Note di Matematica Vol. XI, 157-175 (1991)

A GENERAL VERSION OF BREIMAN'S MINIMAX FILTER

JÜRGEN FRANKE

Dedicated to the memory of Professor Gottfried Köthe

1. INTRODUCTION

The Wiener-Kolmogorov theory of filtering and prediction of stationary processes represents a functional analytic approach with far-reaching applications in the engineering sciences, in particular in signal analysis. The main problem consists in extracting or predicting a signal from observations corrupted by noise. We consider the discrete-time case where both signal and noise are modelled as weakly stationary stochastic processes which are both imbedded in a separable Hilbert space of real random variables with finite second moment. Classically, the filtering problem is translated into a projection problem, and its solution can principally be constructed provided the autocovariances of the processes involved, i.e. the values of the pairwise scalar products, are known. Equivalently, one can specify the Fourier transforms of the autocovariances, the so-called spectral measures. We restrict ourselves to spectral measures which are absolutely continuous with respect to Lebesgue measure, and we call their densities the spectral densities of signal and noise respectively.

In practice, the spectral densities of the signal and noise process are not known precisely. We investigate the filtering in a situation where there are particular few informations on the signal. By borrowing some ideas from robust statistics and applying familiar methods for solving convex extremum problems in function spaces, we are able to derive filters which exhibit a minimax property, i.e. they are in a certain sense optimal in view of the incomplete knowledge of the process characteristics. The results unify and extend approaches in the statistical and engineering literature and illustrate the power of functional analytic methods applied to some practical problems.

We first consider the classical filtering problem. Let, for integer time index t,

$$\{X_t\}, \{S_t\}, \{N_t\}, \qquad -\infty < t < \infty$$

be weakly stationary stochastic processes with spectral densities f_x, f_s, f_n such that

$$X_t = S_t + N_t.$$

We assume that *signal* $\{S_t\}$ and *noise* $\{N_t\}$ are uncorrelated. Given the observations $\{X_t\}$ we want to estimate the signal by means of a *linear filter*

$$\widehat{S}_t = \sum_{k=-\infty}^{\infty} h_k X_{t-k}.$$

We assume that the *filter coefficients* $\{h_k\}$ are square-summable, and we write

$$H(\omega) = \sum_{k=-\infty}^{\infty} h_k e^{-ik\omega}, \qquad -\pi < \omega \leq \pi,$$

for the so-called *filter transfer function*, i.e. for the Fourier transform of the $\{h_k\}$. We measure the performance of such a filter by the mean-square error, which depends on signal and noise spectral densities and filter transfer function only. It is given by the well-known formula

$$E\left(S_t - \widehat{S}_t\right)^2 = \frac{1}{2\pi}\int \left\{|1 - H(\omega)| f_s(\omega) + |H(\omega)|^2 f_n(\omega)\right\} d\omega = \tag{1.1}$$
$$=: e\left(f_s, f_n; H\right).$$

We call the function e the *error function.*

Traditional Wiener-Kolmogorov filtering theory is based on complete knowledge of f_s and f_n. By standard Hilbert space arguments, there exists a unique filter with transfer function H_w, which we call the *Wiener filter* with respect to f_s, f_n, minimizing the mean-square error, i.e.

$$e\left(f_s, f_n; H_w\right) = \min_H e\left(f_s, f_n; H\right).$$

As is well-known from the literature, e.g. Hannan (1970), the transfer function H_w is given by

$$H_w(\omega) = f_s(\omega) / \{f_s(\omega) + f_n(\omega)\} \qquad \text{if} \quad f_s(\omega) + f_n(\omega) > 0, \tag{1.2}$$

and it can be chosen arbitrarily else.

In this paper, however, we consider the situation, which usually occurs in practice, where signal and noise spectral densities are not completely known. We assume instead that we have some partial information of f_s, f_n which can be summarized in the statement that (f_s, f_n) is contained in a given set S of pairs of spectral densities. We call S the *spectral information set.* We follow the minimax approach to this kind of filtering problem under spectral uncertainty. We do not search for a filter which is optimal for one particular pair of signal and noise spectral densities, but we are interested in a filter for which the error function is uniformly bounded over S and for which the uniform bound on the mean-square error is as small as possible. A filter with transfer function H_r, satisfying these requirements, i.e.

$$\sup_{(f_s,f_n)\in S} e\left(f_s, f_n; H_r\right) = \min_H \sup_{(f_s,f_n)\in S} e\left(f_s, f_n; H\right),$$

is called a *minimax-robust filter* with respect to S .

The minimax approach to linear filtering, prediction and interpolation has found some interest in the past. Breiman (1973), Kassam and Lim (1977), and Cimini and Kassam (1980) have discussed the filtering problem under special types of spectral information. Vastola and Poor (1983) have investigated the merit of the minimax procedure compared with traditional Wiener-Kolmogorov filtering for a nominal pair of signal and noise spectral densities, and they have demonstrated with some examples that the latter may result in dangerous losses in performance. Hosoya (1978), Taniguchi (1981) and Franke (1984) have considered the corresponding problem of prediction and interpolation under spectral uncertainty. Poor (1980), Franke (1985) and Kassam (1983) have proposed general formulations for the problems of minimax filtering, prediction and interpolation and have described how to determine explicitly minimax-robust filters for large classes of spectral information sets.

Vastola and Poor (1984) have investigated the general filtering problem with arbitrary restrictions on the filters considered, e.g. requiring causality, which encompasses the noncausal filtering problem, described above, as well as prediction and interpolation. They have formulated theorems on the existence and characterization of minimax-robust filters. Franke and Poor (1984) have generalized these results and have described how to use them for explicitly finding the desired filters which perform uniformly well under spectral uncertainty.

In this paper we consider in more detail a situation where we have rather few informations on the signal spectral density f_s. We assume that we only know bounds $c_0, \dots, c_m$ on the integrals of some functions $p_o, \dots, p_m$ with respect to f_s, i.e.

$$\frac{1}{2\pi}\int p_k(\omega) f_s(\omega)\,d\omega \le c_k, \qquad k = 0, \dots, m. \tag{1.3}$$

Breiman (1973) has given the motivation for this type of spectral information, and he has determined the minimax-robust filter in the case where the noise spectral density is completely known. We do not want to repeat the complete arguments given by Breiman, but in short the background for this problem is the following:

The observed process $\{X_t\}$ consists of a low frequency signal $\{S_t\}$ and of a broadband, e.g. white, noise. The spectral characteristics of the noise are well-known. Think of situations where the same type of noise is present over and over again whereas the signal changes from measurement to measurement, or where the noise is due to recording equipment and has been throughly investigated in the absence of signals. On the other hand, all that is known about the signal are the observations on the $\{X_t\}$ process and some vague prior information that f_s is mainly concentrated in the low frequencies.

