

Focused Detection via Multiresolution Analysis

Chuang He, *Member, IEEE*, and José M. F. Moura, *Fellow, IEEE*

Abstract—In many applications, there are strong discrepancies between the signal models assumed in the design phase and the actual signals encountered in the field. These discrepancies penalize significantly the performance of the matched filter that is fine tuned to the preassumed conditions.

We propose a geometric framework that designs, via wavelet multiresolution-based techniques, a receiver whose performance is to a large degree insensitive to these mismatches. We say that the receiver is a focused detector. The approach defines a signal set \mathcal{S} that identifies the class of diverse conditions that are expected to arise. We illustrate the method in the context of multipath problems. The matched filter, which is a simple receiver, assumes that \mathcal{S} is a singleton. When this is not the case, the matched filter experiences strong degradation. On the other hand, the optimal receiver for the signal set \mathcal{S} is practically infeasible since it requires a multidimensional nonlinear optimization.

The paper designs the focused receiver as a good compromise between these two extremes. We replace the signal set \mathcal{S} by a linear subspace \mathcal{G} —the representation subspace—that minimizes a measure of similarity with \mathcal{S} . We choose \mathcal{G} to be a multiresolution subspace. This choice resolves to satisfaction several issues: The subspace design is reduced to the design of a single shiftable scaling function, the similarity between \mathcal{S} and \mathcal{G} can be computed explicitly, and the focused receiver that computes the energy of the orthogonal projection on \mathcal{G} is implemented by a bank of correlators matched to scaled/delayed versions of the reshaped scaling function followed by an energy detector.

We assume that the transmitted signal is a sample of a random process. The signal set \mathcal{S} becomes an ensemble of linear spaces. We introduce the modified deflection as the appropriate similarity measure. The paper details our algorithm, describes how to compute the modified deflection, and illustrates the performance results that can be obtained.

Index Terms—Deflection, focused detection, gap metric, maximum likelihood, mutiresolution, robust detection, shiftable wavelet.

I. INTRODUCTION

IN MOST applications, *optimally* designed receivers run under conditions that are at considerable variance with the assumptions underlying their design. These discrepancies between model and field conditions penalize significantly the performance of *optimal* receivers. In this paper, we design *robust* receivers, i.e., receivers that are resilient to conspicuous departures from their nominal working conditions.

Our strategy is to identify the type of diverse conditions that are expected to occur in actual operation. As an example, consider a communications propagation channel where multi-

ple paths from sender to receiver often beset the quality of a digital communication system due to energy spreading and the pattern of interference between the direct line of sight path and secondary paths. The multipath effect may be very complex as a result of the reflecting obstacles in a wireless digital communications urban environment or of the bottom/surface interactions in shallow water. The propagation channel is specific to the relative configuration of each sender–receiver pair and may vary if this configuration changes due, for example, to the mobility of the sender or the receiver. In this paper, we present our approach in the context of multipath problems.

When multipath is present, the filter matched to the single transmitted waveform is no longer optimal, and its performance deteriorates appreciably. We say that this matched filter is defocused. In contrast, the optimal receiver estimates the channel distorted signal, i.e., the multipath signal, and correlates this estimate with the received signal. We refer to this receiver as the channel matched detector (CMD). The CMD is practically infeasible since the channel estimation step requires a costly multidimensional nonlinear optimization. To focus the receiver without incurring the costs of the CMD, we design a good compromise between these two extremes: i) a simple but unreliable defocused matched filter and ii) a complex infeasible albeit high performance CMD.

The following are the basic tenets of our design:

- 1) *Signal set \mathcal{S}* : The set \mathcal{S} describes the class of signals that are expected to be received after being distorted by the channel. The complexity of the set \mathcal{S} depends on the nature of the transmitted signal $s(t)$. If $s(t)$ is a known deterministic waveform, then \mathcal{S} is a linear subspace. This problem was considered in [1] and [2]. The optimal CMD constructs the orthogonal projection of the actually received signal $r(t)$ on \mathcal{S} and uses the energy of this orthogonal projection on \mathcal{S} as the test statistic. Finding the orthogonal projection of $r(t)$ on \mathcal{S} is equivalent to the multidimensional nonlinear optimization problem referred to above. In [1] and [2], we developed an approach that avoided the multidimensional nonlinear optimization by approximating \mathcal{S} with a representation subspace \mathcal{G} whose orthogonal projection is easily computed. We then used the energy of the orthogonal projection of the received signal $r(t)$ on \mathcal{G} as the test statistic. In this paper, we assume that $s(t)$ is a sample function of a random process. In this case, the signal set \mathcal{S} is more complex. As we will see in the next section, it is essentially an ensemble of linear subspaces.
- 2) *Representation subspace \mathcal{G}* : We design a linear subspace \mathcal{G} —the representation subspace—that is close to the signal set \mathcal{S} and whose orthogonal projection is easily

Manuscript received February 15, 1997; revised November 30, 1997. This work was supported in part by DARPA through Grants ONR N00014-91-J183 and AFOSR F49620-96-1-0436 and ONR Grant N00014-97-0040.

The authors are with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA.

Publisher Item Identifier S 1053-587X(98)02499-4.

computed. Our receiver projects the received signal $r(t)$ on \mathcal{G} and uses the energy of the orthogonal projection of $r(t)$ on \mathcal{G} as the test statistic.

- 3) *Similarity measure $\hat{\delta}_{\text{mod}}$ between \mathcal{S} and \mathcal{G}* : To measure how close \mathcal{G} is to \mathcal{S} , we introduce in Section II a similarity measure: the modified deflection. Our approach designs \mathcal{G} by optimizing this similarity measure $\hat{\delta}_{\text{mod}}$.
- 4) *Multiresolution design of \mathcal{G}* : There are three hurdles that need to be resolved:
 - i) subspace design is in general difficult;
 - ii) calculating and optimizing the modified deflection $\hat{\delta}_{\text{mod}}$ raises conceptual as well as computational difficulties;
 - iii) complexity of the receiver implementation.

We resolve satisfactorily these issues by resorting to a multiresolution analysis [3], [4]. The multiresolution subspaces $\{\mathcal{G}_i\}$ are generated from a single scaling function $g(t)$ so that the subspace design is reduced to the design of this single scaling function. By choosing a multiresolution subspace as the representation subspace, we are able to compute analytically (see Theorem III.1 in Section III) the modified deflection $\hat{\delta}_{\text{mod}}$. Finally, shifted replicas of the scaling function $g(t)$ (or its scaled versions) provide a straightforward basis for the representation subspace \mathcal{G} so that the focused receiver is reduced to a bank of correlators followed by an energy detector.

There is a further subtlety with the modified deflection that arises because the modified deflection is a one-sided measure; see Section II. Because of this, we need to design the “smallest” subspace that minimizes the modified deflection. This makes it natural to choose a multiresolution subspace as the representation subspace \mathcal{G} . With multiresolution subspaces, our goal becomes finding the lowest scale multiresolution subspace that minimizes the modified deflection. We present a multiresolution subspace design algorithm that achieves this goal. Once we have found the desired representation subspace, the receiver is as simple as computing the energy of the orthogonal projection of the received signal on the representation subspace and comparing it with a threshold.

The presentation of this paper is organized as follows. In Section II, we give the formulation of the detection problem and introduce the modified deflection. In Section III, we describe in detail the multiresolution subspace design algorithm. In Section IV, we provide an example that illustrates the gains in performance afforded by the focused receiver with respect to alternative receivers. Finally, in Section V, we summarize the results.

II. FOCUSED DETECTION AND MODIFIED DEFLECTION

This section formulates the focused detection problem in the context of multipath, describes our geometric approach, and introduces the modified deflection. Finally, a brief review of multiresolution analysis theory is given.

A. Geometric Interpretation of Focused Detection

The focused detection problem is presented in the context of multipath. The problem we consider is the standard binary

hypothesis testing problem

$$H_1: r(t) = s_m(t) + n(t) \quad 0 \leq t \leq T \quad (1)$$

$$= \sum_{k=1}^K \alpha_k s(t - \tau_k) + n(t) \quad (2)$$

$$H_0: r(t) = n(t) \quad (3)$$

where $s_m(t)$ is the multipath noise free signal, and T is the observation interval. Throughout the paper, we assume that the number of paths K , the attenuation coefficients $\{\alpha_k\}$, and the delays $\{\tau_k\}$ are all deterministic unknown. For simplicity, we also assume that the additive noise $n(t)$ is white and Gaussian. The detection goal is to decide, from a single realization of the received signal $r(t)$, which hypothesis (H_1 or H_0) is true.

In [1] and [2], we dealt with the case where the transmitted signal $s(t)$ is known. In this paper, we assume that $s(t)$ is a finite energy random process with autocorrelation function $K(t, u)$. The autocorrelation function $K(t, u)$ is either known or can be estimated from training samples of the received signal.

