ROBUST WIENER FILTERS FOR CORRELATED SIGNALS AND NOISE. (U)

1980  G. MOUSTAKIDES, S. A. KASSAN

AFOSR-77-3154

AFOSR-TR-80-0426

UNCLASSIFIED
**ROBUST WIENER FILTERS FOR CORRELATED SIGNALS AND NOISE**

**AUTHOR(s)**
G. Moustakides and S. A. Kassam

**PERFORMING ORGANIZATION NAME AND ADDRESS**
University of Pennsylvania
Systems Engineering Dept.
Philadelphia, PA 19104

**CONTROLLING OFFICE NAME AND ADDRESS**
Air Force Office of Scientific Research/NM
Bolling AFB, Washington, DC 20332

**REPORT DATE**
1980

**NUMBER OF PAGES**
Six

**DISTRIBUTION STATEMENT (of this Report)**
Approved for public release; distribution unlimited.

**DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)**

**SUPPLEMENTARY NOTES**
Conference on Information Sciences and Systems, Princeton University, March 1980; paper to appear in proceedings

**KEY WORDS (Continue on reverse side if necessary and identify by block number)**
Wiener Filters, Robust Filters, Spectral Density

**ABSTRACT (Continue on reverse side if necessary and identify by block number)**
We consider the Wiener filtering problem, when the cross spectral density matrix of the signal and noise is not exactly known. We seek filters which are saddle point solutions for the criterion of performance (mean square error, MSE) over the classes of allowable density matrices. Solutions for various classes are given.
ABSTRACT

We consider the Wiener filtering problem, when the cross spectral density matrix of the signal and noise is not exactly known. We obtain filters which are saddle point solutions for the criterion of performance (mean square error, MSE) over the classes of allowable density matrices. Solutions for various classes are given.

I. INTRODUCTION

In classical Wiener filtering, we need to assume exact knowledge of the spectral density matrix of signal and noise. In many applications this assumption of exact knowledge is unreasonable. A more realistic assumption is that our matrix belongs to a class of density matrices. This class can be defined according to our knowledge of the true density matrix. For this problem we will derive a filter that performs in an optimum way over the whole class.

Nahi and Weiss [1, 2] derived the bounding filter. This filter is a Wiener filter for some density matrix $D_b$ and if it is applied to any matrix from the class the MSE error is bounded by the minimum MSE for $D_b$. Because $D_b$ does not usually belong to the class, there is no matrix in the class that can reach the bound of the MSE. This means that there is a possibility for better performance.

Kassam and Lin [3] derived the robust Wiener filter, when signal and noise are uncorrelated (density matrix diagonal). This filter sets a bound on the MSE and the bound can be reached by some matrix from the class. In [4] and [5], Poor generalized some of this work. The present paper extends the above idea to the correlated signal and noise case.

Presence of correlation is possible in many applications. An example is a multipath channel with a strong signal component, weak unwanted multipath signals and regular noise. The total "noise" is correlated with the signal.

II. ROBUST WIENER FILTERS

Let us assume that our processes are real, wide sense stationary and zero mean and the noise process is additive. We also assume that there exists a spectral density matrix $D$ for the signal, noise processes, given by:

$$ D = \begin{bmatrix} D_s(w) & D_{sn}(w) \\ D_{sn}(w) & D_n(w) \end{bmatrix} $$

where $s$ is for signal, $n$ for noise and $\bar{\cdot}$ for complex conjugate. The properties that characterize such a matrix $D$ for real processes are:

1. $D_s(w)$, $D_n(w)$ real, even, non-negative functions
2. $|D_{sn}(w)|^2 \leq D_s(w)D_n(w)$

So $D$ is non-negative definite, and diagonal elements are even functions of $w$.

Given random processes $s(t)$ and $n(t)$ with density matrix $D$ and a filter $h(t)$ with Fourier transform $H(w)$, the MSE for signal estimation using this filter is

$$ e(D,H) = E \left[ \int |h(v)x(t-v)|dv \right]^2 $$

We have the saddle point relation

$$ e(D,H) = e(D_{op}(H),H) $$

where $x(t) = s(t) + n(t)$ and $R_{ij}(t) = \text{cov}(x(t),x(t-t))$ is the cross correlation between $i$ and $j$. Using Parseval's theorem, (2) can be written as:

$$ e(D,H) = \frac{1}{2\pi} \int [D_s(w)-2D_{sn}(w)+D_n(w)]dw $$

The optimum filter for $D$ is given by

$$ H_{op} = \frac{D_{sn}(w)}{D_n(w)} $$

If we substitute (4) into (3) we have that the optimum MSE is given by:

$$ e_{op}(D) = \frac{1}{2\pi} \int \frac{D_s(w)-2D_{sn}(w)+D_n(w)}{D_n(w)} dw $$

2.1 Definition of Robust Filter

Assume that a class $A$ of density matrices is given. A robust filter $H_r$ is defined by the following properties:

a. $H_r$ is an optimum filter for some matrix $D^r cA$

b. For any $D c A$ we have: $e(D,H_r) \leq e(D^r,H_r)$

Combining a and b we have the saddle point relation

$$ e(D,H_r) \leq e_{op}(D^r) \leq e(D^r,H_r) $$
for any $DcA$ and any filter $H$.

