Modal group delay estimation for dispersive modes in broadband Arctic tomography

F. Hunter Akins, Peter F. Worcester *Fellow, IEEE*, Hanne Sagen, and Matthew A. Dzieciuch

***Abstract*—The ocean is a geometrically dispersive medium for acoustic waves due to the depth dependence of the index of refraction. For broadband signals, dispersion results in signal distortion that increases with the distance from the transmitter. This paper considers the estimation of modal group delay from broadband tomographic transmissions in the presence of distortion caused by dispersion. The received signal is characterized by delay and dispersiveness parameters which are both considered to be unknown. The modal group delay can be then estimated in a straightforward manner by maximizing the output of a matched filter over a set of replica wave forms. This is directly analogous to the common signal processing problem of joint delay and Doppler estimation. The replica energy is independent of the dispersiveness parameter, so for an additive white Gaussian noise model the maximum of the matched filter output is the maximum likelihood estimator of delay (and dispersiveness). An improved signal-to-noise ratio of approximately 6 dB at the receiver and unbiased modal group delay estimation is demonstrated on a 35 Hz broadband source in simulations and on data from the Coordinated Arctic Acoustic Thermometry Experiment (CAATEX).**

***Index Terms*—Acoustical oceanography, ocean remote sensing, sound propagation and scattering**

I. Introduction

OCEAN acoustic tomography measures the acoustic field from controlled signal transmissions to invert for the properties of the ocean along the transmission path. In most tomography experiments, time delay estimates corresponding to distinct arrivals are extracted from the signal and used in the inversion. Depending on the specific conditions of an experiment, these arrivals may be ray-like, in which case the arrivals correspond to distinct eigenrays, or modal, where the arrivals are associated with normal modes of the waveguide. For a broadband source, the delay may vary substantially over the source bandwidth due to waveguide dispersion. As a result, there is no longer a single group delay that applies to the transmitted wave form, and estimation of a single delay via peak picking or matched filtering is no longer justified. In order to resolve this ambiguity, this work presents a modal group delay estimation method relevant for ocean acoustic tomography in the Arctic Ocean in which the dispersion is characterized by a single unknown parameter. The unknown parameter is jointly estimated along with the modal group delay at a specified frequency.

F. H. Akins, P. F. Worcester, and M. A. Dzieciuch are with the Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA.

H. Sagen is with the Nansen Environmental and Remote Sensing Center, Bergen, Norway.

The normal mode representation of the acoustic field has been used extensively for geophysical inversion in underwater acoustics. The use of a derived normal mode property for inversion is in contrast with full-wave inversion [1], [2], which uses all possible information about the channel in the inversion. The latter approach comes at significant computational cost, the requirement of more sophisticated optimization methods, and precise knowledge of the transmitted wave form. In many cases, it is therefore expedient to first estimate a property of the normal mode field from the data and subsequently use this property in a simpler inversion. Depending on the experiment, this property may be the modal wave number for a single frequency, the modal group delay, the normal mode vertical shape, or the broadband modal dispersion. For example, using a sinusoidal source and a large horizontal array, the modal wave number can be extracted using the Hankel transform [3] and inverted using perturbative methods derived by Rajan, Frisk and Lynch [4]. The work reported here seeks to extract the modal group delay for the center frequency of a transmitted pulse of known shape for subsequent use in tomographic sound-speed profile inversion. In the presence of modal dispersion, however, the received wave form differs from the transmitted wave form, which complicates the estimation of modal group delay.

The acoustics community has developed a number of signal processing approaches to estimate a frequency-dependent modal group delay from measurements of a broadband impulsive source traveling in a dispersive waveguide. One such method is the dispersion-based short-time Fourier transform (D-STFT), which uses a known dispersion curve to define a transform with improved time-frequency tiling for dispersive signals when compared to the standard short-time Fourier transform [5]. The authors of the D-STFT also describe an adaptive, iterative implementation that alternately estimates the dispersion curve from the D-STFT output, then recomputes the D-STFT using the updated dispersion curve estimate until a convergence criterion is met. This adaptive implementation was sucessfully applied to the estimation of modal group delays of a combustive source in shallow water, with an added intermediate step in the loop that fit the delay estimates to a seabed model [6]. Another method, called warping, aims to isolate overlapping, dispersive modal arrivals from a broadband source on a single hydrophone, often for use in geoacoustic inversion [7]–[9]. Warping applies a transformation, determined by an approximate dispersion curve, to the data to map modal arrivals to either a horizontal (time-warping) or a vertical (frequency-warping) line in the time-frequency plane. The modal arrivals can be separated in a straightforward manner in the warped domain and then unwarped back. The techniques behind these methods, which were developed for the estimation of a dispersion curve with an impulsive source, bear similarities to the method presented here. In this tomographic application, the source wave form is well-known and the source bandwidth is relatively small in comparison with e.g. a combustive source. Therefore, the method presented here makes use of the more restrictive assumption that the dispersion of a given mode is well-approximated by a Taylor expansion about the transmitted pulse center frequency. Under this assumption, the received signal wave form is determined by two unknown, deterministic parameters: the group delay τ at the center frequency and a single dispersiveness parameter β. The delay estimation problem is therefore simply a parameter estimation problem analogous to the classical problem of delay estimation in the presence of unknown target Doppler, and the maximum likelihood estimate is obtained by comparing the outputs of a bank of matched filters corresponding to the possible received wave forms [10].

The deterministic, parametric approach presented here takes advantage of the relatively low internal wave energy level in the Arctic Ocean, which is a consequence of its low stratification and minimal wind-driven surface forcing [11]. In contrast, mid-latitude tomography must contend with internal-wave-induced scattering, which randomizes the modal arrivals and spreads them in time except at very low acoustic frequencies [12], [13]. In that case, a stochastic approach is better suited to the tomographic modal group delay estimation problem [14].

After developing the signal processing procedure described in this manuscript, the authors discovered that Poplawski and Birdsall had independently developed a similar signal processing procedure some years ago in the context of the Acoustic Thermometry of Ocean Climate (ATOC) project [15], [16]. This work was never published in the peer-reviewed literature. Further, the ATOC transmissions occurred in the mid-latitude North Pacific Ocean where internal-wave-induced scattering is important, and dispersion removal alone would not have removed the spread. The dispersion correction procedure was never applied to the long-range propagation data obtained during the ATOC project.

This work is organized as follows. Section II describes broadband modal propagation in a dispersive environment. Section III introduces the signal model and the group delay estimator. Section IV illustrates the method on a synthetic modal pulse in a representative Arctic sound-speed profile. Section V applies the method to mode-filtered transmissions from the Coordinated Arctic Acoustic Thermometry Experiment (CAATEX). Concluding remarks follow in Section VI.

