Sequential filtering for dispersion tracking and sediment sound speed inversion

Nattapol Aunsri and Zoi-Heleni Michalopoulou[a)]
Department of Mathematical Sciences, New Jersey Institute of Technology, Newark, New Jersey 07102

(Received 31 May 2014; revised 18 August 2014; accepted 22 September 2014)

Dispersion curves in ocean environments are accurately estimated from received signals through the extraction of instantaneous modal frequencies and corresponding arrival times for long-range propagation. The ultimate goal is to estimate sediment sound speed using the extracted dispersion pattern. The approach extends work previously conducted in dispersion tracking with sequential filtering, improving on the latter technique. The sequential state-space method that is developed for the extraction of time-frequency information from specific time instances relies on a representation of those as a sum of elemental pulses, resulting from analysis of the received field. The method is tested on synthetic noisy data with different noise levels. After dispersion probability density functions are estimated via a particle filter, they are subsequently employed for sound speed inversion. Correct mode identification is a challenge impacting inversion; this is demonstrated through two examples and a way to remedy the problem is discussed. © *2014 Acoustical Society of America*.
[http://dx.doi.org/10.1121/1.4897400]

PACS number(s): 43.60.Hj, 43.60.Jn, 43.30.Pc [PJL] Pages: 2665–2674

I. INTRODUCTION

As discussed in Ref. 1, evolution of acoustic signal information in the time-frequency domain contains critical information to be explored by inversion methods for the estimation of parameters that affect sound propagation in the ocean. Among other papers, work in Refs. 2–9 demonstrates the potential of dispersion analysis—the way modal frequencies change with time—for such estimation.

More specifically, for a single frequency, the acoustic field can be decomposed in a set of normal modes. When a broadband source emits sound, normal modes can be calculated for each source signal frequency. Modal frequencies travel with different group velocities and different modes travel with distinct velocities as well. It is these differences, which are environment-dependent, that generate the dispersion patterns that are exploited toward inversion for properties of the medium.

Using dispersion curves for inversion requires the extraction of accurate modal frequencies and their arrival times. Such a task has in the past been approached with short time Fourier transforms (STFTs) and spectrograms or wavelet techniques; warping techniques have also been applied with success even with short range data.[7,8] In this work, we propose an approach that extends previous efforts,[1] combining sequential Bayesian filtering (specifically, particle filtering), spectrograms, and mathematical models describing them, to accurately extract frequency arrival information from time-frequency representations and quantify the uncertainty in the estimation process. Additionally, we estimate the number of modes present, amplitudes of arriving modal frequencies, and noise variance. Particle filtering has been employed in multiple applications in ocean acoustics with significant success.[10–13] Particle filtering for tracking time-frequency signal information attracted a lot of attention in recent years in a number of problems.[14–18] The advantages of particle filtering handling non-linear relations between observations and unknown parameters, complex noise mechanisms, and unknown and varying model order as well as the concept of mode identification with state-space models in ocean acoustics presented in Refs. 1 and 19 motivate our choice of such an approach for time-frequency tracking in an underwater sound propagation problem.

In summary, we first compute spectrograms from acoustic time-series at one hydrophone and describe time "slices" of the spectrogram with a mathematical model of sound propagation in the ocean. Using particle filtering, we then track dispersion of distinct modal frequencies with time with the assistance of the spectrogram model. At the output of the filter, probability density functions (PDFs) of modal frequencies at arrival times are available which are then propagated backward through a normal modes model for sediment compressional sound speed estimation. Because the input to the inverse process consists of PDFs, we obtain at the end of the process complete PDFs of compressional sound speed, allowing us to calculate point estimates as well as quantify the uncertainty in the estimation process. An outline of the process is illustrated schematically in Fig. 1.

The paper is organized as follows. Section II discusses the acoustic model that is employed for describing spectrogram slices. Section III presents the foundations of particle filtering necessary for our work. Section IV demonstrates how the acoustic model and particle filter (PF) can be combined to track dispersion curves for a signal propagating in an ocean waveguide. Modal tracking results are demonstrated in this section along with sediment sound speed estimation results obtained using the modal trajectories. Conclusions follow in Sec. V.

a)Author to whom correspondence should be addressed. Electronic mail: michalop@njit.edu

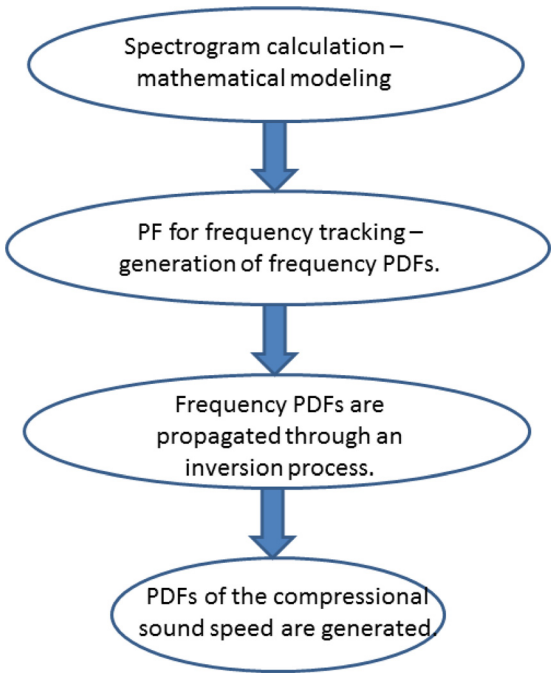

FIG. 1. (Color online) A diagram of the tracking and inversion process.

II. MODELING THE ACOUSTIC FIELD IN THE FREQUENCY DOMAIN

We consider a broadband acoustic signal received at a hydrophone in the ocean. As discussed in Ref. 20, the sound pressure vs time can be written as

$$p(r,z,z_r,t)=\frac{1}{2\pi}\sum_n\int_{-\infty}^{+\infty}\mu(\omega')\,G_n(r,z,z_r,\omega')\times\exp\left\{i\left(\omega' t-k_n r-\frac{\pi}{4}\right)\right\}d\omega'. \quad (1)$$

Quantity r represents the distance between source and receiver, z and z_r are the source and receiver depths, respectively, k_n is the modal wave number, μ is the source spectrum, $\omega=2\pi f$, where f is frequency. Also

$$G_n(r,z,z_r,\omega)=\frac{i\sqrt{\pi}}{\rho(z_r)\sqrt{2k_n r}}\Upsilon_n(z)\Upsilon_n(z_r), \quad (2)$$

where Υ_n are normal modes and $\rho(z_r)$ is density. Although our signals are multimodal, we initially focus on a single mode for simplicity. This restriction will be relaxed later.

The frequency spectrum of a finite time segment of the signal is provided in Refs. 20 and 21 and can be expressed as follows:

$$P_n(\omega,t)=\int_{t-\Delta t/2}^{t+\Delta t/2}p_n(r,z,z_r,t)e^{-i\omega\tau}d\tau, \quad (3)$$

where the segment starts and ends at $t-\Delta t/2$ and $t+\Delta t/2$, respectively. Substituting the nth term of $p(r,z,z_r,t)$ into the previous equation, interchanging the order within the integral, and integrating over time, we obtain

$$P_n(\omega,t)=\frac{e^{-i\omega t}}{\pi}\int_{\infty}^{\infty}\mu(\omega')\,G_n(r,z,z_r,\omega')\times\frac{\sin(\omega'-\omega)\Delta t}{\omega'-\omega}\exp\left\{i\left(\omega' t-k_n r-\frac{\pi}{4}\right)\right\}d\omega'. \quad (4)$$

By applying the stationary phase approximation and squaring P_n, we obtain

$$|P_n(\omega,t)|^2=\frac{\pi}{|k_n''|^2}|\mu(\omega_n)\,G_n(r,z,z_r,\omega_n)|^2\times\left|\frac{\sin(\omega-\omega_n)\Delta t}{\omega-\omega_n}\right|^2, \quad (5)$$

for $|\omega-\omega_n|<\pi/\Delta t$.