2.2 Theorem 0.

Let $A$ be a convex class of density matrices. Then the pair $(D^*, H^*)$ is a saddle point solution for MSE over class $A$ and the class of all linear filters (it satisfies (6)), if and only if:

$$e_{op}(D^*) = \max_{DcA} e_{op}(D)$$

The proof is given in appendix A.

Theorem 0 is the key point in our search for the robust filter, because based on it we have only to maximize $e_{op}(D)$.

2.3 Maximization of $e_{op}(D)$

From (5) if we write the error in terms of $s$ and $n$ we get:

$$e_{op}(D) = \frac{1}{2\pi} \int \frac{D_s(w)D_n(w) - |D_{sn}(w)|^2}{D_s(w)D_n(w) + 2|D_{sn}(w)|^2} dw$$

(7)

For every $w$ and for given $D_s(w), D_n(w), |D_{sn}(w)|$, the worst $Re[D_{sn}(w)]$ is $-|D_{sn}(w)|$, because it minimizes the denominator. This gives:

$$e_{op}(D) = \frac{1}{2\pi} \int \frac{D_s(w)D_n(w) - |D_{sn}(w)|^2}{D_s(w)+D_n(w)+2Re[D_{sn}(w)]} dw$$

(8)

For given $D_s(w), D_n(w)$, expression (8) as a function of $|D_{sn}(w)|$ is:

a. increasing for:

$$0 \leq |D_{sn}(w)| \leq \min(D_s(w), D_n(w))$$

b. decreasing for:

$$\min(D_s(w), D_n(w)) \leq |D_{sn}(w)| \leq \sqrt{D_s(w)D_n(w)}$$

From condition a and b we can say, given $D_s(w), D_n(w)$, the worst $|D_{sn}(w)|$ is the one that is as close as possible to $\min(D_s(w), D_n(w))$. The worst $D_{sn}(w)$ is $D_{sn}(w) = -|D_{sn}(w)|$.

III. APPLICATIONS

1. $D_s(w), D_n(w)$ given with upper and lower bounds on $|D_{sn}(w)|$.

Let $0 \leq L(w) \leq |D_n(w)| \leq U(w)$, with $L(w), U(w)$ given. If we define:

$$A(w) = \min(D_s(w), D_n(w))$$

then the worst-case characteristic $|D^*_{sn}(w)|$ is given by

$$|D^*_{sn}(w)| = \begin{cases} L(w) & \text{if } A(w) \leq L(w) \\ A(w) & \text{if } L(w) < A(w) \leq U(w) \\ U(w) & \text{if } U(w) \leq A(w) \end{cases}$$

where $A(w)$ defined in (9).

From section 2.2 we have that $D^*_{sn}(w) = -D^*_{sn}(w)$.

Figure 1 illustrates this case.

2. Upper and lower bounds on $D_s(w)$ and $D_n(w).

We assume that bounds $L_s(w), U_s(w)$ for $D_s(w)$ are given:

$$0 \leq L_s(w) \leq D_s(w) \leq U_s(w) \quad i=s,n$$

In addition we will assume knowledge of the total power of signal and noise,

$$\int D_s(w) dw = 2\sigma^2$$

where $\sigma^2$ are known.

A. If there are no bounds on $|D_{sn}(w)|$, it can reach the value $\min(D_s(w), D_n(w))$. Under this condition the error is:

$$e_{op}(D) = \frac{1}{2\pi} \int \min(D_s(w), D_n(w)) dw$$

(10)

Because of the power constraints, there are several sub-cases. First we will give some definitions.

$$A_s(w) = \min(U_s(w), \max(L_s(w), L_n(w)))$$

$$A_n(w) = \min(U_n(w), \max(L_s(w), U_n(w)))$$

$$B_s(w) = \min(U_s(w), \max(L_s(w), U_n(w)))$$

$$B_n(w) = \min(U_n(w), \max(L_s(w), L_n(w)))$$

Figure 2 illustrates the definition of $A_s(w), B_s(w)$.