For this situation, Breiman proposes to do a preliminary rough smoothing of the data and, then, getting rough estimates of the variances of some of the differences of the signal. More precisely, let Δ denote the difference operator:

$$\Delta S_t = S_{t+1} - S_t.$$

An estimated upper bound c_1 on the variance $E(\Delta S_t)^2$ translates into the following integral condition of f_s:

$$E\left(\Delta S_t\right)^2 = \frac{1}{2\pi}\int |e^{i\omega} - 1|^2 f_s(\omega)\,d\omega \leq c_1^2. \tag{1.4}$$

Analogously, one can consider bounds on the variances of higher difference processes $\{\Delta^k S_t\}$, too. How many constraints

$$E\left(\Delta^k S_t\right)^2 \leq c_k^2$$

are used in characterizing the spectral information set depends on two considerations. First, how many can easily and reliably be determined from the data? Second, how large an upper bound on the filtering error is tolerable? This error decreases if the spectral information set gets smaller, and it can be calculated explicitly from the following results.

Another kind of spectral information, which also is characterized by integral constraints of the form (1.3), has been investigated by Cimini and Kassam (1980). Their so-called p-point classes are motivated by situations where there is no information on the shape of the signal spectral density but where one can measure the fractional power of the signal in certain frequency bands rather accurately. This kind of spectral information amounts to choosing the functions p_j of (1.3) as indicator functions of certain frequency intervals.

For a number of applications, one has some knowledge about the size of the *signal-to-noise ratio* (SNR) f_s/f_n in certain frequency regions, even if there is no additional information on the spectral characteristics of the signal. Breiman, e.g. assumes a low frequency signal in broadband noise. This kind of information can be quantified by specifying a, perhaps crude, lower bound on the SNR for low frequencies and an upper bound for high frequencies. We propose to exploit this information in minimax filtering, i.e. to characterize the spectral information set by an additional constraint of the form

$$\beta_l(\omega) f_n(\omega) \leq f_s(\omega) \leq \beta_u(\omega) f_n(\omega) \qquad \text{a.e.}, \tag{1.5}$$

where β_l, β_u are known bounds, in addition to the integral constraints (1.3). Using this additional information can cause a considerable imrovement in the performance of minimax-robust filters.

In chapter 2 we investigate the existence and calculation of minimax-robust filters for spectral information of type (1.3) and (1.5). In chapter 3 we illustrate the performance of minimax-robust filters with an example, and we compare them with an autoregressive-type filter derived from a traditional approach to filtering under uncertainty.

2. THE MINIMAX-ROBUST FILTER

As we do not want to discuss some degenerate situations explicitly, we consider only spectral information sets where the total signal power is uniformly bounded. This amounts to choosing one of the functions p_k identical to 1. Furthermore, we require all $p_k,\ k = 0, \ldots, m,$ to be essentially bounded. Henceforth, we shall assume

$$(2.1) \qquad p_o(\omega) = 1 \quad \text{a.e.}; \qquad p_1, \ldots, p_m \in L^\infty,$$

where L^∞ denotes the space of bounded, measurable functions on $(-\pi, \pi]$. Breiman (1973) has discussed only integral constraints with nonnegative, continuous functions p_k. To include the p-point classes of Cimini and Kassam (1980) in our general model we allow for discontinuous p_k. This does not cause difficulties as, unlike Breiman, we consider only stationary processes with absolutely continuous spectral measures. As we dispense with nonnegativity of the p_k the following results are, in particular, applicable to situations where upper and lower bounds on integrals are available, e.g. for some p:

$$c' \leq \frac{1}{2\pi} \int p(\omega) f_s(\omega) d\omega \leq c".$$

Breiman also assumes that f_n is known precisely. This may not always be the case, in particular if the nominal noise spectral density is actually an estimate. To investigate the effect of uncertainty in the noise spectrum on the minimax filtering procedure, we allow for some variation of f_n, too. To have a specific situation, we concentrate on the band-model of uncertainty which has been first discussed in the context of minimax filtering by Kassam and Lim (1977). We assume that

$$(2.2) \qquad g_1(\omega) \leq f_n(\omega) \leq g_u(\omega) \qquad \text{a.e.},$$

where g_1, g_u are known lower and upper bounds on the noise spectral density satisfying

$$(2.3) \qquad 0 \leq g_1(\omega) \leq g_u(\omega) \quad \text{a.e.}; \qquad 0 < g_u(\omega) \quad \text{a.e.}; \qquad g_1 \in L^1.$$

L^1 denotes the space of integrable functions on $(-\pi, \pi]$. g_u may assume the value ∞ on sets of nonvanishing Lebesgue measure which corresponds to the situation where we have no upper bound on the noise spectral density. Additionally, we restrict the total noise power:

$$(2.4) \qquad \frac{1}{2\pi} \int f_n(\omega) d\omega \leq c.$$

The bounds β_1, β_u on the SNR have to satisfy the following conditions:

$$(2.5) \qquad 0 \leq \beta_1(\omega) \leq \beta_u(\omega) \quad \text{a.e.}; \qquad \beta_1 \in L^\infty; \qquad \beta_u 1_{\{\beta_u < \infty\}} \in L^\infty.$$

The boundedness assumptions on β_1, β_u are not necessary but make a simpler formulation of the following results possible. On the set $\{\beta_u = \infty\}$ we have no upper bound on the SNR; accordingly, we have to interpret the right inequality of (1.5) as $f_s(\omega) < \infty$ if $\beta_u(\omega) = \infty$, even if $f_n(\omega) = 0$.

Finally, we exclude a degenerate case by assuming that there is a $\Delta > 0$ such that

$$(2.6) \qquad \{1 + \beta_u(\omega)\} g_u(\omega) \geq \Delta > 0 \quad \text{a.e.}.$$

The following theorem guarantees the existence of a minimax-robust filter, and it provides a criterion which may be used for explicitly determining the required filter. This result follows immediately from the general Theorems 2.1, 2.2 and Corollary 3.1 of Franke and Poor (1984), where (2.6), in particular, guarantees that assumption (ii) of Theorem 2.1 is satisfied. Before we can state the result we have to introduce the notion of a least favorable pair:

Definition. $(f_s^L, f_n^L) \in \mathrm{S}$ *is a least favorable pair for the spectral information set* S *iff*

$$(2.7) \qquad \min_H e\left(f_s^L, f_n^L; H\right) = \max_{(f_s^L, f_n^L) \in \mathrm{S}} \min_H e\left(f_s^L, f_n^L; H\right).$$

Theorem 1. *Let* $p_o, \ldots, p_m, \beta_1, \beta_u, g_1, g_u$ *be extended real-valued measurable functions satisfying (2.1), (2.3), (2.5), (2.6). Let* S *be the spectral information set consisting of all pairs* (f_s, f_n) *of spectral densities which satisfy (1.3), (1.5), (2.2) and (2.4).*

a) There exists a minimax-robust filter with respect to S.

b) Let $(f_s^L, f_n^L) \in \mathrm{S}$ *and let* H_w^L *be the transfer function of the Wiener filter with respect to* f_s^L, f_n^L. (f_s^L, f_n^L) *is a lest favorable pair for* S *iff*

$$(2.8) \qquad e\left(f_s, f_n; H_w^L\right) \leq e\left(f_s^L, f_n^L; H_w^L\right) \quad \text{for all} \quad (f_s, f_n) \in \mathrm{S}.$$

In particular, the Wiener filter with respect to a least favorable pair for S *is minimax-robust with respect to* S.