From the stationary phase condition we can also derive an important relation between group velocity of mode n as a function of frequency ω_n and range and time. Using V_{gn} as the symbol for group velocity of mode n, it can be shown that $V_{gn}(\omega_n)=r/t$, where r is the distance that the mode has traveled and t is the time it took for this propagation. This shows how group velocity, frequency, and time are tightly related, a concept that will be exploited below using spectrograms that connect all three.

From the analysis that led to Eq. (5), it appears that the squared magnitude of the Fourier transform (FT) can be approximated by a squared sinc pulse, weighted by the squared amplitude of the modal arrival. Moreover, the superposition of these pulses provides an approximation of the squared spectrum of a multiple-mode signal. Although this assumption ignores cross-terms resulting from the squaring operation, it leads to a valid representation of the acoustic field for the part of the spectrum that we will be considering. Specifically, the absence of the cross terms implies a zero inner product between the spectra for the single modes. With the central mode frequencies being well separated for the most part and a suitable window length for the spectrogram calculation, this condition is met.

Instead of Eq. (5), we could work with Eq. (4) and P_n instead, where we would still obtain a model involving sincs, which would now not be squared. The motivation for the squaring operation stems from the fact that spectrograms have been conventionally used for inversion using dispersion curves.[2,7,8,20]

The squared spectrum expressed in Eq. (5) has a peak at the modal frequency ω_n. There is a correspondence between ω_n and mode n; that is, at a given time, there should be a unique "central frequency" ω_n for a particular mode. We can trace the peaks of the instantaneous power spectra to identify the modal frequencies of the acoustic signal. Instead of ω, in order to follow conventional notation for instantaneous frequency estimation, we will be using from now on symbol f for frequency.

One more assumption enters our model at this point. The FT of Eq. (3) is valid for a rectangular window. In our work, we will be employing a Hamming window for its smoothing effect. The Hamming process applied in the time domain will create a transformed sinc^2 pulse (very similar to

the original one). We will refer to this wavelet as u^2, which will be illustrated in Sec. IV A.

III. PARTICLE FILTERING FOR DISPERSION EXTRACTION

Working in a Bayesian sequential filtering framework, we apply PFs for tracking the evolution of modal frequencies with time. We cannot apply simple Kalman filters to our case, because of a nonlinearity in the relationship between data and unknowns and the fact that the number of the actual unknowns (the model order, as will be discussed) is uncertain and varies with time.

Summarizing, the aim of this section is the estimation of the multiple modal frequencies within a received acoustic time-series using the model of Sec. II, as these frequencies evolve with time. Concurrently, we estimate the amplitudes of each mode and the number of dispersion tracks at a given time. To estimate those quantities, we need to calculate the joint PDF of the frequencies and all other unknowns for each time step.

To construct the state-space representation, necessary for our tracking problem, let $\mathbf{x}_{k,r_k}$ be the vector containing all unknown frequencies at time k, where r_k is the number of modes (model order) at slice k mentioned earlier. Also $\mathbf{a}_{k,r_k}$ is the vector of amplitudes corresponding to the modal frequencies. The particle filter we are building in our work is based on sequential importance resampling (SIR).[22] The filter combines three steps: (i) prediction, (ii) update, and (iii) resampling.

Omitting the subscript r_k for simplicity, the state or transition equations for frequencies on which the "predict" step of a PF is based, are as follows:

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \dot{\mathbf{x}}_{k-1}dt + \mathbf{v}_{1k-1} \tag{6}$$

and

$$\dot{\mathbf{x}}_k = \dot{\mathbf{x}}_{k-1} + \mathbf{v}_{2k-1}, \tag{7}$$

where $\mathbf{x}_k$ is a vector containing modal frequencies as previously stated.

The first and third terms at the right hand side of Eq. (6) are the standard components in a conventional state/transition equation in sequential filtering and predicts frequencies at a particular time given a cloud of particles (samples) available from the previous time instance. The cloud forms an estimate of the frequency PDFs at that time. We here consider a third component, $\dot{\mathbf{x}}_{k-1}dt$, which is equivalent to velocity estimation in tracking a moving source. Considering all components of Eq. (6), the frequency particles are perturbed by a component that relates to velocity or, more appropriately for our case, frequency gradient $\dot{\mathbf{x}}_k$ and a small additive quantity $\mathbf{v}_{1k}$, which is drawn from a zero-mean Gaussian density with an empirically selected variance; the variance is here constant with k. Quantity dt is the time between consecutive spectrogram slices.

Equation (7) predicts $\dot{\mathbf{x}}_k$. Similarly to $\mathbf{v}_{1k}$, $\mathbf{v}_{2k}$ takes values from a zero mean Gaussian density. The reasoning behind using gradient in Eqs. (6) and (7) is as follows. Component $\mathbf{v}_{1k-1}$ in Eq. (6) is empirically selected. If gradient is not used and $\mathbf{v}_{1k-1}$ is small, there is a possibility that the interval at state k may not contain the true frequency location in time. If, to avoid that, $\mathbf{v}_{1k-1}$ is selected to be large, particles are generated in a broad range, increasing the uncertainty in the process, which is evidenced via a spread in the PDF. Adding a gradient component increases the width of the particle generation only as needed, determining it in a more structured way through Eq. (6). The impact of using gradient information with its significant advantage in tracking is discussed in more detail in Ref. 13.

The dimension of both $\mathbf{x}_k$ and $\dot{\mathbf{x}}_k$ is r_k. Since r_k is also unknown, an additional state equation is required. The state equation for r_k makes use of a transition probability matrix,[23] which contains the probabilities of order changes (or not); that is, trajectories could enter or exit at each state k and the transition probabilities describe stochastically such changes.

These probabilities describe the possible movement from a current state to the one following. In our case

$$r_{k+1} = \begin{cases} r_k, & \text{with probability } p \\ r_k + 1, & \text{with probability } (1-p)/2 \\ r_k - 1, & \text{with probability } (1-p)/2, \end{cases} \tag{8}$$

where $p < 1$. We have chosen $p = 0.6$. If $r(k) = R_{\min}$, where $R_{\min}$ is the minimum allowed order, we have r_{k+1} staying the same with probability p or increasing with probability $1-p$. Similarly, when $r(k) = R_{\max}$, where $R_{\max}$ is the maximum considered order, r_{k+1} remains the same with probability p or decreases with probability $1-p$. There could be additional state equations for the unknown amplitudes and variance. This issue will be discussed later.

