x_k|z^{k-1}) f_{Z_k|X_k}(z_k|x_k)}{f_{Z_k|Z^{k-1}}(z_k|z^{k-1})} \tag{12}$$

$$f_{X_k|Z^{k-1}}(x_k|z^{k-1}) = \int f_{X_{k-1}|Z^{k-1}}(x_{k-1}|z^{k-1}) f_{X_k|X_{k-1}}(x_k|x_{k-1}) dx_{k-1} \tag{13}$$

$$f_{Z_k|Z^{k-1}}(z_k|z^{k-1}) = \int f_{X_k|Z^{k-1}}(x_k|z^{k-1}) f_{Z_k|X_k}(z_k|x_k) dx_k \tag{14}$$

with the initial density

$$f_{X_1|Z_1}(x_1|z_1) = \frac{f_{Z_1|X_1}(z_1|x_1) f_{x_1}(x_1)}{f_{Z_1}(z_1)} \tag{15}$$

where $f_{Z_k|X_k}(z_k|x_k)$ in (12) can be determined by the signal density $f_V(v_k)$ and (2), and $f_{X_k|X_{k-1}}(x_k|x_{k-1})$ in (13) is defined by the noise density $f_W(w_k)$ and (1). In general, it is impossible to get the closed form of $f_{X_k|Z^{k-1}}(x_k|z^{k-1})$, so the *a posteriori* density can not be obtained for most applications. However, for the system described by (1) and (2), the following statement holds.

Theorem 4.1 *Let the complex linear system be defined by (1) and (2). Given the initial a priori density $f_{X_1}(x_1^R, x_1^I)$, the a posteriori densities at each time k are uniquely determined by the recursions*

$$f_{X_k|Z^{k-1}}(x_k|z^{k-1}) = \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} r_{k-1,ij} f_W(x_k^R - A_{k-1}s_{k-1,i}, x_k^I - A_{k-1}t_{k-1,j})}{\sum_{i=1}^{I} \sum_{j=1}^{J} r_{k-1,ij}} \tag{16}$$

$$k = 2, 3, \ldots$$

$$f_{X_k|Z^k}(x_k|z^k) = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} \ r_{k,ij}\ \delta(x_k^R - s_{k,i})\delta(x_k^I - t_{k,j})}{\sum_{i=1}^{I}\sum_{j=1}^{J} \ r_{k,ij}} \tag{17}$$

$$k = 1, 2, ...$$

where

$$s_{k,i} = \frac{z_k^R - c_i}{H_k}, \quad t_{k,j} = \frac{z_k^I - d_j}{H_k}, \quad r_{k,ij} = q_{ij}\, f_{X_k|Z^{k-1}}(s_{k,i}, t_{k,j}|z^{k-1}) \tag{18}$$

$$i = 1, 2, ..., I, \ \ j = 1, 2, ..., J, \ \ k = 1, 2, ...$$

Proof: Notice that the density function for the complex variable $V_k = V_k^R + jV_k^I$ is

$$f_V(v_k^R, v_k^I) = \sum_{i=1}^{I}\sum_{j=1}^{J} q_{ij}\ \delta(v_k^R - c_i)\delta(v_k^I - d_j) \tag{19}$$

where $\delta(\cdot)$ is the Dirac's delta function. By (2), (19) and the property for the delta function $\delta(\cdot)$, i.e., $\delta(at) = \frac{\delta(t)}{|a|}$, $\forall a \neq 0$, we have

$$f_{Z_k|X_k}(z_k|x_k) = f_V(z_k^R - H_k x_k^R, z_k^I - H_k x_k^I) = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} q_{ij}\ \delta(x_k^R - s_{k,i})\delta(x_k^I - t_{k,j})}{H_k^2} \tag{20}$$

where $s_{k,i}$ and $t_{k,j}$ are defined in the above theorem. From the above equation and (14),

$$f_{Z_k|Z^{k-1}}(z_k|z^{k-1}) = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} q_{ij}\ f_{X_k|Z^{k-1}}(s_{k,i}, t_{k,j}|z^{k-1})}{H_k^2} \tag{21}$$

and from (1),

$$f_{X_k|X_{k-1}}(x_k|x_{k-1}) = f_W(x_k^R - A_{k-1}x_{k-1}^R, x_k^I - A_{k-1}x_{k-1}^I) \tag{22}$$

Finally, (17) is obtained from (12), (20), (21), and (13), (17), (22) lead to (16).

Q.E.D.

Based on Theorem 4.1, we will consider how to compute the *a posteriori* densities of the state X_k at any time k. It is assumed that we know the initial *a priori* density $f_{X_1}(x_1)$, the noise density $f_W(w_k)$, the measurements $z^k = \{z_1, z_2, ..., z_k\}$, and the levels $\{c_i : \ i = 1, 2, ..., I\}$ and $\{d_j : \ j = 1, 2, ..., J\}$ of the signal sequence $\{V_k, k = 1, 2, ...\}$. In order to compute the *a posteriori* density functions $f_{X_k|Z^k}(x_k|z^k)$ and $f_{X_{k+1}|Z^k}(x_{k+1}|z^k)$ at time k, the density $f_{X_k|Z^{k-1}}(x_k|z^{k-1})$ at time $(k-1)$ has to be known. It means that the whole function $f_{X_k|Z^{k-1}}(x_k|z^{k-1})$ has to be stored at time $(k-1)$ for retrieving at time k. It would demand a large amount of computing memory and also require longer computing time. Fortunately, by carefully examining (16) in