II. Modal propagation

The pressure field produced by a point source of frequency ω located at a depth z_s on the z-axis of a cylindrical coordinate system in an azimuthally symmetric, constant density acoustic waveguide with sound speed $c(r,z)$ satisfies the Helmholtz equation

$$\begin{aligned}\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{\partial^2}{\partial z^2}+\frac{\omega^2}{c^2(r,z)}\right)&g_\omega(r,z;z_s)\\ &=\frac{-\delta(r)\delta(z-z_s)}{2\pi r}\end{aligned} \tag{1}$$

along with boundary conditions at the surface and seafloor. For a sound-speed profile that is independent of range and azimuth, the solution in the far field can be expressed as a discrete sum of normal modes

$$g_\omega(r,z;z_s)=\frac{e^{-j\pi/4}}{\sqrt{8\pi}}\sum_m \Psi_m(z_s)\Psi_m(z)\frac{e^{-jk_{rm}r}}{\sqrt{k_{rm}r}} \tag{2}$$

where the normal mode depth functions Ψ_m satisfy the depth-separated equation in terms of the corresponding horizontal wave number k_{rm}

$$\frac{\mathrm{d}^2\Psi_m}{\mathrm{d}z^2}+\left(\frac{\omega^2}{c^2(z)}-k_{rm}^2\right)\Psi_m(z)=0 \tag{3}$$

with boundary conditions at $z=0$ and at the seafloor $z=Z_{sb}$. In this work, the surface is assumed to be pressure release ($\Psi_m(0)=0$), and the bottom is assumed to be a fluid halfspace with speed $c_b=1600$ m/s and $\rho_b=2.0$ g/cm^3. The sound-speed profile $c(z)$ is specified as the linear interpolation of the profile at a set of sample depths that span the water column. The horizontal wave numbers and normal mode functions Ψ_m are then computed numerically by solving (3) using the methods in the code KRAKEN [17].

The adiabatic approximation extends the range-independent solution (2) to slowly varying range-dependent environments common in underwater propagation [18]. The adiabatic solution is

$$g_\omega(r,z;z_s)=\frac{e^{-j\pi/4}}{\sqrt{8\pi}}\sum_m \Psi_m(z_s;0)\Psi_m(z;r)\frac{e^{-j\bar{k}_{rm}r}}{\sqrt{\bar{k}_{rm}r}} \tag{4}$$

where $\Psi_m(z;r)$ and $k_{rm}(r)$ denote the solution to (3) for the sound-speed profile $c(r,z)$ at a given range r, and $\bar{k}_{rm}=\frac{1}{r}\int_0^r k_{rm}(r')dr'$ denotes the average modal wavenumber along the transmission path. The following analysis uses the range-independent notation in (2); however, an adiabatic range-dependent interpretation is obtained simply by replacing the modal wavenumber $k_{rm}(\omega)$ with the range-averaged wavenumber $\bar{k}_{rm}(\omega)$ and the mode function Ψ_m with the local mode function at either the source or receiver location. The method developed in this paper is not applicable to range-dependent environments with mode coupling.

The field produced by a broadband source at $r=0$ and $z=z_s$ with wave form $s(t)$ can be computed using the Helmholz equation solution $g_\omega(r,z;z_s)$ by the inverse Fourier transform

$$p(r,z,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}S(\omega)g_\omega(r,z;z_s)e^{j\omega t}d\omega, \tag{5}$$

where $S(\omega)$ is the Fourier transform of the source wave form. Inserting the modal solution (2) above and exchanging summation and integration, the field is seen to be a sum of modal time series

$$p(r,z,t)=\sum_m p_m(r,z,t) \tag{6}$$

where each modal time series is given in terms of the modal transfer function

$$H_m(r,z,\omega) = \frac{e^{-j\pi/4}}{\sqrt{8\pi k_{rm} r}} \Psi_m(z_s)\Psi_m(z) e^{-jk_{rm}(\omega)r} \quad (7)$$

by

$$p_m(r,z,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega) H_m(r,z,\omega) e^{j\omega t} d\omega. \quad (8)$$

The explicit dependence of H_m and p_m on the source depth z_s is suppressed for notational convenience.

Equation (8) produces the received modal wave form in terms of the modal transfer function $H_m(r,z,\omega)$. To gain some intuition into the effect of modal propagation on the received wave form, first assume that $|H_m(r,z,\omega)|$ is constant over the source bandwidth and the horizontal wave number is linear in frequency. For a linear horizontal wave number $k_{rm}(\omega) = s_m \cdot (\omega - \omega_0)$ (constant s_m) and constant magnitude $|H_m(r,z)|$, the modal transfer function is linear phase with slope s_m. Therefore, $p_m(r,z,t)$ is simply a scaled and time-delayed copy of the transmitted wave form

$$p_m(r,z,t) \propto s(t - s_m r). \quad (9)$$

The wave number slope s_m is the modal slowness, and $\tau_m = s_m r$ is the modal delay time, which in this case is independent of frequency. The delay τ_m can therefore be estimated by matched filtering the received signal with the transmitted wave form.

In many propagation cases of interest, it is reasonable to assume that $|H_m(r,z,\omega)|$ varies slowly over the source bandwidth and the horizontal wave number is non-linear in frequency. At a large range $r \gg 0$, for most times t the phase in the modal transfer function, and therefore the phase in (8), varies rapidly with ω and the received pressure field is very small (approaching 0 for $r \to \infty$ by the Riemann-Lebesgue lemma). The exception to this occurs at times $\tau_m(\omega) = \frac{\mathrm{d}k_{rm}}{\mathrm{d}\omega} r$ for ω in the source bandwidth. At such times, the integral in (8) can be approximated using the method of stationary phase [19], [20]:

$$p_m(r,z,\tau_m(\omega)) \approx \frac{|H_m(r,z,\omega)| S(\omega) e^{-j\left(\mathrm{sign}\left(\frac{\mathrm{d}^2 k_{rm}}{\mathrm{d}\omega^2}\right)\pi/4\right)}}{2\pi\sqrt{r\left|\frac{\mathrm{d}^2 k_{rm}}{\mathrm{d}\omega^2}\right|}} e^{j(\omega\tau_m(\omega) - k_{rm}(\omega)r)}. \quad (10)$$

The quantity $s_m(\omega) \equiv \frac{\mathrm{d}k_{rm}}{\mathrm{d}\omega}$ is again the modal group slowness, which now depends on frequency ω due to the assumed non-linearity of $k_{rm}(\omega)$. Each frequency in the transmitted signal arrives at its own time $\tau_m(\omega) = s_m(\omega) r$, resulting in dispersion of the signal. The degree of dispersion is quantified by the derivative of the delay with respect to frequency

$$\frac{\mathrm{d}\tau}{\mathrm{d}\omega} = \frac{\mathrm{d}^2 k_{rm}}{\mathrm{d}\omega^2} r, \quad (11)$$

and increases with both range and the non-linearity of the wave number.

