



# **MINIMAX LINEAR SMOOTHING FOR CAPACITIES**

H. VINCENT POOR

APPROVED FOR PUBLIC RELEASE. DISTRIBUTION UNLIMITED.

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MINIMAX LINEAR SMOOTHING FOR CAPACITIES		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER R-909;UILU-ENG 81-2240
7. AUTHOR(s) H. Vincent Poor		8. CONTRACT OR GRANT NUMBER(s) N00014-81-K-0014; ECS79-16453
9. PERFORMING ORGANIZATION NAME AND ADDRESS Coordinated Science Laboratory, 1101 W. Springfield University of Illinois at Urbana-Champaign Urbana, Illinois 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research National Science Foundation		12. REPORT DATE May 1981
		13. NUMBER OF PAGES 12
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  Minimax smoothing, Choquet capacities, random fields		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  Minimax linear smoothers are considered for the problem of estimating a homogeneous signal field in an additive orthogonal noise field. A minimax game with the quadratic-mean estimation error as an objective function is used to formulate this problem. Uncertainty in signal and noise field spectra is modeled using general nonparametric classes of measures proposed by Huber and Strassen for the problem of minimax hypothesis testing. These classes, which are described in terms of Choquet alternating capacities of order 2, include the conventional models for spectral uncertainty and admit a general solution to the minimax linear smoothing problem.		





## FOREWORD

This report is a preprint of a paper with the same title which is to appear in the Annals of Probability. This work was presented in preliminary form at the 18th IEEE Conference on Decision and Control, Ft. Lauderdale, FL, December 12-14, 1979.

MINIMAX LINEAR SMOOTHING FOR CAPACITIES

by

H. Vincent Poor

This work was supported in part by the National Science Foundation under Grant ECS 79-16453 and in part by the Office of Naval Research under Contract N00014-81-K-0014.

Reproduction in whole or in part is permitted for any purpose of the United States Government.

Approved for public release. Distribution unlimited.

1. Introduction. Suppose we observe the random field  $\{Y_z; z \in \mathbb{R}^n\}$  given for each  $z \in \mathbb{R}^n$  by  $Y_z = (S_z + N_z)$  where  $\{S_z; z \in \mathbb{R}^n\}$  and  $\{N_z; z \in \mathbb{R}^n\}$  are orthogonal random fields, each of which is second order, homogeneous, and quadratic-mean continuous. Suppose further that  $h$  is a complex-valued Borel-measurable function on  $\mathbb{R}^n$ , and that  $\hat{S}_z$  denotes that the linear estimate of  $S_z$  based on  $\{Y_z; z \in \mathbb{R}^n\}$  which has transfer function  $h$ . Then the quadratic-mean estimation error associated with  $\hat{S}_z$  is given by

$$E\{|S_z - \hat{S}_z|^2\} = (2\pi)^{-n} \left[ \int_{\mathbb{R}^n} |1-h|^2 d m_S + \int_{\mathbb{R}^n} |h|^2 d m_N \right] \triangleq e(h; m_S, m_N) \quad (1)$$

where  $m_S$  and  $m_N$  are the spectral measures on  $(\mathbb{R}^n, \mathcal{B}^n)$  associated (via Bochner's theorem [1, p. 245]) with  $\{S_z; z \in \mathbb{R}^n\}$  and  $\{N_z; z \in \mathbb{R}^n\}$ , respectively. For fixed  $m_S$  and  $m_N$ , the minimum possible value of  $e(h; m_S, m_N)$  is achieved by the estimate with transfer function  $\hat{h} = dm_S / d(m_S + m_N)$  and this minimum value is given by  $(2\pi)^{-n} \int_{\mathbb{R}^n} \hat{h} dm_N$ .<sup>1</sup> If, on the other hand,  $m_S$  and  $m_N$  are known only to be in classes  $\mathcal{M}_S$  and  $\mathcal{M}_N$ , respectively, of spectral measures on  $(\mathbb{R}^n, \mathcal{B}^n)$ , then a reasonable design strategy is to find a linear estimate whose transfer function minimizes  $\sup_{\mathcal{M}_S \times \mathcal{M}_N} e(h; m_S, m_N)$ . Such an estimate will be a minimax linear smoother for  $\mathcal{M}_S$  and  $\mathcal{M}_N$ . Certain aspects of this problem have been considered by Kassam and Lim [2] and by the author [3]. In this paper we consider the minimax linear smoothing problem for the situation in which the measure classes  $\mathcal{M}_S$  and  $\mathcal{M}_N$  are of the type generated by 2-alternating capacities as considered by Huber and Strassen [4] in the context of minimax hypothesis testing. Examples of this type of class include mixtures, Prohorov and Kolmogorov (variational) neighborhoods, and other previously considered models for spectral uncertainty.

