



**HAL**  
open science

## Time-Frequency Characterization and Receiver Waveform Design for Shallow Water Environments

Jun Zhang, Antonia Papandreou-Suppappola, Bertrand Gottin, Cornel Ioana

► **To cite this version:**

Jun Zhang, Antonia Papandreou-Suppappola, Bertrand Gottin, Cornel Ioana. Time-Frequency Characterization and Receiver Waveform Design for Shallow Water Environments. *IEEE Transactions on Signal Processing*, 2009, 57 (8), pp.2973-2985. 10.1109/TSP.2009.2020363 . hal-00349535

**HAL Id: hal-00349535**

**<https://hal.science/hal-00349535>**

Submitted on 31 Dec 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Time-Frequency Characterization and Receiver Waveform Design for Shallow Water Environments

Jun Zhang, *Student Member, IEEE*, Antonia Papandreou-Suppappola, *Senior Member, IEEE*,  
Bertrand Gottin, *Student Member, IEEE*, and Cornel Ioana, *Member, IEEE*

## Abstract

The shallow water environment can be considered as a time-dispersive system whose time-varying impulse response can be expressed as a superposition of time-frequency components with dispersive characteristics. In this paper, we propose a frequency-domain characterization of the shallow water system based on the normal-mode model that treats the system as an acoustic medium. After studying the dispersive characteristics of this system, a blind time-frequency processing technique is employed to separate the normal-mode components without knowledge of the relevant environment parameters. This technique is based on first approximating the time-frequency structure of the received signal and then designing time-frequency separation filters based on warping techniques. Following this method, two types of receivers are developed to exploit the diversity of the shallow water channel and to improve underwater communications performance. Simulation results demonstrate the dispersive characterization of this system and the improved processing performance of the receiver schemes.

## Index Terms

Time-frequency analysis, dispersive channels, time-frequency mode separation, underwater acoustic communications, time-dispersion diversity

# Time-Frequency Characterization and Receiver Waveform Design for Shallow Water Environments

## I. INTRODUCTION

Underwater acoustic signal processing is challenging as the transmitted signal interacts with the ocean bottom and surface, and the water medium causes a dense dispersion effect due to the time-varying (TV) changes of the ocean environment. Dispersive linear time-varying (LTV) systems can cause different frequencies to be shifted in time by different amounts [1], [2]. Such dispersive signal transformations are specific to the nature of the system or environment that the signal propagates through. Specifically, the dispersive effect introduced by shallow water environments can severely limit the performance of underwater acoustic applications such as sonar and communications.

In underwater communication applications, incoherent communication systems were initially employed, such systems lacked the ability to adjust waveform parameters in order to adapt to environment changes, and they were often highly inefficient in bandwidth and power requirements [3]. More recently, phase-coherent systems were considered as they can adaptively track the time and frequency spread of the environment changes and can correct inter-symbol interference (ISI) thus providing higher data rates [3]–[6]. Space-time techniques were also explored, as they can achieve spacial diversity [7]–[10]. Currently, time-reversal (or phase-conjugated) techniques use parameters that do not depend on the environment and, unlike equalization, do not require intensive computation [11]–[16]. The work in [15] improved the performance of the decision-feedback equalizer using passive-phase conjugation. In [16], time-reversal was applied to multiple-input multiple-output (MIMO) shallow water communications. On the other hand, in [1], a characterization of the shallow water system was considered using a time-frequency approach. In particular, this characterization matched the dispersive signal transformation caused on the transmitted waveforms. This characterization was successfully used for shallow water communications to obtain time-dispersion diversity. However, this characterization was only applicable to signals with very high bandwidth as it assumed that the transmitted waveform was an impulse.

It can be shown that the propagation characteristics of the shallow water environment can be determined by specific nonlinear functions that define the dispersion in this environment and provide a means of modeling the environment according to how it distorts the transmitted signal. Thus, signals can be used to exploit the potential diversity suggested by the model when the receiver is appropriately designed to match these nonlinear functions. In this paper, we propose a general signal characterization based on the normal-mode model discussed in [17] for shallow water environments that is applicable to a large class of signals. This characterization is based on a frequency

domain formulation and can be used with narrowband as well as wideband signals. The normal-mode characterization assumes perfect waveguide conditions, and as a result, it consists of a homogeneous fluid layer with a soft top and rigid bottom. It also follows the Pekeris model, which assumes a pressure-release surface and fluid boundaries [18]. Based on these assumptions, we propose a new time-frequency receiver design which operates on the individual modes of the multi-component received signal. This receiver requires accurate environment information (such as bathymetry, sound speed profile, attenuation and density) which is often unavailable or inaccurate, in order to obtain closed form expressions of the environment characterization. As a result, we also employ a blind method for separating the time-frequency (TF) components of the received signal that relate to the nonlinear functions causing dispersion in the system. The blind separation method first identifies the TF structures of the received signal [19], and then it separates them using a TF based nonunitary warping technique [20]. After the separation of each mode component, we use a pilot-aided communication scenario with an appropriately designed transmitted waveform and receiver structure to obtain time-dispersion diversity. Specifically, both the transmitter and receiver are designed to match the dispersive characteristics of the underwater medium based on the analysis of the time-frequency characteristics of the proposed model.

The paper is organized as follows. In Section II, we formulate the general waveguide model for shallow water environments, and we discuss two important simplified shallow water models: the isovelocity model and the Pekeris model. In Section III, we present the time-frequency characteristics of the Pekeris model, and we investigate the impact of the environment parameters on the transmitted signal dispersive transformations. In Section IV, we use a TF mode separation technique to separate each component of the received signal. In Section V, we design a new receiver structure with a corresponding optimal detector. Numerical results of bit-error-rate (BER) performance illustrate our improved performance in terms of diversity order.

## II. SHALLOW WATER ENVIRONMENT MODELING

### A. General Shallow Water Modeling

The problem of modeling shallow water environments is equivalent to calculating the response of an isotropic point source in a stratified acoustic medium. The scenario is shown in Fig. 1. In the water column layer, we can obtain the received signal by solving the *Helmholtz* equation in two dimensions with sound speed and density depending only on the depth  $z$  [17]. This equation is given by

$$\nabla \left( \frac{1}{\rho(z)} \nabla y(r, z, t) \right) - \frac{1}{\rho(z)c^2(z)} \frac{\partial^2 y(r, z, t)}{\partial^2 t} = -x(t) \frac{\delta(z - z_0)\delta(r)}{2\pi r} \quad (1)$$

where  $\nabla$  denotes the gradient operator,  $z_0$  is the depth of the source, and  $y(r, z, t)$  denotes the acoustic pressure as a function of the receiver depth  $z$ , range  $r$ , and time  $t$ . The isotropic point source is  $x(t)$ ,  $\rho(z)$  is the density,  $c(z)$  is the sound speed as a function of depth, and  $\delta(\cdot)$  is the Dirac delta function.

If we assume that  $z$ ,  $z_0$  and  $r$  are known, we can rewrite  $y(r, z, t)$  as  $y(t)$ . Denoting the Fourier transform (FT) of  $x(t)$  as  $X(f)$ , we can obtain the following solution for the frequency domain spectrum  $Y(f)$  of the received

signal  $y(t)$  in the far acoustic field:

$$Y(f) = X(f) \frac{C}{\rho(z_0)} \sum_{n=0}^{\infty} C_n(f) \frac{e^{jk_{r,n}(f)r}}{\sqrt{k_{r,n}(f)r}}. \quad (2)$$

Here,  $C$  is a constant,  $C_n(f)$  is the shape function of the  $n$ th mode, and  $k_{r,n}(f)$  is the horizontal propagation constant that depends on the frequency  $f$ , the range  $r$  and mode  $m$ . This model is called the *normal-mode model*. It treats the ocean as a stratified acoustic medium, representing the acoustic field in the ocean medium as a sum of normal modes. While normal modes are more commonly used as a computational tool, they are useful in the present context for their analytical properties and rich time-frequency (TF) structures. In the following, we discuss two important simplified shallow water environment models, the isovelocity model and the Pekeris model.

### B. Isovelocity Model

We first discuss the isovelocity model following [17]. The simplified waveguide model of the ocean is shown in Fig. 2 using the coordinate system  $(r, z)$ , where Medium I, II and III correspond to air, ocean water and ocean bottom, respectively. An omnidirectional point source  $x(t)$  with spectrum  $X(f)$  is located in Medium II at  $r = 0$  and  $z = z_0$ , and the ocean is  $D$  meters deep. We consider the sound speed in Medium II as a constant  $c$  m/s.

The ocean surface (at  $z = 0$ ) should be modeled as an ideal pressure release boundary and the ocean bottom at  $z = D$  as an ideal rigid boundary. However, the ocean bottom cannot be modeled as an ideal rigid boundary due to the roughness and scattering properties of the medium. The ocean waveguide problem involves the derivation of an expression for the velocity potential in Medium II (ocean water); this expression is the solution of the wave equation and satisfies all the boundary conditions, including the boundary condition at the source.

After a detailed derivation in [17], under the assumption of perfect waveguide, the received signal spectrum excited by  $X(f)$  at location  $(r, z)$  is given in terms of  $N_m$  modes as:

$$Y_{isovelocity}(f) = X(f) \sum_{n=0}^{N_m-1} C_n \Theta_n(f). \quad (3)$$

The  $n$ th mode is characterized<sup>1</sup> by the phase function  $\eta_n(f) = \sqrt{f^2 - f_n^2}$ ,  $f > f_n$ , since the group velocity function  $C_n \Theta_n(f)$  is given by:

$$C_n \Theta_n(f) = C_n \sqrt{\frac{1}{\frac{\sqrt{f^2 - f_n^2}}{c} r}} e^{-j2\pi \frac{z}{c} \sqrt{f^2 - f_n^2}}, \quad f > f_n. \quad (4)$$

Here,  $f_n = f_r \frac{(2n+1)c}{4D}$  is the cutoff frequency of the  $n$ th mode and  $N_m$  is the largest mode number. The parameter

$$C_n = e^{-j\frac{\pi}{4}} \frac{1}{2\pi d} \sin\left(\frac{(2n+1)\pi z_0}{2D}\right) \sin\left(\frac{(2n+1)\pi z}{2D}\right),$$

is constant for any given fixed coordinates  $(r, z)$ .

From (3) and (4), we can see that the transmitted signal experiences a group delay shift given by  $\zeta_n(f) = \frac{d}{df} \frac{r}{c} \sqrt{f^2 - f_n^2}$  at the  $n$ th mode, and the received signal is the summation of all the mode contributions.

<sup>1</sup>Note that here, and for the remainder of the manuscript, we assume that the frequencies  $f$  are normalized using a reference frequency  $f_r > 0$ .

### C. Pekeris Model

The Pekeris model treats the shallow water environments with pressure-release surface and fluid seabed following the work in [17], [18], [21]. The Pekeris waveguide model of the ocean is shown in Fig. 3 using the coordinate system  $(r, z)$ , where Medium I, II and III correspond to air, ocean water and seabed, respectively. An omnidirectional point source with signal spectrum  $X(f)$  is located in the ocean at  $r = 0$  and  $z = z_0$ . The ocean is  $D$  m deep. We consider the sound speed in the ocean as a constant  $c$  m/s and density  $\rho$  kg/m<sup>3</sup>; the corresponding parameters for the seabed are  $c_B$  m/s and  $\rho_B$  kg/m<sup>3</sup>, respectively.

The ocean surface (at  $z = 0$ ) is modeled realistically as an ideal pressure release boundary and the ocean bottom (at  $z = D$ ) is modeled as a boundary between two different fluid media. The normal mode model is given by the solution of this ocean waveguide problem, which is determined by the environment parameters and satisfies all boundary conditions, including the boundary condition at the source.