Theorem 4.1, we find that the function $f_{X_k|Z^{k-1}}(x_k|z^{k-1})$ can be completely defined at time $(k-1)$ by using only $(I + J + I \cdot J)$ numbers, i.e., $\{s_{k-1,i} : i = 1, 2, ..., I\}$, $\{t_{k-1,j},\ j = 1, 2, ..., J\}$ and $\{f_{X_{k-1}|Z^{k-2}}(s_{k-1,i}, t_{k-1,j}|z^{k-2}) :\ i = 1, 2, ..., I,\ j = 1, 2, ..., J\}$, together with the known levels $\{c_i :\ i = 1, 2, ..., I\}$, $\{d_j :\ j = 1, 2, ..., J\}$ and the function $f_W(\cdot)$, where $s_{k-1,i}$ and $t_{k-1,j}$ fix the "locations" of the $f_W(\cdot)$ functions and $f_{X_{k-1}|Z^{k-2}}(s_{k-1,i}, t_{k-1,j}|z^{k-2})$ serve as the "weights". So, the following algorithm is obtained.

Algorithm 4.1 Computation of the *a posteriori* densities

For $k = 1, 2, ..., N$, do the following steps:

1. From the known measurement z_k and the level set $\{c_i : i = 1, 2, ..., I\}$ and $\{d_j : j = 1, 2, ..., J\}$ of the multi-level complex signal, obtain the measurement set $T_k = \{(s_{k,i}, t_{k,j}) :\ i = 1, 2, ..., I,\ j = 1, 2, ..., J\}$.
2. Use (16) to compute weights $f_{X_k|Z^{k-1}}(s_{k,i}, t_{k,j}|z^{k-1})$, $i = 1, 2, ..., I$, $j = 1, 2, ..., J$
3. By (16) and (17), define functions $f_{X_k|Z^k}(\cdot|z^k)$, $f_{X_{k+1}|Z^k}(\cdot|z^k)$

with initial conditions $(k = 1)$: $f_{X_1}(s_{1,i}, t_{1,j})$, $i = 1, 2, ..., I$, $j = 1, 2, ..., J$

///

B. Optimal GLM Estimators

By (17) in Theorem 4.1, the objective function of (10) becomes

$$R(\hat{x}_k) = \int_{\mathcal{R}^2} \|x_k - \hat{x}_k\|_p^q \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\ r_{k,ij}\ \delta(x_k^R - s_{k,i})\delta(x_k^I - t_{k,j})}{\sum_{i=1}^{I}\sum_{j=1}^{J}\ r_{k,ij}} dx_k$$

Since $\sum_{i=1}^{I}\sum_{j=1}^{J}\ r_{k,ij}$ is a constant in the above integral, by the definition of $\|\cdot\|_p$ and the basic properties of the $\delta(\cdot)$ function, we can formulate the GLM minimization problem as

$$\hat{x}_k^{opt} = \hat{x}_k^{R\ opt} + j\hat{x}_k^{I\ opt} :$$

$$\min \left\{ \sum_{i=1}^{I}\sum_{j=1}^{J}\ r_{k,ij}\ \left[|s_{k,i} - \hat{x}_k^R|^p + |t_{k,j} - \hat{x}_k^I|^p\right]^{(1\wedge 1/p)q} :\ \hat{x}_k^R,\ \hat{x}_k^I \in \mathcal{R} \right\}. \tag{23}$$

For different values of $p, q \in (0, \infty)$, the above general objective function will take various forms. To simplify the problem, now we only consider the p, q values which belong to the curve specified on the (p, q)-plane such that $(1 \wedge 1/p)q = 1$ (see the figure below).

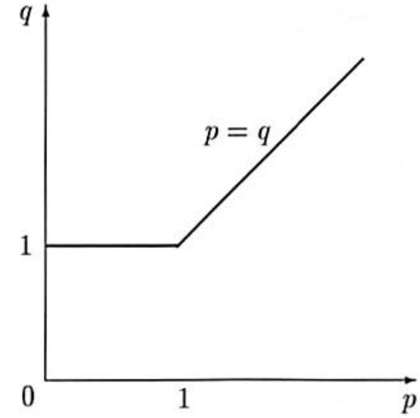


Figure 1: The p and q values on the (p, q)-plane.

Then, the GLM optimization problem becomes

$$\hat{x}_k^{opt} = \hat{x}_k^{R\,opt} + j\hat{x}_k^{I\,opt} :$$

$$\min\left\{\sum_{i=1}^{I}\sum_{j=1}^{J} r_{k,ij}\ \left[|s_{k,i} - \hat{x}_k^R|^p + |t_{k,j} - \hat{x}_k^I|^p\right] :\ \hat{x}_k^R,\ \hat{x}_k^I \in \mathcal{R},\ p > 0\right\}. \qquad (24)$$

This is a 2-D unconstrained nonlinear programming (NLP) problem and the optimal solution $\hat{x}_k^{opt}$ always exits. An example of the above NLP objective function is shown in Figure 2.