<sup>1</sup>Note that  $e(h; m_S, m_N) = (2\pi)^{-n} \int_{\mathbb{R}^n} \hat{h} dm_N + (2\pi)^{-n} \int_{\mathbb{R}^n} |\hat{h} - h|^2 d(m_S + m_N)$ .



Here we apply the results of Huber and Strassen to find the structure of minimax linear smoothers for general models of this type.

2. The minimax smoother for capacity classes. In the following,  $\Omega$  denotes a fixed subset of  $\mathbb{R}^n$ ,  $\mathcal{A}$  denotes the Borel  $\sigma$ -algebra on  $\Omega$ , and  $\mathcal{M}$  denotes the class of all finite measures on  $(\Omega, \mathcal{A})$ . Recall that a finite set function  $v$  on  $\mathcal{A}$  is a 2-alternating capacity (see Choquet [5]) on  $(\Omega, \mathcal{A})$  if it is increasing, continuous from below, continuous from above for closed sets, and if it satisfies  $v(\emptyset) = 0$  and  $v(A \cup B) + v(A \cap B) \leq v(A) + v(B)$  for all  $A, B \in \mathcal{A}$ . For a 2-alternating capacity  $v$  on  $(\Omega, \mathcal{A})$  define the set  $\mathcal{M}_v$  by

$$\mathcal{M}_v = \{m \in \mathcal{M} \mid m(A) \leq v(A) \text{ for all } A \in \mathcal{A}, \text{ and } m(\Omega) = v(\Omega)\}. \quad (2)$$

A number of properties of classes of the form of (2) have been developed by Huber and Strassen [4]. Note, for example, that  $\mathcal{M}_v$  is weakly compact and that, if  $v$  is a measure, then  $\mathcal{M}_v = \{v\}$ .

For any pair  $(v_0, v_1)$  of 2-alternating capacities on  $(\Omega, \mathcal{A})$  there exists a Radon-Nikodym derivative  $dv_1/dv_0$ , introduced in [4], which has the defining property that, for each  $t \in [0, \infty]$ ,

$$r_t(\{dv_1/dv_0 > t\}) = \inf_{A \in \mathcal{A}} r_t(A) \quad (3)$$

where  $r_t(A) \triangleq (1+t)^{-1}[tv_0(A) + v_1(A^c)]$ . This derivative (which is a family of functions having the defining property (3)) is the basis for the minimax tests between capacity classes of the form of (2) as considered in [4]. Further properties and a generalization of this derivative have been considered by Rieder [6]. In this context we state the following result which is Theorem 4.1 of [4]:

Lemma 2.1 (Huber-Strassen): Suppose  $v_S$  and  $v_N$  are 2-alternating capacities and  $\pi_0$  is a version  $dv_S/dv_N$ . Then there exist measures  $q_S \in \mathcal{M}_{v_S}$  and  $q_N \in \mathcal{M}_{v_N}$  such that  $\pi_0 \in dq_S/dq_N$  and such that

$$q_S(\{\pi_0 < t\}) = v_S(\{\pi_0 < t\})$$

and

$$q_N(\{\pi_0 > t\}) = v_N(\{\pi_0 > t\})$$

for all  $t \in [0, \infty]$ .

Let  $\mathcal{K}$  denote the class of all complex-valued  $\mathcal{A}$ -measurable functions on  $\Omega$ . Lemma 2.1 leads to the following theorem:

Theorem 2.2: Suppose  $v_S$  and  $v_N$  are 2-alternating capacities on  $(\Omega, \mathcal{A})$ . Let  $\pi_0$  be a version of  $dv_S/dv_N$  and choose  $(q_S, q_N)$  as in Lemma 2.1. Define  $h_0 = \pi_0(1 + \pi_0)^{-1}$ . Then  $[h_0, (q_S, q_N)]$  is a saddle-point solution to the game

$$\min_{h \in \mathcal{K}} \sup_{(m_S, m_N) \in \mathcal{M}_{v_S} \times \mathcal{M}_{v_N}} e(h; m_S, m_N)$$

where  $e$  is defined in (1), and thus  $h_0$  is a minimax linear smoother for  $\mathcal{M}_{v_S}$  and  $\mathcal{M}_{v_N}$ .