After a detailed derivation in [18] and [21], the received signal spectrum excited by  $X(f)$  at location  $(r, z)$ , is given by the Pekeris waveguide model:

$$Y_{Pekeris}(f) = X(f) \sum_{n=0}^{N_m-1} C_n(f) \Theta_n(f). \quad (5)$$

Although this appears similar to the model in (3), the dispersion characteristics differ. Specifically, without the assumption of the ideal waveguide condition as in [22], the  $n$ th mode is characterized by

$$\Theta_n(f) = \sqrt{\frac{1}{k_n(f) r}} e^{-jk_n(f) r},$$

where

$$k_n(f) = \frac{2\pi}{c} \left( f^2 - \frac{n^2 c^2}{4D^2} \left( 1 + \frac{\rho_B c}{\rho D} \frac{1}{2\pi f \sqrt{1 - \frac{c^2}{c_B^2}}} \right)^{-2} \right)^{\frac{1}{2}} \quad (6)$$

is the wave number of the  $n$ th mode and  $N_m$  is the largest mode number. The parameter

$$C_n(f) = A_n^2(f) \sin(k_{zn}(f) z_0) \sin(k_{zn}(f) z)$$

is a function of frequency where

$$A_n(f) = \sqrt{2} \left[ \frac{1}{\rho} \left( D - \frac{\sin(2k_{zn}(f)D)}{2k_{zn}(f)} \right) - \frac{\rho}{\rho_B^2} \frac{\tan(k_{zn}(f)D) \sin^2(k_{zn}(f)D)}{k_{zn}(f)} \right]^{-\frac{1}{2}} \quad (7)$$

and  $k_{zn}(f) \approx \frac{n\pi}{D} \left( 1 + \frac{\rho_B c}{\rho D} \frac{1}{2\pi f \sqrt{1 - \frac{c^2}{c_B^2}}} \right)^{-1}$ . Thus, the dispersive characteristic in  $X(f)$  follows from the group delay shift by  $\zeta_n(f) = \frac{d}{df} r k_n(f)$  at the  $n$ th mode.

## III. TIME-FREQUENCY CHARACTERISTICS OF SHALLOW WATER ENVIRONMENTS

### A. Isovelocity Model

From (3) and (4), the acoustic shallow water propagation can be represented as the summation of all the mode contributions. Specifically, each mode introduces a different dispersive transformation to the transmitted signal by

causing a shift in its group delay function. The group delay shift of the  $n$ th mode is given by

$$\xi_n(f) = \frac{d}{df} \frac{r}{c} \sqrt{f^2 - f_n^2} = \frac{r}{c} \frac{f}{\sqrt{f^2 - f_n^2}}, \quad n = 0, \dots, N_m - 1. \quad (8)$$

In Fig. 4, we plot the group delay shifts for  $N_m = 6$  modes, when the transmitted signal  $X(f) = 1$  for  $f \in (f_l, f_h)$  is an bandlimited impulse in time, where  $f_l$  and  $f_h$  are the lowest frequency and highest frequency of signal spectrum, respectively. As we can see in Fig. 4, the group delay shift of each mode is dispersive as the different frequency components of a mode travel at different speeds; a lower frequency experiences a larger shift in time. As a result, for isovelocity acoustic waveguides, the transmitted signal scatters in such a way that it experiences group delay shifts within each mode, and the received signal is the summation of multiple group delay shifted versions of the transmitted signal.

As our objective is to separate the different mode contributions, we want to transmit a signal  $X(f)$  that will enable such a process. In particular, we want a signal that, when dispersively group delay shifted by the  $n$ th mode, will result in a frequency domain sinusoid. As a result, we design the transmitted signal spectrum as  $X(f) = \sqrt{f}$  for  $f > f_n$ ; then we can apply a unitary warping operation to the received signal that is specific to each mode. The unitary warping operator  $\mathcal{U}_{\zeta_n}$  is defined as

$$(\mathcal{U}_{\zeta_n} Y)(f) = |\zeta'_n(\zeta_n^{-1}(f))|^{-1/2} Y(\zeta_n^{-1}(f)) \quad (9)$$

where  $(\mathcal{U}_{\zeta_n}^{-1} \mathcal{U}_{\zeta_n} Y)(f) = Y(f)$ ,  $\zeta_n^{-1}(\zeta_n(f)) = f$ ,  $\zeta'_n(f) = \frac{d}{df} \zeta_n(f)$ ,  $\zeta_n(f) = \sqrt{f^2 - f_n^2}$ ,  $f > f_n$ , and  $\zeta_n^{-1}(f) = \sqrt{f^2 + f_n^2}$ ,  $f > 0$ . In particular, for the  $n$ th mode, we define  $Y_n(f) = X(f)\Theta_n(f) = \sqrt{f}\Theta_n(f)$ ,  $f > f_n$  to which we apply the warping operator  $\mathcal{U}_{\zeta_n}$

$$(\mathcal{U}_{\zeta_n} Y_n)(f) = \frac{f^{\frac{1}{2}}}{(f^2 + f_n^2)^{\frac{1}{4}}} Y_n((f^2 + f_n^2)^{\frac{1}{2}}) = \frac{f^{\frac{1}{2}}}{(f^2 + f_n^2)^{\frac{1}{4}}} X((f^2 + f_n^2)^{\frac{1}{2}}) \Theta_n((f^2 + f_n^2)^{\frac{1}{2}}).$$

Note that, using (3),  $\Theta_n(\zeta_n^{-1}(f)) = \frac{1}{\sqrt{f}} \sqrt{\frac{c}{r}} e^{-j2\pi \frac{r}{c} f}$ . Also,  $X(\zeta_n^{-1}(f)) = (f^2 + f_n^2)^{\frac{1}{4}}$ . Replacing both of these in the above equation, we obtain

$$(\mathcal{U}_{\zeta_n} Y_n)(f) = \sqrt{\frac{c}{r}} e^{-j2\pi f \frac{r}{c}}. \quad (10)$$

This illustrates that after warping, the  $n$ th output mode becomes a single sinusoid in frequency with sinusoidal frequency  $\frac{r}{c}$ . This corresponds to an impulse in time  $\sqrt{\frac{c}{r}} \delta(t - \frac{r}{c})$ . When the dispersive warping operation  $\mathcal{U}_{\zeta_n}$  with  $n = n_0$  is applied to the signal  $Y(f) = Y_{isovelocity}(f)$ , where  $Y(f)$  is the noiseless received signal in (3), we have

$$\begin{aligned} \mathcal{Y}_{n_0}(f) &= (\mathcal{U}_{\zeta_{n_0}} Y)(f) \\ &= C_{n_0} \sqrt{\frac{c}{r}} e^{-j2\pi f \frac{r}{c}} + \sum_{\substack{n=0 \\ n \neq n_0}}^{N_m-1} (\mathcal{U}_{\zeta_n} X C_n \Theta_n)(f). \end{aligned} \quad (11)$$

The dispersive group delay nature of received signal in (3) can be demonstrated using simulated data from *KRAKEN* [23]. *KRAKEN* is a software that was written to model underwater acoustic propagation and can generate a received signal given a set of environment parameters and the spectrum of transmitted signal  $X(f)$ . An example of a *KRAKEN* simulated  $Y_{isovelocity}(f)$  in (3) is shown represented by its spectrogram (squared magnitude of the

short-time Fourier transform in (12)) in Fig. 5(a) for  $N_m = 3$  modes. As it can be seen, lower frequencies are shifted in time by larger amounts than higher frequencies for each mode. For the received signal  $y(t)$  with Fourier spectrum  $Y(f)$ , the STFT, with window  $g(t)$  whose Fourier transform is  $G(f)$ , can be written as

$$\text{STFT}_Y(t, f; g) = \int_v Y(v)G(v - f)e^{j2\pi vt} dv. \quad (12)$$

The squared magnitude of the STFT or spectrogram is given by  $S_Y(t, f; g) = |\text{STFT}_Y(t, f; g)|^2$ . The STFT of the noiseless received signal  $Y(f)$  can be obtained from (3) and (12). An appropriate window length for  $g(t)$  was chosen to obtain good time-frequency resolution based on this set of environment parameters. The STFT of the warped noiseless received signal  $\mathcal{Y}_{\zeta_n}(f) = (\mathcal{U}_{\zeta_n}Y)(f)$ ,  $n = 0, 1, 2$  in (11) where the warping operator  $\mathcal{U}_{\zeta_n}$  is defined in (10) is shown in Fig. 5(b) 5(c) and 5(d), respectively.

Fig. 5(a) shows each mode as a dispersive (nonlinear) curve in the TF plane. The three modes in this representation are not easily separable. However, when we obtain the spectrogram of the warped signal and look at each mode separately, the corresponding mode appears as a wideband pulse. Note that the other modes (e.g., mode 1 and 2 in Fig. 5(b)) are still dispersive especially in the low frequency band. This is because, from (11),  $\mathcal{Y}_0(f) = C_0\sqrt{\frac{r}{c}}e^{-j2\pi f\frac{r}{c}} + C_1(\mathcal{U}_{\zeta_0}X\Theta_1)(f) + C_2(\mathcal{U}_{\zeta_0}X\Theta_2)(f)$ . Thus, the last two terms are still dispersive as the warping only simplified the  $n = 0$  mode. In the higher frequency region, the three modes are not separable. This is because as  $f$  increases, all three modes appear as impulses since for  $f \gg f_n$ ,  $e^{-j2\pi(\frac{\sqrt{f^2 - f_n^2}}{u}c)} \approx e^{-j2\pi f\frac{r}{c}}$ .

As we demonstrated, in the low frequency region, each mode can be discriminated by applying a corresponding matched warping. However, whether the modes can be separated or not also depends on the transmission band and the distance between the transmitter and receiver. Generally speaking, a longer transmission distance or a lower transmission frequency band can cause the signal to become more dispersive, so the modes are easier to discriminate. If the signal is transmitted in a high frequency band, the modes are closer to each other, and are more difficult to separate.

### B. Pekeris Model

The TF characteristic of the acoustic signal using Pekeris model is determined by the modal group velocity (MGV)  $g_n(f) = \frac{d}{df}k_{r,n}(f)$  in (6). Specifically, the propagating group delay in the  $n$ th mode is determined by

$$\tau_n(f) = \frac{r}{g_n(f)}, \quad (13)$$

where  $r$  is the range from transmitter to the receiver. The modal group velocity is shown in Fig. 6 using  $c = 1,500$  m/s,  $c_B = 1,800$  m/s,  $\rho = 1$  kg/m<sup>3</sup>,  $\rho_B = 1.8$  kg/m<sup>3</sup> and  $D = 100$  m. From Fig. 6, we notice that, for the waveguide model with fluid seabed, the MGV approaches  $c$  (the velocity of sound in the ocean) when the frequency approaches infinity, and it approaches  $c_B$  (the velocity of sound in the seabed) when the frequency approaches the cutoff frequency of the mode.

Using the Pekeris model, we further analyzed further the TF characteristics of the received signal using the spectrogram and investigated the effect of changing the parameters of the environment. Fig. 7(a) represents the

spectrogram of the received signal with  $N_m = 3$  modes and environment parameters:  $D = 100$  m and  $r = 15$  km. Fig. 7(b) uses the same set of parameters except that  $r = 30$  km. Fig. 7(a) shows each mode as a dispersive curve in the TF plane. We note that the curves stop at the cutoff frequency with no asymptotical behavior to the cutoff frequency. Fig. 7(b) shows that a longer range between the receiver and transmitter causes more dispersion in the received signal. When the ocean depth is decreased to  $D = 30$  m, in Fig. 8(b), the modal cutoff frequency is higher and the dispersive effect is more visible at higher frequencies.

#### IV. SHALLOW ENVIRONMENT BLIND IDENTIFICATION USING TIME-FREQUENCY SEPARATION

In many cases, the environment parameters are not known and we need to blindly identify the corresponding time-frequency characteristics. We are considering the problem of identifying the TF or group delay signatures of the received signal after it propagates over a shallow water environment. Following the isovelocity normal-mode model in (3), we expect the resulting TF characteristics to vary dispersively with frequency according to the functions  $\xi_n(f), n = 0, \dots, N_m - 1$  in (8), where  $N_m$  is the total number of modes. As discussed in Section II, the normal-mode model treats the ocean as a waveguide with plane, parallel boundaries, representing the acoustic field in the ocean medium as a sum of normal modes; each mode can be seen as a TF signature component.

Specifically, we denote group velocity function (GVF) of the  $n$ th mode as

$$M_n(f) = C_n \Theta_n(f). \quad (14)$$

In order to blindly identify the TF characteristics of the mode, we will first estimate the GVFs as a linear combination of linear chirps and then we will design TF separation to separate one mode from another.