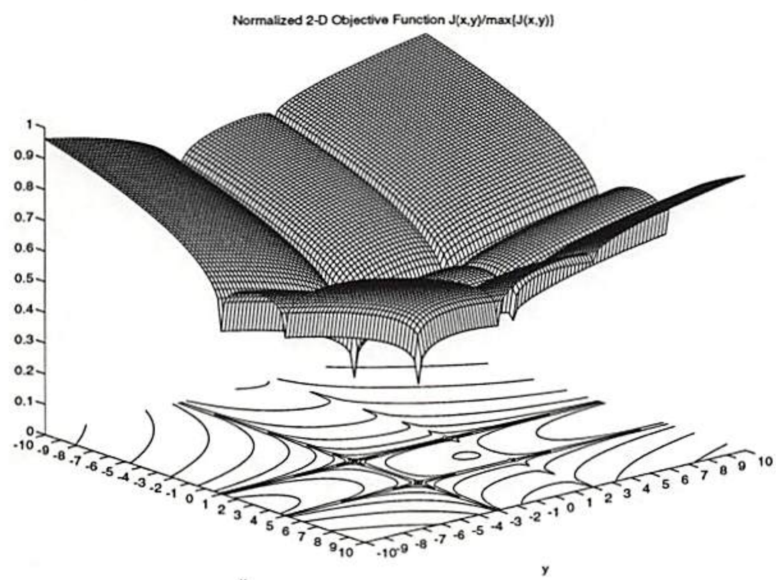


Figure 2: An example of the 2-dimensional normalized objective function $J(x, y) = \sum_{i=1}^{n}[a_i|s_i - x|^p + b_i|t_i - y|^p]^{(1\wedge 1/p)q}$. Parameters: $n = 2$, $p = 0.3$, $q = 1.0$, $a_1 = 5.0$, $a_2 = 4.0$, $b_1 = 1.5$, $b_2 = 4.0$, $s_1 = 1.0$, $s_2 = 5.0$, $t_1 = 2.0$ and $t_2 = -3.0$.

Remarks:

1. In the literature, there has been great amount of research devoted to the NLP problem similar to (24). Generally, the optimal solution (global minimum) $\hat{x}_k^{opt}$ of (24) has no closed form and requires special techniques to approximate the minima (usually the local minima) [16].

2. For $p \to \infty$, $\hat{x}_k^{opt}$ becomes the so-called Min-Max solution. If $p = q = 1$, then $\langle\!\langle \cdot \rangle\!\rangle_{(1,1)}$ is the ordinary $\mathcal{L}^1$-norm and GLM criterion is the same as the absolute value criterion (AVC) [9].

3. The optimization problems (23) and (24) are not equivalent if $(1 \wedge 1/p)q \neq 1$. It is easy to see the influence of the factor $(1 \wedge 1/p)q$ for the 1-D case [9].

Based on (24) and the discussion above, let us consider the interference estimator for the important special cases where $p \in (0,1]$ and $p = 2$, i.e., the optimal $\mathcal{L}^{(p,1)}$-metric, $p \in (0,1]$, and $\mathcal{L}^{(2,2)}$-metric/norm solutions.

a) $\mathcal{L}^{(2,2)}$-metric/norm estimators

Suppose that the $\mathcal{L}^{(2,2)}$-metric/norm exits for the given random variable, by taking the first derivative of (24) with respect to $\hat{x}_k^R$ and $\hat{x}_k^I$ and setting the results to zero, we can easily obtain the following optimal solution, which is the complex version of the general conditional mean (GCM) [9] estimator $E[X_k|Z^k] = E[X_k^R|Z^k] + jE[X_k^I|Z^k]$. From (17) we have the following complex GCM interference estimate

$$\hat{x}_k = E[X_k|Z^k = z^k] = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} r_{k,ij}s_{k,i}}{\sum_{i=1}^{I}\sum_{j=1}^{J} r_{k,ij}} + j\,\frac{\sum_{i=1}^{I}\sum_{j=1}^{J} r_{k,ij}t_{k,j}}{\sum_{i=1}^{I}\sum_{j=1}^{J} r_{k,ij}}, \tag{25}$$

where $s_{k,i}$, $t_{k,j}$ and $r_{k,ij}$ are defined in (18). Since $\overline{X}_{k+1} = E[X_{k+1}|Z^k]$, by (1) the complex GCM predicted interference estimate is

$$\overline{x}_{k+1} = A_k\hat{x}_k, \tag{26}$$

where it is assumed, without loss of generality, that the mean of the i.i.d. noise sequence $\{W_k, k = 1, 2, ...\}$ is zero. Known $\hat{x}_k$ in (25) and by the definition of $s_{k,i}$ and $t_{k,j}$, from (2) we can also estimate the value of the signal V_k using the complex GCM signal estimate

$$\hat{v}_k = z_k - H_k\hat{x}_k = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} r_{k,ij}c_i}{\sum_{i=1}^{I}\sum_{j=1}^{J} r_{k,ij}} + j\,\frac{\sum_{i=1}^{I}\sum_{j=1}^{J} r_{k,ij}d_j}{\sum_{i=1}^{I}\sum_{j=1}^{J} r_{k,ij}}. \tag{27}$$

All the above estimators can be easily implemented by the following algorithm.

Algorithm 4.2 Complex GCM estimators

For $k = 1, 2, ..., N$, do the following steps:

1. Use Algorithm 4.1 to compute the *a posteriori* densities at each time k;
2. Use (25), (26) and (27) to compute estimates $\hat{x}_k$, $\overline{x}_{k+1}$, $\hat{v}_k$.

with initial conditions ($k = 1$): $f_{X_1}(s_{1,i}, t_{1,j})$, $i = 1, 2, ..., I$, $j = 1, 2, ..., J$

///

b) $\mathcal{L}^{(p,1)}$-metric, $0 < p \leq 1$, estimators

Now we will consider how to find the optimal solution of (24) for $0 < p \leq 1$.

