

SRC TR 87-155

**Minimax Robust Matched Filters
for Noise Uncertainty Within 2-
Alternating Capacity Classes**

by

E. Geraniotis

MINIMAX ROBUST MATCHED FILTERS FOR NOISE UNCERTAINTY WITHIN 2-ALTERNATING CAPACITY CLASSES

Evangelos Geraniotis

Department of Electrical Engineering
and Systems Research Center
University of Maryland
College Park, MD 20742

ABSTRACT

In this paper, we address the problem of designing matched filters which are robust against uncertainty in the statistics of the noise process. The design is based on a game-theoretic approach in which a filter is sought that has the maximum worst-case output signal-to-noise ratio possible over the class of allowable statistics, that is the design is maximin signal-to-noise ratio. The problem is formulated and solved for both discrete-time and continuous-time matched filters with uncertainty in either the autocorrelation function or the spectral measure of the noise. For uncertainty models determined by 2-alternating Choquet capacities explicit solutions are obtained which are characterized by the Huber-Strassen derivative of the capacity generating the class with respect to a Lebesgue-like measure on a suitable interval.

I. Introduction

Robust signal processing techniques have received considerable attention in the last 15 years (see the tutorial in [1]). In particular, robust matched filtering problems have been formulated and partially solved in [2] for continuous-time and [3] for discrete-time observations.

In these problems there is uncertainty in the statistics of either the signal, or of the noise, or of both the signal and the noise. Several uncertainty models for the signal and the noise were considered in [2] and [3].

In this paper we focus on uncertainty in the noise autocorrelation function (time domain) or the noise spectral measure (frequency domain) and consider uncertainty classes determined by 2-alternating Choquet capacities [4]-[6] which were not examined in [2]-[3]. Both discrete-time and continuous-time formulations of the robust matched filtering problem with noise uncertainty within capacity classes are considered.

The 2-alternating Choquet capacities classes include several useful uncertainty models like the ϵ -contaminated class [4], the total variation class [4], the band class [5] and the p-point class [6], which have been very popular among the statisticians.

The robust matched filtering problem for noise uncertainty involves the identification of a worst-case noise statistic over the allowable uncertainty class and the derivation of the filter matched to this worst-case noise. Then, when this "robust" filter is used, for any other noise statistic in the uncertainty class the signal-to-noise ratio is guaranteed to be better than for the worst-case statistic.

This paper is organized as follows. In Section II the notation, the basic concepts and some general results for uncertainty classes generated by 2-alternating capacities are cited. In Section III four problems of robust matched filtering with noise uncertainty are

formulated. In Section IV a complete characterization of maximin robust matched filters for the aforementioned four problems is provided.

II. Uncertainty Classes Generated by Choquet Capacities

Suppose that Ω is a compact set and F is the σ -algebra generated by its subsets. We assume that the measures m are only known to lie in a convex class generated by a Choquet 2-alternating capacity [7]

$$M_v = \{m \in M \mid m(A) \leq v(A), A \in F, m(\Omega) = v(\Omega)\} \quad (1)$$

where M denotes the class of all measures on (Ω, F) and v is a 2-alternating capacity on (Ω, F) .

A Choquet 2-alternating capacity [7] on (Ω, F) is a finite set function, which is increasing, continuous from below, continuous from above on closed sets, and satisfies $v(\emptyset) = 0$ and $v(A \cap B) + v(A \cup B) \leq v(A) + v(B)$ for all $A, B \in F$. Notice that any finite measure v is a 2-alternating capacity; in this case the uncertainty class generated by (1) reduces to $M_v = \{v\}$.

As an example of a 2-alternating capacity we cite the ϵ -contaminated mixture model [4] defined by

$$M_v = \{m \in M \mid m(A) = (1-\epsilon)m_0(A) + \epsilon\bar{m}(A), A \in F, \bar{m}(\Omega) = m_0(\Omega)\}, \quad (2)$$

where m_0 is a known measure and the number ϵ in $(0,1)$ is the degree of uncertainty in the model. Eq. (2) can be expressed in the form (1) if we set

$$v(A) = (1-\epsilon)m_0(A) + \epsilon. \quad (3)$$

The total variation model [4], the band class [5], and the p-point class [5] are other popular examples of uncertainty classes generated by 2-alternating capacities.

