Mode filters and energy conservation

Ilya A. Udovydchenkov[a)] and Irina I. Rypina
Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543
ilya@whoi.edu, irypina@whoi.edu

Michael G. Brown
Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida, 33149
mbrown@rsmas.miami.edu

Abstract: The discrete form of the mode filtering problem is considered. The relevant equations constitute a linear inverse problem. Solutions to problems of this type are subject to a well-known trade-off between resolution and precision. But unlike the typical linear inverse problem, the correctly formulated mode filtering problem is subject to an energy conservation constraint. This letter focuses on the importance of satisfying, approximately at least, the energy conservation constraint when mode filtering is performed.
© 2010 Acoustical Society of America

1. Introduction

In the modal description of the underwater sound field, the energy in the field at each frequency is partitioned among a discrete set of normal modes. In generic range-dependent environments where energy is exchanged among modes, much can be learned about the underlying propagation physics by decomposing the energy in the wavefield into energy contained in the local normal modes. The process of estimating complex-valued modal amplitudes from a set of measurements of complex pressure at many depths at a fixed range is often referred to as mode filtering. The discrete form of the equations to be solved constitutes a system of linear equations whose solution is the desired vector of modal amplitudes. The choice of the solution involves consideration of a trade-off between precision and resolution. But the mode filtering problem is subject to an additional constraint—an energy conservation constraint—that plays no role in most linear inverse problems. The importance of satisfying the energy conservation constraint has been overlooked in previous approaches to the design of mode filters. In this letter, it is argued that the energy conservation constraint is an important consideration in the design of a mode filter.

We shall assume that the ocean environment can be modeled as a fluid with constant density and we shall neglect attenuation. For convenience, we focus on cw acoustic fields; the issues that we discuss apply to transient wavefields after Fourier decomposition. At each angular frequency $\omega = 2\pi f$, the acoustic normal modes $\psi_m(z)$ satisfy the equation

$$\psi_m''(z) + \left(\omega^2 c^{-2}(z) - k_m^2\right)\psi_m(z) = 0, \tag{1}$$

together with an appropriate pair of boundary conditions. We shall assume that the boundary conditions are such that together with Eq. (1) they define a Sturm–Liouville problem, thereby guaranteeing that the normal modes constitute a complete orthogonal set of functions over the relevant depth domain (the contribution from evanescent modes is assumed to be negligibly small). The constant k_m in Eq. (1) is a separation constant and $c(z)$ is the sound speed profile at the range at which the modes are defined. At this range (for simplicity we have chosen to suppress the dependence on both r and ω in the equations presented), the depth structure of the complex-valued acoustic pressure field can be expanded as a weighted sum of the locally de-

––––––––––––––––––––––

[a)]Author to whom correspondence should be addressed.

fined normal modes, $p(z) = \Sigma_m d_m \psi_m(z)$. By orthonormality of the modes, $\int \psi_m(z) \psi_n(z) dz = \delta_{nm}$, $d_n = \int p(z) \psi_n(z) dz$. The modal amplitudes d_m are constrained by a generalized Parseval relation (which expresses energy conservation), $\int |p(z)|^2 dz = \Sigma_m \Sigma_n d_m d_n^* \int \psi_m(z) \psi_n(z) dz = \Sigma_m |d_m|^2$. Here d_m^* is the complex conjugate of d_m, and vertical bars denote modulus.

Assume that the pressure field and the normal modes are sampled with uniform spacing Δz. Then fill the column vector $\mathbf{p}$ with sequential samples of the pressure field multiplied by $\sqrt{\Delta z}$ and similarly for each column vector $\boldsymbol{\psi}_n$. (The assumption of uniform spacing between samples is made for convenience. Similarly, multiplication of samples by $\sqrt{\Delta z}$ is a convenient, but not necessary, normalization.) Let $\boldsymbol{\Psi}$ denote the real-valued matrix whose nth column is $\boldsymbol{\psi}_n$. And let $\mathbf{d}$ denote a column vector of complex-valued modal amplitudes.