If $s(t)$ is known as in [1] and [2], the generalized likelihood ratio test (GLRT) statistic is the norm square of the orthogonal projection of the received signal on the signal subspace, i.e.,

$$L = \|P_{\mathcal{S}}[r]\|_2^2 \quad (4)$$

where

$$\mathcal{S} = \left\{ s_m(t; \theta) = \sum_{k=1}^K \alpha_k s(t - \tau_k), K \in \mathbb{Z}^+, \alpha_k, \tau_k \in \mathbb{R} \right\} \quad (5)$$

is the signal subspace, and $P_{\mathcal{S}}$ is the orthogonal projection operator on \mathcal{S} . The GLRT detector is the channel matched detector (CMD) mentioned in the Introduction. The parameter $\theta = \{K, \{a_k\}, \{\tau_k\}\}$ is the collection of all the channel parameters. The orthogonal projection ($P_{\mathcal{S}}[r]$)(t) of the received signal $r(t)$ on the signal subspace \mathcal{S} is also the maximum likelihood (ML) estimate $\hat{s}_m^{\text{ML}}(t)$ of the multipath noise-free signal $s_m(t)$, which is written explicitly as

$$\hat{s}_m^{\text{ML}}(t) = (P_{\mathcal{S}}[r])(t) = s_m(t; \theta^*) \\ \hat{\theta}_{\text{ML}} = \theta^* = \arg \min_{\theta \in \Theta} \|r(\cdot) - s_m(\cdot; \theta)\|_2^2 \quad (6)$$

where Θ is the set of all possible values of the parameter θ . Solving (6) requires a multidimensional nonlinear optimization because $s_m(\cdot; \theta)$ is a nonlinear function of the channel delays $\{\tau_k\}$ and the number of paths K . In general, the CMD is not feasible for practical applications. Instead, in [1] and [2], we developed an algorithm where we approximate the signal subspace \mathcal{S} by a representation subspace \mathcal{G} such that the orthogonal projection on \mathcal{G} is easily computed. To obtain good performance, we required that \mathcal{G} be close to \mathcal{S} in some sense. We used the gap metric [5] to measure the similarity between the subspaces \mathcal{S} and \mathcal{G} . The representation subspace \mathcal{G} was designed to minimize the gap between \mathcal{S} and \mathcal{G} . Once we have designed the representation subspace \mathcal{G} , we used the energy of the orthogonal projection on \mathcal{G}

$$L' = \|P_{\mathcal{G}}[r]\|_2^2 \quad (7)$$

as the test statistic.

When the transmitted signal $s(t)$ is a random process, the signal set \mathcal{S} of interest is no longer a subspace in $L_2(\mathbb{R})$ (the space of all finite energy functions). By noting that, for each sample realization of the signal $s(t)$, the signal set \mathcal{S} is a linear subspace in $L_2(\mathbb{R})$, we conclude that the signal set \mathcal{S} is now an ensemble of subspaces. Therefore, we need to design a representation subspace \mathcal{G} that approximates \mathcal{S} in an ensemble average sense. Because \mathcal{S} is no longer a subspace in $L_2(\mathbb{R})$, the gap metric cannot be used. In this paper, we introduce the novel concept of the *modified deflection* to adjust to the special structure of the new signal set \mathcal{S} . In the next subsection, we introduce the definition of the modified deflection.

B. Modified Deflection

The modified deflection is an ensemble average measure of the deviation of a set with finite energy random elements from a linear subspace in $L_2(\mathbb{R})$. It generalizes the deflection [6], [7]. The deflection measures the deviation of a *set* in $L_2(\mathbb{R})$ from a *linear subspace* in $L_2(\mathbb{R})$. The deflection itself is a generalization of one side of the gap metric used in [1] and [2]. In this subsection, we recall the definition of the deflection and introduce the modified deflection.

1) *Deflection* [6], [7]: Let \mathcal{S} be a set in $L_2(\mathbb{R})$, $S_{\mathcal{S}} = \{u \in \mathcal{S}, \|u\|_2 = 1\}$, and \mathcal{G} a linear subspace in $L_2(\mathbb{R})$. The deflection of \mathcal{S} from \mathcal{G} is

$$\hat{\delta}(\mathcal{S}, \mathcal{G}) = \sup_{u \in S_{\mathcal{S}}} \text{dist}(u, \mathcal{G}) \quad (8)$$

where

$$\text{dist}(u, \mathcal{G}) = \inf_{v \in \mathcal{G}} \|u - v\|_2 \quad (9)$$

is the distance from the vector u to the subspace \mathcal{G} .

2) *Modified Deflection*: When the set \mathcal{S} has finite energy random elements, we modify the deflection measure. Let \mathcal{S} be a set with finite energy random elements and \mathcal{G} a linear subspace in $L_2(\mathbb{R})$. The modified deflection of \mathcal{S} from \mathcal{G} is

$$\hat{\delta}_{\text{mod}}(\mathcal{S}, \mathcal{G}) = \sqrt{\sup E\{\text{dist}^2(u, \mathcal{G})\}} \quad (10)$$

where $E\{\cdot\}$ is the expectation operator taken with respect to u , and the supreme is subject to

$$E\{\|u\|_2^2\} = 1. \quad (11)$$

The modified deflection takes values between 0 and 1. The major difference between (8) and (10) is that (10) involves the expectation operator. This definition is a natural generalization of the deflection. It is essentially a worst-case mean square error measure. If the modified deflection of \mathcal{S} from \mathcal{G} is equal to zero, then all the elements in \mathcal{S} are well represented by \mathcal{G} in the mean square sense. Another possible measure would be

$$\sup E\{\text{dist}(u, \mathcal{G})\}. \quad (12)$$

Equation (10) has the advantage over (12) in that it leads to an analytical treatment of the modified deflection; see Theorem III.1 in Section II-C. This is similar to using the mean squared error $E\{\|\cdot\|_2^2\}$ rather than $E\{\|\cdot\|_2\}$.

Both the deflection and the modified deflection are one-sided measures, i.e., they address only the question of how

\mathcal{S} is “included” in \mathcal{G} but not vice versa. Specifically, if the modified deflection of \mathcal{S} from \mathcal{G} is zero, then the modified deflection of \mathcal{S} from any subspace that includes \mathcal{G} is also zero. An important example is the modified deflection of any set \mathcal{S} (with finite energy random elements) from $L_2(\mathbb{R})$. This modified deflection is zero. In this case, the design problem is trivial since $L_2(\mathbb{R})$ is the space minimizing the modified deflection. However, $L_2(\mathbb{R})$ is not a desirable representation space because it is too large. In fact, what is needed is the “smallest” subspace \mathcal{G} that minimizes the modified deflection of \mathcal{S} from \mathcal{G} . The “smaller” the representation subspace, the smaller the number of coefficients needed to represent the orthogonal projection on the representation subspace. The smaller the number of coefficients, the smaller the number of degrees of freedom for the test statistic and, thus, the better the detection performance.

Minimizing the modified deflection alone does not guarantee that we find the “smallest” subspace. This is in contrast with the gap metric used in [1] and [2], which is a two-sided metric. Minimizing the gap leads to the desired smallest subspace. Here, with the signal set \mathcal{S} , we need to make explicit the additional constraint that the representation subspace is the “smallest” subspace minimizing the modified deflection.

We are left with the issue of designing the “smallest” subspace. This is where the role of a multiresolution analysis comes naturally into play. With multiresolution subspaces (see next subsection), our problem is reduced to finding the lowest scale multiresolution subspace that minimizes the modified deflection of \mathcal{S} from \mathcal{G} .

C. Multiresolution Subspaces

In this section, we give a brief introduction to multiresolution analysis theory. An orthonormal multiresolution analysis is an increasing sequence of closed linear subspaces [4]

$$\cdots \subset \mathcal{G}_{-2} \subset \mathcal{G}_{-1} \subset \mathcal{G}_0 \subset \mathcal{G}_1 \subset \mathcal{G}_2 \subset \cdots$$

that satisfy the following properties.

- 1) $\bigcup_i \mathcal{G}_i = L_2(\mathbb{R})$ and $\bigcap_i \mathcal{G}_i = \{0\}$.
- 2) $g(2^i t) \in \mathcal{G}_i$ if and only if $g(2^{i+1} t) \in \mathcal{G}_{i+1}$, $i \in \mathbb{Z}$.
- 3) $g(2^i t) \in \mathcal{G}_i$ implies $g(2^i t - k) \in \mathcal{G}_i$, $i, k \in \mathbb{Z}$.
- 4) $\{g(t - k), k \in \mathbb{Z}\}$ is an orthonormal basis for the subspace \mathcal{G}_0 and $g(t)$ is called the scaling function.