The particles resulting from the state equations need to be refined/updated given the new data at time k. The updating step relies on the relationship between the received data (here, the spectrogram) and the unknown parameters. This relationship is termed the measurement or observation equation. This equation along with a model for the noise contaminating the data allows us to formulate the likelihood for the parameters to be estimated. The discussion of Sec. II provides us with the measurement equation necessary for the likelihood formulation:

$$\mathbf{y}_k = \sum_{j=1}^{r_k} a_{kj}[\mathrm{sinc}(f - x_{kj})]^2 + \mathbf{w}_k. \tag{9}$$

As mentioned, $\mathbf{y}_k$ is a spectrogram slice (squared magnitude of the FT) of the acoustic time series at time k. Quantity $\mathbf{w}_k$ is considered to be additive Gaussian noise in the data: $\mathbf{w}_k \sim N(\mathbf{0}, w_k^2)$ where w_k varies with time as will be demonstrated later. In reality, the noise is Gaussian in the time domain, creating complex Gaussian noise components after the FT calculation. Squaring the spectrum for the spectrogram calculation results in a non-additive χ^2 noise effect. Despite this analysis, a Gaussian noise model has been selected in the past for instantaneous frequency tracking.[1,18,24] This selection facilitates the filter implementation, creating convenient density functions for efficient sampling. Although suboptimal, it will be shown below that the employed model does not compromise the effective extraction of the parameters of interest: A replica signal calculated with extracted frequencies, amplitudes, and order via the PF matches very closely the true signal.

Because of the Hamming window employed in the FT calculation, we rewrite the observation equation as follows:

$$\mathbf{y}_k = \sum_{j=1}^{r_k} a_{kj}[u(f - x_{kj})]^2 + \mathbf{w}_k, \tag{10}$$

where u has been previously defined.

The likelihood for the center frequencies x_{kj} along with amplitudes, number of modal frequencies, and variance, is then

$$l(\mathbf{x_k}, \mathbf{a_k}, r_k, w_k^2; \mathbf{y}_k) = \frac{1}{w_k^L} \frac{1}{(2\pi)^{L/2}} \exp\left\{ -\frac{1}{2w_k^2} \left\| \mathbf{y}_k - \sum_{j=1}^{r_k} a_{kj}[u(f - x_{kj})]^2 \right\|^2 \right\}, \tag{11}$$

where L is the length of each spectrogram slice at a given time.

In order to understand the importance of order $r(k)$ in the frequency estimation process, we experimented by fixing the number of modes to a maximum value rather than treating it as an unknown. The result was that some noisy peaks were reported by the processor as additional arriving frequencies. In other words, we overfitted the data (that is, we fitted the noise in addition to the true modes). As expected, setting the order to a small value had as a result the estimation of fewer modes than those present.

Amplitudes a_k are unknown as well and here are estimated within the PF following the approach of Refs. 13, 23, and 25. Specifically, since amplitudes are elements of the state vector, they could be treated similarly to the unknown frequencies $\mathbf{x}_k$ and perturbed and updated at each time using a transition equation. This would generate additional state equations, significantly increasing the number of necessary particles for achieving a specific accuracy than when only frequencies are estimated (there is approximately a linear relation between the number of unknowns and the number of particles required). As discussed in Refs. 13, 23, and 25, for a more efficient scheme, a maximum likelihood (ML) or a maximum *a posteriori* (MAP) estimator can be used for amplitudes rather than a sampling procedure at each state. This is possible because of the assumption that the noise in the measurement equation [Eq. (10)] is considered to be additive white Gaussian. Then the conditional PDFs of amplitudes on modal frequencies can be shown to be Gaussian as well; drawing samples from such densities is simple and circumvents prediction and update of amplitudes using the standard PF process. Similarly to Refs. 13, 23, and 25, we compute amplitude ML/MAP estimates conditional on the modal frequency particles. The covariance matrix of these conditional probability densities can also be computed. With the available means (equivalent to MAP values here) and covariance, the normal densities are fully described, providing us with enough information for drawing samples. These samples form the marginal posterior PDFs of the amplitudes at each state. The PDFs are used at the next state for the prediction of the new set of frequencies at a specific arrival time and, consequently, of corresponding amplitudes.

Repeating the procedure followed in Ref. 13 for completeness, for a modal frequency particle $\mathbf{x}^i = [x_1^i, \ldots, x_{r_k}^i]$ at the kth time, the MAP estimate $\mathbf{A}_{\mathrm{MAP},k}^i$, a vector consisting of the amplitude MAP estimates from the amplitude PDF conditional on that frequency particle vector, can be calculated as

$$\mathbf{A}_{\mathrm{MAP},k}^i = \Lambda_k^{-1} \psi_k, \tag{12}$$

where

$$\psi_k = \sum_{f=1}^{L} s(f - x_{kr_k}^i)\mathbf{y}_k(f). \tag{13}$$

Vector $\mathbf{y}_k(f)$ is the spectrogram slice that we are considering and $s(f)$ is the considered u^2 wavelet in the measurement equation. Also,

$$\Lambda_k = \begin{pmatrix} \lambda_{k11} & \lambda_{k12} & \cdots & \lambda_{k1r_k} \\ \lambda_{k21} & \lambda_{k22} & \cdots & \lambda_{k2r_k} \\ \cdots & \cdots & \cdots & \cdots \\ \lambda_{kr_k1} & \lambda_{kr_k2} & \cdots & \lambda_{kr_kr_k} \end{pmatrix}, \tag{14}$$

where

$$\lambda_{ki_r j_r} = \sum_{f=1}^{L} s(f - x_{ki_r}^i)s(f - x_{kj_r}^i), \quad i_r, j_r = 1, \ldots, r_k. \tag{15}$$

To form the joint PDF of all unknown parameters, using Bayes theorem we need to multiply the likelihood of Eq. (11) and the priors for all unknown parameters. Prior densities for arriving frequencies, amplitudes, and model order are assumed to be uniform. We have

$$p(f_{rk}) = \frac{1}{L} \tag{16}$$

for each modal frequency,

$$p(a_{r_k}) = M \tag{17}$$

for each amplitude, and

$$p(r_k) = \frac{1}{R_{\max} - R_{\min} + 1} \tag{18}$$

for the order, where $R_{\max}$ and $R_{\min}$ are maximum and minimum order values, respectively. These values here are 12 and six.

Because we have r_k arrivals, r_k uniform priors $1/L$ are multiplied, providing a complete prior of $1/L^{r(k)}$ for the arriving frequencies. Prior $1/L^{r(k)}$, which obtains small values when $r(k)$ is large, serves as a penalization factor that prevents the process from favoring a large order (similar to the Schwartz criterion).[26]

Since noise variance w_k^2 is also unknown, we need to estimate it at every state k. Variance can also be included in the state equations drawing particles at each time slice and following the predict and update steps. Alternatively, the conditional posterior PDF for the variance on all other unknowns can be traced and variance samples can be drawn from there at every state. To form the posterior, a prior density is needed. A typical prior PDF for an unknown variance is[27]

$$p\left(w_k^2\right) = \frac{1}{w_k^2}, \tag{19}$$

which is what we use here. Such a selection creates a conditional inverse χ^2 density with L degrees of freedom for the variance, from which it is straightforward to sample for each set of frequency and amplitude particles. Variance estimation plays an important role in the frequency and amplitude estimation process. Assuming that the variance is known but has a smaller value than the true one typically leads to overestimation, since data peaks are interpreted as parts of the true trajectories (because the noise level is assumed to be lower than the actual one). On the other hand, assumption of a higher variance value than the true one results in underestimation.