From the preceding discussion, we can identify the waveguide "dispersiveness" for a single modal time series with the function $\frac{\mathrm{d}^2 k_{rm}}{\mathrm{d}\omega^2}$, or, equivalently, $\frac{\mathrm{d}s_m}{\mathrm{d}\omega}$. Also note that from (10), the received amplitude is proportional to the inverse of the square root of this quantity, so that signal-to-noise ratio degrades with increasing dispersion. The received wave form is no longer a delayed copy of the transmitted wave form, and there is not a single group delay that applies to the entire received wave form.

III. Group delay estimation in a dispersive waveguide

This work is interested in the estimation of the modal group delay $\tau_m(\omega_0) = \frac{\mathrm{d}k_{rm}}{\mathrm{d}\omega}\big|_{\omega_0} r$ at the frequency ω_0. Assume that we have isolated a single modal time series $p_m(t)$, for example by windowing the received time series on a time-separated modal arrival or by mode filtering with a vertical array. The measured time series is corrupted with additive white Gaussian noise $n(t)$

$$y_m(t) = p_m(t) + n(t). \quad (12)$$

Also assume that the modal amplitude is approximately independent of frequency, so that $p_m(t)$ has Fourier transform

$$P_m(\omega) \propto S(\omega) e^{-jk_{rm}(\omega)r}. \quad (13)$$

In order to estimate the modal group delay, Taylor expand $k_{rm}(\omega)$ about ω_0 and retain terms to second order

$$k_{rm}(\omega) = k_{rm}(\omega_0) + s_m(\omega - \omega_0) + \beta_m(\omega - \omega_0)^2 \quad (14)$$

where $s_m = \frac{\mathrm{d}k_{rm}}{\mathrm{d}\omega}\big|_{\omega_0}$ is the slowness and $\beta_m = \frac{1}{2}\left.\frac{\mathrm{d}^2 k_{rm}}{\mathrm{d}\omega^2}\right|_{\omega_0}$ is the dispersiveness parameter. This second order truncation will not be valid for a very broadband source, but is reasonable for the source considered in this work. Inserting this into (13), we obtain

$$P_m(\omega) \propto S(\omega) e^{-j[k_{rm}(\omega_0) + s_m(\omega-\omega_0) + \beta_m(\omega-\omega_0)^2]r}, \quad (15)$$

which is seen to be equivalent to a time delayed copy of a distorted wave form

$$p_m(t) = \tilde{s}(t - \tau_m; \beta_m) \quad (16)$$

where the Fourier transform of the distorted wave form $\tilde{s}$ is

$$\tilde{S}(\omega) = A S(\omega) e^{-j\beta_m(\omega-\omega_0)^2 r}. \quad (17)$$

A is a complex constant determined by the modal amplitude and phase, and the delay is $\tau_m = s_m r$. The measured time series is therefore

$$y_m(t) = \tilde{s}(t - \tau_m; \beta_m) + n(t), \quad (18)$$

from which an estimate of τ_m is sought.

In this work, τ_m and β_m are jointly estimated with maximum likelihood estimation. It is shown in Appendix A that for additive white Gaussian noise this is equivalent to maximizing the modulus of the matched filter output $|R(\tau,\beta)|$

$$|R(\tau,\beta)| = \left| \int_{-\infty}^{\infty} y_m(t) \tilde{s}^*(t - \tau; \beta) dt \right| \quad (19)$$

over τ and β.

Alternatively, one can view the proposed method as first applying an all-pass filter with phase $e^{j\beta(\omega-\omega_0)^2 r}$ in order to correct the received data before matched filtering with the original source wave form. To see this, write the matched filter in the frequency domain

$$R(\tau,\beta)=\frac{1}{2\pi}\int_{-\infty}^{\infty}Y_m(\omega)\tilde{S}^*(\omega;\beta)e^{j\omega\tau}d\omega, \tag{20}$$

which can be rewritten using (17)

$$R(\tau_m,\beta)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(Y_m(\omega)e^{j\beta(\omega-\omega_0)^2r}\right)S^*(\omega)e^{j\omega\tau_m}d\omega. \tag{21}$$

The term $\left(Y_m(\omega)e^{j\beta(\omega-\omega_0)^2r}\right)$ is the all-pass filtered received data in the frequency domain. This is similar to the mode reversal method presented by Bonnel et al. [21], [22] which applies an all-pass filter with phase $e^{j\hat{k}_{rm}(\omega)r}$ to isolated modal time series. It also bears similarities to the D-STFT method, which adds a quadratic phase term to the STFT, though the analog to β in the D-STFT is a function of both time and frequency in the transformation.

In this section the modal amplitude was assumed to be constant over frequency, which is a good approximation for the results presented in this work. Appendix B provides the matched filter in the more general case where the modal amplitude varies over frequency. It also shows that the simpler matched filter (21) will produce an unbiased delay estimate as long as the modal amplitude does not change sign over the source bandwidth.

IV. Simulation

The modal group delay estimation method of the previous section is now applied to synthetic data representative of transmissions made during CAATEX. During the experiment, sources transmitted 35 Hz broadband signals which were received at a number of moorings. The simulation is designed to represent a ~877 km transmission path across the Canada Basin from the transceiver mooring SIO1 to to the receiver mooring SIO3. The sound-speed profile used in the simulation, typical of the Canada Basin, is shown in Fig. 1. The corresponding modal dispersion over the source bandwidth along the path from SIO1 to SIO3 is shown in Fig. 2. The simulation focuses on mode 3 because it is the most dispersive mode along this transmission path.

The transducer in the experiment is modeled as the output of a linear damped harmonic oscillator with resonant frequency $f_r = 35$ Hz and Q factor 8.0 driven by a sinusoid at $f_0 = 35$ Hz for 8 cycles. We generate 16384 samples of the source wave form at an 1120 Hz sample rate (14.63 s duration). The transmitted signal is shown in Fig. 3, along with the envelope obtained by complex basebanding the wave form at 35 Hz and low-pass filtering. The bandwidth of the signal, defined as f_0/Q, is $\sim$ 4 Hz. The source wave form is advanced in order to align the envelope peak at the time origin.