##### A. Approximation of GVF

Our aim is to approximate the real part of  $M_n(f)$  over the frequency region  $f \in [i\Delta f_n, (i+1)\Delta f_n]$ ,  $i = 1, \dots, Q_n - 1$ , where  $\Delta f_n$  and  $Q_n\Delta f_n$  are the starting and ending frequency of the  $n$ th mode. Using the real frequency-modulated (FM) linear chirps  $L_{n,i}(f) = \cos(\phi_{n,i}(f))$ , where

$$\phi_{n,i}(f) = a_{n,i} + b_{n,i}f + c_{n,i}f^2. \quad (15)$$

Here,  $a_{n,i}$ ,  $b_{n,i}$  and  $c_{n,i}$  are the constant coefficients of the quadratic phase of the  $i$ th linear chirp for the  $n$ th mode. Specifically, we are trying to approximate  $M_n(f)$  as

$$M_n(f) = \sum_{i=1}^{Q_n-1} L_{n,i}(f)(u(f - i\Delta f_n) - u(f - (i+1)\Delta f_n)), \quad (16)$$

where  $u(f)$  is the frequency domain unit step function. Since the GVF is a continuous and monotonic function, one can expect that the linear chirp  $L_{n,i}(f)$  is related with the linear chirp  $L_{n,i+1}(f)$  by the following continuity constraints. Since the analytic part of the linear chirp has a continuous and monotonic group delay function given by  $\frac{d}{df}\phi_{n,i} = b_{n,i} + 2c_{n,i}f$ , then it can be shown that

$$b_{n,i+1} = b_{n,i} + 2(i+1)\Delta f_n(c_{n,i} - c_{n,i+1}) \quad (17)$$

From the continuous and monotonic phase  $\phi_{n,i}(f)$ , we obtain

$$a_{n,i+1} = a_{n,i} + (i+1)\Delta f_n(b_{n,i} - b_{n,i+1}) + (i+1)\Delta f_n^2(c_{n,i} - c_{n,i+1}). \quad (18)$$

Let  $L_{n,i}(f)$  be the linear chirp which best approximates the component  $M_n(f)$  over the frequency interval  $[i\Delta f_n, (i+1)\Delta f_n]$ ; then the next chirp  $L_{n,i+1}(f)$  on the frequency interval  $[(i+1)\Delta f_n, (i+2)\Delta f_n]$  is chosen to be the one that best matches  $M_n(f)$  over that interval following the continuity constraints in (17) and (18).

The problem of finding  $Q_n - 1$  linear chirps  $L_{n,i}(f)$ ,  $i = 1, \dots, Q_n - 1$  can be formulated as a multi-hypothesis detection problem, which can be solved using quadrature matched filtering [24]. We denote the  $K$  candidates for the linear chirp phase as  $\phi_{n,i}^k(f)$ ,  $k = 1, \dots, K$ . The best of the candidates,  $\phi_{n,i}^{\hat{k}}(f)$ , can be found by the following maximization problem

$$\hat{k} = \arg \max_{k=1, \dots, K} \frac{n_s^k (x_c^k)^2 - 2n_{cs}^k x_c^k x_s^k + n_c (x_s^k)^2}{2n_c^k n_x^k - 2(n_{cs}^k)^2}, \quad (19)$$

subject to (17) and (18), where

$$\begin{aligned} x_c^k &= \int_{i\Delta f_n}^{(i+1)\Delta f_n} Y(f) \cos(\phi_{n,i}^k(f)) dt, \\ x_s^k &= \int_{i\Delta f}^{(i+1)\Delta f} Y(f) \sin(\phi_{n,i}^k(f)) dt, \\ n_c^k &= \int_{i\Delta f_n}^{(i+1)\Delta f} \cos^2(\phi_{n,i}^k(f)) dt, \\ n_s^k &= \int_{i\Delta f_n}^{(i+1)\Delta f} \sin^2(\phi_{n,i}^k(f)) dt, \\ n_{cs}^k &= \int_{i\Delta f}^{(i+1)\Delta f} \cos(\phi_{n,i}^k(f)) \sin(\phi_{n,i}^k(f)) dt. \end{aligned}$$

Thus, the solution to the above constrained programming problem can give the optimal estimation of  $M_n(f)$  between the time interval  $[i\Delta f, (i+1)\Delta f]$ ,  $i = 1, 2, \dots, Q_n - 1$  in additive white Gaussian noise.

### B. Mode Separation

Fig. 7 illustrates that each mode of the received signal appears as a distinct dispersive curve in the TF plane, which makes the separation of the modes possible. After approximating the  $N_m$  GVFs, we will now use a TF mode separation technique based on warping to separate the modes. Specifically, we want to design  $(N_m - 1)$  GVF curve separators in the TF plane, denoted by  $e_n(f)$ ,  $n = 0, \dots, N_m - 2$ . A GVF curve separator is a curve situated between two successive modes in the TF plane. Knowing the TF structures, we appropriately set  $M_n$  TF points  $(t_m, f_m)$ ,  $m = 1, 2, \dots, M_n$  in the middle of the space between two successive modes. The  $M_n$  points between the  $n$ th mode and the  $(n+1)$ th mode will constitute the GVF separator  $e_n(f)$ . As the GVF curves will be processed with the received signal, the separation lines must have the same duration as the received signal. To ensure that, all the GVF curves are assumed to start at the initial TF point of the received signal and are extended from the final TF point to the time axis using a vertical line.

To separate the TF components using the estimated GVF curves [19], we use a non-unitary warping operation defined as

$$(\mathcal{W}_{\mu_n} Y)(f) = Y(\mu_n(f)), \quad (20)$$

where  $\mu_n(f)$  is obtained from the  $n$ th GVF separator  $e_n(f)$  using  $\mu_n = \frac{1}{\lambda_r} \int_{-\infty}^f e_n(v) dv$ , where  $\lambda_r > 0$  is a normalization constant. Note that the non-unitary operator overcomes the spreading effect of unitary warping operators as no frequency-dependent amplitude modulation is needed as in [20]. It can be shown that

$$\mathcal{W}_{\mu_n} e^{j2\pi\lambda_r\mu^{-1}(f)} = e^{j2\pi f\lambda}. \quad (21)$$

Assume that the received noisy signal in (5) is  $r(t)$  with Fourier transform (FT)  $R(f) = Y(f) + W(f)$ , where  $W(f)$  is additive white Gaussian noise. We compute the  $n$ th generalized FT of  $R(f)$  using the  $n$ th warping function  $\mu_n(f)$  as [25]

$$\mathcal{G}_R^n(\lambda) = \int_{\wp} R(f) \frac{d\mu_n(f)}{df} e^{j2\pi\lambda \mu_n(f)} df, \quad (22)$$

where  $\wp$  contains the values of  $f$  in the domain of the warping function  $\mu_n(f)$ , and  $\lambda$  is a real and unitless parameter.

To obtain the first mode, we compute the inverse  $n = 0$  generalized FT as follows:

$$R_0(f) = \int_{-\infty}^{\lambda_r} \mathcal{G}_R^0(\lambda) e^{-j2\pi\lambda\mu_0(f)} d\lambda, \quad (23)$$

that can be written as

$$R_0(f) = \int_{-\infty}^{\infty} h(\lambda) \mathcal{G}_R^0(\lambda) e^{-j2\pi\lambda\mu_0(f)} d\lambda, \quad (24)$$

where  $h(\lambda) = 1$  if  $\lambda \in (-\infty, \lambda_r]$  and  $h(\lambda) = 0$  if  $\lambda \in (\lambda_r, \infty]$ . Inserting (22) into (24), one can show that

$$\begin{aligned} R_0(f) &= \int_{-\infty}^{\infty} \int_{\wp} R(v) h(\lambda) \frac{d\mu_0(v)}{dv} e^{j2\pi\lambda(\mu_0(v) - \mu_0(f))} dv d\lambda \\ &= \int_{\wp} R(v) \frac{d\zeta_0(v)}{dv} H(\mu_0(v) - \mu_0(f)) dv, \end{aligned} \quad (25)$$

where  $H(f)$  is the inverse FT of the lowpass filter  $h(\lambda)$ . If we subtract the first mode from the received signal, i.e.,  $\hat{R}_1(f) = R(f) - R_0(f)$ , then the remaining modes are contained in  $\hat{R}_1(f)$ . To obtain the second mode, we apply the above procedure to  $\hat{R}_1(f)$  using the GVF curve separator  $e_1(f)$ . Repeating the above procedure for each GVF curve separator, we can then separate each mode from the received signal.

### C. Example of Blind Identification and TF Separation

Fig. 9(a) represents the spectrogram of the received signal for  $N_m = 3$  modes when we consider the environment parameters to be  $D = 100$  m, ocean depth and  $r = 15$  km range between the transmitter and receiver. Each mode is shown as a dispersive component in the TF plane, and we notice that the curves stop at the cutoff frequency  $f_n$  of the  $n$ th mode with no asymptotical behavior to the cutoff frequency. For the example in Fig. 9, we use the mode separation technique on the received signal excited by the waveform  $X(f) = \sqrt{f}$ ,  $f > 0$  for  $N_m = 3$  modes using Pekeris model. The GVF curve separators are shown in Fig. 9(c) and the separated components are shown in Fig. 9(d).

## V. TIME-DISPERSION DIVERSITY RECEIVER DESIGN

Although the shallow water environment models provide an inherent frequency domain system model, in realistic shallow water environments, many factors can cause distortion in the signal propagation. These include water fluctuations on the ocean surface and the roughness of the ocean bottom that can affect the signal reflections. Hence, it is reasonable to introduce randomness into the channel model. We model this distortion by introducing random fading to the data generated by the normal-mode modeling software *KRAKEN* for both the isovelocity and Pekeris models [26]. Specifically, we model the randomness in the shallow water environment by [27]

$$D_n = \alpha_n + \Delta\rho_n \quad (26)$$

where  $\alpha_n$  is the deterministic mean amplitude of the  $n$ th mode and  $\Delta\rho_n$  is random fading distortion. In practice, can be obtained using system identification techniques. Additive noise is also introduced in the channel model due to the random disturbance in the ocean environment. Therefore, the received signal spectrum is expressed as

$$R(f) = Y(f) + W(f) = X(f) \sum_{n=0}^{N_m-1} D_n C_n(f) \Theta_n(f) + W(f), \quad (27)$$

where  $W(f)$  is zero-mean additive white Gaussian noise whose samples have variance  $\sigma_W^2$ , and  $Y(f)$  is given for the isovelocity model with  $C_n(f) = C_n$  in (3) and (26), and for the Pekeris model in (5).

### A. Receiver Design without TF Separation

We propose a filter bank receiver scheme which can exploit the frequency domain dispersion diversity in the normal modes. In the following of this section, we discuss the receiver design in the context of isovelocity model, but this design can be also used for Pekeris model. The received signal spectrum  $R(f)$  in (27) is processed using a matched filter to obtain all the modes, and then the resulting matched filtered outputs are combined in the minimum-error-probability sense to obtain the optimal detection rule for the transmitted information symbol  $b$ . This receiver scheme is shown in Fig. 10. Note

As stated in Section III-A, we can design the transmitted signal  $X(f) = \sqrt{f}$ ,  $f > f_n$ ,  $n = 0, \dots, N_m - 1$ , and then employ frequency domain warping and matched filtering to the received signal. We define  $U_n(f) = X(f)C_n(f)\Theta_n(f)$  for the  $n$ th mode. Then the noiseless output of the  $n$ th matched filter can be expressed as

$$Z_n = \langle Y, U_n \rangle = \int_f Y(f) U_n^*(f) df. \quad (28)$$

If we use the operator  $\mathcal{U}_{\zeta_n}$  in (9), then we can rewrite (28) as  $\langle Y, U_{\zeta_n} \rangle = \langle \mathcal{U}_{\zeta_n} Y, \mathcal{U}_{\zeta_n} U_n \rangle$ , due to the unitarity of  $\mathcal{U}_{\zeta_n}$

Using (10) and (11), the output of the  $n$ th matched filter in (28) can also be written as

$$Z_n = \int_f (\mathcal{U}_{\zeta_n} Y)(f) G(f) e^{j2\pi f \frac{z}{c}} df, \quad (29)$$

where  $(\mathcal{U}_{\zeta_n} U_n)(f) = (\mathcal{U}_{\zeta_n} C_n)(f) (\mathcal{U}_{\zeta_n} X \Theta_n)(f) = \mathcal{C}(f) \sqrt{\frac{z}{r}} e^{j2\pi f \frac{z}{c}}$ . This can be seen as the short-time Fourier transform of the warped signal  $(\mathcal{U}_{\zeta_n} Y)(f)$  at epoch  $\frac{z}{c}$  using an analysis window  $G(f) = 1$ . Using the warping

technique in Section III-A that transforms the received signal of each mode at a time to a sinusoid, the modes can be easily discriminated in the TF plane.