Combining the likelihood and priors, the joint PDF of all unknowns (frequencies, amplitudes, number of modes, and variance) given the observed data is as follows:

$$p\left(\mathbf{x_k}, \mathbf{a_k}, \mathbf{r_k}, w_k^2 | \mathbf{y}_k\right) \propto \frac{1}{w_k^{L+2}} \frac{1}{(2\pi)^{L/2}} \frac{1}{L^{r_k}} \exp\left\{ -\frac{1}{2w_k^2} \left\| \mathbf{y}_k - \sum_{j=1}^{r_k} a_{kj}[u(f - x_{kj})] \right\|^2 \right\}. \tag{20}$$

Evaluation of this joint density for distinct particles provides weights for each particle within a cloud representing how important these are, given the available data. The calculation of the posterior PDF for the unknown parameters is the foundation for the "update" step of the PF.

The prediction and update steps are two of the three building blocks of an SIR PF, where, in our case, frequency, amplitude, order, and variance values are predicted using samples from the previous state and are updated based on how they fit in the PDF calculation, given new data entering in the current state. The joint PDF resulting from the two steps could present us with the problem of degeneracy, where only a few particles obtained from the filtering process have large weights, with most particles having negligible values. A resampling step after the update can be employed (the third within SIR) to address this complication.[22] The cumulative density function (CDF) corresponding to the calculated PDF is computed. Via the CDF, particles are resampled, replacing those obtained after the update step; the particle weights/probabilities determine the frequency of each new particle. That is, the particles with larger weights may be chosen multiple times and samples with low weights may not be selected at all. The new set of particles forms a better representation of the posterior PDF.

IV. TRACKING DISPERSION CURVES AND SOUND SPEED ESTIMATION

A. Estimating dispersion tracks

We consider a source that transmits a broadband signal with frequency content between 200 and 600 Hz which propagates in the ocean and is received at a hydrophone located 20 km away from the source. The ocean depth is 111 m and we consider a thin sandy sediment of 3.5 m over limestone. The sediment sound speed is 1670 m/s. The environment is similar to that of the Gulf of Mexico experiment[28] and the sampling rate is 2000 Hz. The synthetic signal is shown in Fig. 2.

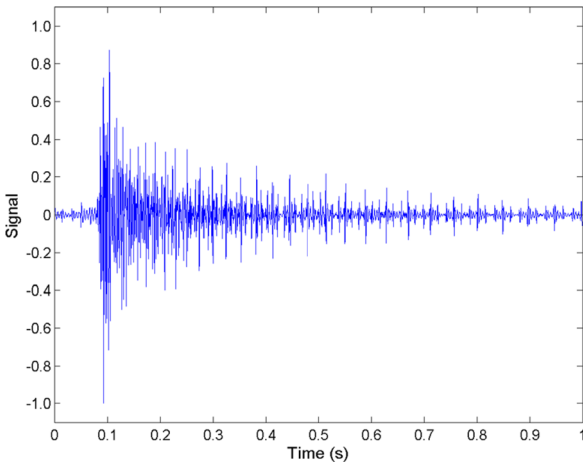

FIG. 2. (Color online) Synthetic time-series generated for the Gulf of Mexico experiment environment.

Before implementing the PF, it is important to determine the width of the wavelet function in Eq. (10). This width depends on the length Δt (or length L) of the FT which leads to the spread of the modal arrival. The corresponding squared sinc pulse is shown in Fig. 3(a) and the resulting wavelet u^2 is illustrated in Fig. 3(b).

The spectrogram of the received signal obtained via an STFT calculation is shown in Fig. 4(a). We identify dispersion patterns, showing how the frequencies of distinct modes arrive at the receiver and also how different frequencies within a mode arrive. From the arrival times on the horizontal axis we can see that some frequencies/modes arrive faster than others. Also frequencies within modes arrive at different times. The whole dispersion pattern characterizes the properties of the waveguide. It is these intermodal and intramodal dispersion effects that we want to exploit for inversion, and, toward this goal, we want to identify the dispersion pattern as accurately as possible. We should ideally extract the numerically

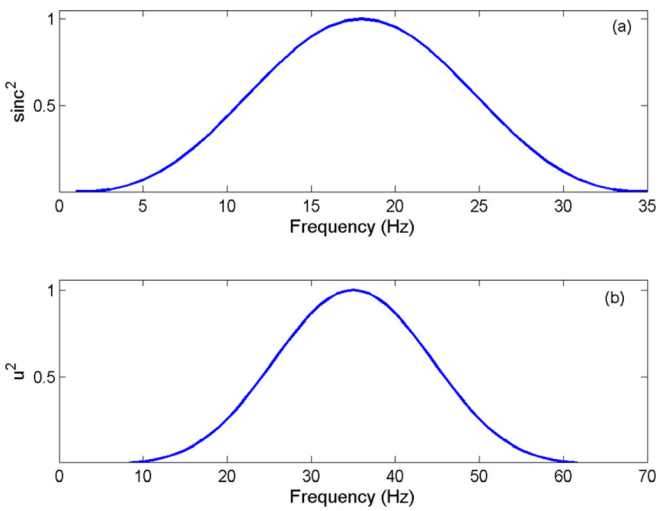

FIG. 3. (Color online) The source wavelet function employed in (a) Eq. (9) and (b) Eq. (10). The latter is used in our observation equation.

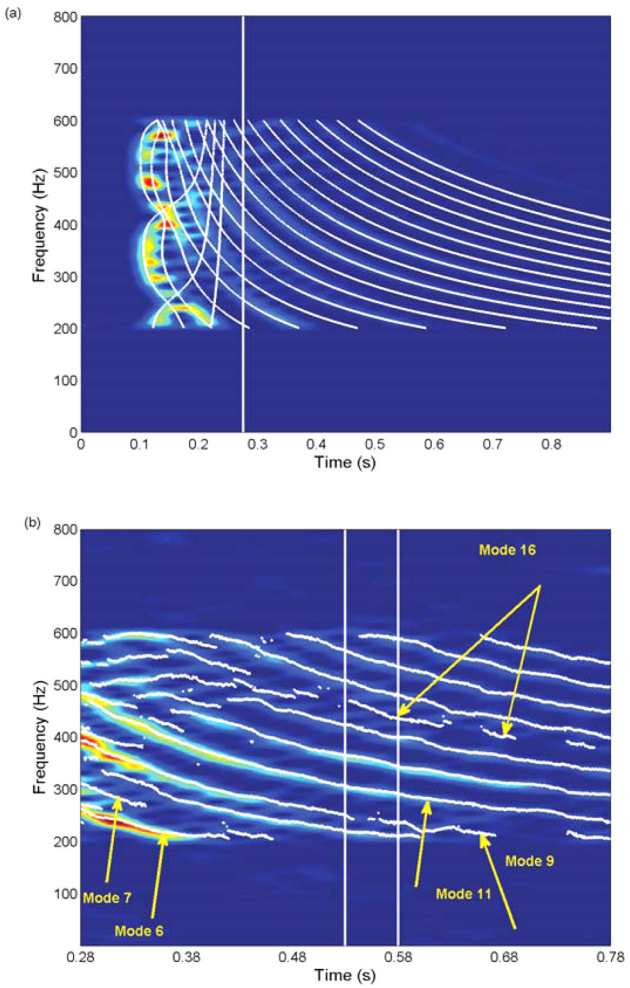

FIG. 4. (Color online) (a) The spectrogram of an acoustic signal that has propagated 20 km away from the source with the numerically calculated dispersion curves superimposed. (b) Frequency tracks as estimated by the PF.

calculated dispersion curves for the considered waveguide that we have superimposed on the spectrogram (solid-line curves); these have been computed using a normal modes model.[29] Noise and FT artifacts are problems that often hinder accurate dispersion curve estimation.