To model single mode broadband pulse propagation, the depth-separated wave equation (3) is solved for the horizontal wave numbers $k_{rm}(f_i)$ and the modal functions $\Psi_m(z,f_i)$

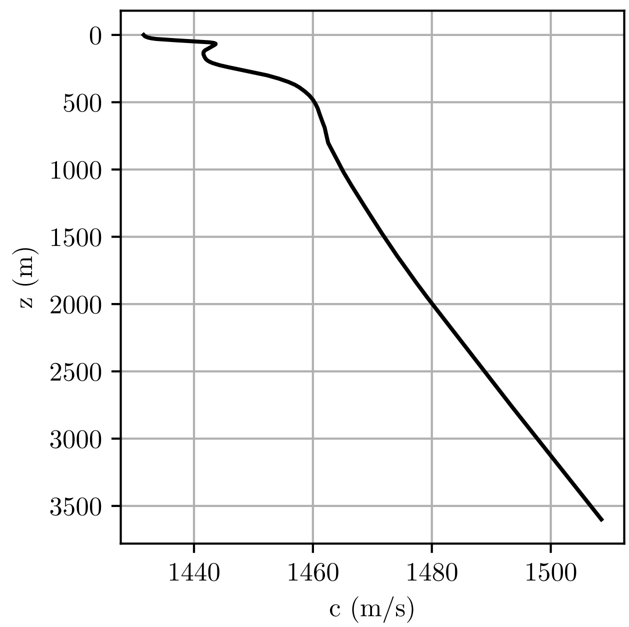

Fig. 1. Canada Basin sound-speed profile used for simulation.

on a discrete grid of frequencies. A single modal pulse is synthesized using an inverse discrete Fourier transform (DFT)

$$r_m[t_k]= C\sum_{i=0}^{N-1}\frac{S[f_i]\Psi_m(z_s;f_i)\Psi_m(z_r;f_i)}{\sqrt{R_s k_{rm}[f_i]}}e^{-jk_{rm}[f_i]R_s}e^{j2\pi f_i t_k}, \tag{22}$$

where $S[f_i]$ is the DFT of the transmitted waveform, $R_s = 877$ km is the source-receiver range, $z_s = 60$ m, and $z_r = 100$ m. The constant $C = \frac{\sqrt{R_s}}{N}\times[\max_{f_i}\Psi_m(z_s;f_i)\Psi_m(z_r;f_i)/\sqrt{k_{rm}[f_i]}]^{-1}$ was chosen so that the propagation is approximately unity gain.

To test the group delay estimator, the matched filtered output $R(\tau,\beta)$ in (19) (equivalently (21)) is computed for mode 3 using the known source wave form and over a range of values of β. The maximum likelihood estimates of $\hat{\tau}_m$ and $\hat{\beta}_m$ are the values which maximize $|R(\tau,\beta)|$. For comparison with the estimates, the correct values of $\tau_m = \left.\frac{\mathrm{d}k_{rm}}{\mathrm{d}\omega}\right|_{\omega_0} R_s$ and $\beta_m = \frac{1}{2}\left.\frac{\mathrm{d}^2k_{rm}}{\mathrm{d}\omega^2}\right|_{\omega_0}$ at $\omega_0 = 2\pi\times 35$ s^{-1} are computed using a centered finite difference approximation. A comparison of the matched filter output $|R(\tau,\beta)|$ using no correction ($\beta = 0$) and using the correct dispersiveness value $\beta = \beta_m$ is shown in Fig. 4. Using the correct value of β, the matched filter output magnitude increases by a factor of 2 over the naive matched filter ($\beta = 0$), corresponding to an increase of 6 dB signal-to-noise ratio. Additionally, the naive matched filter output has a small bias of approximately 30 ms due to mismatch between the received wave form and the source wave form.

In real data, the parameter β_m is not known *a priori* and must be jointly estimated along with the group delay τ_m. The matched filter output using a range of values of $\beta = 0.0$ to

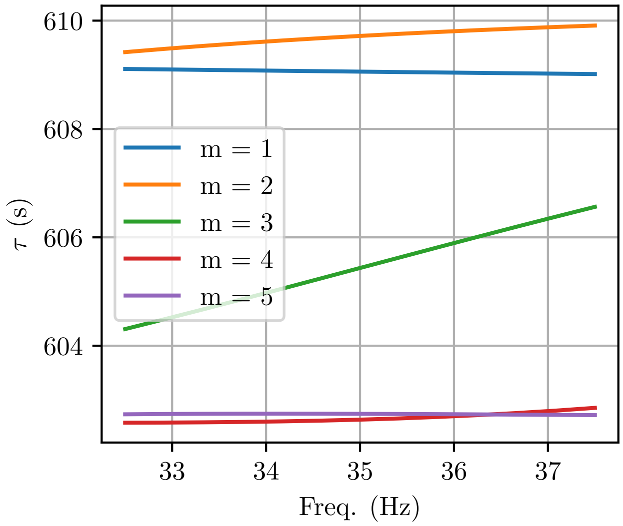

Fig. 2. Modal group delay as a function of frequency at a range of 877 km for the sound-speed profile used in the simulation results (shown in Fig. 1).

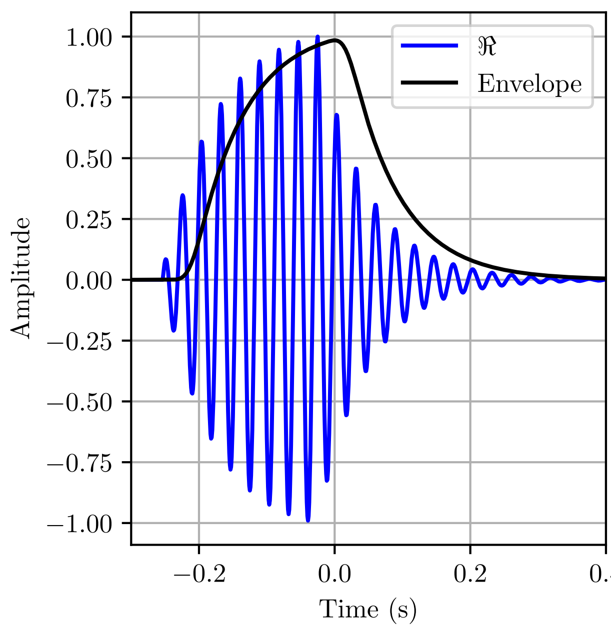

Fig. 3. Simulated transmitted wave form (blue) and envelope after demodulation (black). The peak of the envelope, which has a small filter delay, is aligned with $t = 0.0$.

$\beta = 2\beta_m$ is shown in Fig. 5. It is seen that the maximum matched filter output occurs for the correct value of $\beta = \beta_m$ and the correct group delay $\tau = \tau_m$. The slight tilt of the cost function surface with respect to β results in the biased delay estimate for the uncorrected wave form ($\beta = 0$).