In the sequel we will need the following results which are special cases for the theory developed by Huber and Strassen in [7]:

Lemma 1: If v is a 2-alternating capacity on (Ω, F) , M_v is a convex class of probability measures determined by it as in (1), and λ denotes the Lebesgue measure on Ω , then there exists a unique Lebesgue measurable function $\pi_v: \Omega \rightarrow [0, \infty]$ with the defining property that for all $x \in [0, \infty]$ and A_x defined by $A_x = \{\pi_v > x\}$

$$x \lambda(A_x) + v(A_x^c) \leq x \lambda(A) + v(A^c), A \in F. \quad (4)$$

Furthermore there exists a measure \hat{m} in M_v such that for all $x \in [0, \infty]$

$$\hat{m}(\{\pi_v \leq x\}) = v(\{\pi_v \leq x\}), \quad (5)$$

which means that \hat{m} makes π_v **stochastically smallest** over all m in M_v , and π_v is equal to $d\hat{m}/d\lambda$, the generalized Radon-Nikodym (R-N) derivative of \hat{m} with respect to λ .

The function π_v is termed the **Huber-Strassen derivative** of v with respect to λ (v may not be a measure). The measure \hat{m} singled out by Lemma 1 is termed the **least-favorable** measure of the class M_v . For the ϵ -contaminated example above we have

$$\frac{d\hat{m}}{d\lambda}(u) = \max \left\{ (1-\epsilon) \frac{dm_0}{d\lambda}(u), c \right\}, \quad (6)$$

where c is chosen so that $\hat{m}(\Omega) = m_0(\Omega)$.

We would like to mention at this point that if Ω is a discrete set, then \hat{m} becomes a cumulative distribution function (cdf) and all the results involving the capacities described above still hold, provided that we replace the integrals with sums and the R-N derivatives with probability mass functions (pmf's). This duality follows from an appli-

cation of the results of [7] (especially Lemmas 3.1 and 3.2 and Theorem 4.1) to the following case: The underlying space Ω is discrete, the first capacity class consists of measures m which are singular with respect to the Lebesgue measure, and the second class consists of a single element: the measure μ which assigns equal mass to all elements of Ω .

III. Formulation of Four Matched Filtering Problems with Noise Uncertainty

A general matched-filter design situation can be described by the following formulation [2]: Let H be a separable Hilbert space (e.g., L_2 or R^n) with product $\langle \cdot, \cdot \rangle$, and let \bar{H} denote a space of bounded (self-adjoint) nonnegative linear operators mapping H to itself. A matched-filtering problem on H involves three quantities: a signal quantity $s \in H$ (e.g., a signal spectrum or waveform); a noise quantity $n \in \bar{H}$ (e.g., a noise spectrum, autocorrelation function, or covariance matrix); and a filter quantity $h \in H$ (e.g., a filter transfer function or impulse response). The design criterion for the filtering problem is based on a functional $\rho: H \times H \times \bar{H} \rightarrow R$ defined by

$$\rho(h; s, n) = \frac{|\langle h, s \rangle|^2}{\langle h, nh \rangle}, \quad h \in H, s \in H, n \in \bar{H} \quad (7)$$

and representing a signal-to-noise ratio.

Within the above formulation, the matched-filtering problem for fixed s and n is given by

$$\max_{h \in H} \rho(h; s, n)$$

and the **matched filter** is $h_0 \in H$, any solution of the equation $nh_0 = s$, when one exists. Then

$$\max_{h \in H} \rho(h; s, n) = \rho(h_0; s, n) \quad (8)$$

Note that if n is invertible we have $h_0 = n^{-1}s$. Four cases of interest when there is uncertainty about the noise quantity which is determined by a 2-alternating capacity are described next.