Proof: Noting that  $h_0 \in dq_S/d(q_S + q_N)$ , we have directly that

$$e(h_0; q_S, q_N) \leq e(h; q_S, q_N)$$

for all  $h \in \mathcal{K}$ . Thus, it is sufficient to show

$$e(h_0; m_S, m_N) \leq e(h_0; q_S, q_N) \tag{4}$$



for all  $(m_S, m_N) \in \mathcal{M}_S \times \mathcal{M}_N$ . Lemma 2.1 asserts that  $\pi_0$  is stochastically smallest over  $\mathcal{M}_{V_S}$  under  $q_S$  and is stochastically largest over  $\mathcal{M}_{V_N}$  under  $q_N$ . Thus, since  $|1 - h_0|^2 = (1 + \pi_0)^{-2}$  is decreasing in  $\pi_0$  and  $|h_0|^2 = \pi_0^2(1 + \pi_0)^{-2}$  is increasing in  $\pi_0$ , we have

$$\int_{\Omega} |1 - h_0|^2 dm_S \leq \int_{\Omega} |1 - h_0|^2 dq_S$$

and

$$\int_{\Omega} |h_0|^2 dm_N \leq \int_{\Omega} |h_0|^2 dq_N$$

for all  $(m_S, m_N) \in \mathcal{M}_{V_S} \times \mathcal{M}_{V_N}$ . Equation (4) and hence Theorem 2.1 follow.

Note that, in view of Theorem 2.1, the pair  $(q_S, q_N)$  singled out by Lemma 2.1 can be thought of a least-favorable pair of spectral measures for minimax linear smoothing. Concerning this pair of measures, we may also state the following property:

**Theorem 2.3:** The pair  $(q_S, q_N) \in \mathcal{M}_{V_S} \times \mathcal{M}_{V_N}$  satisfies the conclusion of Lemma 2.1 if and only if it maximizes

$$\min_{h \in \mathcal{K}} e(h; m_S, m_N) = (2\pi)^{-n} \int_{\Omega} [dm_S / d(m_S + m_N)] dm_N$$

over all  $(m_S, m_N) \in \mathcal{M}_{V_S} \times \mathcal{M}_{V_N}$ .

**Proof:** Define  $f = dm_N / d(m_S + m_N)$ . Then

$$\min_{h \in \mathcal{K}} e(h; m_S, m_N) = (2\pi)^{-n} \int_{\Omega} f dm_S = (2\pi)^{-n} \int_{\Omega} (f - f^2) d(m_S + m_N).$$

Since  $C[x] = (x - x^2)$  is concave and twice continuously differentiable on  $[0, 1]$ , Theorem 2.3 follows from Theorem 6.1 of [4].

3. Discussion. Theorem 2.2 gives the general solution to the minimax linear smoothing problem for signal and noise uncertainty classes of the form of (2). Several useful examples of classes of this type are given by Huber and Strassen in [4], and other useful examples are given by Rieder [6], Strassen [7], and Vastola and Poor [8]. Some of the most commonly used examples of classes of the form  $\mathcal{M}_V$  can be written as  $\epsilon$ -neighborhoods of some nominal measure  $\mu$ . Examples of capacity classes that have this structure are contaminated mixtures, variational neighborhoods, and Prohorov neighborhoods (see [4]). For this type of class, an uncertainty model will consist of a nominal pair  $(\mu_S, \mu_N)$  of signal and noise spectral measures with respective degrees  $\epsilon_S$  and  $\epsilon_N$  of uncertainty placed on the nominal measures. The derivative between capacities generating classes of this type is often of the form (see Huber [9, 10] and Rieder [6])

$$\pi_0(\omega) = \max\{c', \min\{c'', \lambda(\omega)\}\}, \omega \in \Omega, \quad (5)$$

where  $\lambda$  is the Radon-Nikodym derivative between the nominal pair of measures (i.e.,  $\lambda \in d\mu_S/d\mu_N$ ) and  $c'$  and  $c''$  are nonnegative constants with  $c' \leq c''$ .