V. Data

We now apply the method to estimation of modal group delay to measured transmissions from CAATEX. We look at transmissions from the transceiver moorings SIO1 and

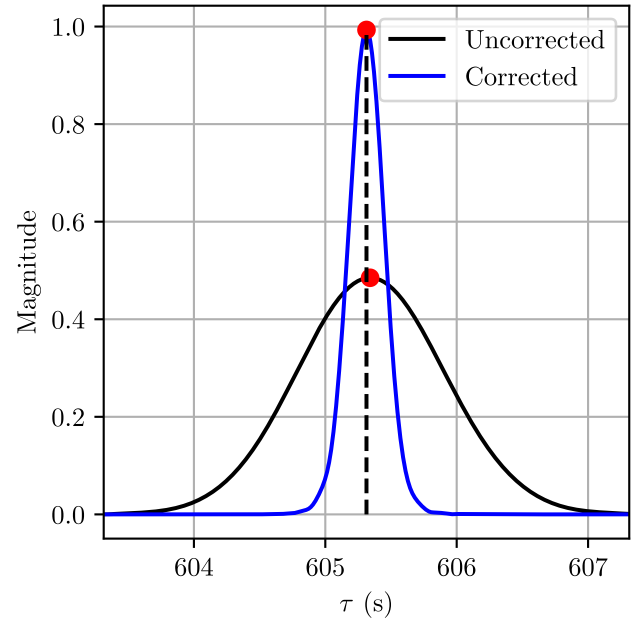

Fig. 4. Matched filter output for the simulated mode 3 wave form with (blue line) and without (black line) the dispersion correction. The group delay at the center frequency 35 Hz is shown as a dashed black line. The peak of the matched filter output without accounting for dispersion is biased from the true value by around 30 ms. The input and replica signal wave forms have been normalized to unit energy, so that the matched filter output magnitude can be interpreted as a correlation coefficient. The peak of the blue curve is the maximum likelihood group delay estimate.

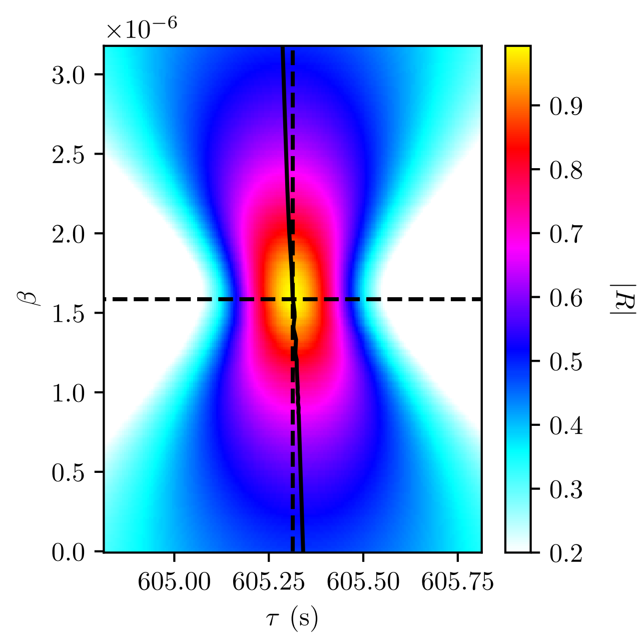

Fig. 5. Simulated matched filter output contoured for different values of dispersiveness β and modal group delay τ at the center frequency 35 Hz. Dashed black lines indicate the true values β_m and τ_m. Solid black line shows maximum likelihood delay estimate for each β. Note the bias in this estimate for incorrect choice of β.

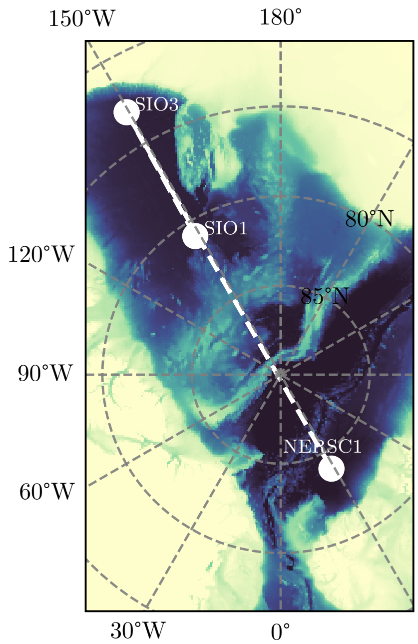

Fig. 6. The locations of the transceiver moorings SIO1 and NERSC1 and the receiver mooring SIO3, along with geodesics (white dashed lines) from SIO1 to SIO3 and from NERSC1 to SIO3. The colorscale indicates water depth, with a maximum value of 4000 m.

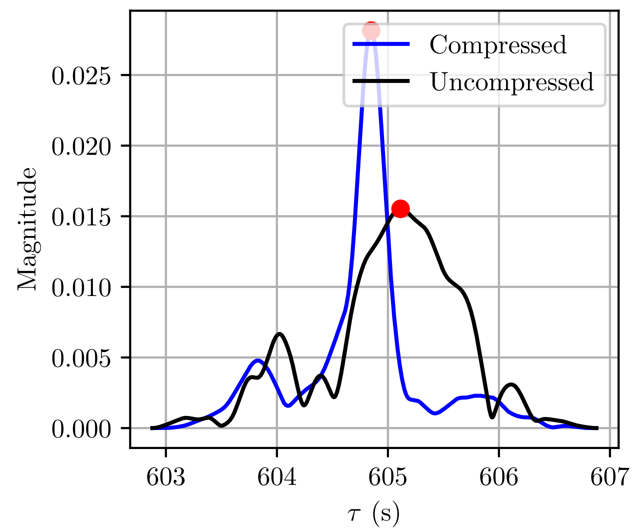

Fig. 7. Comparisons of dispersion corrected wave form to uncorrected wave form from a transmission during CAATEX from SIO1 to SIO3. The blue wave form corresponds to the matched filter output magnitude with the replica wave form that uses the maximum likelihood estimate of β. The black wave form corresponds to the matched filter output magnitude using the transmitted signal wave form.

NERSC1 measured on the receiver mooring SIO3. For each source, results are presented for the most dispersive mode, which is mode 3 for SIO1 to SIO3 and mode 2 for NERSC1 to SIO3 as predicted by a reference sound-speed profile for the path obtained from the World Ocean Atlas [23]. Although results for the most dispersive modes are presented here, modal group delays for all resolvable modes would be used in inversions for the ocean sound-speed field. The three moorings are shown in Fig. 6. The transmission paths from SIO1 to SIO3 and from NERSC1 to SIO3 have lengths of 877 km and 2566 km, respectively. The sources on the moorings SIO1 and NERSC1 were at depths of about 60 m. The mooring SIO3 has 40 hydrophones at 30-m spacing from $\sim$ 50 m depth to $\sim$ 1200 m depth.