Using the sequential method of Sec. III, where state and observation equations are derived, we estimate modal trajectories present in the spectrogram. We only process the spectrogram after the first 0.28 s [indicated by the solid line in Fig. 4(a)], because in this segment a downward trend of the dispersion curves is evident, facilitating the tracking process. Also, the modes are better separated in frequency after the first few ms of the FT, during which the modes can overlap and cross at some frequencies. This violates the assumption of the absence of cross-terms made in Sec. II. Additionally, we are interested in modes that have interacted with the bottom sediment, which may not be present in the early part of the spectrogram. Point estimates of elements of frequency vector $\mathbf{x}_j^k$ present in modal trajectory j at time k are calculated as

$$\hat{x}_{kij} = \mathrm{MAP}(x_{kij}), \quad j = 1, 2, \ldots, r_k, \tag{21}$$

where $\mathrm{MAP}(x_{kij})$ is the most frequent value of x_{kij} for mode j at time k among all particles i.

The MAP trajectory estimates obtained for the PDFs calculated by the PF are shown in Fig. 4(b) with dots that are superimposed on the spectrogram. The considered signal is the same as the one used for the calculation of the spectrogram of Fig. 4(a), but now white Gaussian noise has been added to the time-series before the STFT and spectrogram computation. The results demonstrate a much higher resolution than the resolution of the results presented in Ref. 1, which were already superior to simple spectrogram maximization. Figure 5 demonstrates the estimated model order (number of modal frequencies). In the beginning of the considered segment, because of noise and FT effects, 12 modes appear present. As we move along in the tracking, the PF identifies seven to eight modes. Noisy peaks in the spectrogram cause the PF to occasionally add more trajectories which exit rapidly, as there is no data to further support the new tracks. It is interesting to observe that mode seven is detected although it is very faint. The same applies to mode 16, which is extracted for many of the time slices. Although some dispersion curves (or parts of those) are "crisply" identified, other estimates appear poor. For example, mode 9, which is indicated in the figure, becomes "wavy" (and interrupted) toward its end, because is it distorted by noise. Also several modes are not detected, because they are not excited or are very weak.

To further investigate the validity of our model, we demonstrate in Fig. 6 a slice of the spectrogram at a particular time (solid line) and the corresponding squared spectrum we obtain using MAP estimates for arriving frequencies and amplitudes at the same time ("stars"). Although assumptions have been made both in terms of the noise model and the mathematical model describing arriving modes, the match is excellent with the MAP spectrum almost coinciding with the squared FT magnitude of the data.

Tracking results are fairly robust with respect to noise. The noise level cannot be expressed here in terms of a single signal-to-noise ratio (SNR) value. As the time progresses, we can see in Fig. 4 that the spectrogram becomes weaker, because the signal attenuates with time. We calculate the SNR as a function of time, taking into account this effect.

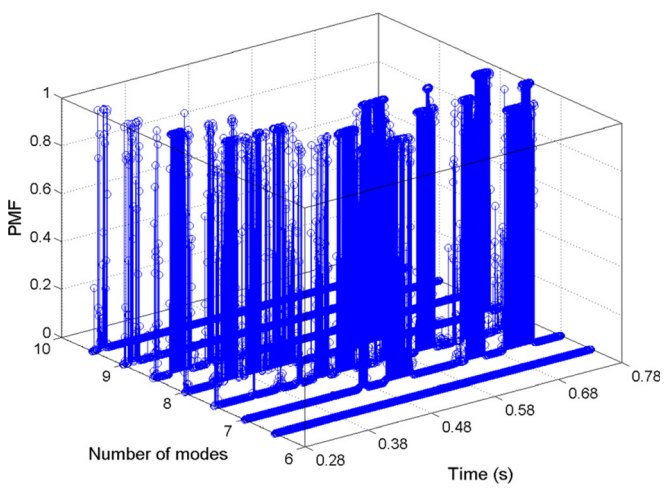

FIG. 5. (Color online) The probability mass function (PMF) of the number of modes.

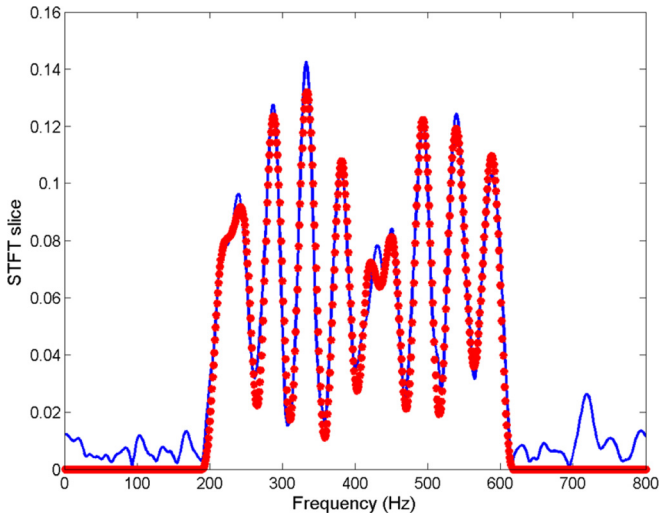

FIG. 6. (Color online) A slice of the spectrogram for a particular time (solid line) with the spectrum constructed using the MAP estimates for the PF superimposed ("stars").

The SNRs vs time for the two noise levels are shown in Fig. 7. Figure 8 shows two spectrograms with the superimposed frequency MAP estimates for the SNRs of Fig. 7 [Fig. 8(a) demonstrates the same results as those in Fig. 4(b); the results are reproduced for an easier comparison]. In terms of MAP estimates, the frequency trajectories follow well the dispersion curves when the noise is increased. The quality is maintained even for the times when the signal is weak. For the noisier circumstances [Fig. 7(b)], however, there is higher uncertainty in the frequency estimation—albeit, very small. This is demonstrated in Fig. 9, where frequency PDFs are shown for the same mode for the two cases. The spread increase when the noise level is raised is small but evident. As expected, as the noise rises further, so does the uncertainty in the frequency estimation.

Finally, in Fig. 10 we show the spectrogram of the noisy realization of the time-series in Fig. 2 when a rectangular window is used in the FT computation, which is directly comparable to that of Fig. 4(b). Although we can still see

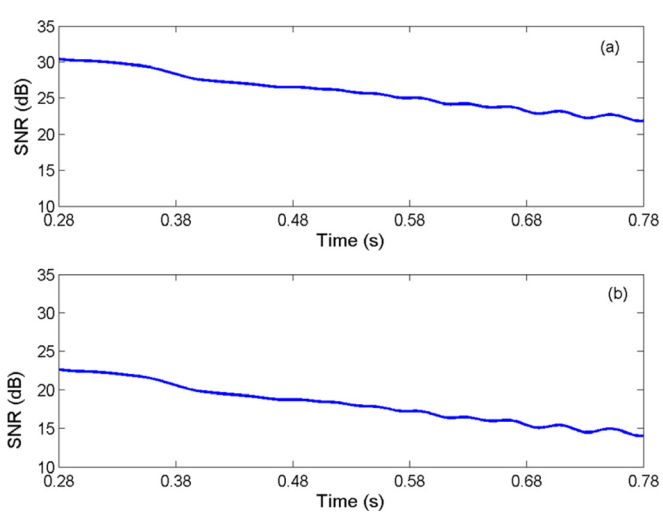

FIG. 7. (Color online) The SNR for two different noise levels as a function of time.