Lemma 4.1 *For any two vectors* $\mathbf{x} = (x_1, x_2, ..., x_n)$, $\mathbf{y} = (y_1, y_2, ..., y_n) \in \mathcal{R}^n$ *and non-negative real numbers* $w_i \in \mathcal{R}^+$, $i = 1, 2, ...n$, *the function* $d_p : \mathcal{R}^n \times \mathcal{R}^n \to \mathcal{R}^+$ *defined as*

$$d_p(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} w_i |x_i - y_i|^p, \quad \forall p \in (0, 1] \tag{28}$$

is a metric on $\mathcal{R}^n$.

Proof We first show that the function d_p satisfyes the triangular inequality. Fix the vectors $\mathbf{x} = (x_1, x_2, ..., x_n)$, $\mathbf{y} = (y_1, y_2, ..., y_n)$, $\mathbf{z} = (z_1, z_2, ..., z_n) \in \mathcal{R}^n$,

$$d_p(\mathbf{x}, \mathbf{z}) + d_p(\mathbf{z}, \mathbf{y}) = \sum_{i=1}^{n} w_i [|x_i - z_i|^p + |z_i - y_i|^p].$$

By (4), we have $|x_i - z_i|^p + |z_i - y_i|^p \geq |x_i - y_i|^p$, $p \in (0, 1]$, $\forall i$. Thus,

$$d_p(\mathbf{x}, \mathbf{z}) + d_p(\mathbf{z}, \mathbf{y}) \geq \sum_{i=1}^{n} w_i |x_i - y_i|^p = d_p(\mathbf{x}, \mathbf{y}).$$

It is also easy to verify that d_p meets the other axioms of a metric.

Q.E.D.

Lemma 4.2 [17, 18] *For the fixed set* $S = \{s_i : s_i \in \mathcal{R},\ i = 1, 2, ..., n\}$, *let* $\mathbf{s} = (s_1, s_2, ..., s_n)$ *and* $\mathbf{x} = (x, x, ..., x)$ *be two n-tuples in the metric space* $(\mathcal{R}^n, d_p)$ *defined in Lemma 4.1. Then,* $\exists\, \mathbf{x}_o = (x_o, x_o, ..., x_o) \in \mathcal{R}^n$ *and* $x_o \in S$ *such that*

$$d_p(\mathbf{s}, \mathbf{x}_o) = \min\{d_p(\mathbf{s}, \mathbf{x}_c) : \mathbf{x}_c = (x, x, ..., x) \in \mathcal{R}^n\}.$$

We can extend Lemma 4.2 to get the following theorem which will be used to derive the optimal $\mathcal{L}^{(p,1)}$-metric, $p \in (0,1]$, interference estimator.

Theorem 4.2 *For the fixed sets* $S = \{s_i : s_i \in \mathcal{R},\ i = 1,2,...,n\}$, $T = \{t_i : t_i \in \mathcal{R},\ i = 1,2,...,n\}$ *and non-negative real numbers* $a_i, b_i \in \mathcal{R}^+$, $i = 1,2,...,n$, *define the function* $J : \mathcal{R}^2 \to \mathcal{R}^+$ *as*

$$J(x,y) = \sum_{i=1}^{n} [a_i|s_i - x|^p + b_i|t_i - y|^p], \quad \forall p \in (0,1], \tag{29}$$

then $\exists (x_o, y_o) \in T \times S$ *such that*

$$J(x_o, y_o) = \min\{J(x,y) :\ x, y \in \mathcal{R}\}. \tag{30}$$

Proof: For the given subsets $S, T \subseteq \mathcal{R}$ and any fixed real numbers $x, y \in \mathcal{R}$, define the four n-tuples $\mathbf{s} = (s_1, s_2, ..., s_n)$, $\mathbf{t} = (t_1, t_2, ..., t_n)$, $\mathbf{x}_a = (x, x, ..., x)$, $\mathbf{y}_b = (y, y, ..., y) \in \mathcal{R}^n$. By Lemma 4.1, choose the two metrics on $\mathcal{R}^n$ as

$$d_a(\mathbf{s}, \mathbf{x}_a) = \sum_{i=1}^{n} a_i|s_i - x|^p \quad \text{and} \quad d_b(\mathbf{t}, \mathbf{y}_b) = \sum_{i=1}^{n} b_i|t_i - y|^p, \quad \forall p \in (0,1],$$

then

$$J(x,y) = d_a(\mathbf{s}, \mathbf{x}_a) + d_b(\mathbf{t}, \mathbf{y}_b).$$

So,

$$\min\{J(x,y) :\ x, y \in \mathcal{R}\}$$

$$= \min\{d_a(\mathbf{s}, \mathbf{x}_a) + d_b(\mathbf{t}, \mathbf{y}_b) :\ \mathbf{x}_a = (x, x, ..., x),\ \mathbf{y}_b = (y, y, ..., y) \in \mathcal{R}^n\}$$

$$\geq \min\{d_a(\mathbf{s}, \mathbf{x}_a) :\ \mathbf{x}_a = (x, x, ..., x) \in \mathcal{R}^n\} + \min\{d_b(\mathbf{t}, \mathbf{y}_b) :\ \mathbf{y}_b = (y, y, ..., y) \in \mathcal{R}^n\}.$$

By Lemma 4.2, for the fixed $\mathbf{s}$ and $\mathbf{t}$, $\exists\ \mathbf{x}_o = (x_o, x_o, ..., x_o)$, $\mathbf{y}_o = (y_o, y_o, ..., y_o) \in \mathcal{R}^n$ and $(x_o, y_o) \in S \times T$ such that

$$J(x_o, y_o) = \min\{J(x,y) :\ x, y \in \mathcal{R}\} = d_a(\mathbf{s}, \mathbf{x}_o) + d_b(\mathbf{t}, \mathbf{y}_o),$$

where

$$d_a(\mathbf{s}, \mathbf{x}_o) = \min\{d_a(\mathbf{s}, \mathbf{x}_a) :\ \mathbf{x}_a = (x, x, ..., x) \in \mathcal{R}^n\},$$

$$d_b(\mathbf{t}, \mathbf{y}_o) = \min\{d_b(\mathbf{t}, \mathbf{y}_b) : \mathbf{y}_b = (y, y, ..., y) \in \mathcal{R}^n\}.$$

Q.E.D.