Modal time series are obtained by first windowing $\pm$ 2 s around the expected arrival times with a Tukey window ($\alpha = 0.5$ as defined in Harris [24]). A spatial mode filter is then applied to the windowed data to obtain gain over noise and reject any other modes received in the window. A simple linear mode filter using the sampled mode shapes is used [25], for which the modal time series is

$$y_m[t_k] = \sum_{n=1}^{N_r} \Psi_m[z_n] y_n[t_k] \tag{23}$$

where n denotes the hydrophone index, $N_r = 40$ is the number of hydrophones, $\Psi_m[z_n]$ is the mode shape (at the receiver location) sampled at the array depths, and $y_n[t_k]$ is the sampled time series of the nth hydrophone. Windowing the reception in time prior to spatial filtering avoids crosstalk from high order modes that are not well resolved by the array [12]. Local temperature measurements made on the receiver mooring at the hydrophone depths were used to estimate the local sound-speed profile, which was subsequently used to compute the mode shape Ψ_m required for the mode filter in (23). The mode filters for modes 2 and 3, analyzed in this section, are insensitive to the exact sound-speed profile used and have no sensitivity to the sound-speed below a depth of $\sim$ 800 and $\sim$ 1000 m, respectively.

As in the simulation results, the group delay and dispersiveness parameter β are estimated from the mode filtered time series by maximizing the matched filter modulus over τ and β using a simple grid search. Fig. 7 shows matched filter outputs $|R(\tau, \beta)|$ for mode 3 from SIO1 to SIO3 at $\beta = 0$ (naive delay estimation) and at the maximum likelihood estimate $\hat{\beta}$. The proposed method increases the peak magnitude by nearly a factor of 2, i.e. 6 dB, and the difference in the matched filter peak output location is on the order of a few hundred milliseconds.

We also compare the performance of the proposed delay estimator over the roughly one year experiment duration. For multiple transmissions, a variety of joint estimation approaches may be taken that incorporate prior correlations on the delay τ and dispersiveness β. A simple approach is taken here: the dispersiveness β is assumed to be constant over the experiment, and no correlation of τ is assumed from one transmission to the next. With these assumptions, the maximum likelihood estimate $\hat{\beta}$, applicable to the full experiment, is obtained by maximizing the sum of the squared matched filter peak output over all transmissions. We compare the matched filter output power for all transmissions at $\beta = 0$ and $\beta = \hat{\beta}$ on mode 3 from SIO1 to SIO3 and mode 2 from NERSC1 to SIO3, shown in Fig. 8 and Fig. 9, respectively. The dispersion correction

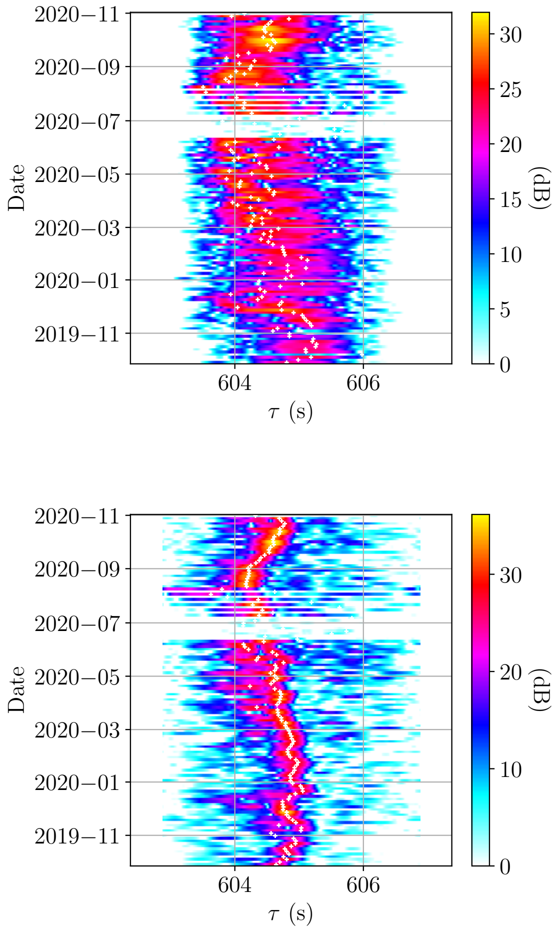

Fig. 8. Matched filter output power $|R(\tau,\beta)|^2$ in decibels for Mode 3 receptions from SIO1 to SIO3 for $\beta = 0$ (top) and using the estimate $\hat{\beta}$ that maximizes the peak matched filter output power averaged over the whole experiment period (bottom). The delay estimate for each transmission is shown as a white cross. The gap in receptions in June and July is due to a source malfunction. The colorscale is normalized to be interpreted as signal-to-noise ratio.

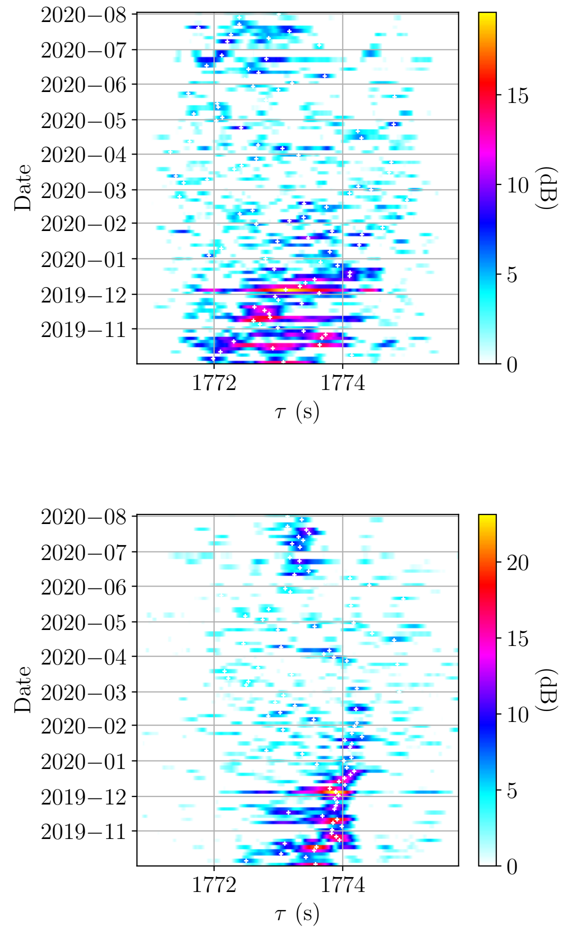

Fig. 9. Matched filter output power $|R(\tau,\beta)|^2$ in decibels for Mode 2 receptions from NERSC1 to SIO3 for $\beta = 0$ (top) and using the estimate $\hat{\beta}$ that maximizes the peak matched filter output power averaged over the whole experiment period (bottom). The delay estimate for each transmission is shown as a white cross. The colorscale is normalized to be interpreted as signal-to-noise ratio.

sharpens up the received signal peak and increases the power of the matched filter output by approximately 6 dB. The delay estimates for the estimated β value demonstrate much greater correlation from one transmission to the next and also lower overall variance when compared to the delay estimates using the transmitted source wave form.