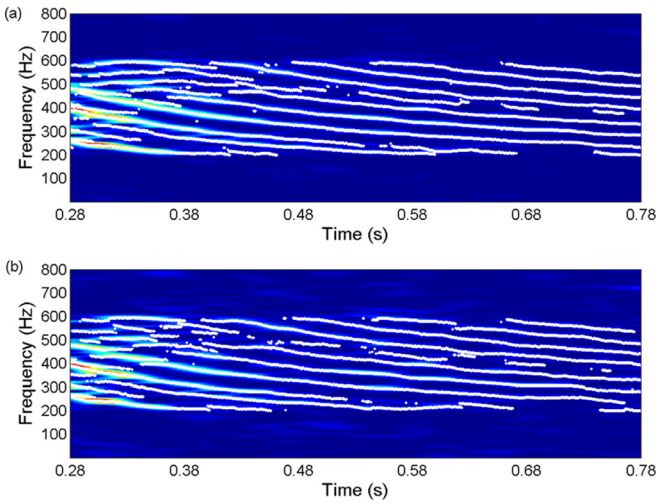

FIG. 8. (Color online) Tracking for two different noise levels.

most modes traced correctly in the new results, the estimated tracks with the rectangular window appear to be inferior to those obtained using a Hamming window.

B. Inversion for sediment sound speed

Frequency PDFs can be used for inversion for sound speed in the sediment. Frequencies are associated with arrival times, which are directly related to group velocities via the propagation range. Specific group velocity values and modal arrival times are the result of the properties of the waveguide, and, thus, of sediment sound speed.

To solve the inverse problem of estimating environmental parameters of the propagation medium, matched field processing (MFP) has often been employed.[30] This method uses search algorithms that navigate a large parameter space to seek parameter values which generate synthetic fields (replicas) that best fit the data. MFP requires a combination of wave propagation modeling for the generation of replica fields at receiving phones and a decision rule that estimates model parameters entering the replica calculation. Inversion

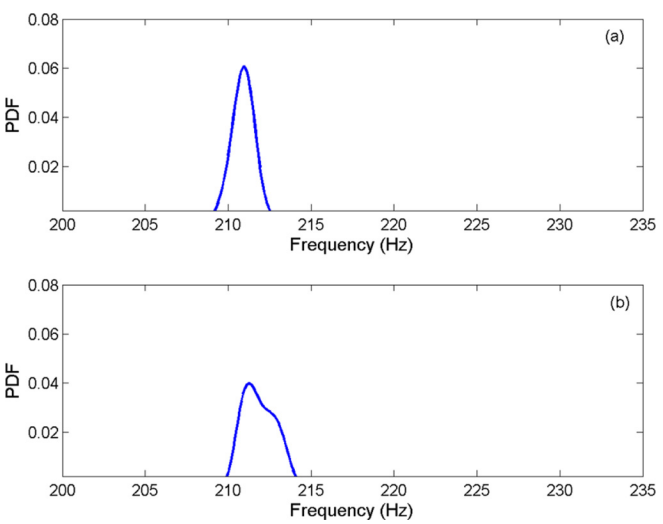

FIG. 9. (Color online) The frequency PDFs for the same mode for two different SNRs.

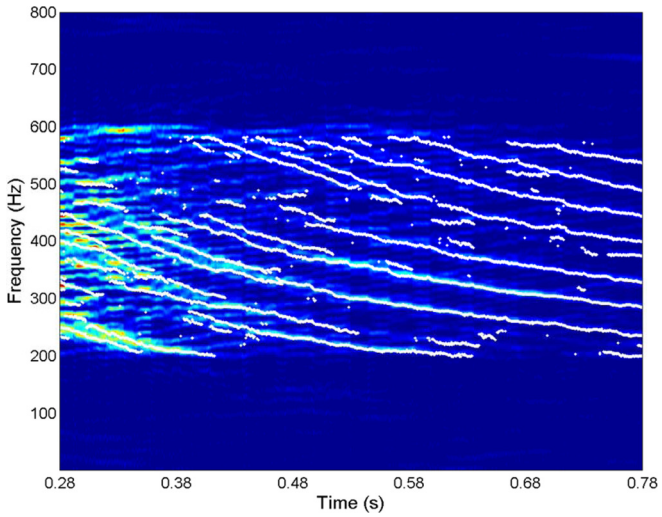

FIG. 10. (Color online) The spectrogram of an acoustic signal using a rectangular window.

is performed by identifying those values of the model parameters that maximize a similarity measure between replica and true acoustic fields. MFP was originally used for source localization and was later adapted and applied to estimation of environmental parameters.[31]

Here, we follow an approach similar to MFP. However, instead of creating full-field replicas that we then match to the acoustic signal, we only work with arrival times or, equivalently, group velocities. To focus on the effect of dispersion tracking on compressional sound speed estimation, all other parameters (source location, bathymetry, sediment thickness, density, attenuation, and half-space properties) are considered known and fixed in our calculations. Using a sound propagation model,[29] we first calculate the acoustic field with normal modes for frequencies ranging between 200 and 600 Hz—with 1 Hz spacing—for different values of compressional velocity c_p. The step between considered c_p values is 1 m/s. The normal mode modeling allows us to calculate modal wavenumbers k_n; using those as well as frequency information, we calculate group velocities,

$$V_g = \frac{d\omega}{dk_n}, \tag{22}$$

for each mode and for each value of c_p. These are "replica" group velocities which can be readily related to replica arrival times because the range is known (20 000 m).

Here, we select a specific mode and the PDFs of its frequencies at specific times as obtained by the PF. Since the arrival time is different for the selected modal arrivals, the corresponding frequencies have different group velocities. We focus on mode 11 and we define arrival times $t_{11,1}$ and $t_{11,2}$ as

$$t_{11,1} = t_0 + 20\,000/V_{g11}(x_{11t_1}, c_p), \tag{23}$$

$$t_{11,2} = t_0 + 20\,000/V_{g11}(x_{11t_2}, c_p), \tag{24}$$

where t_0 is the transmission time, which is unknown, x_{11t_1} and x_{11t_2} are the arriving frequencies, and $V_{g11}(x_{11t_1}, c_p)$ and $V_{g11}(x_{11t_2}, c_p)$ are the group velocities for frequencies x_{11t_1} and x_{11t_2} for the 11th mode.