The scaling function $g(t)$ satisfies the two-scale equation

$$g(t) = \sum_k c_k g(2t - k) \quad (13)$$

where $\{c_k\}$ are the scaling coefficients that satisfy given conditions [3], [8]. Daubechies [3] provides a complete characterization of the power spectrum of the sequences of scaling coefficients that lead to compactly supported orthonormal scaling functions and wavelets. Zou and Tewfik [9] provide a further parameterization of these coefficients. They show that for a given support $2P - 1$, there is an infinite number of scaling functions and mother wavelets. They also show that all the scaling functions and wavelets of support $2P - 1$ are parameterizable by choosing just $P - 1$ parameters $(\zeta_1, \zeta_2, \dots, \zeta_{P-1})$ over $[0, 2\pi]^{P-1}$.

The structure of the subspaces $\{\mathcal{G}_i\}$ is given by

$$\mathcal{G}_i = \left\{ \sum_{n=-\infty}^{+\infty} 2^{i/2} \beta_n g(2^i t - n), \beta_n \in \mathbb{R} \right\}. \quad (14)$$

These multiresolution subspaces $\{\mathcal{G}_i\}$ are the candidates for our representation subspace. Specifically, we look for the subspace \mathcal{G}_i that minimizes the modified deflection of \mathcal{S} from \mathcal{G}_i with the smallest scale index i .

III. MULTIREOLUTION SUBSPACE DESIGN

In this section, we describe the multiresolution subspace design in detail. In Section III-A, we rewrite the signal set \mathcal{S} in a more explicit form using the truncated Karhunen–Loève (K–L) expansion, which we will use to compute the modified deflection. In Section III-B, we present the design algorithm. In Section III-C, we provide an explicit formula for the modified deflection. Finally, in Section III-D, we discuss how to minimize the modified deflection.

A. Signal Set \mathcal{S}

When the transmitted signal $s(t)$ is a random process, the signal set

$$\mathcal{S} = \left\{ s_m(t : \theta) = \sum_{k=1}^K \alpha_k s(t - \tau_k), K \in \mathbb{Z}^+, \alpha_k, \tau_k \in \mathbb{R} \right\} \quad (15)$$

is a set of random processes. We assume that all the elements in \mathcal{S} have finite energy; in addition, the autocorrelation function $K(t, u)$ of the transmitted signal $s(t)$ is either known or is estimated from the available data. With knowledge of the autocorrelation function $K(t, u)$, we express the signal set \mathcal{S} in a more explicit form.

From the K–L expansion, we decompose $s(t)$ as [10]

$$s(t) = \sum_{m=1}^{+\infty} \gamma_m h_m(t) \quad (16)$$

where the functions $\{h_m(t), m \geq 1\}$ are the orthonormal eigenfunctions of $K(t, u)$ satisfying the eigenequation

$$\int_0^{T'} K(t, u) h_m(u) du = \sigma_m^2 h_m(t), \quad m = 1 \dots \quad (17)$$

where $\{\sigma_m^2, m \geq 1\}$ are the corresponding eigenvalues, and T' is the time interval within which the K–L expansion is performed. It can be equal to the observation interval T or smaller than the observation interval if the transmitted signal $s(t)$ is a transient. In our simulations, T' is estimated from the training data. The coefficients $\{\gamma_m, m \geq 1\}$ are statistically orthogonal random variables [11]. They are given by

$$\gamma_m = \langle s(\cdot), h_m(\cdot) \rangle \quad m \geq 1 \quad (18)$$

and

$$E\{\gamma_m^2\} = \sigma_m^2 \quad m \geq 1. \quad (19)$$

Without loss of generality, we assume that $\sigma_1^2 > \sigma_2^2 > \dots > \sigma_n^2 > \dots$. Since the coefficients $\{\gamma_m, m \geq 1\}$ are statistically

orthogonal, for practical reasons, we truncate the expansion in (16) to a finite sum

$$s(t) \approx \sum_{m=1}^M \gamma_m h_m(t). \quad (20)$$

The choice of M depends on the energy distribution represented by the eigenvalues $\{\sigma_m^2, m \geq 1\}$ and the desired approximation accuracy.

With the truncated K–L expansion, the signal set \mathcal{S} can be written as

$$\mathcal{S} = \left\{ s_m(t : \theta) = \sum_{k=1}^K \alpha_k s(t - \tau_k), K \in \mathbb{Z}^+, \alpha_k, \tau_k \in \mathbb{R} \right\} \quad (21)$$

$$\approx \left\{ \sum_{k=1}^K \alpha_k \sum_{m=1}^M \gamma_m h_m(t - \tau_k), K \in \mathbb{Z}^+, \alpha_k, \tau_k \in \mathbb{R} \right\}. \quad (22)$$

In most applications, we can only estimate the autocorrelation function $K(t, u)$. The training data set used to estimate the autocorrelation function $K(t, u)$ is often corrupted by noise. In these situations, the truncated K–L expansion is preferred because of its noise reduction property.

B. Subspace Design Algorithm

Our goal is to design the lowest scale multiresolution subspace that minimizes the modified deflection of the signal set \mathcal{S} given by (22) from the representation subspace. In other words, we want to find the multiresolution subspace with the smallest scale index i

$$\mathcal{G}_i = \left\{ \sum_{n=-\infty}^{+\infty} 2^{i/2} \beta_n g(2^i t - n), \beta_n \in \mathbb{R} \right\} \quad (23)$$

that minimizes the modified deflection. With the structure of the multiresolution subspaces given in (23), the subspace design problem is reduced to the design of the single function $g(t)$. We call $g(t)$ the generating function. This will be a reshaped scaling function.

Solving this design problem directly is still a difficult task. Just calculating the modified deflection of \mathcal{S} from \mathcal{G}_i is a computationally intensive multidimensional nonlinear optimization problem, let alone minimizing it. Like we did in [1] and [2], we attack the problem in two major steps.

1) *Integer Shift Signal Set \mathcal{S}_{int}* : In the first step, we design the representation subspace \mathcal{G}_i that matches in the modified deflection sense the integer shift signal set

$$\begin{aligned} \mathcal{S}_{int} &= \left\{ \sum_{n=-\infty}^{+\infty} \alpha_n s(t - n), \alpha_n \in \mathbb{R} \right\} \\ &\approx \left\{ \sum_{n=-\infty}^{+\infty} \alpha_n \sum_{m=1}^M \gamma_m h_m(t - n), \alpha_n \in \mathbb{R} \right\}. \end{aligned} \quad (24)$$

The design goal of the first step is to design the generating function $g(t)$ such that the modified deflection of \mathcal{S}_{int} from \mathcal{G}_i is minimized with the smallest scale index i . We denote

the resulting optimal generating function as $g_{i^*}^*(t)$ and the smallest scale index as i^* .

By restricting the channel delays to an integer lattice, we are able to compute explicitly the modified deflection of \mathcal{S}_{int} from \mathcal{G}_i because it is no longer a nonlinear optimization problem. This is achieved in the theorem presented in the next subsection.

To perform the minimization, we restrict the generating function $g(t)$ to be a compactly supported orthonormal scaling function so that we can use the parameterization given in [9]. This reduces the functional optimization to a finite parameter optimization. In addition, compactly supported orthonormal scaling functions provide an advantage when dealing with practical detection problems. In practice, only signals of finite duration are of interest. The focused detector we develop in this paper is implemented by a bank of correlators, each of which are matched to an integer time-shifted replica of the generating function (or its scaled version), which is a reshaped compactly supported orthonormal scaling function; see Section IV. Due to the compact support of the received signal $r(t)$ and of the generating function $g(t)$, there is a maximum integer shift of a scaled version of $g(t)$, beyond which $r(t)$ and the shifted replica of the scaled version of $g(t)$ have zero overlap. This is not the case if $g(t)$ is not compactly supported. If the generating function is not of compact support, more correlators will, in general, be needed at the receiver than when the generating function is of compact support. This increases the number of degrees of freedom of the detection test statistic and degrades the detection performance accordingly.

Once the minimization is accomplished, the resulting subspace

$$\mathcal{G}_{i^*}^* = \left\{ \sum_{n=-\infty}^{+\infty} 2^{i^*/2} \beta_n g_{i^*}^*(2^{i^*} t - n), \beta_n \in \mathbb{R} \right\} \quad (25)$$

is a good approximation to \mathcal{S}_{int} . However, \mathcal{S}_{int} is not \mathcal{S} . The original signal set \mathcal{S} is much richer than \mathcal{S}_{int} since it contains linear combinations of arbitrary real shifts of the transmitted signal. We address this with the second step.