VI. Conclusion

This work models the received modal wave form from a known tomographic signal in terms of a delay τ and dispersiveness β. Both τ and β are treated as deterministic unknown parameters, and can therefore be estimated in a straightforward manner from a measured signal by maximum likelihood. The maximum likelihood estimate is obtained by matched filtering the received signal with a bank of replicas corresponding to the possible values of the dispersiveness β. This is analogous to the joint estimation of delay and Doppler common in radar. The method was applied to both simulation and acoustic tomography data from CAATEX. The simulation indicated the presence of a small bias (~ 30 ms) in group delay estimation that does not consider the modal dispersion. On data, the method provides a nearly 6 dB increase in matched filter output power and a dramatic improvement in the estimation of mode 3 group delay across the Canada Basin and mode 2 group delay along a trans-Arctic path.

Acknowledgments

This material is based upon work supported by the Research Council of Norway Award No. 280531 and the Office of Naval Research under Award No. N00014-18-1-2698. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the Office of Naval Research.

References

[1] Y.-T. Lin, J. Bonnel, D. P. Knobles, and P. S. Wilson, "Broadband waveform geoacoustic inversions with absolute travel time," *IEEE Journal of Oceanic Engineering*, vol. 45, no. 1, pp. 174–188, 2020.

[2] J. Virieux, A. Asnaashari, R. Brossier, L. Métivier, A. Ribodetti, and W. Zhou, *6. An introduction to full waveform inversion*, 2017, pp. R1–1–R1–40. [Online]. Available: https://library.seg.org/doi/abs/10.1190/1.9781560803027.entry6

[3] G. V. Frisk and J. F. Lynch, "Shallow water waveguide characterization using the Hankel transform," *The Journal of the Acoustical Society of America*, vol. 76, no. 1, pp. 205–216, 07 1984. [Online]. Available: https://doi.org/10.1121/1.391098

[4] S. D. Rajan, J. F. Lynch, and G. V. Frisk, "Perturbative inversion methods for obtaining bottom geoacoustic parameters in shallow water," *The Journal of the Acoustical Society of America*, vol. 82, no. 3, pp. 998–1017, 09 1987. [Online]. Available: https://doi.org/10.1121/1.395300

[5] J.-C. Hong, K. H. Sun, and Y. Y. Kim, "Dispersion-based short-time Fourier transform applied to dispersive wave analysis," *The Journal of the Acoustical Society of America*, vol. 117, no. 5, pp. 2949–2960, 04 2005. [Online]. Available: https://doi.org/10.1121/1.1893265

[6] G. R. Potty, J. H. Miller, P. S. Wilson, J. F. Lynch, and A. Newhall, "Geoacoustic inversion using combustive sound source signals," *The Journal of the Acoustical Society of America*, vol. 124, no. 3, pp. EL146–EL150, 08 2008. [Online]. Available: https://doi.org/10.1121/1.2960974

[7] J. Bonnel, B. Nicolas, J. I. Mars, and S. C. Walker, "Estimation of modal group velocities with a single receiver for geoacoustic inversion in shallow water," *The Journal of the Acoustical Society of America*, vol. 128, no. 2, pp. 719–727, 08 2010. [Online]. Available: https://doi.org/10.1121/1.3459855

[8] J. Bonnel, A. Thode, D. Wright, and R. Chapman, "Nonlinear time-warping made simple: A step-by-step tutorial on underwater acoustic modal separation with a single hydrophone," *The Journal of the Acoustical Society of America*, vol. 147, no. 3, pp. 1897–1926, 03 2020. [Online]. Available: https://doi.org/10.1121/10.0000937

[9] J. Bonnel, S. E. Dosso, D. Eleftherakis, and N. R. Chapman, "Trans-Dimensional Inversion of Modal Dispersion Data on the New England Mud Patch," *IEEE Journal of Oceanic Engineering*, vol. 45, no. 1, pp. 116–130, 2020.

[10] C. W. Helstrom, D. W. Fry, L. Costrell, and K. Kandiah, *Statistical Theory of Signal Detection: International Series of Monographs in Electronics and Instrumentation*, 2nd ed. Chantilly: Elsevier Science & Technology, 1968.

[11] M. Kucukosmanoglu, J. A. Colosi, P. F. Worcester, M. A. Dzieciuch, and D. J. Torres, "Observations of sound-speed fluctuations in the Beaufort Sea from summer 2016 to summer 2017," *The Journal of the Acoustical Society of America*, vol. 149, no. 3, pp. 1536–1548, 03 2021. [Online]. Available: https://doi.org/10.1121/10.0003601

[12] K. E. Wage, A. B. Baggeroer, and J. C. Preisig, "Modal analysis of broadband acoustic receptions at 3515-km range in the North Pacific using short-time Fourier techniques," *The Journal of the Acoustical Society of America*, vol. 113, no. 2, pp. 801–817, 01 2003. [Online]. Available: https://doi.org/10.1121/1.1530615

[13] K. E. Wage, M. A. Dzieciuch, P. F. Worcester, B. M. Howe, and J. A. Mercer, "Mode coherence at megameter ranges in the North Pacific Ocean," *The Journal of the Acoustical Society of America*, vol. 117, no. 3, pp. 1565–1581, 03 2005. [Online]. Available: https://doi.org/10.1121/1.1854551

[14] T. K. Chandrayadula, K. E. Wage, P. F. Worcester, M. A. Dzieciuch, J. A. Mercer, R. K. Andrew, and B. M. Howe, "Reduced rank models for travel time estimation of low order mode pulses," *The Journal of the Acoustical Society of America*, vol. 134, no. 4, pp. 3332–3346, 10 2013. [Online]. Available: https://doi.org/10.1121/1.4818847

[15] J. E. Poplawski and T. G. Birdsall, "Multimodal dispersion identification," ATOC Project, University of California San Diego, Tech. Rep. Occasional Note Number 34, Mar. 1996, reprinted in *Acoustic Thermometry of Ocean Climate: ATOC, Volume III, Arctic ATOC and ATOC Occasional Notes* (UCSD Geisel Library, Special Collections and Archives), 2003.

[16] J. E. Poplawski, "Dispersion correction and identification of ocean acoustic normal modes," Ph.D. dissertation, University of Michigan.