Intuitively, a least favorable pair corresponds to the worst possible situation which is still compatible with the spectral information represented by S. The smallest achieveable mean-square error is largest in S for such a pair. By Theorem 1, we get the desired uniformly well-behaved filter by choosing the best filter, i.e. the Wiener filter, for the worst possible situation, i.e. for the least favorable pair. The following theorem provides a criterion for such a least favorable pair which follows from the general condition (2.8) by exploiting the special structure of the spectral information set S.

Theorem 2. *Let* S *be the spectral information set of Theorem 1, and let the assumptions of Theorem 1 be satisfied. Let* h_1, h_u *be given by*

$$h_1(\omega) = \beta_1(\omega) / \{1 + \beta_1(\omega)\}; \ h_u(\omega) = \beta_u(\omega) / \{1 + \beta_u(\omega)\}$$

where, in particular, $h_u(\omega) = 1$ *if* $\beta_u(\omega) = \infty$. *Let* $(f_s^L, f_n^L) \in$ S *such that* $f_s^L(\omega) + f_n^L(\omega) > 0$ *a.e., and let* H_w^L *be the transfer function of theWiener filter with respect to* f_s^L, f_n^L.

Then, (f_s^L, f_n^L) *is least favorable for* S *iff there exist Lagrange multipliers* $\lambda, \lambda_0, \ldots, \lambda_m \geq 0$ *satisfying*

$$\lambda = 0 \quad \text{if} \quad \frac{1}{2\pi} \int f_n^L(\omega) \, d\omega < c$$

$$\lambda_j = 0 \quad \text{if} \quad \frac{1}{2\pi} \int p_j(\omega) f_s^L(\omega) \, d\omega < c_j, \qquad j = 0, \ldots, m,$$

such that the following three conditions are satisfied:

1) *The Wiener filter with respect to* f_s^L, f_n^L *is given by*

$$H_w^L(\omega) = \max \left\{ \min \left\{ 1 - \sqrt{\Delta^+(\omega)}, h_u(\omega) \right\}, h_1(\omega) \right\} \tag{2.9}$$

where Δ^+ *denotes the positive part of the function*

$$\Delta(\omega) = \sum_{j=0}^{m} \lambda_j p_j(\omega).$$

2) *Let* $\Delta(\omega) \neq 0$ *or* $\beta_u(\omega) < \infty$. *Then the function*

$$J(\omega) = H_w^L(\omega) \left\{ 1 - \frac{\Delta(\omega)}{1 - H_w^L(\omega)} \right\}$$

is well-defined, and

$$f_n^L(\omega) = g_u(\omega) \quad \text{if} \quad J(\omega) > \lambda$$

$$f_n^L(\omega) = g_1(\omega) \quad \text{if} \quad J(\omega) < \lambda,$$

whereas $f_n^L(\omega)$ *can be chosen arbitrarily between* $g_1(\omega)$ *and* $g_u(\omega)$ *if* $J(\omega) = \lambda$.

3) $f_n^L(\omega) = 0$ *iff* $\Delta(\omega) = 0,\ \beta_u(\omega) = \infty$

$$f_s^L(\omega) = \frac{f_n^L(\omega)\, H_w^L(\omega)}{1 - H_w^L(\omega)} \quad \text{if} \quad f_n^L(\omega) > 0 .$$

Due to its technicality we defer the proof of Theorem 2 to the appendix. The important assertion is (2.9) which, together with Theorem 1, gives the typical form of the minimax-robust filter. It depends on $m+1$ nonnegative parameters $\lambda_0, \dots, \lambda_m$. If λ_j is not vanishing it has to be determined from the relation

$$\frac{1}{2\pi} \int p_j(\omega) f_s^L(\omega)\, \mathrm{d}\omega = c_j .$$

Theorem 2 contains Breiman's result which corresponds to the special choice

$$\beta_1(\omega) = 0, \quad \beta_u(\omega) = \infty, \quad g_1(\omega) = g_u(\omega) \quad \text{for all} \quad \omega, \lambda = \infty,$$

i.e. there are no SNR-bounds, and the noise spectral density is fixed. Then, the minimax filter is given by

$$H_w^L(\omega) = \left\{ 1 - \sqrt{\Delta^+(\omega)} \right\}^+ . \tag{2.10}$$

We call a filter of this form the *Breiman filter* corresponding to the integral constraints (1.3).

The constraint (1.5) of f_s, f_n is, with exception of frequencies for which f_n is vanishing, equivalent to the following constraint on the corresponding Wiener filter

$$h_1(\omega) \le H_w(\omega) \le h_u(\omega) .$$

Together with the form (2.10) of the minimax filter in the absence of SNR-bounds this consideration provides an intuitive interpretation of the minimax filter (2.9): it has the same form as the Breiman filter, but it is trimmed at its upper and lower bounds h_1, h_u which are imposed by the SNR-constraint (1.5). Therefore, we call (2.9) a *trimmed Breiman filter* corresponding to the integral constraints (1.3) and the SNR-bounds (1.5). Mark, however, that imposing SNR-bounds does not only cause trimming of the minimax filter but usually changes the parameters $\lambda_0, \dots, \lambda_m$, too.

Condition 2) of Theorem 2 implies that, with the exception of some degenerate situations, $f_n^L(\omega)$ coincides a.e. with the upper or lower bound imposed on the noise spectral densities. This behavior is typical for least favorable spectral densities in band models (compare Kassam and Lim (1977)). Condition 3) of Theorem 2 takes account of some degenerate cases where f_n^L may vanish. Furthermore, the form of f_s^L as a function of f_n^L and the Wiener filter is stated explicitly, which is an immediate consequence of (1.2).

A careful look at the proof of Theorem 2 shows that similar results hold for other types of uncertainty of the noise spectral density instead of (2.2). In particular, the minimax filter will be a trimmed Breiman filter though the parameters $\lambda_0, \dots, \lambda_m$ depend implicitly on the constraints on f_n.

3. THE PERFORMANCE OF MINIMAX FILTERS - AN EXAMPLE

In this chapter, an examples serves to illustrate the performance of minimax-robust filters and to compare them with a more traditional approach to filtering under uncertainty.

Following Vastola and Poor (1983), we measure the performance of a filter with transfer function H by means of the *total output signal-to-noise ratio:*

$$\rho_{out}\left(f_s, f_n; H\right) = \frac{ES_t^2}{E\left(S_t - \widehat{S}_t\right)^2} = \frac{ES_t^2}{e\left(f_s, f_n; H\right)}.$$

The terminology is motivated by the decomposition $\widehat{S}_t = S_t + (\widehat{S}_t - S_t)$, i.e. the output $\widehat{S}_t$ of the filter can also be interpreted as a noisy version of the signal S_t, where the estimation error acts as noise. Analogously, we measure the relative contribution of signal and noise to the observed process X_t, i.e. to the input of the estimation filter, by means of the *total input signal-to-noise ratio:*

$$\rho_{in} = \frac{ES_t^2}{EN_t^2}.$$

In the following figures, we plot ρ_{out} against ρ_{in}, where both are measured in logarithmic dB-scale.

We consider the situation described by Breiman (1973), i.e. we assume that the noise spectral density f_n coincides with a given spectral density, say g, and that $E(S_{t+1} - S_t)^2$ is bounded from above by a given c_1^2.