Remark: Theorem 4.2 can be extended to the m-dimensional function $J : \mathcal{R}^m \to \mathcal{R}^+$.

Now, for the optimization problem of (24) with $p \in (0, 1]$, at time k, suppose that $\hat{x}_k^{opt}$ is the optimal solution. Let the data sets be $S_k = \{s_{k,i} : i = 1, 2, ..., n\}$ and $T_k = \{t_{k,i} : i = 1, 2, ..., n\}$. Then, according to Theorem 4.2, we known that $\hat{x}_k^{opt} \in S_k \times T_k$, i.e., on the 2-dimensional complex plane $\mathcal{C}$, the global minimum point belongs to the set of complex data points. So, the optimal $\mathcal{L}^{(p,1)}$-metric, $p \in (0, 1]$, estimates can be easily obtained using the following algorithm, which is the complex version of the absolute value criterion (AVC) [9] estimators.

Algorithm 4.3 Complex AVC estimators

For $k = 1, 2, ..., N$, do the following steps:

1. From the known measurement z_k and the level sets $\{c_i : i = 1, 2, ..., I\}$ and $\{d_j : j = 1, 2, ..., J\}$ of the multi-level complex signal, obtain the measurement set $S_k = \{s_{k,i} : i = 1, 2, ..., I\}$ and $T_k = \{t_{k,j} : j = 1, 2, ..., J\}$ by (18).
2. Use Algorithm 4.1 to compute the *a posteriori* densities at each time k.
3. Evaluate the objective function of (24)

$$J_k(x, y) = \sum_{i=1}^{I} \sum_{j=1}^{J} r_{k,ij} \left[|s_{k,i} - x|^p + |t_{k,j} - y|^p \right]$$

on $S_k \times T_k$. Find the minimum value of $J_k(x, y)$ and the corresponding point in $S_k \times T_k$, say (s_{k,i_o}, t_{k,j_o}) (may not unique), then define the optimal solution $\hat{x}_k^{opt} = s_{k,i_o} + j t_{k,j_o}$.

with initial conditions ($k = 1$): $f_{X_1}(s_{1,i}, t_{1,j},\ i = 1, 2, ..., I,\ j = 1, 2, ..., J$.

///

Remark:

Complex GCM estimator is locally optimal and complex AVC estimator is globally optimal (may not unique). For computational complexity, the later is approximately $I \cdot J$ times more complex.

V. Simulation Results

As mentioned in introduction, the Gaussian distribution is not a proper model for impulsive interference. To simulate the impulsive interference, we assume that the interference has symmetric α-stable (SαS) distribution [10, 11, 12, 13] which has been shown to characterize a general class of man-made and natural impulsive interference. The characteristic function of the SαS distribution is given by $\phi(t) = e^{-\gamma|t|^{\alpha}}$, where $\alpha \in (0, 2]$ is the characteristic exponent and $\gamma \in (0, \infty)$ the dispersion. The signal sequence is chosen to be the most widely used binary random sequence. The proposed algorithms were tested for this special case ($I = J = 2$) and the results were described below.

The system of (1) and (2) are specified as follows. System parameters are fixed to be constants (time invariant), i.e., $A_k = 0.7$, $H_k = 1.0$, $\forall k$. For the complex noise sequence $\{W_k = W_k^R + jW_k^I : k = 1, 2, ..., N\}$, the real part and the imaginary part are both assumed to be SαS processes with the same characteristic exponent $\alpha = 1.2$ and the dispersion $\gamma = 0.5$, and they are mutually independent. The complex signal sequence $\{V_k = V_k^R + jV_k^I : k = 1, 2, ..., N\}$ consists of the two real binary random processes defined as $Pr(V_k^R{=}{+1}) = Pr(V_k^R{=}{-1}) = 0.5$ and $Pr(V_k^I{=}{+1}) = Pr(V_k^I{=}{-1}) = 0.5$, $k = 1, 2, ...N$, where V_k^R and V_k^I are also assumed to be mutually independent. The following approach were taken in the simulation. N samples of the signal sequence $\{V_k : k = 1, 2, ..., N\}$ are randomly generated only once and then fixed. N samples of the interference sequence $\{X_k : k = 1, 2, ..., N\}$ are randomly generated at each independent run while using the same N signal samples. This independent run is repeated M times.

For each independent run, given the measurement data z^k at time k, both complex GCM and complex AVC estimators were applied and the estimated interference $\hat{x}_k$ was computed for both real and imaginary parts. From the experiment we found that these two estimators performed almost equally good under the specified conditions and the results obtained by the complex GCM estimator is plotted (see Figure 3 for a typical realization). To evaluate the performance of the algorithm, the distance between interference realization x_k and its estimate $\hat{x}_k$, and the estimation error $e_k = x_k - \hat{x}_k$ are both shown in Figure 4, which contains the information about the estimation error. The classic performance indices such as the mean, variance and signal-to-noise ratio (SNR) are not used to evaluate the estimators because the second moment of the SαS process is infinity for $\alpha < 2$ and the input "SNR" in our simulation is $-\infty$ dB. Once the estimated interference $\hat{x}_k$ is obtained, it can be removed from the measurement data so that the signal is estimated.