Case 1. (Continuous-Time Matched Filtering with a General Observation Interval and Uncertainty in the Noise Autocorrelation Function): Suppose we observe a random process given by

$$Y_t = N_t + s(t), \quad 0 \leq t \leq T, \quad (9)$$

In (9) $\{N_t; 0 \leq t \leq T\}$ (here we can replace $[0, T]$ with any general observation interval) is a zero-mean real random process with autocorrelation function $\{R_N(u, v) = n(u)\delta(u-v); (u, v) \in [0, T]^2\}$, where $n(u)$ is nonnegative for all $u \in [0, T]$, and $\{s(t); 0 \leq t \leq T\}$ is a real-valued square-integrable deterministic waveform. For fixed t , consider linear filters of the form $\int_0^T h(t, u) Y_u du$ where $\{h(t, u); 0 \leq u \leq T\}$ is a real-valued square-integrable function; the output signal-to-noise ratio at time t is given by

$$SNR = \frac{\left| \int_0^T h(t, u) s(u) du \right|^2}{\int_0^T \int_0^T h(t, u) R_N(u, v) h(t, v) dudv} \quad (10)$$

One can show that this problem fits the general framework developed above with H being $L_2[0, T]$, with the usual L_2 inner product, and with $(nh)_u = \int_0^T R_N(u, v) h(u, v) dv; 0 \leq u \leq T$. The signal, noise and filter quantities of the general problem are identified with the signal waveform $s(t)$, the integral kernel $R_N(u, v) = n(u)\delta(u-v)$, and the filter's impulse response $h(t, u)$, respectively. The optimum filter is given by any L_2 solution to the Fredholm integral equation $\int_0^T R_N(u, v) h(t, v) dv = s(u), \quad 0 \leq u \leq T$, which in this case reduces to

$$h(t, u) = s(u)/n(u), \quad \text{independent of } t \quad (11)$$

Therefore, the signal-to-noise ration in (9) takes the form

$$\rho(h; s, n) = \left| \int_0^T h(u) s(u) du \right|^2 / \int_0^T h^2(u) n(u) du \quad (12)$$

It is assumed that the nonnegative function $n(u)$ is the Radon-Nikodym derivative of a measure m_n , defined on the σ -algebra generated by the interval $[0, T]$, with respect to λ , the Lebesgue measure on $[0, T]$. The measure m_n satisfies the (fixed-noise-variance) constraint $m_n([0, T]) = \sigma^2$ and is known to belong to an uncertainty class of the form (1).

Case 2 (Discrete-Time Matched Filtering with a Finite Number of Observations and Uncertainty in the Covariance Matrix of the Noise): Suppose observe the random vector

$$Y_i = N_i + s_i, \quad i = 1, 2, \dots, k, \quad (13)$$

where $\{s_i; i = 1, 2, \dots, k\}$ is a real deterministic sequence and where $\{N_i; i = 1, 2, \dots, k\}$ is a real random sequence with zero mean and covariance matrix $\underline{\Sigma}$, where $\underline{\Sigma}_{i,j} = n_i \delta_{i,j}$ for n_i nonnegative and $\delta_{i,j} = 1$ if $i=j$ and 0 if $i \neq j$. For filters of the form $\sum_{i=1}^k h_{l,i} Y_i$ the output signal-to-noise ratio at time l is given by

$$SNR = \left| \sum_{i=1}^k h_{l,i} s_i \right|^2 / \sum_{i=1}^k \sum_{j=1}^k h_{l,i} \underline{\Sigma}_{i,j} h_{l,j} \quad (14)$$

Here, of course, $H = R^k$, $s = [s_1, s_2, \dots, s_k]^T$, $h = [h_{l,1}, h_{l,2}, \dots, h_{l,k}]^T$, n is identified with $\underline{\Sigma}$, and the operation nh is premultiplication of the k -vector h by the matrix $\underline{\Sigma}$. If $\underline{\Sigma} > 0$, then the matched filter is given by $\underline{\Sigma}^{-1}s$, which for $\underline{\Sigma}_{i,j} = n_i \delta_{i,j}$ reduces to

$$h_{l,i} = s_i/n_i, \text{ independent of } l \quad (15)$$

Therefore the signal-to-noise ratio in (14) takes the form

$$\rho(h;s,n) = \left| \sum_{i=1}^k h_i s_i \right|^2 / \sum_{i=1}^k h_i^2 n_i \quad (16)$$

It is assumed that the nonnegative function n_i satisfies the (fixed-noise-variance) constraint $\sum_{i=1}^k n_i = \sigma^2$ and defines a measure m_n on the σ -algebra generated by $\{1, 2, \dots, k\}$ which belongs to an uncertainty class of the form (1).