2) *Reshaping and Shiftability*: In the second step of our approach, we deal with arbitrary real shifts of the transmitted signal. We reshape the optimal scaling function $g_{i^*}^*(t)$ obtained from the first step to make it nearly shiftable [12]–[14]. A function $g(t)$ is shiftable if there exists a set of real coefficients $\{\beta_n(\tau), n \in \mathbb{Z}\}$ such that [15]

$$g(t - \tau) = \sum_n \beta_n(\tau) g(t - n) \quad \forall \tau \in [0, 1). \quad (26)$$

Equation (26) still holds when the function $g(t)$ is scaled. Specifically

$$g(2^i t - \tau) = \sum_n \beta_n(\tau) g(2^i t - n) \quad \forall \tau \in [0, 2^{-i}). \quad (27)$$

In practice, shiftability is too strong a requirement. Instead, we work with nearly shiftable functions. A function $g(t)$ is nearly shiftable if the mean square error

$$E(\tau) = \inf_{\{\beta_n(\tau)\}} \left\| g(t - \tau) - \sum_n \beta_n(\tau) g(t - n) \right\|_2^2 \quad (28)$$

TABLE I
SUBSPACE DESIGN ALGORITHM

1. Parameterize the scaling function $g(t)$ using the parameterization given in [9].
2. For a fixed scale index i , find the scaling function $g_i^*(t)$ such that $\hat{\delta}_{\text{mod}}^2(\mathcal{S}_{int}, \mathcal{G}_i)$ is minimized. The modified deflection $\hat{\delta}_{\text{mod}}^2(\mathcal{S}_{int}, \mathcal{G}_i)$ is computed by Theorem III.1 in subsection III-C.
3. Find the smallest scale index i^* such that, for a given error threshold $\epsilon > 0$, $\hat{\delta}_{\text{mod}}^2(\mathcal{S}_{int}, \mathcal{G}_{i^*}) < \epsilon$ and $\hat{\delta}_{\text{mod}}^2(\mathcal{S}_{int}, \mathcal{G}_{i^*-1}) > \epsilon$. The subspace $\mathcal{G}_{i^*}^*$ is given by (25).
4. Reshape $g_{i^*}^*(t)$ to make it nearly shiftable using Benno and Moura's algorithm given in [12].

is small, say, $\forall \tau \in [0, 1)$, $E(\tau)$ is below a threshold η [12], [13]. The authors of [12] and [13] developed an iterative algorithm to reshape a nonshiftable function to make it nearly shiftable.

Hopefully, by making the optimal scaling function $g_{i^*}^*(t)$ nearly shiftable, the subspace generated by the reshaped $g_{i^*}^*(t)$, i.e.,

$$\mathcal{G}_{i^*,r}^* = \left\{ \sum_{n=-\infty}^{+\infty} 2^{i^*/2} \beta_n g_{i^*,r}^*(2^{i^*} t - n), \beta_n \in \mathbb{R} \right\} \quad (29)$$

represents well linear combinations of arbitrary real shifts of the transmitted signal, where $g_{i^*,r}^*(t)$ is the reshaped optimal scaling function.

The subspace design algorithm is summarized in Table I. Items 1–3 together form the first step. Item 4 is the second step. We do not have analytical proof that the modified deflection is necessarily reduced after the reshaping. Our simulation results presented in the next section show that the performance is improved by reshaping.

3) *Convergence on i^* and Choice of ϵ* : In the subspace design algorithm, one may ask if, for a given function $g(t)$ and a given error threshold $\epsilon > 0$, there in fact exists a smallest scale i^* . In other words, is item 3 in the algorithm of Table I well defined?

The modified deflection of any set with finite energy random elements \mathcal{S} from $L_2(\mathbb{R})$ is zero. Therefore, from the principles of mathematical analysis, by letting the scale index i go to infinity, the modified deflection of \mathcal{S} from \mathcal{G}_i goes to zero.

Since the multiresolution subspace \mathcal{G}_i is strictly included in the subspace \mathcal{G}_{i+1} , it is not difficult to show that $\hat{\delta}_{\text{mod}}(\mathcal{S}, \mathcal{G}_i) \leq \hat{\delta}_{\text{mod}}(\mathcal{S}, \mathcal{G}_{i+1})$. Thus, the modified deflection of \mathcal{S} from \mathcal{G}_i is nondecreasing when i is increased. Then, given $\epsilon > 0$, there is a minimum i^* such that $\hat{\delta}_{\text{mod}}^2(\mathcal{S}, \mathcal{G}_i) < \epsilon$ so that indeed, item 3 is well defined.

Although on the surface this may seem counterintuitive, choosing smaller ϵ 's may not necessarily lead to better detection performance. A smaller ϵ may increase the optimal i^* . This increases the number of degrees of freedom of the test statistic. Increasing the number of degrees of freedom may deteriorate the detection performance. A tradeoff needs to be achieved between the representation error and the complexity of the representation subspace. This has to be done on a signal-by-signal basis.

C. Modified Deflection Computation

Basic to the design algorithm of Table I is the computation of the modified deflection $\hat{\delta}_{\text{mod}}(\mathcal{S}_{\text{int}}, \mathcal{G}_i)$. The following theorem computes it analytically.

Theorem III.1: Let \mathcal{S}_{int} be the set given by (24) and

$$\mathcal{G}_i = \left\{ \sum_{n=-\infty}^{+\infty} 2^{i/2} \beta_n g(2^i t - n), \beta_n \in \mathbb{R} \right\} \quad i \geq 0 \quad (30)$$

where $g(t)$ is an orthonormal scaling function. Then, the modified deflection of \mathcal{S}_{int} from \mathcal{G}_i is

$$\hat{\delta}_{\text{mod}}(\mathcal{S}_{\text{int}}, \mathcal{G}_i) = \sqrt{1 - \inf_{f \in [0,1]} C(f)} \quad i \geq 0 \quad (31)$$

where $C(f)$ is

$$C(f) = \frac{\sum_{m=1}^M \sigma_m^2 C_{hg}^m(f)}{\sum_{m=1}^M \sigma_m^2 \sum_l |\mathcal{F}_{h_m}(f+l)|^2} \quad (32)$$

The functions $\{C_{hg}^m(f), m = 1, \dots, M\}$ are given by

$$\sum_{l=0}^{2^i-1} \left| \sum_k \mathcal{F}_{h_m}(f+2^i k+l) \overline{\mathcal{F}_g(2^i f+k+2^i l)} \right|^2 \quad (33)$$

where $\{h_m(t), m = 1, \dots, M\}$, and $\{\sigma_m^2, m = 1, \dots, M\}$ are the eigenfunctions and eigenvalues of the autocorrelation function $K(t, u)$. $\mathcal{F}_{h_m}(f)$ and $\mathcal{F}_g(f)$ are the Fourier transform of $h_m(t)$ and $g(t)$, respectively. $\overline{\mathcal{F}_g(f)}$ is the complex conjugate of $\mathcal{F}_g(f)$. The infimum is taken over the regions where the function $C(f)$ is continuous. We assume that

$$\sum_{m=1}^M \sigma_m^2 \sum_l |\mathcal{F}_{h_m}(f+l)|^2 > 0 \quad \text{a.e.} \quad (34)$$

The proof of the theorem is shown in the Appendix.

Despite its formidable appearance, (32) is very easy to compute. The terms on the numerator $\{C_{hg}^m(f), m = 1, \dots, M\}$ are essentially the discrete time Fourier transform (DTFT) of the downsampled autocorrelation sequences

$$R_m^i[l] = \sum_{k=-\infty}^{+\infty} c_m^i[k] c_m^i[2^i l + k] \quad m = 1, \dots, M. \quad (35)$$

The sequences $\{c_m^i[k], m = 1, \dots, M\}$ are given by

$$c_m^i[k] = \langle h_m(\cdot), 2^{i/2} g(2^i \cdot -k) \rangle \quad m = 1, \dots, M. \quad (36)$$

These are the orthogonal projection coefficients of $\{h_m(t), m = 1, \dots, M\}$ on the subspace \mathcal{G}_i .

In our simulations, we deal with discrete time signals. The inner product in (36) is implemented as a vector inner product. The DTFT is implemented by the fast Fourier transform (FFT).

Note that Theorem III.1 provides only the modified deflection of \mathcal{S} from \mathcal{G}_i for $i \geq 0$. It is very difficult, if not impossible, to come up with an explicit formula for the modified deflection if $i \leq 0$. Theoretically, this may limit the performance of the algorithm. However, in general, we can

intuitively argue that this is not the case. The signal $s(t)$ is a linear combination of its eigenfunctions, whereas $g(t)$ is only a single function. To be able to approximate multiple integer shifted replicas of $s(t)$, which involve multiple integer shifted replicas of several eigenfunctions of $s(t)$ by multiple integer shifted replicas of a single function, we should expect that we need to work with a higher scaled version of $g(t)$.