Additionally to Breiman's approach, we take only signals into consideration which have a total power ES_t^2 bounded by $c > 0$. A perhaps only crude upper bound on the total signal power should be known in practice. Using this information improves the uniform error bound corresponding to the application of the minimax filter, and, moreover, we get the convenient existence of a least favorable pair.

The described knowledge about signal and noise corresponds to the particular spectral information set S which consists of all pairs (f_s, g) where f_s satisfies

$$\frac{1}{2\pi}\int f_s(\omega)\,d\omega \le c, \quad \frac{1}{2\pi}\int (1 - \cos\omega) f_s(\omega)\,d\omega \le c' \tag{3.1}$$

for suitable constants c, c'. $\mathbf{S}$ belongs to the class of spectral information sets investigated in chapter 2 where, in particular, we have to choose $g_1 = g_u = g$ in (2.3).

In the following, let $\{f_s^B, g\}$ denote a least favorable pair with respect to $\mathbf{S}$, and let H_w^B be the transfer function of the Wiener filter with respect to f_s^B, g. By Theorem 2, H_w^B is minimax-robust with respect to $\mathbf{S}$, and it is given as the Breiman filter corresponding to the integral constraints (3.1):

$$H_w^B(\omega) = \left[1 - \{\lambda_0 + \lambda_1(1 - \cos\omega)\}^{1/2}\right]^+ ,$$

where λ_0, λ_1 are chosen such that (3.1) is satisfied with equality for $f_s = f_s^B$.

Breiman started from some vague knowledge that the signal power is concentrated in the low frequencies. In engineering applications, a more traditional approach to that type of filtering problem under uncertainty would be to model the signal as an autoregressive process of order 1 (compare Vastola and Poor, 1983), i.e. to fix f_s as the nominal spectral density

$$f_s^A(\omega) = \frac{\sigma^2}{|1 - \alpha e^{i\omega}|^2},$$

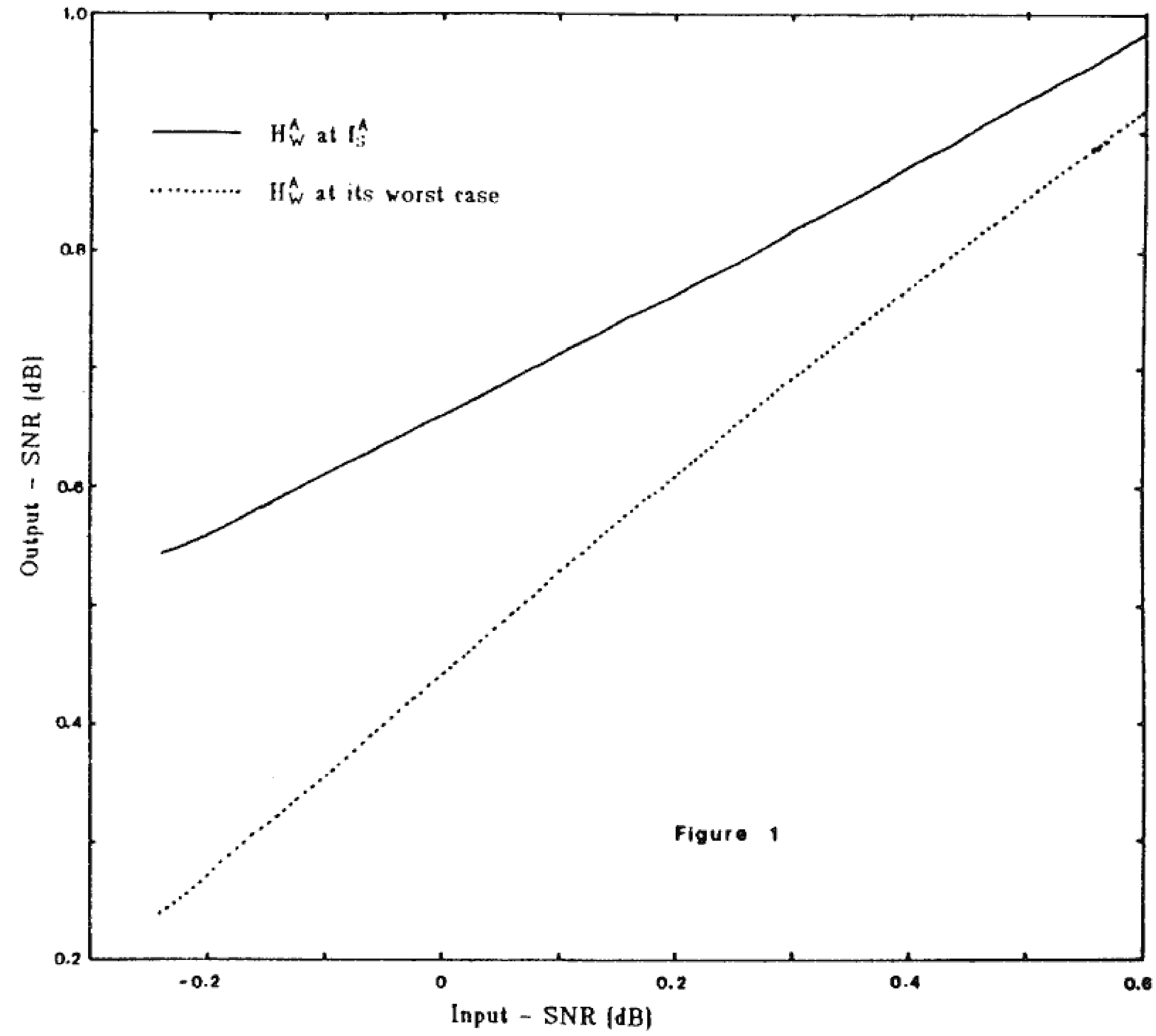
where σ^2, $0 < \alpha < 1$ are chosen to reflect the knowledge about the total signal power and the relative signal power in the low frequencies. To allow for comparisons, we choose α, σ^2 such that f_s^A satisfies (3.1) with equality, i.e. $\alpha = 1 - c'/c$ and $\sigma^2 = (1 - \alpha^2)c$. Let H_w^A denote the Wiener filter with respect to f_s^A, g.