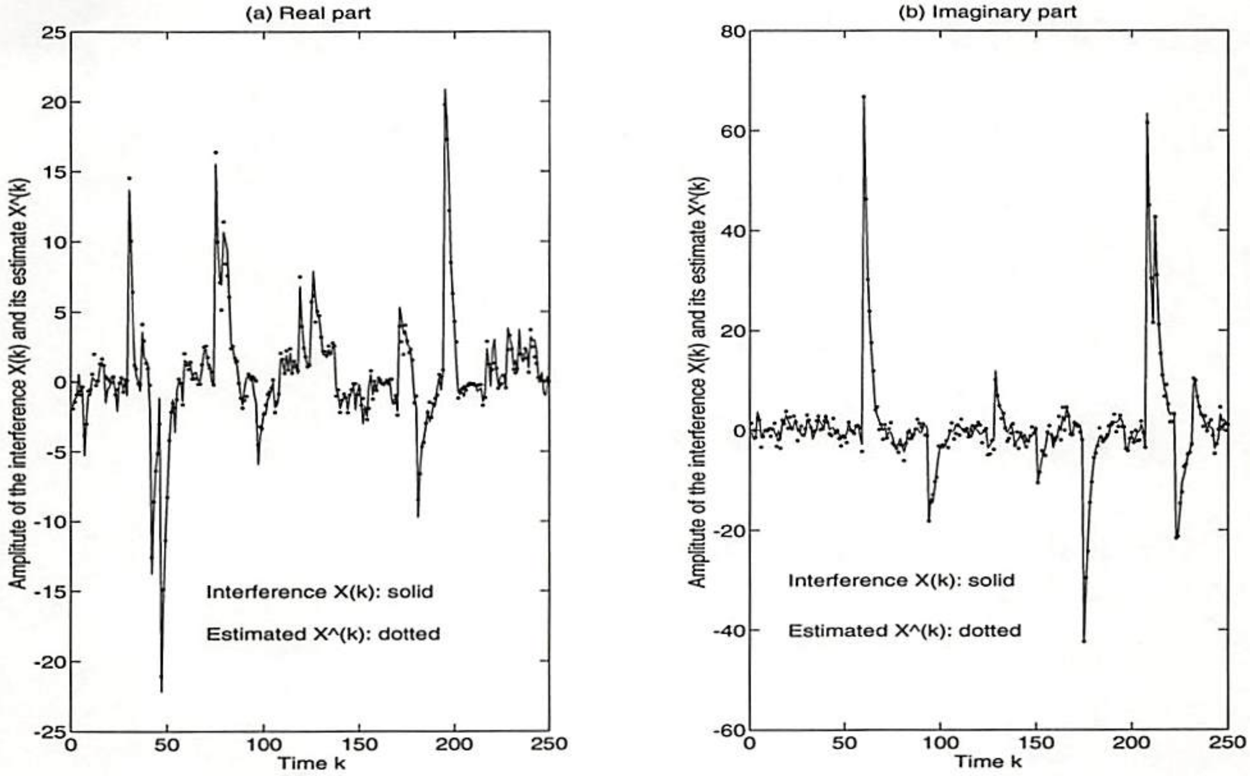
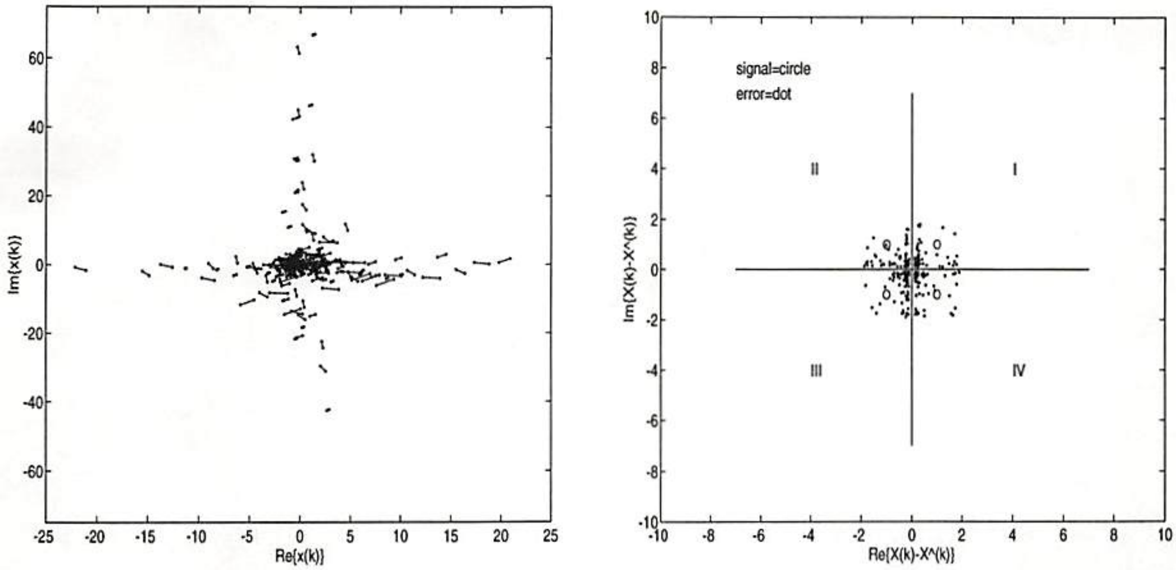


Figure 3: Original interference and its complex GCM estimate: (a) real part, (b) imaginary part.