D. Minimization of the Modified Deflection

To minimize the modified deflection given in (31), we restrict the scaling function $g(t)$ to be a compactly supported orthonormal scaling function so that we may use the parameterization proposed in [9]. With this constraint, the modified deflection of \mathcal{S}_{int} from \mathcal{G}_i given in (31) is a function of the parameter vector $\zeta = [\zeta_1, \dots, \zeta_{P-1}]$ that parameterizes the scaling function $g(t)$ (see Section II-C), i.e.,

$$\hat{\delta}_{\text{mod}}(\mathcal{S}_{\text{int}}, \mathcal{G}_i) = \sqrt{1 - \inf_{f \in [0,1]} C(f, \zeta)}. \quad (37)$$

Minimizing the modified deflection is equivalent to maximizing the quantity

$$\inf_{f \in [0,1]} C(f, \zeta). \quad (38)$$

The actual maximization is done by an exhaustive search of the parameter space of ζ , i.e.,

$$\zeta^* = \arg \max_{\zeta} \left\{ \inf_{f \in [0,1]} C(f, \zeta) \right\}. \quad (39)$$

The optimal scaling function $g_i^*(t)$ at scale i is reconstructed from ζ^* using an algorithm described in [8].

Even though the modified deflection minimization is a multidimensional nonlinear optimization problem, we stress that it is done off line. Once we have the representation subspace, the focused detector is easily implemented by a bank of correlators followed by an energy detector. There is no multidimensional nonlinear optimization involved in the focused detector implementation.

IV. SIMULATION RESULTS

A. Whale Sounds Database

In this section, we test the performance of our new receiver using a small database of whale sounds and compare it with other alternative receivers. The database consists of 50 realizations of whale sounds sampled at 32 kHz/s. We partition the database into two sets: a training set (25 realizations) and a testing set (25 realizations). The training set is used to estimate the autocorrelation function of the signal, and the testing set is used to test the performance. The whale sounds in the database are not actual multipath signals. In our study, we simulate the multipath channel. The multipath channel parameters are generated in the following way. The number of paths K is set to 8, and in our simulation, we set all attenuation factors $\{\alpha_k\}$ to be equal to 1. The set of delays $\{\tau_k, k = 1, \dots, 8\}$ are generated by a random number generator.

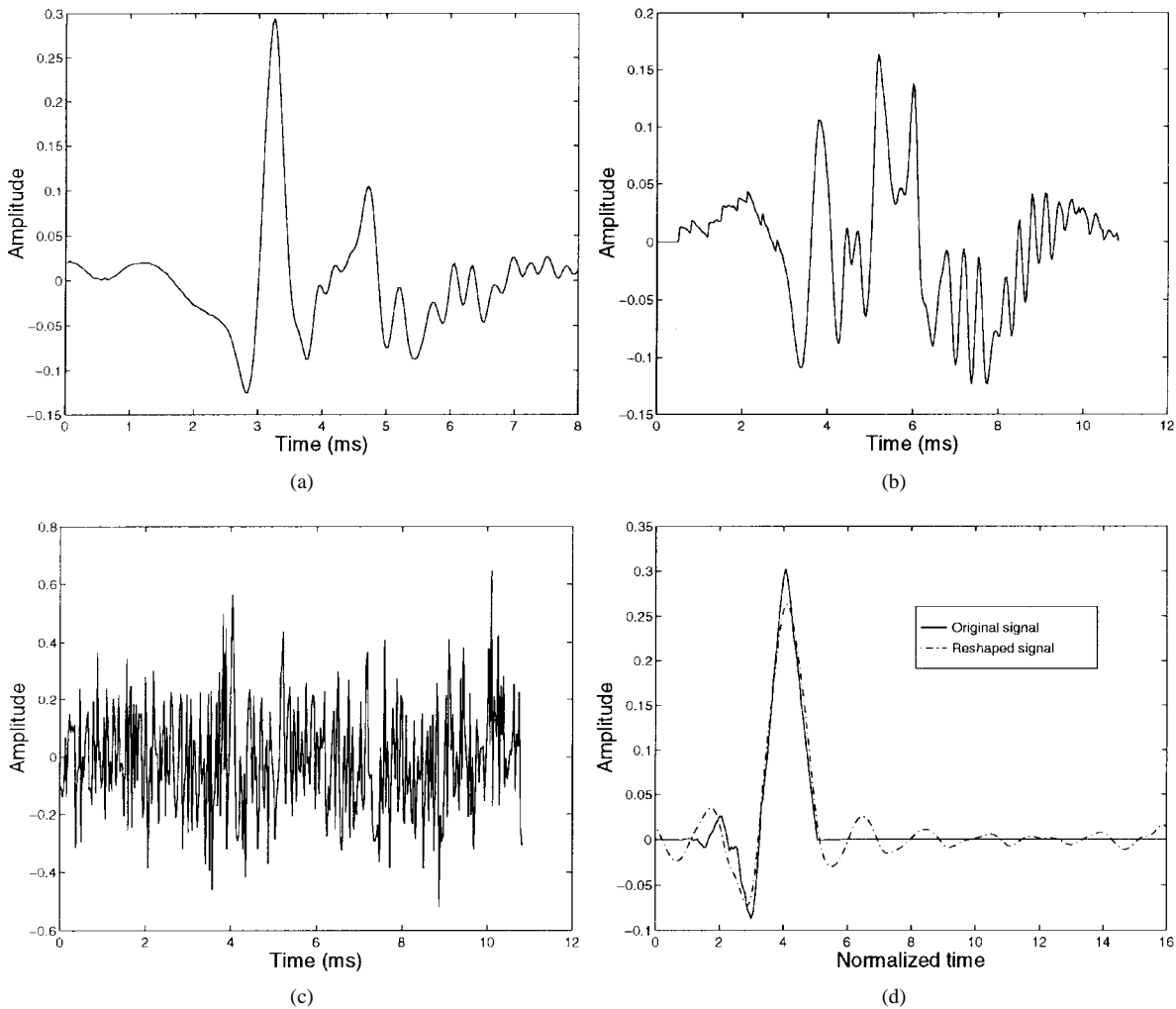


Fig. 1. (a) One realization of the whale sounds in our testing data set. (b) Multipath noise free whale sound (eight paths). (c) Multipath noisy whale sound (eight paths). (d) Optimal scaling function and reshaped optimal scaling function ($P - 1 = 2$).

B. Sample Autocorrelation

Since we do not know the autocorrelation function $K(t, u)$, we first estimate it from the training data set. Because the whale sounds are sampled data, we deal with discrete sequences instead of continuous functions. We stack these sequences in vectors. Denote by $\{s_i, i = 1, \dots, L\}$ the set of training vectors (column vectors). The estimated autocorrelation matrix is

$$R = \frac{1}{L-1} \sum_{i=1}^L s_i s_i^T. \quad (40)$$

We perform the eigenvalue decomposition of R and pick the most significant eigenvectors. For our training data set, 10 significant eigenvectors are kept to represent the transmitted signal. The time interval T' [see (17)] within which the decomposition is performed is 8 ms.

C. Generating Function Design

With these eigenvectors and the corresponding eigenvalues, we use the algorithm described in Section III-B to design the generating function of the representation subspace. We design the generating function for both the two-parameter ($P-1=2$)

case and the three-parameter ($P-1=3$) case. In the two-parameter case, the optimization in item 2 of the algorithm of Table I is done by computing the modified deflection for the subspaces $\{\mathcal{G}_i\}$ generated by each of the scaling functions $g(t)$ parameterized by the values on the uniform two-dimensional grid

$$\zeta_1(l) = \zeta_2(l) = \frac{2\pi l}{50} \quad l = 0, \dots, 49 \quad (41)$$

and finding the parameter vector $\zeta^* = [\zeta_1^*, \zeta_2^*]$ that leads to the minimum value of the modified deflection of \mathcal{S}_{int} from \mathcal{G}_i with the smallest scale index i^* . For the three-parameter case, the optimization process is similar but with an extra dimension to minimize over. We choose ϵ to be equal to 0.40. A smaller ϵ will lead to larger i^* , which will increase the number of degrees of freedom for the detection test statistic. As mentioned before, increasing the number of degrees of freedom has a negative effect on the detection performance. The optimal scaling function $g_{i^*}^*(t)$ is reconstructed from ζ^* using an algorithm described in [8]. Then, we reshape $g_{i^*}^*(t)$ to make it nearly shiftable. The reshaped scaling function is $g_{i^*,r}^*(t)$.

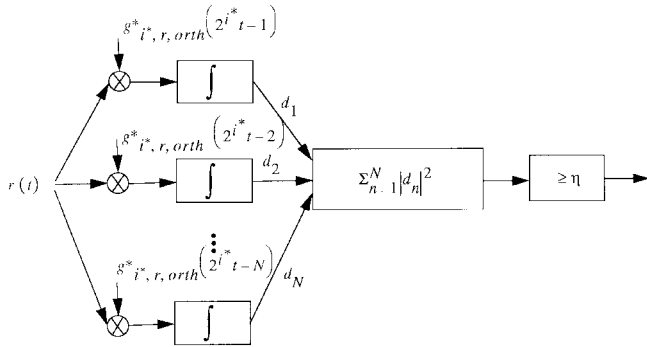


Fig. 2. Structure of the focused receiver.