The receiver processing can be described as follows. Concatenating the signals  $U_n(f)$ , we can express the resulting filter bank using vector  $\mathbf{u}(f) = [U_0(f) \ \cdots \ U_{N_m-1}(f)]^\mathbb{T}$ , where  $\mathbb{T}$  denotes matrix transpose. Similarly, we rewrite the information bit  $b$  to the transmitted as an  $N_m \times 1$  vector  $\mathbf{b} = [b, b, \dots, b]^\mathbb{T}$ . Let  $\mathbf{D} = \text{diag}(D_0, D_1, \dots, D_{N_m-1})$  be the  $N_m \times N_m$  matrix whose diagonal elements are the random channel coefficients in (26). Using the aforementioned vector notation, the received spectrum can also be written as

$$R(f) = \mathbf{u}^\mathbb{T}(f) \mathbf{D} \mathbf{b} + W(f). \quad (30)$$

The output of this filter bank is given by  $\mathbf{z} = \mathbf{P} \mathbf{D} \mathbf{b} + \mathbf{w}$ , where  $\mathbf{P} = \int_f \mathbf{u}^*(f) \mathbf{u}^\mathbb{T}(f) df$  is the matrix of correlations between the different modes, and  $\mathbf{w} = \int_f \mathbf{u}^*(f) W(f) df$  is the noise at the output of the matched filters with covariance  $\sigma_w^2 \mathbf{P}$ .

If binary antipodal symbols are transmitted, i.e.,  $b = +1$  or  $-1$ , and we denote  $\mathbf{d} = \mathbf{D} \mathbf{b}$  using the above notation, then the communication problem stated above can be converted to a classic, detection problem with two hypothesis

$$\mathcal{H}_0 : \mathbf{z} = -\mathbf{P} \mathbf{d} + \mathbf{w}, \quad (31)$$

$$\mathcal{H}_1 : \mathbf{z} = \mathbf{P} \mathbf{d} + \mathbf{w}. \quad (32)$$

To solve the detection problem in (31) and (32), we first perform prewhitening [28]. Denoting the conjugate transpose operation as  $\dagger$ , notice that  $\mathbf{P} = \mathbf{P}^\dagger$  is a Hermitian matrix, thus it can be expressed as  $\mathbf{P} = \mathbf{Q}^\dagger \mathbf{\Lambda} \mathbf{Q}$ , where  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_{N_m})$  is the eigenvalue matrix of  $\mathbf{P}$ , and  $\mathbf{Q}$  is the prewhitening matrix. From the properties of Hermitian matrices, we also know that all the eigenvalues are greater or equal to zero. Hereby we assume  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{N_m-1} \geq \lambda_{N_m} \geq 0$ , and that the first  $N_r$  eigenvalues are greater than 0. Multiplying  $\mathbf{Q}$  to both sides of (31) and (32), we have

$$\begin{aligned} \mathcal{H}_0 : \hat{\mathbf{z}} &= -\mathbf{\Lambda} \hat{\mathbf{d}} + \hat{\mathbf{w}}, \\ \text{where } \hat{\mathbf{d}} &= \mathbf{Q} \mathbf{d}, \hat{\mathbf{w}} = \mathbf{Q} \mathbf{w}, \hat{\mathbf{z}} = \mathbf{Q} \mathbf{z}, \end{aligned} \quad (33)$$

$$\begin{aligned} \mathcal{H}_1 : \hat{\mathbf{z}} &= \mathbf{\Lambda} \hat{\mathbf{d}} + \hat{\mathbf{w}}, \\ \text{where } \hat{\mathbf{d}} &= \mathbf{Q} \mathbf{d}, \hat{\mathbf{w}} = \mathbf{Q} \mathbf{w}, \hat{\mathbf{z}} = \mathbf{Q} \mathbf{z}. \end{aligned} \quad (34)$$

Now we assume that the rank of  $\mathbf{P}$  is  $N_r$ . Then, according to the Neyman-Pearson Theorem, the detector decides  $\mathcal{H}_1$  if

$$L(\hat{\mathbf{z}}) = \left( \prod_{i=1}^{N_r} e^{-\frac{|\hat{z}_i - \lambda_i \hat{d}_i|^2}{2\sigma_w^2 \lambda_i}} \right) / \left( \prod_{i=1}^{N_r} e^{-\frac{|\hat{z}_i + \lambda_i \hat{d}_i|^2}{2\sigma_w^2 \lambda_i}} \right) > \gamma, \quad (35)$$

where  $L(\hat{\mathbf{z}})$  is the likelihood function for this detection problem,  $\hat{z}_i$  is the  $i$ th element of vector  $\hat{\mathbf{z}}$ ,  $\hat{d}_i$  is the  $i$ th element of vector  $\hat{\mathbf{d}}$ , and  $\lambda_1, \dots, \lambda_{N_r}$  are  $N_r$  nonzero eigenvalues of matrix  $\mathbf{P}$ .

We can use the Bayesian approach to minimize the probability of error in the received symbols. If the probabilities of transmitting  $+1$  and  $-1$  are equal, we can choose  $\gamma = 1$  to obtain the minimum bit-error-rate (BER). After

simplification of (35), the detector decides  $\mathcal{H}_1$  if

$$\sum_{i=1}^{N_r} \text{Re}\{\hat{z}_i \hat{d}_i^*\} > 0 \quad (36)$$

where  $\text{Re}\{\cdot\}$  denotes the real part.

We use the first  $N_r$  rows of  $\mathbf{Q}$  to form the new matrix  $\mathbf{Q}_r$  and the first  $N_r$  rows of  $\mathbf{\Delta}$  to form  $\mathbf{\Delta}_r$ . The first  $N_r$  elements of  $\hat{\mathbf{z}}$  form a new vector  $\hat{\mathbf{z}}_r$ , and the first  $N_r$  elements of  $\hat{\mathbf{d}}$  form a new vector  $\hat{\mathbf{c}}$ . The following relationship exists among those variables:  $\hat{\mathbf{z}}_r = \mathbf{Q}_r \mathbf{z}$  and  $\hat{\mathbf{c}} = \mathbf{Q}_r \mathbf{d}$ . Notice that  $\sum_{i=1}^{N_r} \text{Re}\{\hat{z}_i \hat{d}_i^*\} = \text{Re}\{\hat{\mathbf{c}}^\dagger \hat{\mathbf{z}}_r\} = \text{Re}\{\mathbf{d}^\dagger \mathbf{Q}_r^\dagger \mathbf{Q}_r \mathbf{z}\}$ . The minimum error probability detector can then be expressed as, decide  $\mathcal{H}_1$  if

$$\hat{b} = \text{Re}\{\mathbf{d}^\dagger \mathbf{Q}_r^\dagger \mathbf{Q}_r \mathbf{z}\} > 0. \quad (37)$$

### B. Receiver Design with TF Component Separation

Using the warping technique as in Section III-A, although the mode we choose to process is warped using a function that is well-matched to it, the same function also warps the other modes that can then act as interference. In order to avoid the presence of this interference, we use the TF mode separation discussed in Section IV. After the TF mode-components are separated, each component can be treated as a subchannel of the shallow water communication channel; hence diversity can be obtained if the receiver is properly designed. Our proposed receiver design for diversity is shown in Fig. 11. To exploit the potential diversity, we first transmit a pilot signal to separate the modes. Then, the received signal corresponding to the pilot signal is analyzed, and each component is separated and used jointly with the channel coefficients as the matched filter for the received signal of the next transmitted symbols. The outputs of the matched filters are combined, and the decisions for the estimated symbols are made using a minimum error probability detector.

This receiver processing can be described as follows. Concatenating the signals  $U_n(f)$ , we can express the resulting separation results using matrix  $\mathbf{U}(f) = \text{diag}\{U_0(f), \dots, U_{N_m-1}(f)\}$ . Note that  $\mathbf{U}(f)$  is a diagonal function matrix, which is different from  $\mathbf{u}(f)$  in (30). We also concatenate noise at each matched filter as  $\mathbf{w}(f) = [W_0(f), \dots, W_{N_m-1}(f)]^\top$ . And using the aforementioned vector notation, the received spectrum after separation can also be written in vector form as

$$\mathbf{r}(f) = \mathbf{U}(f) \mathbf{D} \mathbf{b} + \mathbf{w}(f). \quad (38)$$

The output of this filter bank is given by  $\mathbf{z} = \mathbf{P} \mathbf{D} \mathbf{b} + \mathbf{w}$ , where  $\mathbf{P} = \int_f \mathbf{U}^*(f) \mathbf{U}(f) df$  is the matrix of correlations between different modes, and  $\mathbf{w} = \int_f \mathbf{U}^*(f) \mathbf{w}(f) df$  is the noise at the output of the matched filters with covariance  $\mathbf{C}_{cov} = \text{E}\{\mathbf{w} \mathbf{w}^\dagger\}$ .

If binary antipodal symbols are transmitted, i.e.,  $b = +1$  or  $-1$ , and we denote  $\mathbf{d} = \mathbf{D} \mathbf{b}$  using the above notation, then the communication problem stated above can be converted to the same classic, detection problem in (31) and (32).

We notice that  $\mathbf{P}$  is a diagonal matrix with the power of each mode as its element, thus it can be expressed as  $\mathbf{P} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_{N_m})$ . We also know that all the elements are greater or equal to zero. Hereby we assume

$\lambda_1 \geq \lambda_2 \geq \dots \lambda_{N_m-1} \geq \lambda_{N_m} \geq 0$ . Then according to the *Neyman-Pearson Theorem*, the detector decides  $\mathcal{H}_1$  (or that bit  $b = 1$  was transmitted) if

$$L(\mathbf{z}) = \frac{\prod_{i=1}^{N_m} e^{-\frac{|z_i - \lambda_i d_i|^2}{2\sigma_i^2 \lambda_i}}}{\prod_{i=1}^{N_m} e^{-\frac{|z_i + \lambda_i d_i|^2}{2\sigma_i^2 \lambda_i}}} > \gamma, \quad (39)$$

where  $L(\mathbf{z})$  is the likelihood function for this detection problem,  $z_i$  is the  $i$ th element of vector  $\mathbf{z}$ ,  $d_i$  is the  $i$ th element of vector  $\mathbf{d}$ , and  $\lambda_1, \dots, \lambda_{N_m}$  are  $N_r$  nonzero eigenvalues of matrix  $\mathbf{P}$ .

We can use the Bayesian approach to minimize the probability of error in the received symbols. If the probabilities of transmitting  $+1$  and  $-1$  are equal, we can choose  $\gamma = 1$  to obtain the minimum BER. After simplification of (39), the detector decides  $\mathcal{H}_1$  if

$$\sum_{i=1}^{N_m} \text{Re} \left\{ \frac{z_i d_i^*}{\sigma_i^2} \right\} > 0 \quad (40)$$

### C. Performance Analysis

1) *Performance of Receivers without TF Component Separation*: In this section, the BER and diversity performances of the waveform and receiver design are investigated when the modes are not separated but only warped.

From (33), we know that the  $N_m$  correlated received signals can be transformed into  $N_r$  independent received signals. Without loss of generality, we assume that  $b = 1$  is transmitted, and the  $N_r$  independent signals can be expressed as

$$\hat{z}_i = \lambda_i \mathbf{q}_i \mathbf{d} + \hat{w}_i, \quad i = 1, 2, \dots, N_r, \quad (41)$$

where  $\hat{z}_i$  is the  $i$ th element of  $\hat{\mathbf{z}}$  in (33),  $\mathbf{q}_i$  is the  $i$ th row of unitary matrix  $\mathbf{Q}$ , and  $\hat{w}_i$  is the  $i$ th element of noise  $\hat{\mathbf{w}}$ . According to the minimum error probability detector rule in (37),

$$\hat{z}_i \hat{d}_i = \lambda_i \mathbf{q}_i \mathbf{d} \mathbf{d}^\dagger \mathbf{q}_i^\dagger + \hat{w}_i \mathbf{d}^\dagger \mathbf{v}_i^\dagger, \quad i = 1, 2, \dots, N_r. \quad (42)$$

The signal-to-noise ratio (SNR)  $\gamma_i$  of the  $i$ th received signal in (42) can be expressed as

$$\gamma_i = \frac{\lambda_i^2 |\mathbf{v}_i \mathbf{d}|^4}{\mathbb{E}[(\mathbf{q}_i \mathbf{d}) w_i^* w_i (\mathbf{q}_i \mathbf{d})^\dagger]} = \frac{\lambda_i^2 |\mathbf{q}_i \mathbf{d}|^4}{\lambda_i |\mathbf{q}_i \mathbf{d}|^2} = \lambda_i |\mathbf{q}_i \mathbf{d}|^2. \quad (43)$$

From (26), we know that the  $n$ th component of  $\mathbf{d}$  can be expressed as  $D_n = C_n + \Delta \varrho_n$ . If we reasonably assume that  $\Delta \varrho_n$  is independent for each mode, the covariance matrix of  $\mathbf{d}$  can be given by

$$\begin{aligned} \mathbf{C}_{DD} &= E[(\mathbf{d} - E[\mathbf{d}])(\mathbf{d} - E[\mathbf{d}])^\dagger] \\ &= \text{diag}(\sigma_{D,0}^2, \sigma_{D,1}^2, \dots, \sigma_{D,N_p-1}^2). \end{aligned} \quad (44)$$

If we define  $\Sigma = \frac{1}{\sigma_w^2} \sqrt{\Lambda_r} \mathbf{Q}_r \mathbf{C}_{DD} \mathbf{Q}_r^\dagger \sqrt{\Lambda_r}^\dagger$ , from (43) and (44), we know that  $\Sigma$  is full rank. Thus, using the minimum error probability detector rule in (37), the average BER can be given by [29]:

$$P_b = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \left[ \det \left( \frac{\Sigma}{\sin^2 \theta} + \mathbf{I} \right) \right]^{-1} e^{-\mathbf{m}^\dagger (\Sigma + \sin^2 \theta \mathbf{I})^{-1} \mathbf{m}} d\theta \quad (45)$$

where  $\mathbf{m} = \frac{1}{\sigma_w} \sqrt{\Lambda_r} \mathbf{Q}_r E[\mathbf{d}]$ . The potential diversity order is given by  $N_r$ , which is the rank of  $\mathbf{P}$ .