(a) (b)

Figure 4: (a) Distance between interference x_k and its estimate $\hat{x}_k$, (b) Interference estimation error $e_k = x_k - \hat{x}_k$.

In the simulation, we also compared the complex GCM/AVC estimators with the approximate conditional mean (ACM) estimator [3] and the conventional direct threshold detection (DTD) method [19] by first detecting the signal from the estimate $\hat{v}_k$ and then computing the correct detection probability. The decision regions, according to the nearest neighbor rule, are the four quadrants I-IV for the complex signal (see Figure 4(b)). If the error is within the square region with the four signal points as its vertexes, then there is no detection error. For the DTD scheme, the measurement data z_k is treated as the estimated signal and then fed directly to the threshold detector. The complex version of the ACM estimator, under the given simulation conditions, can be performed by the recursions below for both real(R) and imaginary(I) parts of the time sequence.

$$\hat{X}_k^{R/I} = (A_k - 1)\hat{X}_{k-1}^{R/I} + Z_k^{R/I} - tanh(\frac{Z_k^{R/I} - \hat{X}_{k-1}^{R/I}}{M_k^{R/I}}),$$

$$M_{k+1}^{R/I} = A_k^2 sech^2(\frac{Z_k^{R/I} - \hat{X}_{k-1}^{R/I}}{M_k^{R/I}}) + \sigma_w^2,$$

where $\sigma_w^2 = 2\gamma_w$ when the interference W_k is SαS. The following block diagram (Figure 5) is used to test the complex GCM, AVC, ACM and DTD algorithms. $N = 1000$ samples were taken from each random process realization and the complex GCM, AVC, ACM, DTD algorithms were applied.

$M = 10$ independent runs were executed and the average correct detection probability computed. The results are summarized in Figure 6. Notice that in the worst case when no action is taken, the average correct detection probability will be about 0.25.

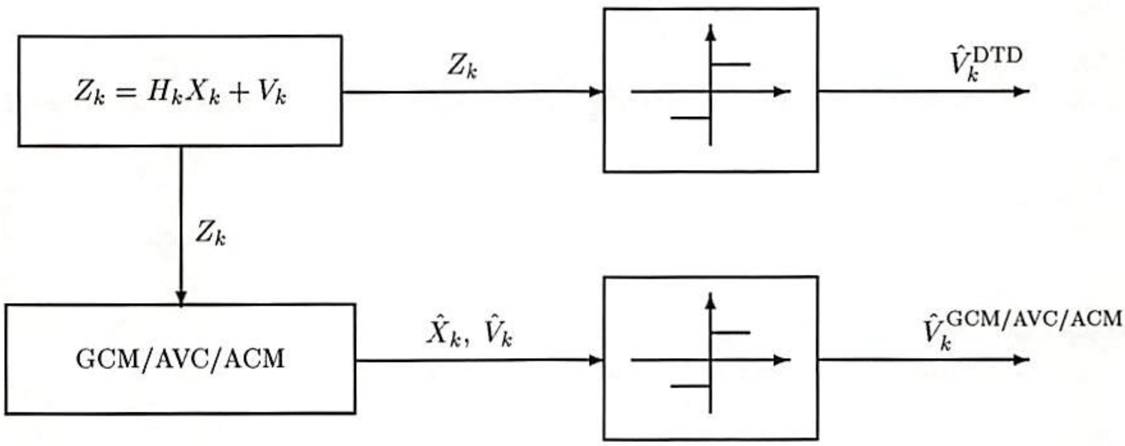
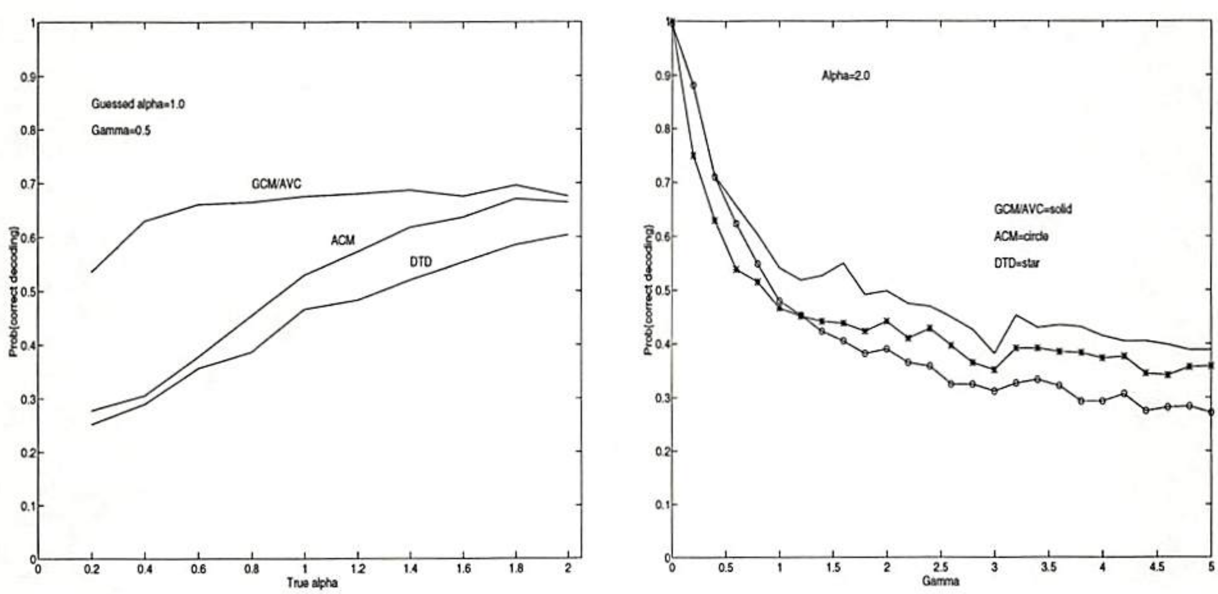


Figure 5: Comparison of the complex GCM, ACM, AVC and DTD algorithms



(a) (b)

Figure 6: (a) Correct detection probability for different values of α, (b) Correct detection probability for different values of γ.