Fig. 1(a) shows one realization of the whale sounds in our testing data set. Fig. 1(b) shows a multipath noise-free whale sound with eight paths. Fig. 1(c) is a plot of the same multipath signal with additive white and Gaussian noise. Fig. 1(d) shows the optimal scaling function and the reshaped optimal scaling function.

D. Focused Detector

The focused receiver uses

$$\tilde{L} = \|P_{G_{i^*, r}^*}[r]\|_2^2 \quad (42)$$

as the test statistic. Since $g_{i^*, r}^*(t)$ is the reshaped scaling function, $\{2^{i^*}/2 g_{i^*, r, orth}^*(2^{i^*} t - n), n \in \mathbb{Z}\}$ are not necessarily orthogonal to each other. We orthonormalize it without changing its span by

$$\mathcal{F}_{g_{i^*, r, orth}^*}(f) = \frac{\mathcal{F}_{g_{i^*, r}^*}(f)}{\sqrt{\sum_l |\mathcal{F}_{g_{i^*, r}^*}(f+l)|^2}} \quad (43)$$

where $\mathcal{F}_{g_{i^*, r, orth}^*}(\cdot)$ and $\mathcal{F}_{g_{i^*, r}^*}(\cdot)$ are the Fourier transforms of $g_{i^*, r, orth}^*(\cdot)$ and $g_{i^*, r}^*(\cdot)$, respectively [8]. With the orthonormality of $\{2^{i^*}/2 g_{i^*, r, orth}^*(2^{i^*} t - n), n \in \mathbb{Z}\}$, $P_{G_{i^*, r}^*}[r]$ is easily computed by taking the inner product of the received signal with $\{2^{i^*}/2 g_{i^*, r, orth}^*(2^{i^*} t - n), n \in \mathbb{Z}\}$. The structure of the focused receiver is shown in Fig. 2. It is essentially a bank of correlators followed by an energy detector (ED).

Since the signals we deal with are of finite duration, we treat the boundary problem by making the signals periodic. In addition, we need only a finite number of correlators.

In general, since we do not know the higher order statistics of $s(t)$, the distribution of the test statistic given in (42) under H_1 is unknown, which precludes a complete analytical performance analysis. However, we realize that for a fixed realization of $s(t)$ and a fixed set of channel parameters, the test statistic in (42) is chi-square distributed under H_0 and noncentral chi-square distributed under H_1 . Thus, for a fixed $s_m(t)$, we can compute the probability of detection (P_D) as a function of the signal-to-noise ratio (SNR) for a fixed false alarm probability (P_F). In our simulation results shown below, we take the average of P_D over all the realizations of the transmitted signal $s(t)$ in the testing set and over all the different channel delay patterns.

We compare the performance of our focused receiver with three other receivers. The first receiver with which we compare

is the correlator receiver (CR); see [2] and [16] for details. With the correlator receiver, since the transmitted signal is random, we correlate the received signal with the most significant eigenfunction of $K(t, u)$ and use the peaks in the correlator output to detect the signal. Second, we compare our receiver with a ‘‘matched filter with integer shifts’’ (MFIS) receiver; see [2] for details. The MFIS receiver has the same structure as our focused receiver. The difference is that it uses the most significant eigenfunction of $K(t, u)$ as the generating function rather than the one we have designed. Finally, we compare our receiver with the ED [17].

Fig. 3 shows the average detection probabilities P_D as a function of the SNR. The average is taken over all the testing data and over all the delay patterns. The false alarm probability P_F is fixed at 0.01. There are seven curves in the figure. The solid line is obtained by assuming perfect knowledge about the transmitted signal $s(t)$, the number of paths K , and the delays $\{\tau_k\}$. Since, in practice, these parameters are in general unknown, it is an overly optimistic performance bound. The dashed line is our focused receiver using the reshaped optimal scaling function for $P - 1 = 2$. The ‘‘*’’ curve represents our focused receiver using the reshaped optimal scaling function for $P - 1 = 3$. The ‘‘o’’ curve is the receiver using the optimal scaling function without reshaping for $P - 1 = 2$. The dash-dotted line and the dotted line represent the performance of the correlator receiver and of the MFIS receiver, respectively. The ‘‘+’’ curve represents the performance of the energy detector.

Fig. 3 shows that, in this study, the correlator receiver (the right-most curve) is outperformed by all others. This is not surprising since we chose the delay patterns so that the delayed replicas of the transmitted signal have considerable overlap. Under these conditions, we expect the correlator receiver to perform poorly. The potential maximum gain over the correlator receiver is given by the distance between the two extreme curves in Fig. 3, namely, the left most curve corresponding to the optimistic upper bound and the right-most curve corresponding to the correlator receiver. This potential gain is about 4.7 dB. We see from Fig. 3 that the focused receiver using the reshaped optimal scaling function ($P - 1 = 3$) recovers about 2.7 dB of these potential 4.7 dB. It also outperforms the MFIS receiver by about 1.3 dB and outperforms the energy detector by about 2.3 dB.

The performance of the receiver using the optimal scaling function without reshaping ($P - 1 = 2$) (‘‘o’’ curve in Fig. 3) is worse than that of the one using the reshaped optimal scaling function ($P - 1 = 2$) (dashed line in Fig. 3). Note that the performance of our focused receiver for $P - 1 = 3$ (‘‘*’’ curve in Fig. 3) is just slightly better than that of the receiver for $P - 1 = 2$ (dashed line in Fig. 3). Increasing the number of parameters used in the parameterization does not significantly improve the detection performance.

Fig. 3(b) is a plot of the detection probabilities at SNR=16.5 dB for all 100 trials. In this case, the robustness of our new detector (for both $P - 1 = 2$ and $P - 1 = 3$) is better than that of the MFIS receiver and much better than that of the correlator receiver. The energy detector is the most robust receiver. However, its average performance is poor.

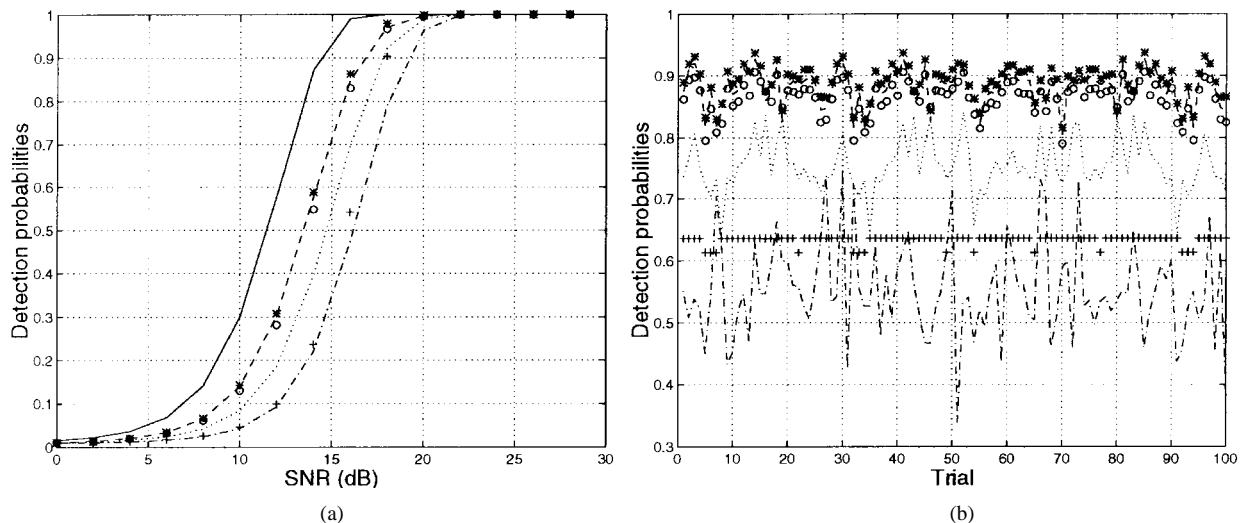


Fig. 3. (a) Average detection probabilities $K = 8$, $P_F = 0.01$. (b) Detection probabilities at SNR = 16.5 dB of the 100 trials. The solid line is a performance bound. “—” is our focussed receiver with $P - 1 = 2$. “o” is the receiver using the optimal scaling function without reshaping for $P - 1 = 2$. “*” is our focussed receiver with $P - 1 = 3$. “-” is the correlator receiver. “...” is the MFIS receiver. “+” is the energy detector.