Case 3 (Continuous-Time Matched Filtering with Infinite Observation Interval and Uncertainty in the Spectral Measure of Bandlimited Noise): Suppose we have the situation in Case 1 with $(0, T) = (-\infty, \infty)$ and with $\{N_t; -\infty < t < \infty\}$ zero-mean, second order stationary with power spectral measure m_N concentrated on $(-\omega_0, \omega_0)$. Suppose further that $h(t, u) = h(t - u)$ for all $t, u \in (-\infty, \infty)$, then the output SNR at any time is given by

$$\rho(H; S, m_N) = \frac{1}{2\pi} \left| \int_{-\omega_0}^{\omega_0} H^*(\omega) S(\omega) d\omega \right|^2 / \int_{-\omega_0}^{\omega_0} |H(\omega)|^2 dm_N(\omega) \quad (17)$$

where H is the transfer function of the filter and S is the Fourier transform of the signal waveform which is also assumed to be bandlimited in $[-\omega_0, \omega_0]$. Here we identify the signal, noise and filter quantities of the general formulation with S , m_N , and H , respectively. We thus have H being complex $L_2(-\infty, \infty)$ with the usual inner product and $hdn(\omega) = H(\omega)dm_N(\omega)$, $-\omega_0 < \omega < \omega_0$. The optimum filter is given by

$$H(\omega) = S(\omega)/N(\omega), \quad -\omega_0 < \omega < \omega_0 \quad (18)$$

where $N(\omega) = \frac{dm_N}{d\lambda}(\omega)$ is the R-N derivative with respect to λ , the Lebesgue measure

on $[-\omega_0, \omega_0]$. It is assumed that the spectral measure m_N satisfies the (fixed-noise-power) constraint $m_N([-\omega_0, \omega_0]) = P_N$ and is known to belong to an uncertainty class of the form (1).

Case 4 (Discrete-Time Matched Filtering with Infinite Observation Interval and Uncertainty in the Spectral Measure of the Noise): Suppose we have the discrete-time analog of Case 3. Then the problem formulation and the equations for the signal-to-noise ratio and the optimal filter follow from those of the continuous-time case, if we replace the spectral interval $[-\omega_0, \omega_0]$ with $[-\pi, \pi]$.

IV. Characterization of Maximin Robust Filters

Equation (8) indicates that, for known s and n , the matched filter in the general case is given by any solution to $nh = s$ if one exists. When, however, n is known only to be a member of a class $N \subset \bar{H}$, representing uncertainty in the noise statistics, the following minimax (actually maximin) robust formulation is motivated [2];

Definition: Suppose that $s \in H$ is given and $n \in N \subset \bar{H}$. We say that $\hat{h} \in H$ is **robust** over $\{s\} \times N$ if

$$\inf_{n \in N} \rho(\hat{h}; s, n) = \max_{h \in H} \{ \inf_{n \in N} \rho(h; s, n) \} \quad (19)$$

Note that $(\hat{h}; s, \hat{n}) \in H \times H \times \bar{H}$ is a **saddle point** solution to the game of eq. (19) if

$$\inf_{n \in N} \rho(\hat{h}; s, n) \geq \rho(\hat{h}; s, \hat{n}) = \max_{h \in H} \rho(h; s, \hat{n}) \quad (20)$$

Next we prove the following proposition for the four problems formulated above:

Proposition: For each of the four problems formulated in Section III for which the noise statistic belongs to a capacity class of the form (1), the **worst-case noise statistic** \hat{n} of (20) is associated with the **least-favorable element in the capacity class** and

can be singled out by Lemma 1.

Proof: To prove this we need to show that (20) is satisfied. Because of (7), this is equivalent to

$$\langle \hat{h}, n\hat{h} \rangle \leq \langle \hat{h}, \hat{n}\hat{h} \rangle \quad (21)$$

where \hat{h} is the solution to $\hat{n}\hat{h} = s$. We prove that (21) is valid for each of the four cases of interest.