When the output correct detection probability is used as the performance measurement, we observed the following behaviors of the complex GCM/AVC interference estimators (see Figure 6).

- The complex GCM and complex AVC estimators are robust with respect to the value of α. This property is very important since it allows the proposed algorithm to be used in the situation when the impulsiveness of the interference is unknown, which is the case for most of the real physical interference sequences.
- When the interference "power" is weak, i.e, the small γ, all the estimators performed well. And for all the situation, the complex GCM/AVC estimators are superior over the ACM and DTD estimators.

VI. Conclusion

Impulsive interference cancellation problem has been of great concern in the communication community for a long time. Such interference is and will become more and more a major limiting factor in the successful functioning of communication systems which play the crucial role in the modern information era.

To attack the interference mitigation problem in a broad sense, we introduced the general $\mathcal{L}^{(p,q)}$-metric space and derived the $\mathcal{L}^{(p,q)}$-metric optimality criterion. For the interference, we do not constrain ourselves to the specific model. Our new approach to the interference problem has the following advantages. First, the interference can take arbitrary density function, which give us an unlimited flexibility to model the real physical processes as accurately as possible. Second, the *a posteriori* density function has analytical closed form so that the further study of the properties of the estimators and other theoretical concerns become easier. Third, the algorithm to compute the *a posteriori* density function and the corresponding estimators is recursive and straightforward, which is very crucial in the implementation. Forth, the interference algorithm is robust with respect to the impulsiveness of the interference which means the *a priori* of the interference can be unknown. All of these features will allow us to apply the proposed method to a wide range of problems in communications and signal processing.

References

[1] H. V. Poor, "Signal Processing for Wideband Communications," *IEEE Information Theory Society Newsletter*, vol.42, No.2, pp.1-10, June 1992.

[2] R. Vijayan and V. Poor, "Nonlinear Techniques for Interference Suppression in Spread-Spectrum Systems," *IEEE Trans. Commun.*, vol.38, no.7, pp.1060-1065, July 1990.

[3] C. J. Masreliez, "Approximate Non-Gaussian Filtering with Linear State and Observation Relations," *IEEE Trans. Automat. Contr.*, vol.AC-20, pp.107-110, Feb. 1975.

[4] D. Middleton, "Statistical-Physical Models of Electromagnetic Interference," *IEEE Trans. Electromagn. Compat.*, vol. EMC-19, pp.106-127, Aug. 1977.

[5] A. D. Spaulding and D. Middleton, "Optimum Reception in an Impulsive Interference Environment, Part I: Coherent Detection, Part II: Incoherent Reception," *IEEE Trans. Commun.*, vol. COM-25, pp.910-934, Sept. 1977.

[6] S. M. Zabin and H. V. Poor, "Recursive Algorithms for Identification of Impulsive Noise Channels," *IEEE Trans. Inform. Theory*, vol.36, no.3, pp.559-578, May 1990.

[7] H. W. Sorenson and D. L. Alspach, "Recursive Bayesian Estimation Using Gaussian Sums," *Automatica*, vol.7, pp.465-479, 1971.

[8] J. Shen and C. L. Nikias, "Recursive Computation of *a Posteriori* Density Functions for Arbitrary I.I.D. State Noise and Its Application to Impulsive Interference Mitigation," *Conf. Rec. of 27th Asilomar Conf. on Signals, Systems & Computers*, Pacific Grove, CA, Nov. 1-3, 1993.

[9] J. Shen and C. L. Nikias, "Estimation of Multi-Level Digital Signals in the Presence of Arbitrary Impulsive Interference," *IEEE Trans. Signal Process.*, to appear in January 1995.

[10] M. Shao and C. L. Nikias, "Signal Processing with Fractional Lower Order Moments: Stable Processes and Their Applications," *Proceedings of the IEEE*, vol.81, no.7, pp.986-1010, 1993.

[11] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*, New York, John Wiley & Sons, 1995.

[12] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes*, New York: Chapman & Hall, 1994.

[13] V. Zolotarev, *One-Dimensional Stable Distributions*, Providence, RI: American Mathematical Society, 1986.

[14] F. D. Neeser and J. L. Massey, "Proper Complex Random Processes with Applications to Information Theory," *IEEE Trans. Info. Theory*, vol.39, no.4, pp.1293-1302, 1993.

[15] M. Loève, *Probability Theory*, 3rd Edition, Princeton, NJ: D. Van Nostrand, 1963.

[16] R. Gonin and A. H. Money, *Nonlinear L_p-Norm Estimation*, New York: Marcel Dekker, 1989.

[17] I. Barrodale and F. D. K. Roberts, "Application of Mathematical Programming to L_p Approximation," *Nonlinear Programming*, J.B.Rosen, O.L.Mangasarian and K.Ritter (eds.), New York: Academic Press, 1970.

[18] J. Astola and Y. Neuvo, "Optimal Median Type Filters for Exponential Noise Distributions," *Signal Processing*, 17, pp.95-104, 1989.

[19] H. Tong, *Threshold Models in Non-linear Time Series Analysis*, Springer Lecture Notes in Statistics, 21, 1983.