As stated in the previous section, the normal modes are not mutually orthogonal; this leads to the nonorthogonality between the matched-filters for corresponding modes. As a result, matrix  $\mathbf{P}$  may not have full rank. From the analysis of the time-frequency properties of the shallow water modes, the rank of  $\mathbf{P}$  is jointly decided by the environment parameters and the transmission frequency band, as shown in Fig. 12. In this figure, both the axis represent the mode numbers  $n = 0, 1, \dots, 9$ , respectively, and the amplitudes of the correlation between any two modes are plotted. To obtain this figure, the following shallow water environment parameters were used: the transmission distance was 15 km, the sound of speed underwater was 1,500 m/s, and the depth of the ocean was 100 m. In Fig. 12(a), the transmitted frequency band is within 1 kHz – 2 kHz, so an identity like correlation matrix  $\mathbf{P}$  is obtained with an almost full rank of  $N_r = 10$ ; when the transmission frequency band changes, the correlation matrix starts to spread. This means that the modes become more correlated to each other. This leads to the reduction of the rank of  $\mathbf{P}$  as well. In Fig. 12(d), the rank of  $\mathbf{P}$  reduced to 7. In general, for the same set of environment parameters, the modes are more difficult to discriminate in higher frequency bands, thus the rank of  $\mathbf{P}$  decreases as the transmission frequency band increases.

The BER simulation results are shown in Fig. 13, using a set of environment parameters with  $D = 50$  m,  $c = 1,500$  m/s and  $r = 15$  km. Fig. 13 shows the BER performance for 0 dB – 30 dB SNR values. The numerical results show the BER and diversity performances of three transmission frequency bands: 1–2 kHz, 5–6 kHz and 7–8 kHz. As we can see, the BER performance deteriorates and the diversity order decreases when the transmission frequency band increases. From Fig. 13, we can observe the approximate diversity orders of 9.8, 7.9 and 5.2 for the three frequency bands. This illustrates that different transmission frequency bands can obtain different diversity orders, corresponding to the rank of  $\mathbf{P}$  in each frequency band.

2) *Performance of Receivers with Separation:* In this section, the BER and diversity performances of the proposed waveform and receiver design are investigated when the modes are separated. From (38), without loss of generality, we assume that  $b = 1$  is transmitted. The  $N_m$  independent signals can be expressed as

$$z_i = \lambda_i d_i + W_i, \quad i = 1, 2, \dots, N_m, \quad (46)$$

where  $z_i$  is the  $i$ th element of  $\mathbf{z}$  in (33), and  $w_i$  is the  $i$ th element of the noise  $\mathbf{w}$ .

According to the minimum error probability detector rule in (37), we have

$$z_i d_i^* = \lambda_i d_i d_i^* + w_i d_i^*, \quad i = 1, 2, \dots, N_m. \quad (47)$$

The SNR  $\gamma_i$  of the  $i$ th received signal in (47) can be expressed as

$$\gamma_i = \frac{\lambda_i^2 |d_i|^4}{\mathbb{E}[(d_i) w_i^* w_i d_i^*]} = \frac{\lambda_i^2 |d_i|^4}{\lambda_i \sigma_i^2 |d_i|^2} = \frac{\lambda_i |d_i|^2}{\sigma_i^2}. \quad (48)$$

From (26), we know that the  $n$ th component of  $\mathbf{d}$  can be expressed as  $D_n = \alpha_n + \Delta \varrho_n$ . If we reasonably assume that  $\Delta \varrho_n$  is independent for each mode, the covariance matrix of  $\mathbf{d}$  can be given by (44).

If we define  $\mathbf{\Sigma} = \mathbf{P}^{\frac{1}{2}} \mathbf{C}_{cov}^{-\frac{1}{2}} \mathbf{C}_{DD} \mathbf{C}_{cov}^{-\frac{1}{2}} \mathbf{P}^{\frac{1}{2}}$ , then from (48), we know that  $\mathbf{\Sigma}$  is full rank. Using the minimum error

probability detector rule in (37), the average BER can be given by [29]:

$$P_b = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \left[ \det \left( \frac{\Sigma}{\sin^2 \theta} + \mathbf{I} \right) \right]^{-1} e^{-\mathbf{m}^\dagger (\Sigma + \sin^2 \theta \mathbf{I})^{-1} \mathbf{m}} d\theta, \quad (49)$$

where  $\mathbf{m} = \mathbf{P}^{\frac{1}{2}} \mathbf{C}_{cov}^{-\frac{1}{2}} E[\mathbf{d}]$ . The potential diversity order, obtained by calculating the slope of the BER curve, is given by  $N_m$ , which is the rank of  $\mathbf{P}$ .

The BER simulation results are shown in Fig. 14 using  $D = 100$  m,  $c = 1,500$  m/s, and  $r = 15$  km. Fig. 14 shows the BER performance for 0 dB – 30 dB SNR. The numerical results show the BER and diversity performances of three different types of receivers: receiver with TF component separation, without TF component separation, and receiver without diversity. The receiver without diversity uses a single matched filter to receive the whole signal, hence, no diversity is obtained. As we can see, the BER performance of the receiver with TF component separation outperforms the other two. This is because this receiver avoids interference between the normal modes. Also the separation procedure performed TF denoising to the received signal, which further improved the received SNR.

## VI. CONCLUSION

We investigated the general normal-mode modeling of shallow water environments by considering the isovelocity and Pekeris model. Following these normal-mode models, and assuming the environment parameters are known, we proposed a frequency domain system model for shallow water environments and designed the corresponding transmission waveform. We constructed a matched filter bank receiver which matched the frequency domain dispersive characteristic function of the model. Furthermore, we developed the optimal detection rule that combines the outputs of the filter bank to obtain the minimum BER. Numerical results demonstrated that this receiver scheme can improve BER and diversity performance, and we can observe that different transmission bands have different impacts on the performance of this detector.

As in many cases, the environment parameters are unknown, we investigated a new method based on a warping technique for the blind separation of the TF mode-components. This TF separation was achieved by adaptively determining the instantaneous frequency curves of each mode, and filtering the TF components of the received signal. As an application example, we developed the corresponding waveform and receiver design to exploit the diversity existing in the system when the modes are separated. Numerical results demonstrated that the diversity and BER performances were improved by the aforementioned waveform and receiver design schemes.

Although the two models, isovelocity model and Pekeris model are still very ideal for the realistic applications, we expect that the methodology that is investigated in this paper is able to be applied on more complicated and sophisticated models, e.g. the multiple stratified layer model, the range-dependent model, using realistic environment parameters obtained in certain ocean area.

APPENDIX I  
DERIVATION OF DIVERSITY ORDER

The diversity order can be obtained as follows. We start with the BER expression from Equation (6) of [29]:

$$P_b = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \left[ \det \left( \frac{\boldsymbol{\Sigma}}{\sin^2 \theta} + \mathbf{I} \right) \right]^{-1} e^{-\mathbf{m}^\dagger (\boldsymbol{\Sigma} + \sin^2 \theta \mathbf{I})^{-1} \mathbf{m}} d\theta. \quad (50)$$

As  $\boldsymbol{\Sigma}$  is a Hermitian matrix,  $\boldsymbol{\Sigma} + \sin^2 \theta \mathbf{I}$  is also a Hermitian matrix. Hence  $(\boldsymbol{\Sigma} + \sin^2 \theta \mathbf{I})^{-1}$  is also a Hermitian matrix, and as a result,  $\mathbf{m}^\dagger (\boldsymbol{\Sigma} + \sin^2 \theta \mathbf{I})^{-1} \mathbf{m} > 0$ . It follows that

$$e^{-\mathbf{m}^\dagger (\boldsymbol{\Sigma} + \sin^2 \theta \mathbf{I})^{-1} \mathbf{m}} < 1. \quad (51)$$

As a result of Equation (7) from [29], if  $\lambda_l$  are the eigenvalues of  $\boldsymbol{\Sigma}$ , we have

$$\left[ \det \left( \frac{\boldsymbol{\Sigma}}{\sin^2 \theta} + \mathbf{I} \right) \right]^{-1} = \prod_{l=1}^{N_r} \left( \frac{\lambda_l}{\sin^2 \theta} + 1 \right)^{-1}. \quad (52)$$

Because  $\boldsymbol{\Sigma}$  is a Hermitian matrix,  $\lambda_l \geq 0$ , so

$$\prod_{l=1}^{N_r} \left( \frac{\lambda_l}{\sin^2 \theta} + 1 \right)^{-1} < \prod_{l=1}^{N_r} (\lambda_l + 1)^{-1}.$$

As a result,

$$\begin{aligned} P_b &= \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \prod_{l=1}^{N_r} \left( \frac{\lambda_l}{\sin^2 \theta} + 1 \right)^{-1} e^{-\mathbf{m}^\dagger (\boldsymbol{\Sigma} + \sin^2 \theta \mathbf{I})^{-1} \mathbf{m}} d\theta \\ &< \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \prod_{l=1}^{N_r} \left( \frac{\lambda_l}{\sin^2 \theta} + 1 \right)^{-1} d\theta \\ &< \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \prod_{l=1}^{N_r} (\lambda_l + 1)^{-1} d\theta \\ &= \frac{1}{2} \prod_{l=1}^{N_r} (\lambda_l + 1)^{-1}. \end{aligned} \quad (53)$$

If we consider that  $\lambda_l$  represents the SNR on the  $l$ th subchannel, then the eigenvalue can be expressed as  $\lambda_l = \frac{E_l}{\sigma_w^2}$ , where  $E_l$  is a constant for each  $l$ . This implies that  $\frac{1}{2} \prod_{l=1}^{N_r} (\lambda_l + 1)^{-1}$  is on the order of  $(\frac{1}{\sigma_w^2})^{-N_r}$ , and, as a result the attainable diversity is  $N_r$ .

REFERENCES

- [1] Y. Jiang and A. Papandreou-Suppappola, "Discrete time-frequency characterizations of dispersive linear time-varying systems," *IEEE Transactions on Signal Processing*, vol. 55, pp. 2066–2076, May 2007.
- [2] C.-S. Chen, J. Miller, G. Bourdreaux-Bartels, G. Potty, and C. Lazauski, "Time-frequency representations for wideband acoustic signals in shallow water," in *Proceedings OCEANS 2003*, vol. 5, 22-26 Sept. 2003, pp. SP2903–SP2907.
- [3] D. B. Kilfoyle and A. B. Baggeroer, "The state of the art in underwater acoustic telemetry," *IEEE Journal of Oceanic Engineering*, vol. 25, no. 1, pp. 4–27, Jan. 2000.
- [4] M. Stojanovic, J. Catipovic, and J. G. Proakis, "Adaptive multichannel combining and equalization for underwater acoustic communications," *Acoustical Society of America Journal*, vol. 94, pp. 1621–1631, Sept. 1993.
- [5] M. Stojanovic, J. A. Catipovic, and J. G. Proakis, "Phase-coherent digital communications for underwater acoustic channels," *IEEE Journal of Oceanic Engineering*, vol. 19, no. 1, pp. 100–111, 1994.