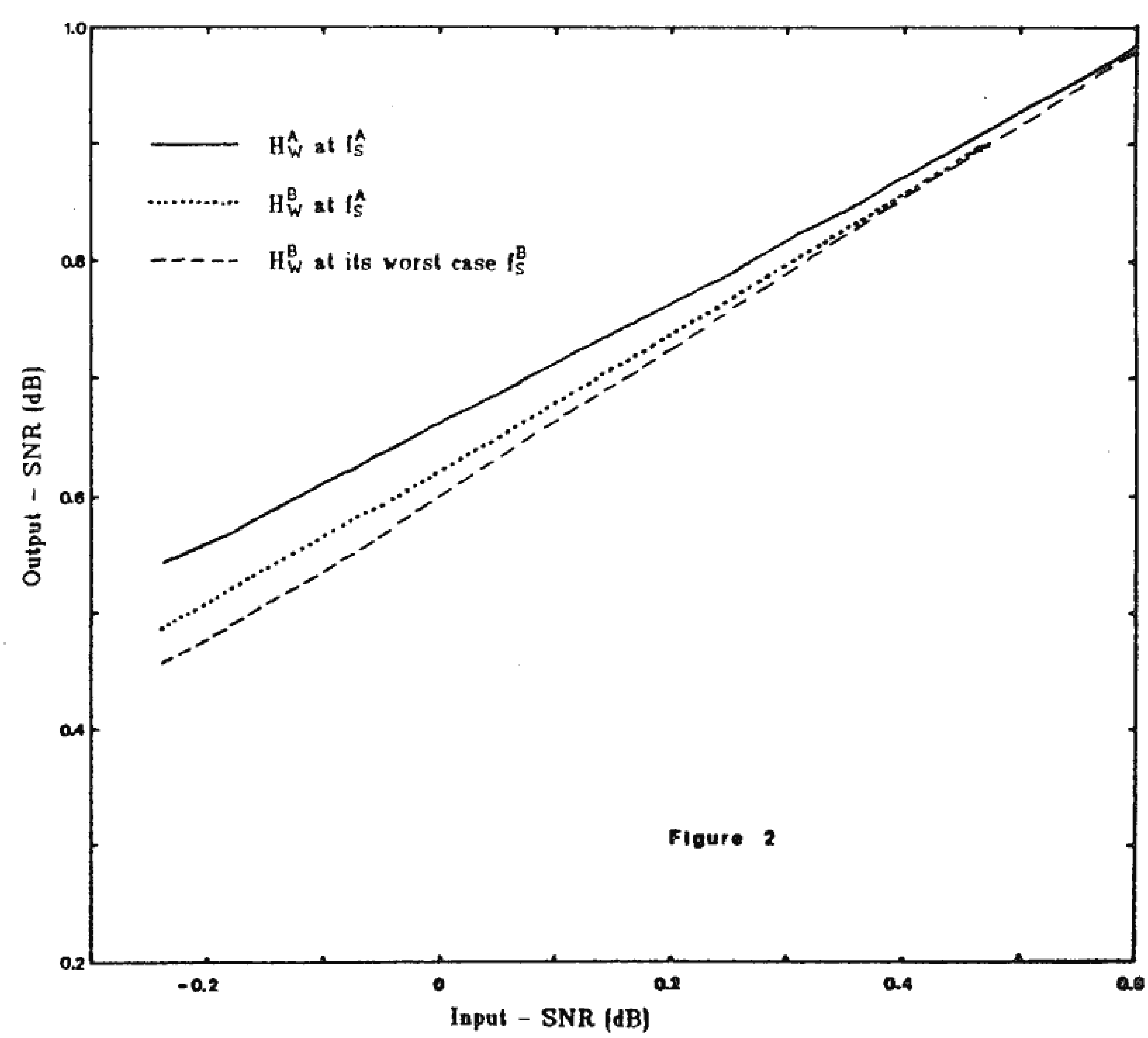
Figure 1 is based on the particular choice $g \equiv 1$, i.e. $\{N_t\}$ is white noise with unit variance, and $c' = c/10$. As $EN_t^2 = 1$, we always have $c = \rho_{in}$, provided f_s satisfies the first relation of (3.1) with equality. The unbroken curve of Figure 1 shows $\rho_{out}(f_s^A, g; H_w^A)$, i.e. the performance of H_w^A, provided f_s^A is really the spectral density of $\{S_t\}$. The dotted curve represents the worst case performance of H_w^A, i.e.

$$\inf_{(f_s,g)\in\mathbf{S}_e} \rho_{out}\left(f_s, g; H_w^A\right),$$

where, for getting the right scale, $\mathbf{S}_e$ consists only of those $(f_s, g) \in \mathbf{S}$ for which f_s satisfies the first relation of (3.1) with equality.

We see that, for low ρ_{in}, the performance of the traditional approach, i.e. of choosing the filter H_w^A can be pretty bad if our imprecise knowledge about the signal allows for variation of (f_s, g) over the whole set $\mathbf{S}$. Similar observations have been made by Vastola and Poor (1983) for other types of spectral information.


1.0
0.8
0.6
0.4
0.2
-0.2
0
0.2
0.4
0.6
Output - SNR (dB)
Input - SNR (dB)
H^A_W at f^A_S
H^A_W at its worst case
Figure 1


1.0
0.8
0.6
0.4
0.2
-0.2
0
0.2
0.4
0.6
Output - SNR (dB)
Input - SNR (dB)
H^A_W at f^A_S
H^B_W at f^A_S
H^B_W at its worst case f^B_S
Figure 2

The unbroken curve of Figure 2 again shows $\rho_{out}(f_s^A, g; H_w^A)$. The broken curve represents $\rho_{out}(f_s^B, g; H_w^B)$, i.e. the performance of H_w^B if $\{S_t\}$ has the spectral density f_s^B. As H_w^B is minimax-robust with respect to $\mathbf{S}$, this curve also describes the worst case performance of H_w^B, i.e. it gives a uniform lower bound of $\rho_{out}(f_s, g; H_w^B)$ where f_s varies over all spectral densities satisfying (3.1). The dotted curve represents $\rho_{out}(f_s^A, g; H_w^B)$: it describes the performance of the robust procedure if F_s^A is the spectral density of the signal $\{S_t\}$. Comparing Figures 1 and 2, we observe that for using the robust approach, i.e. for choosing H_w^B, we have to pay with a small increase in mean-square error if the traditional approach happens to be the optimal one in the sense that H_w^A really is the Wiener filter with respect to the signal and noise spectral densities in question. On the other hand, the robust approach provides a considerable improvement of the worst case performance.

To investigate the effect of additional constraints of f_s of the type (1.5), we consider the spectral information set $\mathbf{S}'$ consisting of all pairs (f_s, g), where f_s satisfies (3.1) and, for suitable $0 < \omega_a < \omega_b < \pi$,

$$(3.2) \qquad \begin{aligned} f_s(\omega) &\geq a \quad \text{for almost all} \quad \omega \in [0, \omega_a], \\ f_s(\omega) &\leq b \quad \text{for almost all} \quad \omega \in [\omega_b, \pi]. \end{aligned}$$

As $g \equiv 1$, a choice of $a > 1$, $b < 1$ corresponds to the idea of a low frequency signal in broadband noise. Let (f_s^L, g) denote a least favorable pair with respect to $\mathbf{S}'$, and let H_w^L denote the Wiener filter with respect to f_s^L, g.

We choose moderate signal-to-noise ratio bounds: $a = 5$, $\omega_a = \pi/10$, $b = 0.1$, $\omega_b = \pi/2$. The unbroken curve of Figure 3 shows $\rho_{out}(f_s^L, g; H_w^L)$ which, due to Theorem 2, represents also the worst case performance of H_w^L:

$$\inf_{(f_s, g) \in \mathbf{S}'_e} \rho_{out}\left(f_s^L, g; H_w^L\right),$$

where $\mathbf{S}'_e$ consists of those $(f_s^L, g) \in \mathbf{S}'$ for which f_s satisfies the first relation of (3.1) with equality. We remark that it is not possible to consider too small $c = \rho_{in}$, as (3.1) and (3.2) contradict each other if $\pi c < a\omega_a$ or $\pi c' < a(\omega_a - \sin(\omega_a))$.