Difference $\Delta t = t_{11,2} - t_{11,1}$ is 0.05 s, because we work with times of 0.53 and 0.58 s as indicated by the solid lines in Fig. 4(b); the unknown transmission time t_0 is removed because of the calculation of a time difference. We seek with a grid search values of c_p that minimize error e, where

$$e = |\Delta t - \Delta t_{\text{rep}}|^2. \tag{25}$$

This is equivalent to selecting an MFP measure of similarity between replica and true fields. To generate Δt_{rep}, we compute replica group velocity V_{g11} using wavenumbers and frequencies, as mentioned earlier, for x_{11i1} and x_{11i2}, obtained by the PF, for different values of c_p within the selected

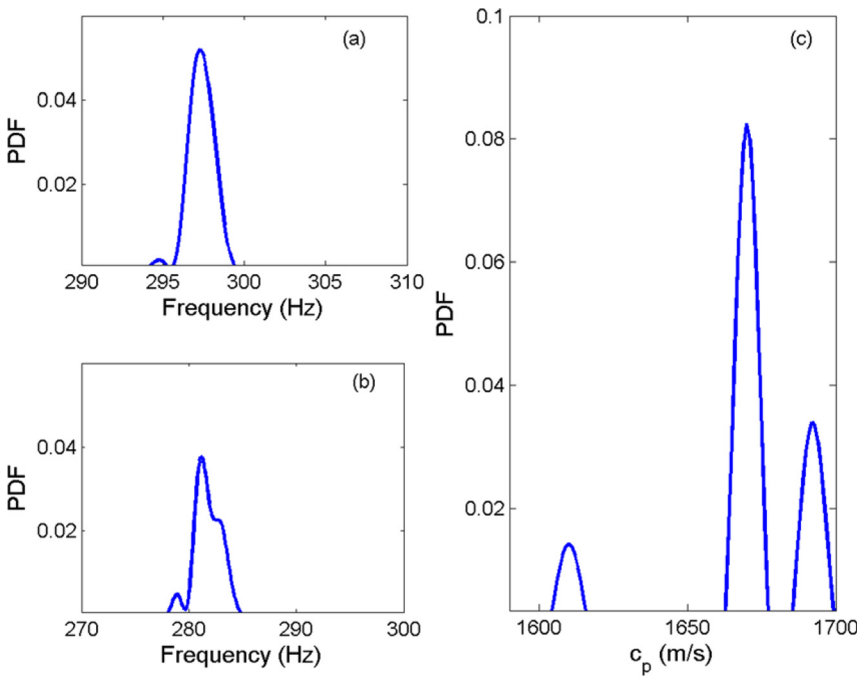

FIG. 11. (Color online) The frequency PDFs for mode 11 (a) at 0.53 s and (b) at 0.58 s at the times slice indicated in Fig. 4(b) with the solid lines. Particles for these frequencies are used for sound speed inversion. (c) The PDFs of the sediment sound speed calculated using the frequency densities of (a) and (b).

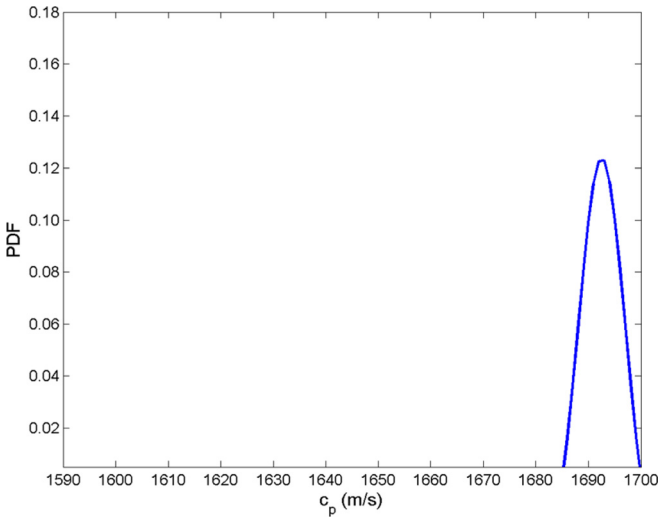

FIG. 12. (Color online) The PDFs for the sediment sound speed calculated using the frequency densities of Figs. 11(a) and 11(b). Mode 11 has been misidentified as mode 10.

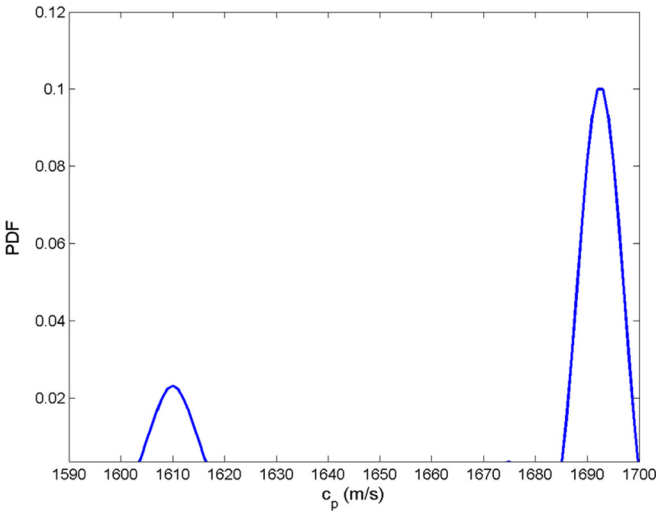

FIG. 13. (Color online) The PDF for the sediment sound speed calculated using the frequency densities of Figs. 11(a) and 11(b). Mode 11 has been misidentified as mode 12.

search interval; x_{11i1} and x_{11i2} are particles drawn for the arriving modal frequencies. In short, the PDFs of frequencies as estimated by the particle filter are propagated backward through a sound propagation model, resulting in PDFs for sediment velocity c_p.

The PDFs estimated with the PF for the two frequencies considered in the inversion are demonstrated in Figs. 11(a) and 11(b). The resulting density for sediment sound speed is shown in Fig. 11(c). Prior information for the sound speed included 1590 and 1700 m/s as minimum and maximum for the search limits. The MAP estimate of the PDF is 1670 m/s, the true value. It is interesting that two secondary modes are present, centered at 1609 and 1692 m/s. These are the result of inversion using particles at the borders of the frequency PDFs. In spite of their presence, which indicates a sensitivity of the method to frequency variation, the PDF is centered at the correct sound speed, with significant probability surrounding this value.

Results presented in Fig. 11 have been calculated based on the assumption that mode 11 has been correctly identified. This is an important issue that has a large impact on the quality and accuracy of the environmental parameter estimates. Figure 12 shows results that are produced after misidentification of the mode for the case illustrated in Fig. 11. Mode 11 is treated as if it is the tenth mode. The sound speed PDF shown in Fig. 12 is peaked at 1692 m/s, which is very close to the upper limit of the original search interval. The relation of the estimate to the prior information, which is fairly broad, indicates that there may be errors in the modal recognition, suggesting that the estimation process should be revisited. The sound velocity PDF when mode 11 is incorrectly labeled as mode 12 is illustrated in Fig. 13. The probability density is peaked again at $c_p = 1692$ m/s with a smaller mode at $c_p = 1609$ m/s. The main mode peaks at the upper limit of the search interval, again indicating that the search should be repeated more carefully with emphasis on mode recognition. More specifically, in addition to repeating the mode identification, more modes and times can be used for the inversion. This was performed here, with results corroborating the estimation of 1670 m/s as the compressional velocity in the sediment.

V. CONCLUSIONS

Modal dispersion curves are extracted combining a mathematical model for arriving modal frequencies at a given time and an approach that tracks frequency behavior in time; this approach relies on particle filtering. The process considers a number of modes present in a signal that has propagated a long distance from the source in a Gaussian noise environment. The number of modes, modal amplitudes, and noise variance are unknown. The filter integrates all these parameters in the frequency estimation process and produces PDFs in addition to point estimates. The estimation process is successful under different noise levels; however, the spread of the PDFs increases with variance, as expected.