- [6] M. Stojanovic, "Recent advances in high-speed underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. 21, no. 2, pp. 125–136, April 1996.
- [7] L. LeBlanc and P.-P. Beaujean, "Spatio-temporal processing of coherent acoustic communication data in shallow water," *IEEE Journal of Oceanic Engineering*, vol. 25, no. 1, pp. 40–51, Jan. 2000.
- [8] P.-P. Beaujean and L. LeBlanc, "Adaptive array processing for high-speed acoustic communication in shallow water," *IEEE Journal of Oceanic Engineering*, vol. 29, no. 3, pp. 807–823, July 2004.
- [9] J. Flynn, J. Ritcey, D. Rouseff, and W. Fox, "Multichannel equalization by decision-directed passive phase conjugation: experimental results," *IEEE Journal of Oceanic Engineering*, vol. 29, no. 3, pp. 824–836, July 2004.
- [10] D. Kilfoyle, J. Preisig, and A. Baggeroer, "Spatial modulation experiments in the underwater acoustic channel," *IEEE Journal of Oceanic Engineering*, vol. 30, no. 2, pp. 406–415, April 2005.
- [11] D. Rouseff, D. Jackson, W. Fox, C. Jones, J. Ritcey, and D. Dowling, "Underwater acoustic communication by passive-phase conjugation: theory and experimental results," *IEEE Journal of Oceanic Engineering*, vol. 26, no. 4, pp. 821–831, Oct. 2001.
- [12] G. Edelmann, T. Akal, W. Hodgkiss, S. Kim, W. Kuperman, and H. C. Song, "An initial demonstration of underwater acoustic communication using time reversal," *IEEE Journal of Oceanic Engineering*, vol. 27, no. 3, pp. 602–609, July 2002.
- [13] T. Yang, "Temporal resolutions of time-reversal and passive-phase conjugation for underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. 28, no. 2, pp. 229–245, April 2003.
- [14] G. F. Edelmann, H. C. Song, S. Kim, W. S. Hodgkiss, W. A. Kuperman, and T. Akal, "Underwater acoustic communications using time reversal," *IEEE Journal of Oceanic Engineering*, vol. 30, no. 4, pp. 852–864, Oct. 2005.
- [15] T. Yang, "Correlation-based decision-feedback equalizer for underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. 30, no. 4, pp. 865–880, Oct. 2005.
- [16] H. Song, P. Roux, W. Hodgkiss, W. Kuperman, T. Akal, and M. Stevenson, "Multiple-input-multiple-output coherent time reversal communications in a shallow-water acoustic channel," *IEEE Journal of Oceanic Engineering*, vol. 31, no. 1, pp. 170–178, Jan. 2006.
- [17] L. J. Ziemek, *Fundamentals of Acoustic Field Theory and Space-Time Signal Processing*. CRC Press, 1999.
- [18] G. V. Frisk, *Ocean and Seabed Acoustics*. Prentice-Hall, 1994.
- [19] A. Jarrot, C. Ioana, C. Gervaise, and A. Quinquis, "A time-frequency characterization framework for signals issued from underwater dispersive environments," in *Proceedings of IEEE International Conference on Acoustic, Speech and Signal Processing*, vol. 3, April 2007, pp. 1145–1148.
- [20] A. Jarrot, C. Ioana, and A. Quinquis, "A class of linear time-varying filters based on non-unitary time-warping operators. Part I: principle and examples," *submitted to IEEE Transactions on Signal Processing*, 2006.
- [21] G. B. Deane, "Internal friction and boundary conditions in lossy fluid seabeds," *The Journal of the Acoustical Society of America*, vol. 101, no. 1, pp. 233–240, January 1997.
- [22] J. Zhang and A. Papandreou-Suppappola, "Time-frequency based waveform and receiver design for shallow water communications," in *Proceedings of IEEE International Conference on Acoustic, Speech and Signal Processing*, vol. 3, April 2007, pp. 1149–1152.
- [23] M. Porter and A. Duncan, "Underwater acoustic propagation modelling software - actui version 1.6," <http://www.cmst.curtin.edu.au/products/actoolbox/>, June 2002.
- [24] É. Chassande-Mottin and A. Pai, "Best chirplet chain: Near-optimal detection of gravitational wave chirps," *Physical Review*, vol. 73, no. 042003, pp. 1–25, Feb. 2006.
- [25] A. Papandreou-Suppappola, "Time-varying processing: Tutorial on principles and practice," in *Applications in Time-Frequency Signal Processing (A. Papandreou-Suppappola, ed.)*. Florida: CRC Press, 2002.
- [26] A. L. Maggi and A. J. Duncan, "Underwater acoustic propagation modelling software," <http://www.cmst.curtin.edu.au/products/actoolbox>.
- [27] R. H. Ferris, "Comparison of measured and calculated normal-mode amplitude functions for acoustic waves in shallow water," *The Journal of the Acoustical Society of America*, vol. 52, no. 3B, pp. 981–989, 1972.
- [28] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall PTR, 1993.
- [29] V. Veeravalli, "On performance analysis for signaling on correlated fading channels," *IEEE Transactions on Communications*, vol. 49, no. 11, pp. 1879–1883, Nov. 2001.

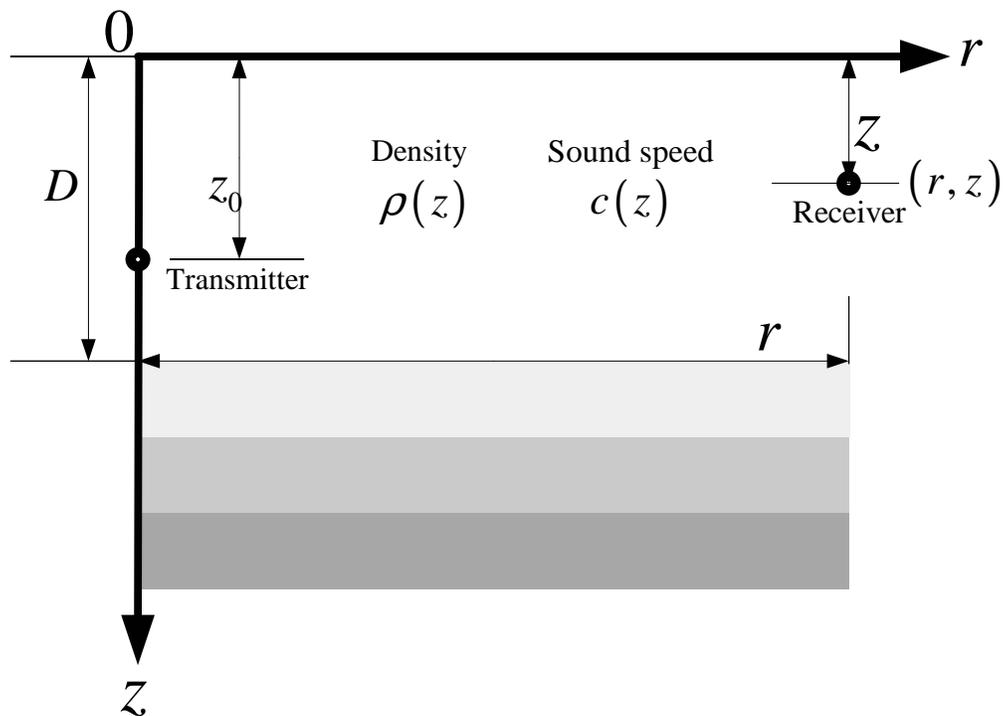


Fig. 1. Stratified acoustic medium model with a point source in the water column at range  $r = 0$  and depth  $z = z_0$ .

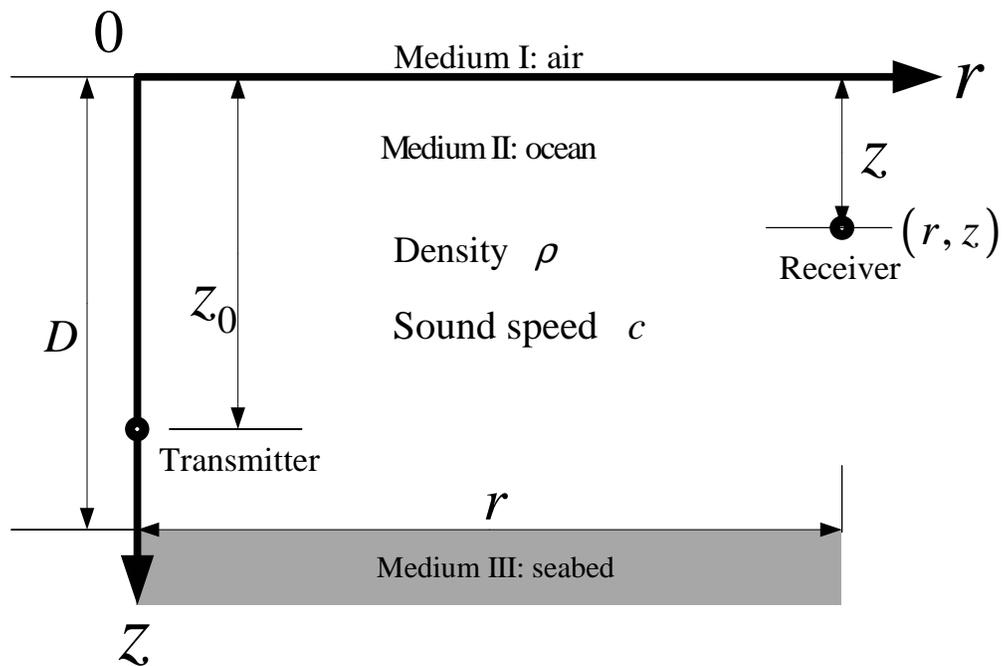


Fig. 2. Waveguide model with point source in Medium II at range  $r = 0$  depth  $z = z_0$ ; the ocean is  $D$  m deep.

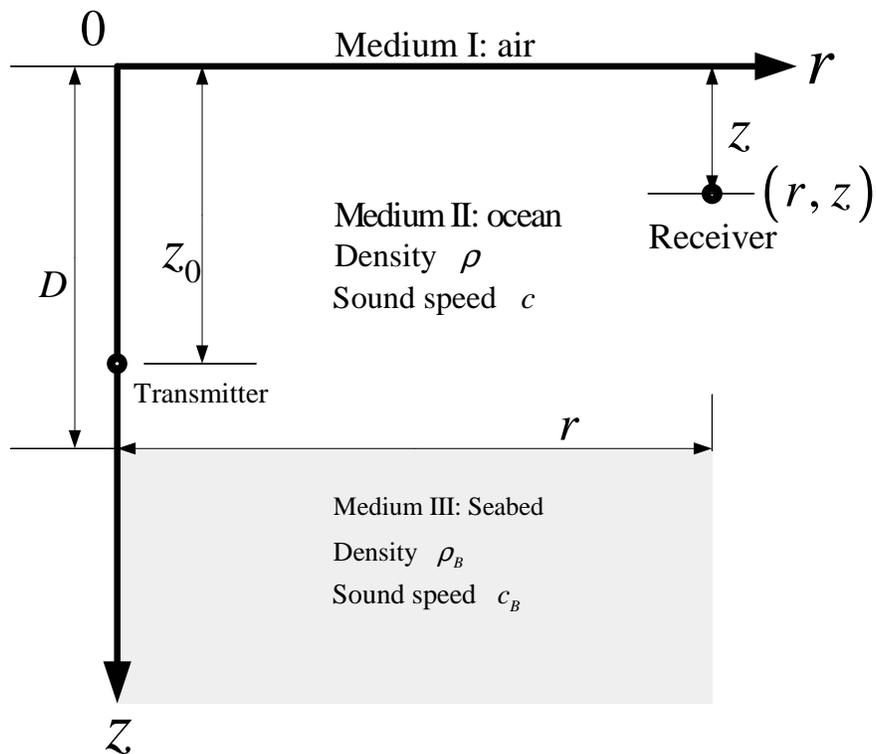


Fig. 3. Waveguide model with point source in Medium II at  $r = 0, z = z_0, D$  m deep. This corresponds to a pressure release surface with a fluid seabed environment [18].