V. SUMMARY

We presented a signal analysis design technique using scaling functions that are robust to translations. We are motivated by detection in multipath environments. Translation-invariant scaling functions are useful in a number of other applications and have been the subject of extensive literature. Our approach should be useful in many of these other areas as well.

Multipath is an example of distortion that degrades the detection performance but is difficult to compensate for. In the geometric setting we described, the optimal detector uses as test statistic the energy of the projection of the received signal on the multipath signal set \mathcal{S} . Computation of this orthogonal projection is, however, a multidimensional optimization that is equivalent to the maximum-likelihood estimation of the multipath itself. In the approach taken, we replace \mathcal{S} with a representation subspace \mathcal{G} that is designed such that the projection $\Pi_{\mathcal{G}}$ on \mathcal{G} is close to the projection on \mathcal{S} , whereas, at the same time, $\Pi_{\mathcal{G}}$ is simple to compute.

In general, the design of \mathcal{G} presents major difficulties. We highlight two of them. First, because \mathcal{S} is an ensemble of subspaces, the similarity between \mathcal{S} and \mathcal{G} is measured by the modified deflection. The modified deflection is a one-sided measure. Therefore, while designing \mathcal{G} to be close to \mathcal{S} , we need \mathcal{G} to be as small as possible; see Section III-B. A second difficulty is the subspace design itself. Subspaces are designed by specifying a basis for the subspace. For separable subspaces, bases are countable. While it is possible in principle to use an arbitrary set of linearly independent functions (e.g., orthogonal functions) as basis functions, this is not a viable strategy. For one thing, the design would now mean designing multiple functions (the subspace basis). On the other hand, the “size” of the subspace becomes coupled with the design of the basis functions, and in general, choosing the “smallest” such \mathcal{G} that is closest to \mathcal{S} in the modified-deflection sense leads possibly to a coupled iterative scheme, where each iteration requires designing all the corresponding basis functions.

We showed how a multiresolution analysis (MRA) avoids all these difficulties and provides a successful approach, as measured by the good detection results in Section IV. The MRA designs \mathcal{G} by the design of a *single* function, namely, the associated scaling function $g(t)$. Second, because the MRA is a nested sequence of subspaces, choosing the “smallest” \mathcal{G} is simply a question of choosing the lowest scale i such that the modified-deflection to the signal set is below a certain threshold. The MRA approach also has practical implications, namely, the simplicity of the receiver (a bank of filters where each filter is matched to an integer translate of the *same* signal) makes its hardware implementation simple and economical (essentially design a *single* filter).

We illustrated the design of the scaling function $g(t)$ by restricting $g(t)$ to be compact. This led to additional advantages that were not readily available with arbitrary sets of functions.

- 1) Using parametrizations for compactly supported scaling functions, the (off-line) subspace design is further simplified from the design of a *single* function to an (off-line) optimization of a finite number of parameters.¹
- 2) Increased detection performance requires a balance between the signal energy and the noise energy in the test statistic. This balance relates to the number of degrees of freedom retained in the detector, i.e., the number of filters in the filter bank. The MRA provides a simple solution. Because \mathcal{G} replicates the structure of the multipath signal set \mathcal{S} and the scaling function has compact support, the number of filters follows from the maximum expected interarrival delay between paths; see the discussions in Sections III-B and IV.

The experimental results in Section IV show that the MRA detector provides significant performance gains over alternative receivers and is robust to the multipath delay pattern.

¹Although we worked with the parametrizations for compact supported scaling functions [9], other finite parametrizations that provide other tradeoffs (e.g., prespecified degree of smoothness; see [3]) can also be used.

APPENDIX

In this Appendix, we prove Theorem III.1.

From the definition of the modified deflection, we need first to compute

$$\text{dist}^2[s_{int}(\cdot), \mathcal{G}_i] = \|s_{int}(\cdot)\|_2^2 - \|P_{\mathcal{G}_i}[s_{int}]\|_2^2 \quad i \geq 0 \quad (44)$$

where

$$s_{int}(\cdot) = \sum_{n=-\infty}^{+\infty} \alpha_n s(\cdot - n) \quad (45)$$

$$= \sum_{n=-\infty}^{+\infty} \alpha_n \sum_{m=1}^M \gamma_m h_m(\cdot - n) \in \mathcal{S}_{int} \quad (46)$$

and $P_{\mathcal{G}_i}[s_{int}]$ is the orthogonal projection of $s_{int}(\cdot)$ on \mathcal{G}_i .

Since $\{2^{i/2}g(2^i \cdot -k), k \in \mathbb{Z}\}$ is an orthonormal basis of \mathcal{G}_i , the orthogonal projection $P_{\mathcal{G}_i}[s_{int}]$ is obtained by taking the inner product of $s_{int}(\cdot)$ with $\{2^{i/2}g(2^i \cdot -k), k \in \mathbb{Z}\}$, i.e.,

$$(P_{\mathcal{G}_i}[s_{int}])(t) = \sum_k \langle s_{int}(\cdot), 2^{i/2}g(2^i \cdot -k) \rangle \cdot 2^{i/2}g(2^i t - k). \quad (47)$$

Taking the Fourier transform of (47) and using Poisson's summation formula, we have

$$\begin{aligned} \mathcal{F}[P_{\mathcal{G}_i}[s_{int}]](f) &= 2^{-i} \mathcal{F}_g(2^{-i} f) \cdot \sum_k \int_{-\infty}^{+\infty} \\ &\quad \cdot \mathcal{F}_{s_{int}}(f_1) \overline{\mathcal{F}_g}(2^{-i} f_1) \\ &\quad \cdot \exp\{-j2\pi k 2^{-i}(f - f_1)\} df_1 \quad (48) \end{aligned}$$

$$\begin{aligned} &= 2^{-i} \mathcal{F}_g(2^{-i} f) \cdot \int_{-\infty}^{+\infty} \mathcal{F}_{s_{int}}(f_1) \overline{\mathcal{F}_g}(2^{-i} f_1) \\ &\quad \cdot \sum_k \exp\{-j2\pi k 2^{-i}(f - f_1)\} df_1 \quad (49) \end{aligned}$$

$$\begin{aligned} &= \mathcal{F}_g(2^{-i} f) \cdot \int_{-\infty}^{+\infty} \mathcal{F}_{s_{int}}(f_1) \overline{\mathcal{F}_g}(2^{-i} f_1) \\ &\quad \cdot \sum_k \delta(f_1 - f - 2^i k) df_1 \quad (50) \end{aligned}$$

$$\begin{aligned} &= \mathcal{F}_g(2^{-i} f) \cdot \left\{ \sum_n \alpha_n \exp\{-j2\pi n f\} \right\} \\ &\quad \cdot \sum_k \mathcal{F}_s(f + 2^i k) \overline{\mathcal{F}_g}(2^{-i} f + k). \quad (51) \end{aligned}$$

Taking the norm square of (51) and using the fact that for orthonormal scaling functions, the following is true [8]

$$\sum_l |\mathcal{F}_g(2^{-i} f + l)|^2 = 1 \quad \text{a.e.} \quad (52)$$

we have

$$\|P_{\mathcal{G}_i}[s_{int}]\|_2^2 = \|\mathcal{F}[P_{\mathcal{G}_i}[s_{int}]]\|_2^2 \quad (53)$$

$$\begin{aligned} &= \int_{-\infty}^{+\infty} \left| \left\{ \sum_n \alpha_n \exp\{-j2\pi n f\} \right\} \cdot \mathcal{F}_g(2^{-i} f) \right. \\ &\quad \left. \cdot \sum_k \mathcal{F}_s(f + 2^i k) \overline{\mathcal{F}_g}(2^{-i} f + k) \right|^2 df \quad (54) \end{aligned}$$

$$\begin{aligned} &= \int_0^{2^i} \left| \sum_n \alpha_n \exp\{-j2\pi n f\} \right|^2 \\ &\quad \cdot \sum_l |\mathcal{F}_g(2^{-i} f + l)|^2 \\ &\quad \cdot \left| \sum_k \mathcal{F}_s(f + 2^i k) \overline{\mathcal{F}_g}(2^{-i} f + k) \right|^2 df \quad (55) \end{aligned}$$

$$\begin{aligned} &= \int_0^{2^i} \left| \sum_n \alpha_n \exp\{-j2\pi n f\} \right|^2 \\ &\quad \cdot \left| \sum_k \mathcal{F}_s(f + 2^i k) \overline{\mathcal{F}_g}(2^{-i} f + k) \right|^2 df \quad (56) \end{aligned}$$

$$\begin{aligned} &= \int_0^1 \left| \sum_n \alpha_n \exp\{-j2\pi n f\} \right|^2 \\ &\quad \cdot \sum_{l=0}^{2^i-1} \left| \sum_k \mathcal{F}_s(f + 2^i k + l) \overline{\mathcal{F}_g}(2^{-i} f + k + 2^{-i} l) \right|^2 df. \quad (57) \end{aligned}$$