1. $A_s(w), B_s(w)$ given with upper and lower bounds on $|D_{sn}(w)|$.

Let $0 \leq L_s(w) \leq |D_n(w)| \leq U_s(w)$, with $L_s(w), U_s(w)$ given. If we define:

$$A_w = \min(D_s(w), D_n(w))$$

then the worst-case characteristic $|D^*_{sn}(w)|$ is given by

$$|D^*_{sn}(w)| = \begin{cases} L_s(w) & \text{if } A_s(w) \leq L_s(w) \\ A_s(w) & \text{if } L_s(w) < A_s(w) \leq U_s(w) \\ U_s(w) & \text{if } U_s(w) \leq A_s(w) \end{cases}$$

where $A_s(w)$ defined in (9).

From section 2.2 we have that $D^*_{sn}(w) = -D^*_{sn}(w)$.

Figure 1 illustrates this case.

A2. 
\[ \int \tilde{A}_s(w) dw > 2\sigma_s^2 \quad \int \tilde{A}_n(w) dw < 2\sigma_n^2 \]

\[ D_{s n}^T(w) \text{ is as in case A1} \]

\[ D_{s n}^T(w) = \begin{cases} U_n(w) \text{ if } U_n(w) = A_n(w) \\ \ell_n(w) \text{ otherwise} \end{cases} \]

where \( \ell_n(w) \) arbitrary function with \( \ell_n(w) \geq A_n(w) \).

\[ e_{\text{op}}(D^T) = \sigma_s^2 - \frac{1}{2\pi} \int \left( L_s(w) - \min(L_s(w), L_n(w)) \right) dw \]

A3. 
\[ \int \tilde{B}_s(w) dw > 2\sigma_s^2 > \int \tilde{A}_s(w) dw \]

\[ \int \tilde{B}_n(w) dw > 2\sigma_n^2 > \int \tilde{A}_n(w) dw \]

also assume \( 2\sigma_n^2 - \int \tilde{A}_n(w) > 2\sigma_n^2 > \int \tilde{A}_s(w) dw \)

\[ D_{s n}^T(w) = \begin{cases} A_s(w) \text{ if } A_s(w) = B_s(w) \\ \ell_s(w) \text{ otherwise} \end{cases} \]

\[ D_{n s}^T(w) = \begin{cases} A_n(w) \text{ if } A_n(w) = B_n(w) \\ \ell_n(w) \text{ otherwise} \end{cases} \]

Where \( A_s(w) \leq L_s(w) \leq B_s(w) \) and \( D_s^T(w) \leq \ell_s(w) \)

\[ e_{\text{op}}(D^T) = \sigma_s^2 - \frac{1}{2\pi} \int \left( L_s(w) - \min(U_s(w), L_n(w)) \right) dw \]

A4. 
\[ \int \tilde{B}_n(w) dw > 2\sigma_n^2 > \int \tilde{A}_n(w) dw, 2\sigma_n^2 > \int \tilde{B}_n(w) dw \]

\[ D_{n s}^T(w) \text{ as in case A3.} \]

\[ D_{n s}^T(w) = \begin{cases} U_n(w) \text{ if } U_n(w) = A_n(w) \\ \ell_n(w) \text{ otherwise} \end{cases} \]

Where \( L_n(w) > D_{n s}^T(w) \)

\[ e_{\text{op}}(D^T) = \sigma_s^2 - \frac{1}{2\pi} \int \left( L_s(w) - \min(U_s(w), L_n(w)) \right) dw \]

A5. 
\[ 2\sigma_n^2 > \int \tilde{B}_n(w) dw, 2\sigma_n^2 > \int \tilde{B}_n(w) dw \]

\[ D_{s n}^T(w) = \begin{cases} U_s(w) \text{ if } U_s(w) = B_s(w) \\ \ell_s(w) \text{ otherwise} \end{cases} \]

\[ D_{n s}^T(w) = \begin{cases} U_n(w) \text{ if } U_n(w) = B_n(w) \\ \ell_n(w) \text{ otherwise} \end{cases} \]

Where \( \ell_s(w) > B_s(w) \) and \( \ell_n(w) > B_n(w) \).

\[ e_{\text{op}}(D^T) = \frac{1}{2\pi} \int \min(U_s(w), U_n(w)) dw \]

In the cases above, interchanging the roles of \( s \) and \( n \) in the conditions gives us similar results. The proof for cases A1 and A3 can be found in Appendix B.