We expect that using the first constraint of (3.2) improves the performance of the minimax-robust filter if the signal is weak, i.e. if ρ_{in} is small. For the example of Figure 3, this conjecture is confirmed. The figure also shows that there is not much difference between the performance of H_w^L and the Breiman filter H_w^B if signal and noise are approximately of the same size. This effect can be explained by observing that, for medium-sized p_{in}, (f_s^B, g) is almost contained in the smaller spectral information set $\mathbf{S}'$. For large ρ_{in}, we expect that using the second constraint of (3.2) cause a better performance of H_w^L compared to H_w^B. This effect can just be recognized at the right end of Figure 3. As $g \equiv 1$, we can conclude that, for $c \to \infty$, H_w^B converges to 1, whereas H_w^L converges to $1_{(0,\omega_b)} + h_b 1_{(\omega_b,\pi)}$ with $h_b = b/(1+b)$. Therefore, for $c \to \infty$,

$$\rho_{out}\left(f_s^B, g; H_w^B\right)/\rho_{in} \to 1,$$

$$\rho_{out}\left(f_s^L, g; H_w^L\right)/\rho_{in} \to \frac{\pi}{\omega_b + (\pi - \omega_b) h_b} > 1.$$

Finally, let us remark that, by changing the fixed ratio c'/c and the parameters a, b, ω_a, ω_b, we arrived qualitatively at the same conclusions as exhibited by Figures 1 to 3.

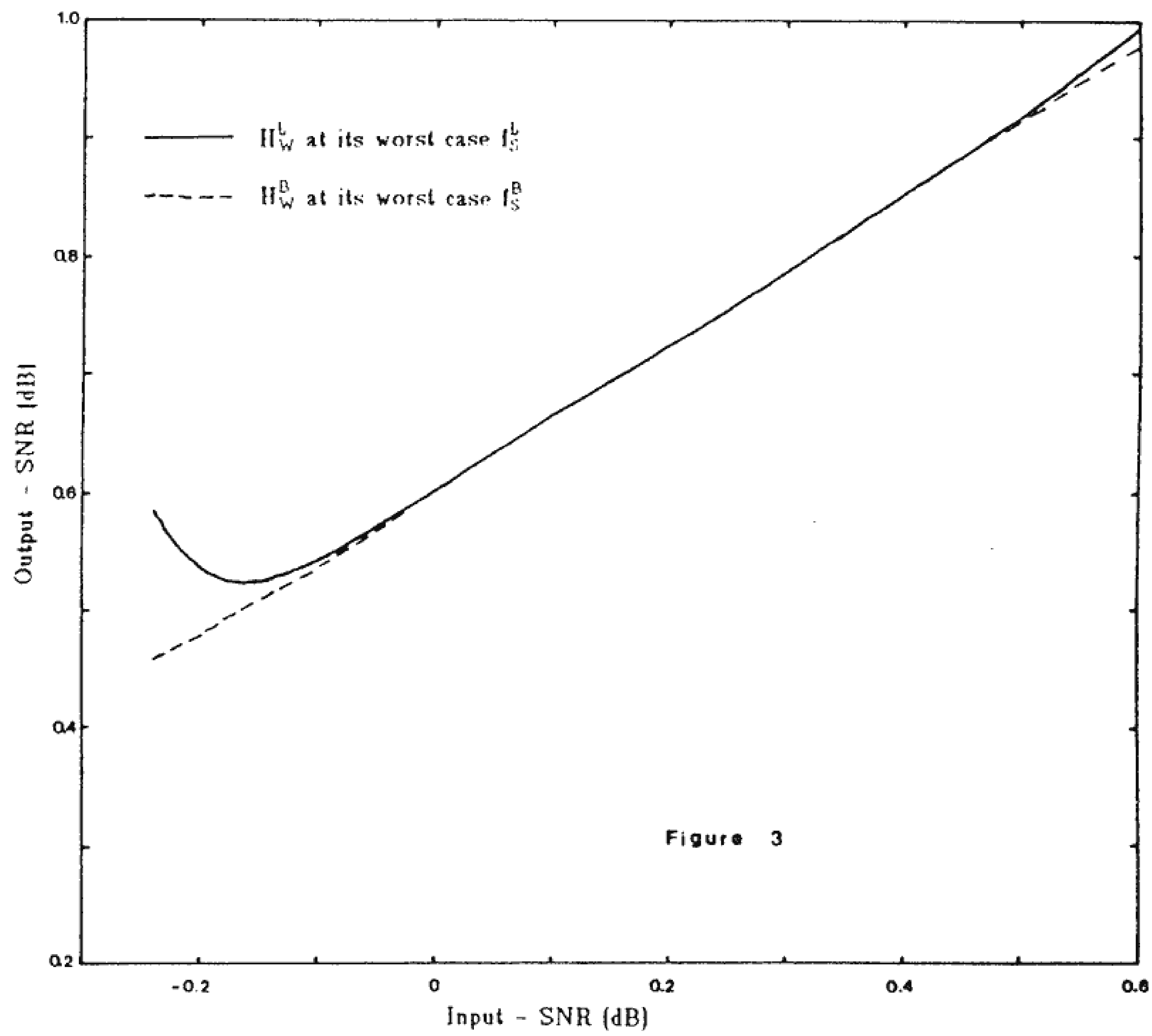


Figure 3

APPENDIX

Proof of Theorem 2. To prove the theorem some auxiliary results are needed. First, we have to derive the general form of supports of the spectral information set S at a point (f_s^L, f_n^L), where such a *support* is defined as a continuous linear functional Φ on $L^1 \times L^1$ satisfying

$$\Phi\left(f_s, f_n\right) \geq \Phi\left(f_s^L, f_n^L\right) \quad \text{for all} \quad \left(f_s, f_n\right) \in \mathrm{S}.$$

Lemma A1. *Let $c, c_0, \dots, c_m$ be positive constants; let $p_1, \dots, p_m$ be bounded measurable functions on $(-\pi, \pi]$; and let $p_0(\omega) \equiv 1$. Let $\mathbf{J}$ denote the set of all pairs (f_s, f_n) of integrable functions on $(-\pi, \pi]$ satisfying*

$$\frac{1}{2\pi}\int f_n(\omega)\,\mathrm{d}\omega \leq c$$

$$\frac{1}{2\pi}\int p_j(\omega) f_s(\omega)\,\mathrm{d}\omega \leq c_j \qquad j = 0, \dots, m.$$

A continuous linear functional Φ on $L^1 \times L^1$ is a support of $\mathbf{J}$ at $(f_s^L, f_n^L) \in \mathbf{J}$ iff there exist nonnegative $\lambda, \lambda_0, \dots, \lambda_m$ satisfying

$$(A1) \qquad \begin{aligned} \lambda &= 0 \quad \text{if} \quad \frac{1}{2\pi}\int f_n^L(\omega)\,\mathrm{d}\omega < c, \\ \lambda_j &= 0 \quad \text{if} \quad \frac{1}{2\pi}\int p_j(\omega) f_s^L(\omega)\,\mathrm{d}\omega < c_j, \end{aligned}$$

such that

$$\Phi\left(f_s, f_n\right) = -\frac{1}{2\pi}\int \left\{ \lambda f_n(\omega) + \sum_{j=0}^{m} \lambda_j p_j(\omega) f_s(\omega) \right\} \mathrm{d}\omega.$$