The particles that are produced at the output of the filter forming the dispersion PDFs are used as input to an inverse propagation model for the estimation of sediment sound speed in the environment. Normal mode modeling is used for calculating modal group velocities that are then matched to estimated velocities and corresponding arrival times using the dispersion analysis. When modes are correctly identified, sediment sound speed is successfully estimated, with the MAP estimate coinciding with the true value. Uncertainty is expressed via the spread of the PDF and, especially, through secondary modes. On the other hand, sound speed is not estimated accurately when the modes are misidentified. In that case, for the example that we show, the resulting sound speed estimates are right at the border of the available prior information, pointing to potential problems with mode selection and identification. The process can then be repeated exploring different combinations of modes.

It should be noted here that the PF was applied to simple spectrograms obtained via STFT calculations. It is conceivable that even better results may be attainable, if an improved time-frequency representation is used before the application of the tracking algorithm.

This approach was followed in Ref. 21 resulting in a good match between data and replica spectra for static estimation. The concern, however, with other time-frequency representations that typically improve SNR is that a smoothing factor may eliminate some modes. Although this issue may not necessarily hinder inversion and may potentially improve it, it is a challenge that should be carefully addressed in the future.

ACKNOWLEDGMENTS

This work is supported by the Office of Naval Research through grants N000141010073 and N000141310077.

[1]I. Zorych and Z.-H. Michalopoulou, "Particle filtering for dispersion curve tracking in ocean acoustics," J. Acoust. Soc. Am. **124**, EL45–EL50 (2008).
[2]C.-S. Chen, J. Miller, F. Boudreaux-Bartels, G. Potty, and C. Lazauski, "Time-frequency representations for wideband acoustic signals in shallow water," in *Proceedings of OCEANS*, San Diego, CA (September 22–26, 2003), pp. 2903–2907.
[3]M. Taroudakis and G. Tzagkarakis, "On the use of the reassigned wavelet transform for mode identification," J. Comput. Acoust. **12**, 175–196 (2004).
[4]G. Potty, J. Miller, P. Dahl, and C. Lazauski, "Geoacoustic inversion results from the ASIAEX East China Sea Experiment," IEEE J. Ocean. Eng. **29**(4), 1000–1010 (2004).
[5]S. Rajan and K. M. Becker, "Inversion for range-dependent sediment compressional-wave-speed profiles from modal dispersion data," IEEE J. Oceanic Eng. **35**, 43–58 (2010).
[6]H. Dong and S. E. Dosso, "Bayesian inversion of interface-wave dispersion for seabed shear-wave speed profiles," IEEE J. Oceanic Eng. **36**, 1–11 (2011).
[7]J. Bonnel and N. R. Chapman, "Geoacoustic inversion in a dispersive waveguide using warping operators," J. Acoust. Soc. Am. **130**, EL101–EL107 (2011).
[8]J. Bonnel, S. E. Dosso, and N. R. Chapman, "Bayesian geoacoustic inversion of single hydrophone light bulb data using warping dispersion analysis," J. Acoust. Soc. Am. **134**, 120–130 (2013).
[9]J. Bonnel and A. Thode, "Using warping processing to range bowhead whale sounds from a single receiver," Proc. Meetings Acoust. **19**, 1–7 (2013).
[10]C. Yardim, Z.-H. Michalopoulou, and P. Gerstoft, "An overview of sequential Bayesian filtering in ocean acoustics," IEEE J. Ocean. Eng. **36**, 71–89 (2011).
[11]C. Yardim, P. Gerstoft, and W. S. Hodgkiss, "Geoacoustic and source tracking using particle filtering: Experimental results," J. Acoust. Soc. Am. **128**, 75–87 (2010).
[12]J. Dettmer, S. E. Dosso, and C. W. Holland, "Sequential trans-dimensional Monte Carlo for range-dependent geoacoustic inversion," J. Acoust. Soc. Am. **129**, 1794–1806 (2011).
[13]Z.-H. Michalopoulou and R. Jain, "Particle filtering for arrival time tracking in space and source localization," J Acoust. Soc. Am. **132**, 3041–3052 (2012).
[14]N. Ikoma, "Estimation of time varying peak of power spectrum based on non-Gaussian nonlinear state space modeling," Signal Process. **49**, 85–95 (1996).
[15]N. Ikoma and H. Maeda, "Nonstationary spectral peak estimation by Monte-Carlo filter," in *IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium* (2000), pp. 245–250.
[16]C. Andrieu, M. Davy, and A. Doucet, "Efficient particle filtering for jump Markov systems. Application to time-varying autoregressions," IEEE Trans. Signal Process. **51**, 1762–1770 (2003).
[17]R. Prado, M. West, and A. Krystal, "Multichannel electroencephalographic analyses via dynamic regression models with time-varying lag-lead structure," J. R. Stat. Soc.: Ser. C (Appl. Statist.) **50**(1), 95–109 (2001).
[18]C. Dubois and M. Davy, "Joint detection and tracking of time-varying harmonic components: A flexible Bayesian approach," IEEE Trans. Audio, Speech, Language Process. **4**, 1283–1295 (2007).
[19]J. Candy and D. Chambers, "Internal wave signal processing: a model-based approach," IEEE J. Ocean. Eng. **21**, 37–52 (1996).
[20]T. C. Yang, "A method for measuring the frequency dispersion for broadband pulses propagated to long ranges," J. Acoust. Soc. Am. **76**, 253–261 (1984).
[21]Z.-H. Michalopoulou, "A Bayesian approach to modal decomposition in ocean acoustics," J. Acoust. Soc. Am. **126**, EL147–EL152 (2009).
[22]B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications* (Artech House, Boston, MA, 2004), pp. 1–291.
[23]J. Larocque, J. Reilly, and W. Ng, "Particle filters for tracking an unknown number of sources," IEEE Trans. Signal Process. **50**(12), 2926–2937 (2002).
[24]C. Dubois, M. Davy, and J. Idier, "Tracking of time-frequency components using particle filtering," in *Proc. IEEE ICASSP*, Philadelphia, PA (March 2005), Vol. 2, pp. 887–891.
[25]C. Andrieu, M. Davy, and A. Doucet, "Improved auxiliary particle filtering: applications to time-varying spectral analysis," in *11th IEEE Signal Processing Workshop on Statistical Signal Processing* (2001), pp. 309–312.
[26]Z.-H. Michalopoulou, C. Yardim, and P. Gerstoft, "Particle filtering for passive fathometer tracking," J. Acoust. Soc. Am. **131**, EL74–EL80 (2012).
[27]G. Box and G. Tiao, *Bayesian Inference in Statistical Analysis* (Wiley-Interscience, New York, 1992), pp. 1–582.
[28]Z.-H. Michalopoulou, M. B. Porter, and J. P. Ianniello, "Broadband source localization in the Gulf of Mexico," J. Comput. Acoust. **4**, 361–370 (1996).
[29]M. B. Porter, "The Kraken normal mode program," SACLANT Undersea Research Centre Memorandum (SM-245) and Naval Research Laboratory Mem. Rep. 6920 (1991).
[30]A. Tolstoy, *Matched Field Processing for Underwater Acoustics* (World Scientific, Singapore, 1993).
[31]E. Livingston and O. Diachok, "Estimation of average under ice reflection amplitudes and phases using matched field processing," J. Acoust. Soc. Am. **86**, 1909–1919 (1989).