[17] M. Porter and E. L. Reiss, "A numerical method for ocean-acoustic normal modes," *The Journal of the Acoustical Society of America*, vol. 76, no. 1, pp. 244–252, 07 1984. [Online]. Available: https://doi.org/10.1121/1.391101

[18] F. Jensen, W. Kuperman, M. Porter, and H. Schmidt, *Computational Ocean Acoustics, Second Edition*, ser. Modern Acoustics and Signal Processing. Berlin: Springer Nature, 2011.

[19] C. M. Bender and S. A. Orszag, *Advanced mathematical methods for scientists and engineers*. New York: Springer, 1999 - 1978.

[20] L. Brillouin, *Wave propagation and group velocity*, ser. Pure and applied physics ; v. 8. New York: Academic Press, 1960.

[21] J. Bonnel, C. Gervaise, B. Nicolas, and J. I. Mars, "Single-receiver geoacoustic inversion using modal reversal," *The Journal of the Acoustical Society of America*, vol. 131, no. 1, pp. 119–128, 01 2012. [Online]. Available: https://doi.org/10.1121/1.3664083

[22] J. Bonnel, S. Caporale, and A. Thode, "Waveguide mode amplitude estimation using warping and phase compensation," *The Journal of the Acoustical Society of America*, vol. 141, no. 3, pp. 2243–2255, 03 2017. [Online]. Available: https://doi.org/10.1121/1.4979057

[23] J. R. Reagan, T. P. Boyer, H. E. García, R. A. Locarnini, O. K. Baranova, C. Bouchard, S. L. Cross, A. V. Mishonov, C. R. Paver, D. Seidov, Z. Wang, and D. Dukhovskoy, "World Ocean Atlas 2023," 2024, Dataset. NCEI Accession 0270533.

[24] F. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83, 1978.

[25] J. R. Buck, J. C. Preisig, and K. E. Wage, "A unified framework for mode filtering and the maximum a posteriori mode filter," *The Journal of the Acoustical Society of America*, vol. 103, no. 4, pp. 1813–1824, 04 1998. [Online]. Available: https://doi.org/10.1121/1.421334

[26] S. M. Kay, *Fundamentals of statistical signal processing : estimation theory*. Prentice Hall International, 1993.

[27] T. G. Birdsall, "On understanding the matched filter in the frequency domain," *IEEE Transactions on Education*, vol. 19, no. 4, pp. 168–169, 1976.

Appendix A
Matched filter for maximum likelihood estimation

The measured modal time series is

$$y_m(t) = A\tilde{s}(t-\tau_m;\beta_m) + n(t) \tag{24}$$

where A, τ_m and β_m are unknown and deterministic amplitude, delay, and dispersiveness parameter, respectively, and $n(t)$ is a white Gaussian noise random process with power spectral density σ^2. The corresponding log-likelihood of the modal time series data $y_m(t)$ conditioned on A, τ_m and β_m

$$\Lambda(\tau_m,\beta_m,A) = -\frac{1}{2\sigma^2}\int_{-\infty}^{\infty} |y_m(t) - A\tilde{s}(t-\tau_m;\beta_m)|^2 dt + C \tag{25}$$

where C is a normalization constant. The maximum likelihood amplitude estimate $\hat{A}_{ML}$ conditioned on τ_m and β_m can be found analytically and is

$$\hat{A}_{ML} = \frac{\int_{-\infty}^{\infty} y_m(t)\tilde{s}^*(t-\tau_m;\beta_m)dt}{\int_{-\infty}^{\infty} |\tilde{s}(t-\tau_m;\beta_m)|^2 dt}. \tag{26}$$

The log-likelihood function at the maximum likelihood estimate of A is, up to a normalization constant, given by

$$\begin{aligned}\Lambda(\tau_m,\beta_m,\hat{A}_{ML}) = &-\frac{1}{2\sigma^2}\int |y_m|^2 dt \\ &+\frac{|\int y_m^*(t)\tilde{s}^*(t-\tau_m;\beta_m)dt|^2}{2\sigma^2\int|\tilde{s}(t-\tau_m;\beta_m)|^2 dt}.\end{aligned} \tag{27}$$

The maximum of the log-likelihood coincides with the maximum of

$$J(\tau_m,\beta_m) = \frac{|\int y_m^*(t)\tilde{s}^*(t-\tau_m;\beta_m)dt|^2}{2\sigma^2\int|\tilde{s}(t-\tau_m;\beta_m|^2 dt}. \tag{28}$$

From (17), the energy of the distorted wave form $\tilde{s}(t;\beta_m)$ is independent of the dispersiveness β_m. Therefore, the maximum of J occurs at the maximum of the numerator, which is seen to be the modulus-squared of the matched filter output

$R(\tau_m, \beta_m)$ (19). For a more in-depth description of maximum likelihood estimation, the reader is referred to the textbook of Kay [26].

APPENDIX B
THE EFFECT OF A FREQUENCY-DEPENDENT MODAL AMPLITUDE

The matched filter in (20) is defined in terms of the distorted source waveform $\tilde{S}$, which for a constant modal amplitude is given by (17). This leads to the matched filter in (21) with transfer function

$$H_{MF}(\omega) = S^*(\omega) e^{j\beta(\omega-\omega_0)^2 r}, \tag{29}$$

which can be written in polar form as

$$H_{MF}(\omega) = |S(\omega)| e^{-j\left[\Phi_S(\omega) - \beta(\omega-\omega_0)^2 r\right]}. \tag{30}$$

With a frequency-dependent modal amplitude, the distorted wave form is

$$\tilde{S}(\omega) = A_m(\omega)|S(\omega)| e^{j\left[\Phi_S(\omega) - \beta(\omega-\omega_0)^2 r\right]} \tag{31}$$

where the modal amplitude $A_m(\omega)$ is real-valued. As an example, for a single receiver at depth z_r,

$$A_m(\omega) = \Psi_m(z_s; \omega)\Psi_m(z_r; \omega)/\sqrt{k_{rm}(\omega)}. \tag{32}$$

The matched filter transfer function for this more general case is

$$H_{MF}(\omega) = A_m(\omega)|S(\omega)| e^{-j\left[\Phi_S(\omega) - \beta(\omega-\omega_0)^2 r\right]}, \tag{33}$$

which is seen to differ from the constant amplitude filter (29) by the real-valued weighting $A_m(\omega)$. Birdsall [27] shows that the multiplication of a matched filter transfer function by a real-valued, non-negative function does not change the expected peak output location (i.e. the peak location in the absence of noise). Therefore, if $A_m(\omega)$ does not change sign over the source bandwidth, the expected delay estimates using (29) and (33) coincide. In this case, $A_m(\omega)$ in (33) functions as an amplitude weighting that maximizes the signal-to-noise ratio but does not move the expected peak location. However, if $A_m(\omega)$ changes sign over the source bandwidth, then (29) is no longer matched phase, so an unbiased delay estimate is no longer guaranteed.