Proof. Let P be the continuous operator from $L^1 \times L^1$ into $\mathbb{R}^{m+2}$ given by

$$\left\{P\left(f_s, f_n\right)\right\}_j = \frac{1}{2\pi}\int p_j(\omega) f_s(\omega)\,\mathrm{d}\omega \qquad j = 0, \dots, m$$

$$\left\{P\left(f_s, f_n\right)\right\}_{m+1} = \frac{1}{2\pi}\int f_n(\omega)\,\mathrm{d}\omega.$$

Let

$$Q = \left\{ y = \left(y_0, \dots, y_{m+1}\right) \in \mathbb{R}^{m+2};\ y_j \leq c_j,\ j = 0, \dots, m, \ \text{and} \ y_{m+1} \leq c \right\}.$$

Then,

$$\mathbf{J} = \left\{ \left(f_s, f_n\right); P\left(f_s, f_n\right) \in Q \right\}.$$

Let ψ be a continuous linear functional on $\mathbb{R}^{m+2}$ given by

$$\psi(y) = -\sum_{j=0}^{m} \lambda_j y_j - \lambda y_{m+1}.$$

By definition ψ is a support of Q at $y^L = P(f_s^L, f_n^L)$ iff

$$(A2) \qquad \psi(y) \geq \psi\left(y^L\right) \quad \text{for all} \quad y \in Q.$$

(A2) is equivalent to requiring

$$(A3) \qquad \begin{aligned} &\lambda_j \geq 0,\ \lambda_j = 0 \ \text{ if } \ y_j^L < c_j, \qquad j = 0, \ldots, m, \\ &\lambda \geq 0,\ \lambda = 0 \ \text{ if } \ y_{m+1}^L < c. \end{aligned}$$

Let P^* be the dual operator of P. By Theorem 4.2.2 of Ioffe and Tihomirov (1979), Φ is a support of $\mathbf{J}$ at (f_s^L, f_n^L) iff there exists a support ψ of Q at y^L such that $\Phi = P^*\psi$. Together with (A3) this remark implies the assertion of the Lemma. ∎

Lemma A2. *Let $g_1, g_u, \beta_1, \beta_u$ be functions satisfying (2.3), (2.5). Let $\mathbf{B}$ denote the set of all pairs (f_s, f_n) of integrable functions on $(-\pi, \pi]$ satisfying*

$$\begin{aligned} &g_1 \leq f_n \leq g_u \qquad \text{a.e.}, \\ &\beta_1 f_n \leq f_s \leq \beta_u f_n \qquad \text{a.e.}. \end{aligned}$$

A continuous linear functional Φ on $L^1 \times L^1$ is a support of $\mathbf{B}$ at $(f_s^L, f_n^L) \in \mathbf{B}$, iff there exist bounded measurable functions φ, γ satisfying

$$(A4) \qquad \begin{aligned} &\varphi \leq 0 \quad \text{a.e. on the set} \quad B_u = \left\{\omega; \beta_1 f_n^L < f_s^L = \beta_u f_n^L\right\} \\ &\varphi = 0 \quad \text{a.e. on the set} \quad B_i = \left\{\omega; \beta_1 f_n^L < f_s^L < \beta_u f_n^L\right\} \\ &\varphi \geq 0 \quad \text{a.e. on the set} \quad B_1 = \left\{\omega; \beta_1 f_n^L = f_s^L < \beta_u f_n^L\right\} \end{aligned}$$

and

$$(A5) \qquad \begin{aligned} &\gamma \leq 0 \quad \text{a.e. on the set} \quad G_u = \left\{\omega; g_1 < f_n^L = g_u\right\} \\ &\gamma = 0 \quad \text{a.e. on the set} \quad G_i = \left\{\omega; g_1 < f_n^L < g_u\right\} \\ &\gamma \geq 0 \quad \text{a.e. on the set} \quad G_1 = \left\{\omega; g_1 = f_n^L < g_u\right\} \end{aligned}$$

such that

$$(A6) \quad \Phi\left(f_s, f_n\right) = -\frac{1}{2\pi} \int \gamma(\omega) f_n(\omega) \, \mathrm{d}\omega + \frac{1}{2\pi} \int \varphi(\omega) \left\{f_s(\omega) - \beta(\omega) f_n(\omega)\right\} \mathrm{d}\omega,$$

where

$$\beta(\omega) = \begin{cases} \beta_u(\omega) & \omega \in B_u \\ \beta_1(\omega) & \omega \notin B_u . \end{cases}$$

Proof. Let Φ be a continuous linear functional on $L^1 \times L^1$ given by

$$\Phi\left(f_s, f_n\right) = -\frac{1}{2\pi} \int \{\varphi(\omega) f_s(\omega) + \psi(\omega) f_n(\omega)\} \,\mathrm{d}\omega,$$

where φ, ψ are bounded functions. By definition, Φ is a support of $\mathbf{B}$ at (f_s^L, f_n^L) iff

$$(A7) \qquad \int \varphi(\omega)\left\{f_s(\omega) - f_s^L(\omega)\right\} \mathrm{d}\omega + \int \psi(\omega)\left\{f_n(\omega) - f_n^L(\omega)\right\} \mathrm{d}\omega \geq 0 \quad \text{for all} \quad \left(f_s, f_n\right) \in \mathbf{B}.$$

Choosing f_s such that $(f_s, f_n^L) \in \mathbf{B}$, (A7) implies in particular

$$\int \varphi(\omega)\left\{f_s(\omega) - f_s^L(\omega)\right\} \mathrm{d}\omega \geq 0 \text{ for all } f_s \in L^1 \text{ with } \beta_1 f_n^L \leq f_s \leq \beta_u f_n^L \text{ a.e.}$$

This condition can only be true if (A4) is satisfied.

We define

$$\gamma = \varphi\beta + \psi.$$

As β_u is finite on B_u we conclude from the assumptions on β_1, β_u that β and, consequently, γ are bounded functions. Let f_n be an arbitrary integrable function satisfying $g_1 \leq f_n \leq g_u$ a.e. . Then the pair $(\beta f_n, f_n)$ is contained in $\mathbf{B}$, and (A7) implies

$$\int \gamma(\omega)\left\{f_n(\omega) - f_n^L(\omega)\right\} \mathrm{d}\omega \geq 0 \text{ for all } f_n \in L^1 \text{ with } g_1 \leq f_n \leq g_u \text{ a.e.}.$$

This condition can only be true if (A5) is satisfied.