B. Suppose we are given a function \( R(w) \) such that

\[ 0 \leq |D_{sn}(w)| \leq R(w) \leq \min(L_s(w), L_n(w)) \]

Because of the maximization problem \( |D_{sn}(w)| \) has to be as close as possible to \( \min(D_s(w), D_n(w)) \), so

\[ |D_{sn}^T(w)| = R(w). \] To find the pair \( D_{s n}^T(w), D_{n s}^T(w) \) notice that:

\[ e_{\text{op}}(D) = \frac{1}{2\pi} \int D_s(w) \cdot D_n(w) - R^2(w) \]

\[ = \frac{1}{2\pi} \int R(w) dw + \frac{1}{2\pi} \int S(w) \cdot N(w) \]

(11)

\[ Where \ S(w) = D_s(w) - R(w) \geq 0 \text{ and } N(w) = D_n(w) - R(w) \leq 0. \] To maximize (11) it is enough to maximize the second term. But this term is the expression for the minimum MSE for uncorrelated signal \( S(w) \) and noise \( N(w) \), with

\[ L_s(w) - R(w) \leq S(w) \leq U_s(w) - R(w) \]

\[ L_n(w) - R(w) \leq N(w) \leq U_n(w) - R(w) \]

and

\[ \int S(w) dw = 2\sigma_s^2 - \int R(w) dw \]

\[ \int N(w) dw = 2\sigma_n^2 - \int R(w) dw. \]

This problem has been solved in [3] and gives the solution to the present case.

3. Given classes for \( D_s(w), D_n(w) \)

We assume again knowledge of the total power of signal and noise. When there is no restriction on \( D_{sn}(w) \) we have seen that the error is given by (10) and it can be written in the following way:

\[ e_{\text{op}}(D) = \frac{1}{2\pi} \int S(x) \cdot D_n(w) \cdot dw \]
where
\[
g(x) = \begin{cases} 
  x & \text{for } 0 \leq x \leq 1 \\
  1 & \text{for } x > 1 
\end{cases}
\]
and
\[
\frac{D_s(w)}{D_n(w)} = \frac{D_s(w)}{D_n(w)}
\]

\(g(x)\) is a convex function of \(x\) for \(x \geq 0\). So (12) can be used as a measure of distance between \(D_s(w)\) and \(D_n(w)\). For given classes of \(D_s(w)\) and \(D_n(w)\) we want the pair \(D'_s(w), D'_n(w)\) that has the maximum distance. It turns out as we can see from \([4,6]\) and \([7]\) that this pair does not depend on the form of \(g(x)\). As long as \(g(x)\) is any convex function the pair is always the same. Solutions to this problem are also given in these works for various classes.

**IV. NUMERICAL RESULTS**

We will assume the case when signal and noise are given as in Figure 3 and there is no restriction on \(|D_{sn}(w)|\). The robust filter is simply:

\[
H_1(w) = \begin{cases} 
  1 & \text{when } D_s(w) \geq D_n(w) \\
  0 & \text{otherwise}
\end{cases}
\]

It turns out that this filter has the same error performance for any \(D_{sn}(w)\). If we use instead of \(H_1(w)\) the filter

\[
H_2(w) = \frac{D_s(w)}{D_s(w)+D_n(w)}
\]

assuming that signal and noise are uncorrelated, then the error is given by (3) and it becomes maximum for \(D_{sn}(w) = -D_s(w)D_n(w)\).

In Table 1 are given some numbers for different \(b\). The error is \(\epsilon\), when we use \(H_2(w)\) and \(e_u\) is the worst error when we use \(H_1(w)\). In the third column we can see the percentage of performance improvement. Also shown is the optimum error \(e_o\) when \(D_{sn}(w) = 0\).

**APPENDIX A**

Before proving Theorem O, we will prove the following lemma.

**Lemma 1.** Let \(D'_s, D'_n, D, D''_s, D''_n\) define \(D^a = -(1-e)D + eD''\) \(0 \leq e \leq 1\) then the expression:

\[
G(e,w) = \frac{D'_s(w)D'_n(w) - D_s(w)D_n(w) - \|D'_{sx}(w)\|^2}{D_s(w)}
\]

is a convex function of \(e\).

**Proof.** It is sufficient to prove that:

\[
G(e,w) = (1-e)G(0,w) + eG(1,w).
\]

Subtracting each side from \(D'_s(w)\) we have to prove that

\[
\frac{|D'_{sx}(w)|^2}{D'_s(w)} < (1-e) \frac{|D'_{sx}(w)|^2}{D_s(w)} + e \frac{|D''_{sx}(w)|^2}{D'_s(w)} \text{ or}
\]

\[
|D(1-e)D'_{sx}(w) + eD''_{sx}(w)|^2 \leq \frac{1}{2} \left( (1-e) \frac{|D'_{sx}(w)|^2}{D_s(w)} + e \frac{|D''_{sx}(w)|^2}{D'_s(w)} \right).
\]

But this is the Schwartz inequality.