If  $\pi_0$  of (5) is a version of  $dv_S/dv_N$ , then Theorem 2.2 implies that a minimax linear smoother for  $\mathcal{M}_{V_S}$  and  $\mathcal{M}_{V_N}$  is given by

$$h_0(\omega) = \max\{k', \min\{k'', h'(\omega)\}\}, \omega \in \Omega \quad (6)$$

where  $k' = c'/(1+c')$ ,  $k'' = c''/(1+c'')$  and  $h' = \lambda/(1+\lambda)$ . Note that  $h'$  is the optimum smoother for the nominal model, and thus the minimax linear smoother for this case desensitizes the nominal smoother (to a degree depending on  $\epsilon_S$  and  $\epsilon_N$ ) in those spectral regions where either  $\mu_S$  or  $\mu_N$  is dominant (i.e., where  $h'$  is near 1 or is near 0).



In the situations for which (5) is valid, (6) gives the transfer function of the minimax linear smoother. Suppose, for example, that  $n=1$ ,  $\Omega = [-b, b]$  for some  $b < \infty$ ,  $c' < c''$ , and  $h'$  is symmetric about  $\omega=0$  and is strictly decreasing on  $[0, b]$ . Then the minimax linear estimate of  $S_z$  determined by  $h_0$  is given explicitly by

$$\hat{S}_z = \int_{-\infty}^{\infty} \bar{h}_0(z-t) Y_t dt$$

where  $\bar{h}_0 \triangleq \mathcal{F}^{-1}\{h_0\}$  is given by

$$\begin{aligned} \bar{h}_0(t) = & \bar{h}'(t) + k' [\sin(bt) - \sin(a't)] / (\pi t) + k'' \sin(a''t) / (\pi t) \\ & - \int_{-\infty}^{\infty} \bar{h}'(t-\tau) [\sin(b\tau) - \sin(a'\tau) + \sin(a''\tau)] (\pi\tau)^{-1} d\tau \end{aligned}$$

with  $\bar{h}' = \mathcal{F}^{-1}\{h'\}$  and with  $a'$  [resp.,  $a''$ ] the positive solution to  $h'(a') = k'$  [resp.,  $h'(a'') = k''$ ].

As a final comment we note that, although we assumed initially that the observation field was a continuous-parameter field, Theorems 2.2 and 2.3 are also directly applicable to the case in which the observation field is a discrete-parameter field (i.e., in which the time set is  $\mathbb{Z}^n$ ) since this latter situation corresponds to the particular case of the analysis of Section 2 in which  $\Omega = [-\pi, \pi]^n$ .

4. Acknowledgement. The author would like to thank K. S. Vastola for several stimulating discussions concerning the techniques used in this paper. Also the helpful comments and suggestions of the anonymous referee are gratefully acknowledged.



# References

- [1] Wong, E. (1971). Stochastic Processes in Information and Dynamical Systems. McGraw-Hill, New York.
- [2] Kassam, S. A. and Lim, T. L. (1977). Robust Wiener filters. J. Franklin Inst. 304 171-185.
- [3] Poor, H. V. (1980). On robust Wiener filtering. IEEE Trans. Automatic Control AC-25 521-526.
- [4] Huber, P. J. and Strassen, V. (1973). Minimax tests and the Neyman-Pearson lemma for capacities. Ann. Statist. 1 251-263.
- [5] Choquet, G. (1953/54). Theory of capacities. Ann. Inst. Fourier 5 131-292.
- [6] Rieder, H. (1977). Least favorable pairs for special capacities. Ann. Statist. 5 909-921.
- [7] Strassen, V. (1965). The existence of probability measures with given marginals. Ann. Math. Statist. 36 423-439.
- [8] Vastola, K. S. and Poor, H. V. (1980). On generalized band models in robust detection and filtering. Proc. 1980 Conf. on Inform. Sciences and Systems, Princeton University, Princeton, NJ, March 1980, 1-5.
- [9] Huber, P. J. (1965). A robust version of the probability ratio test. Ann. Math. Statist. 36 1753-1758.
- [10] Huber, P. J. (1968). Robust confidence limits. Z. Wahr. verw. Geb. 10 269-278.

Coordinated Science Laboratory  
1101 W. Springfield Ave.  
University of Illinois  
Urbana, Illinois 61801

Footnote

<sup>1</sup>Note that  $e(h; m_S, m_N) = (2\pi)^{-n} \int_{\mathbb{R}^n} \hat{h} dm_N + (2\pi)^{-n} \int_{\mathbb{R}^n} |\hat{h} - h|^2 d(m_S + m_N)$ .