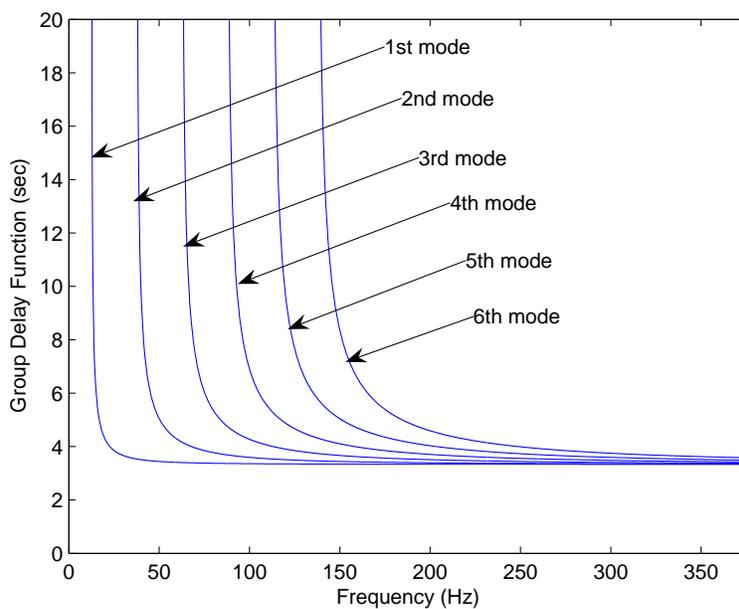


Fig. 4. The group delay function of the first 6 modes of the isovelocity model in (8).

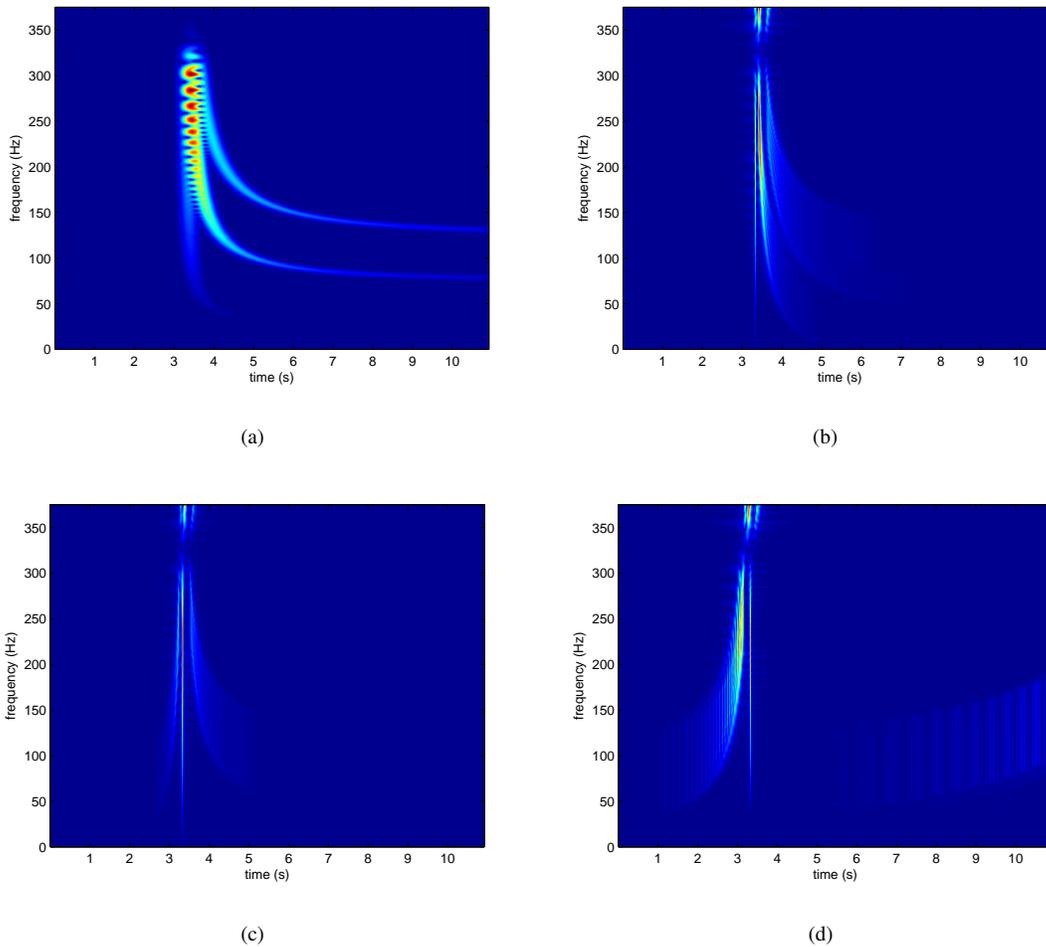


Fig. 5. Spectrogram of (a) a *KRAKEN* simulated received signal  $Y(f)$  using the isovelocity shallow water model in (3). Spectrograms of the warped signals (b)  $(\mathcal{U}_{\zeta_0} Y)(f)$ , (c)  $(\mathcal{U}_{\zeta_1} Y)(f)$ , and (d)  $(\mathcal{U}_{\zeta_2} Y)(f)$ .

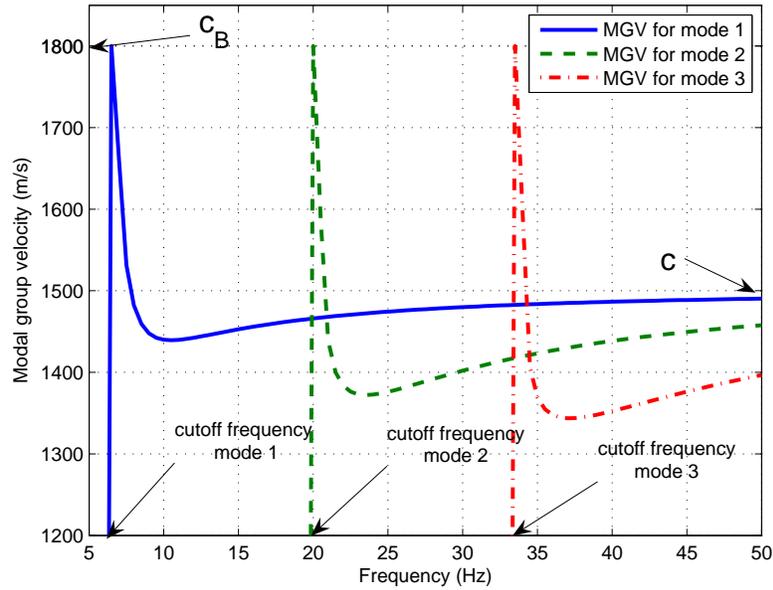


Fig. 6. Modal group velocity (MGV) function  $g_n(f)$  for the  $n$ th mode for the Pekeris model in (6). Here,  $c$  is 1,500 m/s and  $c_B$  is 1,800 m/s.

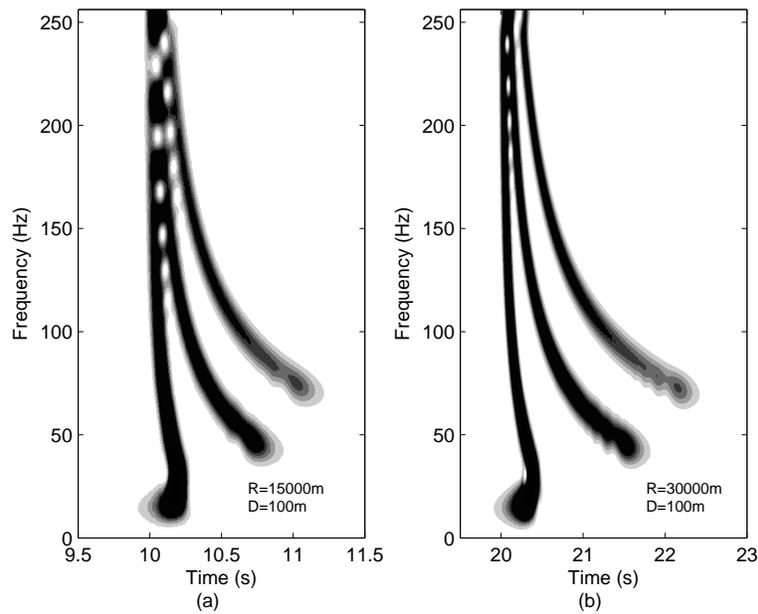


Fig. 7. Spectrograms of the received waveform using the Pekeris model in (5) with environment parameters (a)  $r = 15$  km,  $D = 100$  m, and (b)  $r = 30$  km,  $D = 100$  m.

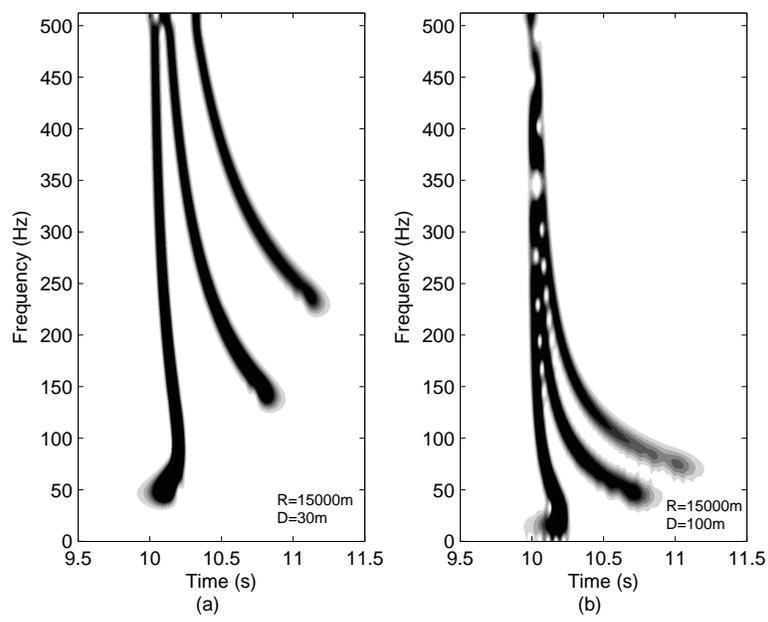


Fig. 8. Spectrograms of the received waveform using Pekeris model in (5) with environment parameters (a)  $r = 15$  km,  $D = 30$  m, and (b)  $r = 15$  km,  $D = 100$  m.

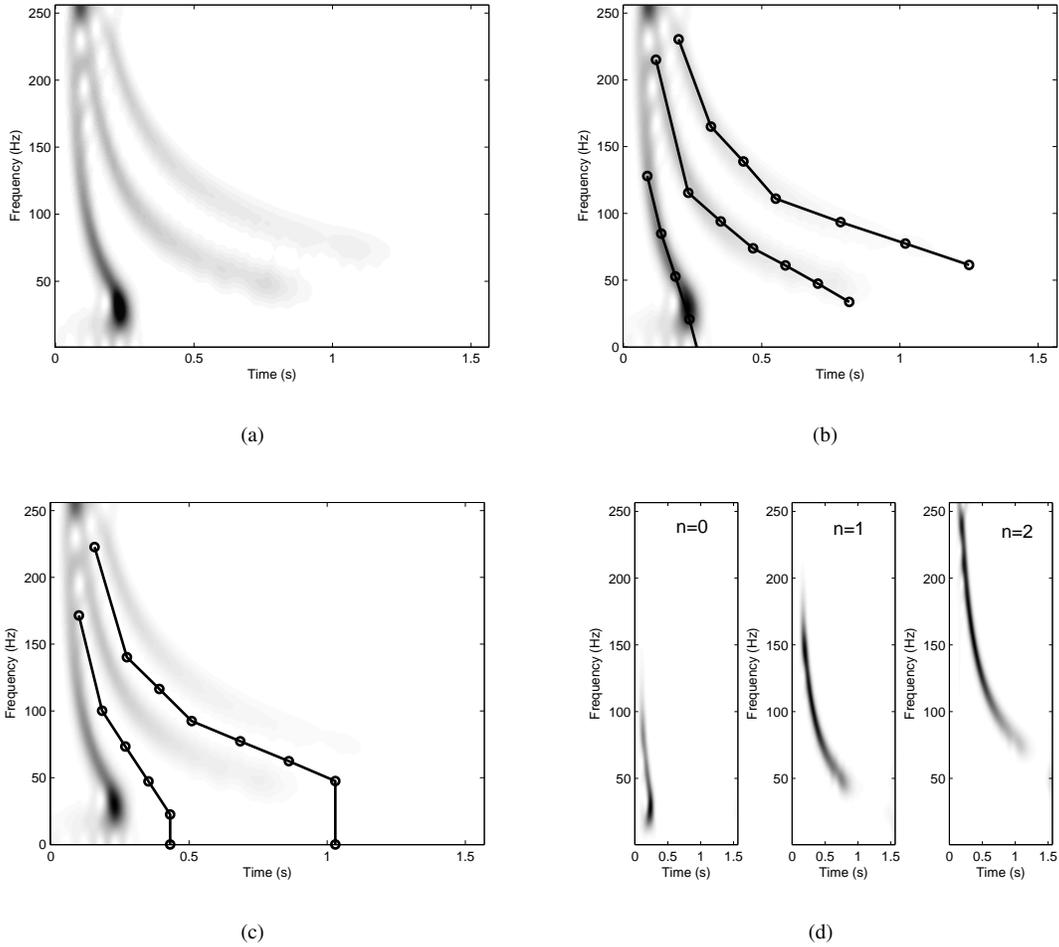


Fig. 9. (a) Spectrogram of the received signal following Pekeris model in (5). (b) Estimated TF signatures for  $N_m = 3$  modes using linear combinations of non-overlapping linear chirps. (c)  $N_m - 1 = 2$  GVF curve separators in the TF plane. (d) Separated TF mode components.