**Proof of Theorem O.**

The only if part is easy. From (6)

\[
e(D,H_r) \leq e(D',H_r) = e_0(D')
\]

but \(e(D,H_r) \geq e_0(D)\) so \(e_0(D') \leq e_0(D)\).

To prove the if part, define \(D^a = -(1-e)D + eD\), \(0 \leq e \leq 1\) such that \(D^a\). Because of the lemma and the fact that

\[
e_0(D') = \frac{1}{2 \pi} \int_D G(e,w) \operatorname{dw}
\]

\(e_0(D')\) is a convex function of \(e\). So:

\[
e_0(D') = (1-e)e_0(D') + e e_0(D)\text{ and}
\]

\[
0 \geq \frac{e_0(D') - e_0(D)}{e} \geq e_0(D) - e_0(D')
\]

Because \(\frac{e_0(D') - e_0(D)}{e}\) is monotonic with respect to \(e\) and bounded as we can see from (13) its limit exists as \(e\to^+\). But:

\[
e_0(D') = e_0(D') = \frac{1}{2 \pi} \int_D G(e,w) - G(0,w) \operatorname{dw}
\]

Now \(\frac{d}{de} G(e,w)\) is also monotonic with respect to \(e\).

From (13) and (14), we have

\[
0 \geq \lim_{e\to^+} \frac{1}{2 \pi} \int_D G(e,w) - G(0,w) \operatorname{dw} =
\]

\[
\frac{1}{2 \pi} \int_{e^o}^{e_\infty} G(e,w) - G(0,w) \operatorname{dw} =
\]

\[
\frac{1}{2 \pi} \int_{e^o}^{e_\infty} \frac{d}{de} G(e,w) \operatorname{dw} +
\]

But the last expression after we take the derivative and make some manipulations, equals \(e(D,H_r) - e_0(D')\). So \(e_0(D') \leq e(D,H_r)\).
APPENDIX B

For the second application we will outline the proofs for subcases A1 and A3. In a similar way, we can prove the rest of the cases.

**Lemma 2.** If \(a, b, c\) non-negative numbers with \(a \geq c\) then:

\[
a - c \geq \min(a, b) - \min(b, c)
\]

**Proof.**

\[
a - c \geq 0 \geq \min(a, b) - b = \min(a, b) - \min(b, c)
\]

From (10) we have that the error is

\[
e_{op}(D) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[ \min(L_s, D_n(w)) \right] dw
\]

**Case A1.**

\[
\min(D_s, D_n) = \min(L_s, D_n) + [\min(D_s, D_n) - \min(L_s, D_n)]
\]

(using Lemma 2) \(\leq \min(L_s, D_n) + D_s - L_s\)

(\(\leq \min(L_s, D_n) + D_s + D_n - L_s - L_n\)

So

\[
e_{op}(D) \leq a^2 + b^2 - \frac{1}{2\pi} \int_{-\infty}^{\infty} [L_s(w) + L_n(w) - \min(L_s(w), L_n(w))] dw
\]

And we have equality when \(D = D^r\).

**Case A3.**

\[
\min(D_s, D_n) = \min(L_s, D_n) + [\min(D_s, D_n) - \min(L_s, D_n)] \leq \min(L_s, D_n) + D_s - L_s
\]

\[
e_{op}(D) \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} [L_s(w) - \min(L_s(w), U_n(w))] dw
\]

with equality when \(D = D^r\).

**REFERENCES**


**FIGURES AND TABLES**

Figure 1 Worst-case \(|D^r_s|\) for case 1, section III
Figure 2 Illustration of functions defined in case 2A, section III
--- for $A_s(w)$
-- for $B_s(w)$

Figure 3 Example spectra of section IV

| Table 1 Performance Comparison of Robust and Nominally Optimum Filter for Example of Figure 3. (No bound on $|D_m(w)|$.) |
|---|---|---|---|---|
| $b$ | $e_R$ | $e_u$ | $z$ | $e_u^b$ |
| 0.10 | 0.09 | 0.12 | 34.5 | 0.07 |
| 1.00 | 0.50 | 0.62 | 25.7 | 0.33 |
| 5.00 | 0.83 | 1.10 | 31.7 | 0.62 |
| 10.0 | 0.91 | 1.20 | 34.5 | 0.73 |