Case 1: We assume that $\hat{m}_n \ll \lambda$ (absolutely continuous with respect to λ) and denote the density by \hat{n} and the optimal filter by $\hat{h} = s/\hat{n}$. The absolute continuity restriction is easy to satisfy; for example, it is satisfied for the ϵ -contaminated and total variation classes if (the nominal) $m_n^0 \ll \lambda$ and for the band class if $m_n^0 \ll \lambda$ and $m_n^1 \ll \lambda$. Then we have

$$\int_0^T \hat{h}^2(t) n(t) dt \leq \int_{[0,T]} \hat{h}^2 dm_n = \int_{[0,T]} s^2 \hat{n}^{-2} dm_n \leq \int_{[0,T]} s^2 \hat{n}^{-2} d\hat{m}_n = \int_0^T \hat{h}^2(t) \hat{n}(t) dt \quad (22)$$

In proving the first inequality in (22) we used the fact that the absolutely continuous part of a measure is no larger than the measure itself; to prove the second inequality in (22) we used the fact that \hat{n} is stochastically smallest under \hat{m}_n over all other elements in the capacity class.

Case 2: Here we make use of the discrete-set version of Lemma 1. The optimal filter is $\hat{h}_i = s_i/\hat{n}_i$ where \hat{n}_i is the pmf corresponding to the measure \hat{m}_n singled out by Lemma 1 when applied to this case. The equivalent form to (21) is

$$\sum_{i=1}^k \hat{h}_i^2 n_i = \sum_{i=1}^k s_i^2 \hat{n}_i^{-2} n_i \leq \sum_{i=1}^k s_i^2 \hat{n}_i^{-2} \hat{n}_i = \sum_{i=1}^k \hat{h}_i^2 \hat{n}_i \quad (23)$$

In proving the inequality in (23) we used the fact that \hat{n} becomes stochastically smallest under \hat{m}_n , singled out by Lemma 1.

Cases 3 and 4: The optimal filter here is $\hat{H}(\omega) = S(\omega)/\hat{N}(\omega)$ where $\hat{N}(\omega)$ is the R-N

derivative of the least-favorable spectral measure \hat{m}_N (singled out by Lemma 1) with respect to λ . Both the continuous-time and the discrete-time problems are represented here with common notation. The equivalent form to (21) is

$$\int | \hat{H} |^2 dm_N = \int | S |^2 \hat{N}^{-2} dm_N \leq \int | S |^2 \hat{N}^{-2} d\hat{m}_N = \int | \hat{H} |^2 d\hat{m}_N \quad (24)$$

In proving the inequality in (24) we used the fact that \hat{N} becomes stochastically smallest under \hat{m}_N singled out by Lemma 1. The integrals in (24) are over $[-\omega_0, \omega_0]$ or over $[-\pi, \pi]$ for the continuous-time and discrete-time cases, respectively.

V. Conclusions

The robust matched filter for uncertainty in the noise autocorrelation function or the noise spectral measure is derived for both continuous-time and discrete-time problems when the uncertainty classes are generated by 2-alternating capacities. In all cases the maximin robust matched filter depends on the inverse of the worst-case noise statistic which is obtained as the Huber-Strassen derivative of the capacity generating the uncertainty class with respect to the Lebesgue (or other equivalent measure) on a suitable interval.

REFERENCES

- [1] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing," *Proc. IEEE*, vol. 73, pp. 433-481, March 1985.
- [2] H. V. Poor, "Robust matched filters," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 677-687, September 1983.
- [3] S. Verdu and H. V. Poor, "Minimax robust discrete-time matched filters," *IEEE Trans. Commun.*, vol. COM-31, pp. 208-215, February 1983.
- [4] P. J. Huber, "A robust version of the probability ratio test," *Ann. Math. Stat.*, vol. 36, pp. 1753-1758, 1965.
- [5] S. A. Kassam, "Robust hypothesis testing for bounded classes of probability densities," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 242-247, 1981.
- [6] K. S. Vastola and H. V. Poor, "On the p-point uncertainty class," *IEEE Trans. Information Theory*, vol. IT-30, pp. 374-376, 1984.
- [7] P. J. Huber and V. Strassen, "Minimax tests and the Neyman-Pearson lemma for capacities," *Ann. Statist.*, vol. 1, pp. 251-265, 1973.