Now, we expand $s(t)$ using the truncated K-L expansion and take the expectation. Since the coefficients $\{\gamma_m, m = 1, \dots, M\}$ are statistically orthogonal random variables, we have

$$\begin{aligned} E\{\|P_{\mathcal{G}_i}[s_{int}]\|_2^2\} &= \int_0^1 \left| \sum_n \alpha_n \exp\{-j2\pi n f\} \right|^2 \\ &\quad \cdot \sum_{m=1}^M \sigma_m^2 C_{hg}^m(f) df \quad (58) \\ &= \int_0^1 \left\{ \left| \sum_n \alpha_n \exp\{-j2\pi n f\} \right|^2 \sum_{m=1}^M \right. \\ &\quad \left. \cdot \sigma_m^2 \sum_l |\mathcal{F}_{h_m}(f + l)|^2 \right\} \cdot C(f) df \quad (59) \end{aligned}$$

where

$$C_{hg}^m(f) = \sum_{l=0}^{2^i-1} \left| \sum_k \mathcal{F}_{h_m}(f + 2^i k + l) \overline{\mathcal{F}_g}(2^{-i} f + k + 2^{-i} l) \right|^2 \quad (60)$$

and

$$C(f) = \frac{\sum_{m=1}^M \sigma_m^2 C_{hg}^m(f)}{\sum_{m=1}^M \sigma_m^2 \sum_l |\mathcal{F}_{h_m}(f + l)|^2}. \quad (61)$$

Once we have $E\{\|P_{\mathcal{G}_i}[s_{int}]\|_2^2\}$, the square of the modified deflection $\hat{\delta}_{\text{mod}}(\mathcal{S}_{int}, \mathcal{G})$ is obtained by taking the supremum, i.e.,

$$\hat{\delta}_{\text{mod}}^2(\mathcal{S}_{int}, \mathcal{G}_i) = \sup_{s_{int}(\cdot) \in \mathcal{S}_{int}} E\{\text{dist}^2[s_{int}(\cdot), \mathcal{G}_i]\} \quad (62)$$

$$= 1 - \inf_{s_{int}(\cdot) \in \mathcal{S}_{int}} E\{\|P_{\mathcal{G}_i}[s_{int}]\|_2^2\} \quad (63)$$

subject to

$$E\{\|s_{int}(\cdot)\|_2^2\} = \int_{-\infty}^{+\infty} E\{|s_{int}(t)|^2\} dt \quad (64)$$

$$= \int_0^1 \left| \sum_n \alpha_n \exp\{-j2\pi n f\} \right|^2 \sum_{m=1}^M \sigma_m^2 \sum_l |\mathcal{F}h_m(f+l)|^2 df \quad (65)$$

$$= 1. \quad (66)$$

Since we assume that

$$\sum_{m=1}^M \sigma_m^2 \sum_l |\mathcal{F}h_m(f+l)|^2 > 0 \quad \text{a.e.} \quad (67)$$

using a limiting argument, we can choose

$$\left| \sum_n \alpha_n \exp\{-j2\pi n f\} \right|^2 \cdot \sum_{m=1}^M \sigma_m^2 \sum_l |\mathcal{F}h_m(f+l)|^2 \quad (68)$$

to be the Dirac delta function located at the frequency where $C(f)$ has its minimum and is continuous. Then, the quantity

$$E\{\|P_{G_i}[s_{int}]\|_2^2\} = \int_0^1 \left| \sum_n \alpha_n \exp\{-j2\pi n f\} \right|^2 \sum_{m=1}^M \sigma_m^2 \sum_l |\mathcal{F}h_m(f+l)|^2 \cdot C(f) df \quad (69)$$

is minimized. The minimum is

$$\inf_{f \in [0,1]} \frac{\sum_{m=1}^M \sigma_m^2 C_{hg}^m(f)}{\sum_{m=1}^M \sigma_m^2 \sum_l |\mathcal{F}h_m(f+l)|^2}. \quad (70)$$

Therefore, the modified deflection $\hat{\delta}_{\text{mod}}(\mathcal{S}_{int}, \mathcal{G}_i)$ is given by

$$\sqrt{1 - \inf_{f \in [0,1]} C(f)}. \quad (71)$$

This proves Theorem III.1.

REFERENCES

- [1] C. He, J. M. F. Moura, and S. A. Benno, "Gap detector for multipath," in *Proc. ICASSP*, May 1996, pp. V-2650-V-2653.
- [2] C. He and J. M. F. Moura, "Robust detection with the gap metric," *IEEE Trans. Signal Processing*, vol. 45, pp. 1591-1604, June 1997.
- [3] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 909-996, Nov. 1988.
- [4] S. Mallat, "A theory for multiresolution signal decomposition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674-693, July 1989.
- [5] T. Kato, *Perturbation Theory for Linear Operators*, 2nd ed. New York: Springer-Verlag, 1976.
- [6] A. Kolmogoroff, *Ann. Math.*, vol. 37, 1936.
- [7] J. Liang and T. W. Parks, "Kolmogorov n -widths and wavelet representations for signal classes," *IEEE Trans. Signal Processing*, vol. 44, pp. 1693-1703, July 1996.
- [8] C. K. Chui, *An Introduction to Wavelets*. New York: Academic, 1992.
- [9] H. Zou and A. H. Tewfik, "Parameterization of compactly supported orthonormal wavelets," *IEEE Trans. Signal Processing*, vol. 41, pp. 1428-1431, Mar. 1993.

- [10] H. L. V. Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968.
- [11] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1992.
- [12] S. A. Benno and J. M. F. Moura, "Nearly shiftable scaling functions," in *Proc. ICASSP*, May 1995, pp. II-1097-II-1100.
- [13] ———, "On translation invariant subspaces and critically sampled wavelet transforms," *Multidimensional Syst. Signal Process.*, vol. 8, pp. 89-110, Jan. 1997.
- [14] ———, "Scaling functions optimally robust to translations," submitted for publication.
- [15] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multiscale transforms," *IEEE Trans. Inform. Theory*, vol. 38, pp. 587-607, Mar. 1992.
- [16] G. L. Turin, "Introduction to spread-spectrum antimultipath techniques and their application to urban digital radio," *Proc. IEEE*, vol. 68, pp. 328-353, Mar. 1980.
- [17] A. D. Whalen, *Detection of Signals in Noise*. New York: Academic, 1971.



Chuang He (S'93-M'98) received the B.S. degree from Tsinghua University, Beijing, China, in 1991, the M.S. degree from Tulane University, New Orleans, LA, in 1993, and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, in 1997.

Since September 1997, he has been with Los Alamos National Laboratory, Albuquerque, NM, working on speech processing. His research interests are digital signal processing, speech processing, wavelet and time-frequency transforms, and communications.

Dr. He is a member of Eta Kappa Nu.



José M. F. Moura (S'71-M'75-SM'90-F'94) received the Engenheiro Electrotécnico degree in 1969 from Instituto Superior Técnico (IST), Lisbon, Portugal, and the M.Sc., E.E., and the D.Sc. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 1973 and 1975, respectively.

He is a Professor of Electrical and Computer Engineering at Carnegie Mellon University (CMU), Pittsburgh, PA, which he joined in 1986. Prior to this, he was on the faculty of IST, where he was an Assistant Professor (1975), Professor Agregado (1978), and Professor Catedrático (1979). He has had visiting appointments at several institutions, including MIT (Genrad Associate Professor of Electrical Engineering and Computer Science from 1984 to 1986) and the University of Southern California, Los Angeles (Research Scholar, Department of Aerospace Engineering, during the summers of 1978 to 1981). His research interests include statistical signal processing (one- and two-dimensional), digital communications, image and video processing, radar and sonar, and multiresolution techniques. He has organized and codirected two international scientific meetings on signal processing theory and applications. He has more than 190 published technical contributions, including invited ones, published in international journals and conference proceedings and is co-editor of two books: *Nonlinear Stochastic Problems* (New York: Reidal, 1983) and *Acoustic Signal Processing for Ocean Exploration* (Boston, MA: Kluwer, 1993).

Dr. Moura is currently the Editor-in-Chief for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, a member of the Underwater Acoustics and the Multimedia Signal Processing Technical Committees of the IEEE Signal Processing Society. He was a member of the IEEE Press Board from 1991 to 1995, a Technical Associate Editor for the IEEE SIGNAL PROCESSING LETTERS from 1993 to 1995, and an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1988 to 1992. He was a program committee member for the IEEE International Conference on Image Processing (ICIP'95) and for the IEEE International Symposium on Information Theory (ISIT'93). He is a corresponding member of the Academy of Sciences of Portugal (Section of Sciences). He is affiliated with several IEEE societies, Sigma Xi, AMS, IMS, and SIAM.