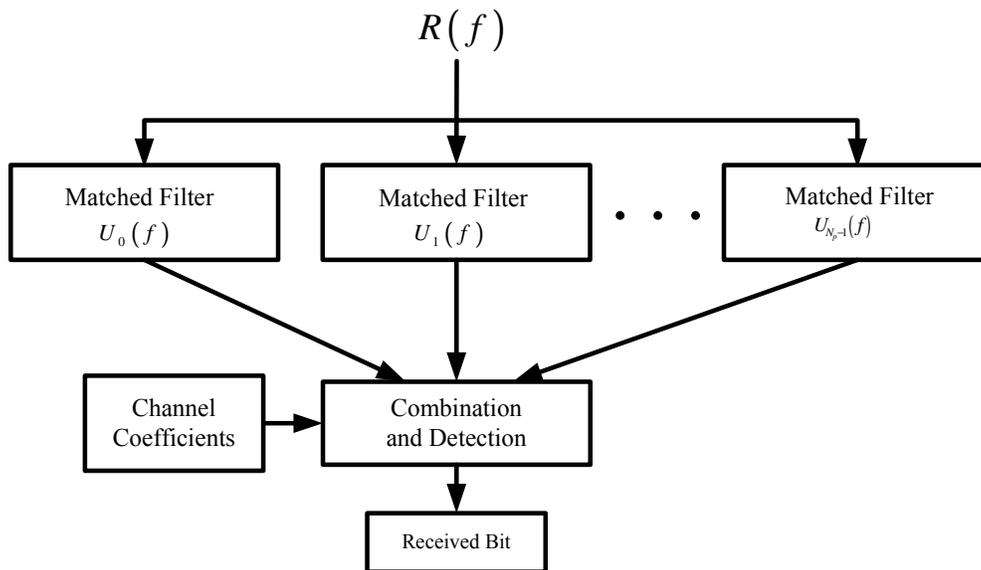


Fig. 10. A filter bank receiver scheme for the shallow water transmitted signal using a matched filter to obtain all the modes, and then the resulting matched filtered outputs are combined in the minimum-error-probability sense.

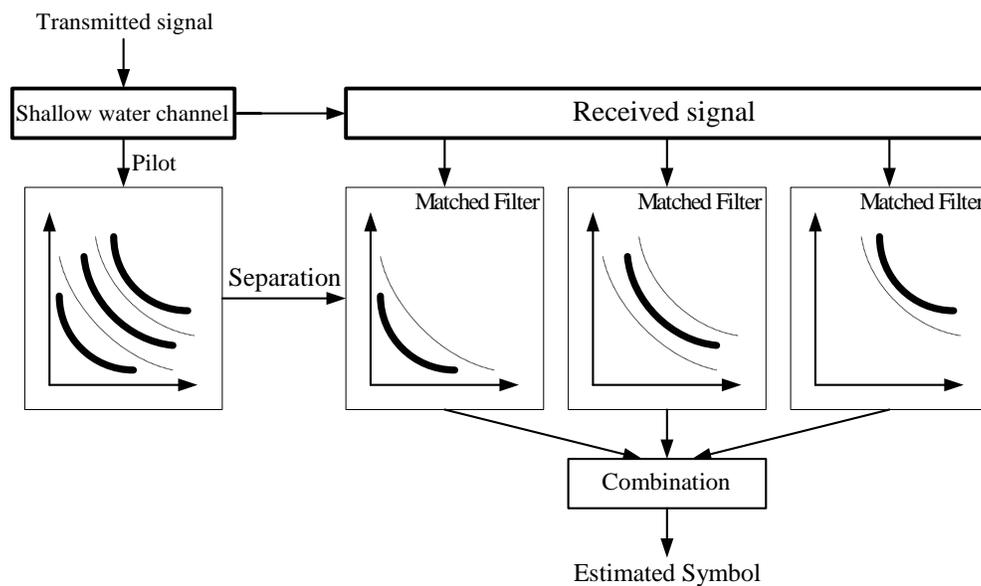


Fig. 11. A receiver design for a shallow water channel for based on TF mode separation to exploit the potential diversity.

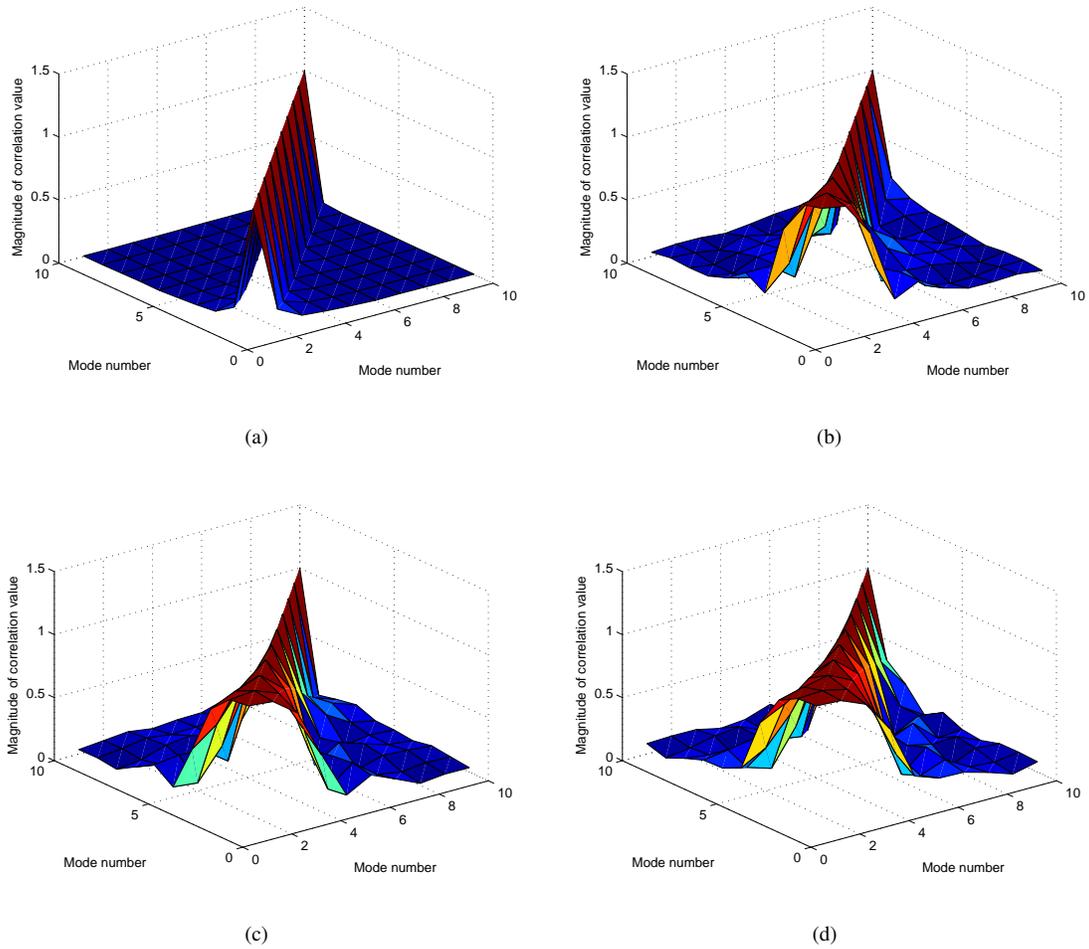


Fig. 12. Correlation matrices for the different modes for different frequency bands: (a) 1 kHz – 2 kHz, resulting correlation matrix rank is 10, (b) 5 kHz – 6 kHz, rank  $\approx 9$ , (c) 6 kHz – 7 kHz, rank  $\approx 8$ , (d) 7 kHz – 8 kHz, rank  $\approx 7$ .

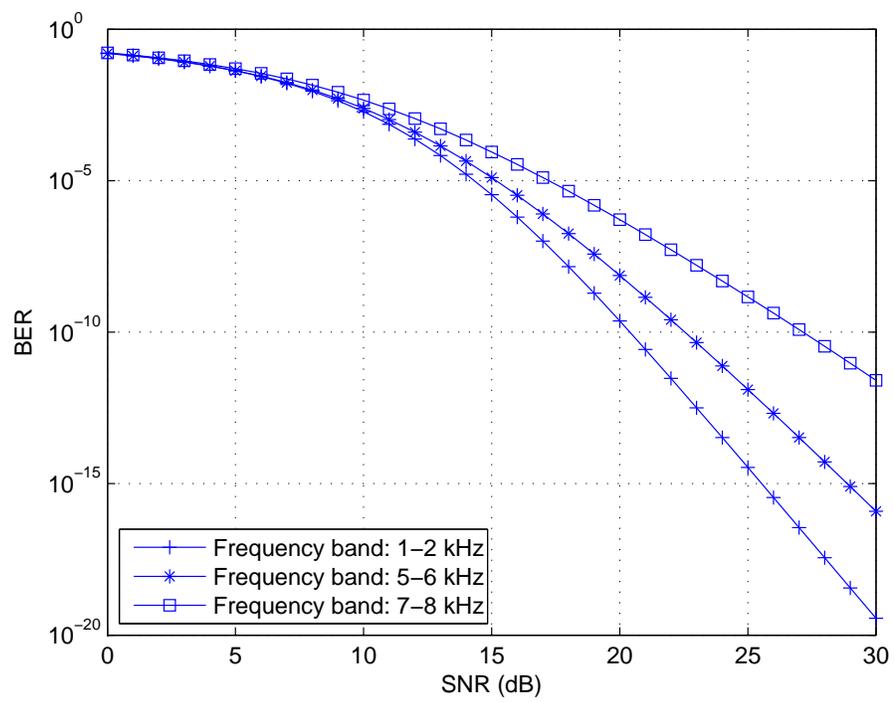


Fig. 13. BER and diversity performance in different transmission bands for the isovelocity model when TF mode separation is not used.

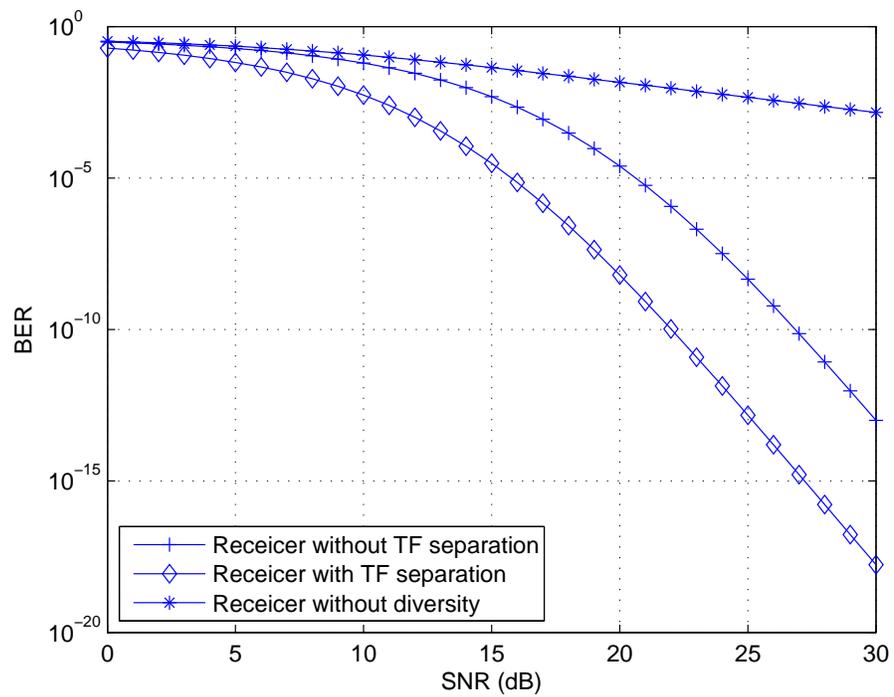


Fig. 14. BER and diversity performance for different receivers for the Pekeris model when TF mode separation is used. The frequency band used for these simulations is 0 – 500 Hz.