We have shown that supports have to be of the form described in the Lemma. Vice versa, each functional of the form (A6), where φ, γ satisfy (A4), (A5), is a support of $\mathbf{B}$ at (f_s^L, f_n^L), as can be checked by a straightforward calculation. ∎

To continue with the proof of Theorem 2 let us first remark that, as H_w^L is bounded, condition (2.8) is equivalent to requiring $-e(.,.; H_w^L)$ to be a support of $\mathbf{S}$ at (f_s^L, f_n^L). As $\mathbf{S} = \mathbf{J} \cap \mathbf{B}$, and as $\mathbf{J}$ has non-empty interior, the Moreau-Rockafellar-Theorem (0.3.3 of

Ioffe and Tihomirov, 1979) implies that the supports of **S** at (f_s^L, f_n^L) are exactly the sums of supports of **J** and **B** at (f_s^L, f_n^L). From Lemma A1, A2, and from (1.1), (2.8) we conclude that a pair $(f_s^L, f_n^L) \in \mathbf{S}$ is least favorable in **S** iff there exist nonnegative $\lambda, \lambda_0, \ldots, \lambda_m$ satisfying (A1) and bounded functions φ, γ satisfying (A4), (A5) such that for all integrable f_s, f_n:

$$\int \{|1 - H_w^L(\omega)|^2 f_s(\omega) + |H_w^L(\omega)|^2 f_n(\omega)\} \, d\omega =$$

$$(A8) \qquad = \int \{\lambda f_n(\omega) + \Delta(\omega) f_s(\omega)\} \, d\omega - \int \gamma(\omega) f_n(\omega) \, d\omega +$$

$$- \int \varphi(\omega) \{f_s(\omega) - \beta(\omega) f_n(\omega)\} \, d\omega,$$

where $\Delta(\omega) = \sum_{j=0}^m \lambda_j p_j(\omega)$.

(A8) is equivalent to the following two conditions on H_w^L

$$(A9) \qquad |1 - H_w^L(\omega)|^2 = \Delta(\omega) - \varphi(\omega) \qquad \text{a.e.}$$

$$(A10) \qquad |H_w^L(\omega)|^2 = \beta(\omega)\varphi(\omega) + \lambda - \gamma(\omega) \qquad \text{a.e.}.$$

First, we use (A9) and the properties of φ to conclude that the minimax robust filter H_w^L has to be a trimmed Breiman-filter of the form given in condition 1) of Theorem 2. On the degenerate set where β_1, β_u coincide the SNR is known precisely, and, therefore,

$$H_w^L(\omega) = h_u(\omega) = h_1(\omega) \quad \text{a.e. on} \quad \{\omega; \beta_u(\omega) = \beta_1(\omega)\}.$$

Therefore, we may now assume $\beta_1(\omega) < \beta_u(\omega)$ and restrict our attention to the sets B_u, B_i, B_1 of (A4). As $1 - H_w^L = f_n^L/(f_s^L + f_n^L)$, we conclude from (A4) and (A9):

$$\Delta(\omega) \leq 1/\left(1 + \beta_u(\omega)\right)^2 = |1 - H_w^L(\omega)|^2 \text{if } \omega \in B_u,$$

$$|1 - H_w^L(\omega)|^2 = 1/\left(1 + \beta_1(\omega)\right)^2 \leq \Delta(\omega) \text{if } \omega \in B_1,$$

$$1/\left(1 + \beta_u(\omega)\right)^2 < \Delta(\omega) = |1 - H_w^L(\omega)|^2 < 1/\left(1 + \beta_1(\omega)\right)^2 \text{if } \omega \in B_i, \ f_n^L(\omega) > 0,$$

$$\beta_u(\omega) = \infty, \ 0 = \Delta(\omega) = |1 - H_w^L(\omega)|^2 \text{if } \omega \in B_i, \ f_n^L(\omega) = 0.$$

Therefore, using the definition of H_w^L (A9) is equivalent to conditions 1) and 3) of the theorem.

Finally, we show that condition 2) of the theorem is equivalent to (A10), provided conditions 1) and 3) are satisfied. We assume $\Delta(\omega) \neq 0$ or $\beta_u(\omega) < \infty$. By condition 3), $f_n^L(\omega) > 0$ and, consequently, $H_w^L(\omega) < 1$. Therefore, $J(\omega)$ is well-defined. (A9) and condition 1) imply

$$|H_w^L(\omega)|^2 - \beta(\omega)\varphi(\omega) = H_w^L(\omega)\left\{1 - \Delta(\omega)/\left(1 - H_w^L(\omega)\right)\right\} = J(\omega).$$

Therefore, we conclude from (A5), (A10):

$$\lambda \leq J(\omega) \text{ if } \omega \in G_u,$$

$$\lambda = J(\omega) \text{ if } \omega \in G_i,$$

$$\lambda \geq J(\omega) \text{ if } \omega \in G_1,$$

i.e. condition 2) is satisfied if $g_1(\omega) < g_u(\omega)$. On the set of frequencies where g_1, g_u coincide condition 2) is satisfied anyhow. ∎

ACKNOWLEDGEMENTS

This work has been supported by the Deutsche Forschungsgemeinschaft as a part of SFB 123 «Stochastische Mathematische Modelle». The whole work on chapter 3 has been done while the author was Visiting Fellow of the Department of Statistics, IAS, of the Australian National University, Canberra. The main part of this research has been done at the Department of Mathematics, University of Frankfurt.

REFERENCES

G. Baxter, *An asymptotic result for the finite predictor*, Math. Scandinavica, **10** (1962), pp. 137-144.

L. Breiman, *A note on minimax filtering*, Ann. Probab., **1** (1973), pp. 175-179.

L.J. Cimini, S.A. Kassam, *Robust and quantized Wiener filters for P-point spectral classes*, Proc. 14th Conf. on Inform. Sciences and Systems (1980), pp. 314-319, Princeton University, Princeton, N.J.

J. Franke, *Minimax-robust prediction of discrete time series*, Z. Wahrscheinlichkeitstheorie verw. Gebiete, **68** (1985), pp. 337-364.

J. Franke, *On the prediction and interpolation of a discrete time series in the presence of correlated noise*, J. Time Series Analysis, **5** (1984), pp. 227-244.

J. Franke, V.H. Poor, *Minimax-robust filtering and finite-length robust predictors*, in «Robust and nonlinear time series analysis», eds. J. Franke, W. Härdle, R.D. Martin. Lecture Notes in Statistics, **26** (1984), Springer, Berlin-Heidelberg-New York.

Y. Hosoya, *Robust linear extrapolation of second-order stationary processes*, Ann. Probab., **6** (1978), pp. 574-584.

A.D. Ioffe, V.M. Tihomirov, *Theory of extremal problems*, North-Holland Publ. Co., Amsterdam, 1979.

S.A. Kassam, *Robust hypothesis testing and robust time series interpolation and regression*, J. Time Series Analysis, **3** (1982), pp. 185-194.

S.A. Kassam, T.M. Lin, *Robust Wiener filters*, J. Franklin Inst., **304** (1977), pp. 171-185.

H.V. Poor, *On robust Wiener filtering*, IEEE Trans. Autom. Contr., **25** (1980), pp. 531-536.

M. Taniguchi, *Robust regression and interpolation for time series*, J. Time Series Analysis, **2** (1981), pp. 53-62.

K.S. Vastola, H.V. Poor, *An analysis of the effects of spectral uncertainty on Wiener filtering*, Automatica, **19** (1983), pp. 289-293.

K.S. Vastola, H.V. Poor, *Robust Wiener-Kolmogorov theory*, IEEE Trans. Inform. Theory, **30** (1984), pp. 316-327.

Received December 22, 1990
Jürgen Franke
Department of Mathematics
University of Kaiserslautern
P.O. Box 3049
D-6750 Kaiserslautern
Germany