Assume, for now, that the vectors $\mathbf{p}$ and $\boldsymbol{\psi}_n$ cover the entire relevant depth domain; this assumption will be relaxed in Sec. 3. Assume, in addition, that Δz is sufficiently small that the condition $\boldsymbol{\psi}_n^T \boldsymbol{\psi}_m = \delta_{nm}$—or, equivalently, $\boldsymbol{\Psi}^T \boldsymbol{\Psi} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix—is satisfied with tolerably small error. Unless stated otherwise, this assumption is made throughout this letter. (Deficient receiving arrays will be discussed in Sec 3.) With these assumptions, the discrete analogs of the above equations are $\mathbf{p} = \boldsymbol{\Psi} \mathbf{d}$, $\mathbf{d} = \boldsymbol{\Psi}^T \mathbf{p}$, and $\mathbf{p}^H \mathbf{p} = (\boldsymbol{\Psi} \mathbf{d})^H (\boldsymbol{\Psi} \mathbf{d}) = \mathbf{d}^H \mathbf{d}$. The superscript H denotes the Hermitian transpose. Because the modes are assumed to be real valued $\boldsymbol{\Psi}^H = \boldsymbol{\Psi}^T$. Note also that $\boldsymbol{\Psi} \boldsymbol{\Psi}^T \mathbf{p} = \mathbf{p}$. We emphasize that validity of the equations $\mathbf{p} = \boldsymbol{\Psi} \mathbf{d}$ and $\boldsymbol{\Psi} \boldsymbol{\Psi}^T \mathbf{p} = \mathbf{p}$ rests on the assumption that the modes are complete, i.e., that the measured $\mathbf{p}$ lies in the space spanned by the columns of $\boldsymbol{\Psi}$.

Mathematically, the condition $\mathbf{p}^H \mathbf{p} = \mathbf{d}^H \mathbf{d}$ is a consequence of the orthonormality of the modes. But the physical importance of this result as a statement of energy conservation must not be overlooked. If an estimate of the vector $\mathbf{d}$ of modal amplitudes clearly violates the energy conservation condition, then that estimate must be a poor approximation to the true $\mathbf{d}$.

2. Mode filters: Resolution, precision, and energy conservation

Real measurements are noisy, so it is natural to consider a slight generalization of $\mathbf{p} = \boldsymbol{\Psi} \mathbf{d}$,

$$\mathbf{p} = \boldsymbol{\Psi} \mathbf{d} + \mathbf{e}. \tag{2}$$

Here $\mathbf{e}$ is a column vector containing complex-valued errors. We seek to solve Eq. (2) for the vector $\mathbf{d}$ given measurements of $\mathbf{p}$, knowledge of $\boldsymbol{\Psi}$, and knowledge of the statistics of $\mathbf{e}$. Because $\mathbf{p}$ is not measured independently of $\mathbf{e}$, the statistics of these quantities are linked. The estimated vector of modal amplitudes, $\hat{\mathbf{d}}$, is an inverse of $\boldsymbol{\Psi}$, $\boldsymbol{\Psi}_{\text{inv}}$, operating on $\mathbf{p}$, $\hat{\mathbf{d}} = \boldsymbol{\Psi}_{\text{inv}} \mathbf{p}$ $= \boldsymbol{\Psi}_{\text{inv}} \boldsymbol{\Psi} \mathbf{d} + \boldsymbol{\Psi}_{\text{inv}} \mathbf{e}$, so $\hat{\mathbf{d}} - \mathbf{d} = (\boldsymbol{\Psi}_{\text{inv}} \boldsymbol{\Psi} - \mathbf{I}) \mathbf{d} + \boldsymbol{\Psi}_{\text{inv}} \mathbf{e}$. An unbiased estimator satisfies $\boldsymbol{\Psi}_{\text{inv}} \boldsymbol{\Psi} = \mathbf{I}$. More generally, the matrix $\boldsymbol{\Psi}_{\text{inv}} \boldsymbol{\Psi}$ is a measure of resolution in mode amplitude space. For an unbiased estimator, the covariance matrix of the estimated modal amplitudes is $\langle (\hat{\mathbf{d}} - \mathbf{d})(\hat{\mathbf{d}} - \mathbf{d})^H \rangle$ $= \boldsymbol{\Psi}_{\text{inv}} \langle \mathbf{e} \mathbf{e}^H \rangle \boldsymbol{\Psi}_{\text{inv}}^T$, where angular brackets denote ensemble average, or (rewriting the same equation using a more convenient notation) $\mathbf{K}_{\hat{d}\hat{d}} = \boldsymbol{\Psi}_{\text{inv}} \mathbf{K}_{ee} \boldsymbol{\Psi}_{\text{inv}}^T$. The diagonal elements of $\mathbf{K}_{\hat{d}\hat{d}}$ are the variances of the modal amplitude estimates, $(\mathbf{K}_{\hat{d}\hat{d}})_{ii} = \sigma_{\hat{d}i}^2$. For an unbiased estimate of $\hat{\mathbf{d}}$, a lower bound on these variances exists,[1,2] $\sigma_{\hat{d}i}^2 \geq [[\boldsymbol{\Psi}^T \mathbf{K}_{ee}^{-1} \boldsymbol{\Psi}]^{-1}]_{ii}$. The lower bound is referred to as the Cramer–Rao lower bound (CRLB). Finally, energy conservation dictates that

$$\mathbf{p}^H \mathbf{p} = \hat{\mathbf{d}}^H \hat{\mathbf{d}} = \mathbf{p}^H \boldsymbol{\Psi}_{\text{inv}}^T \boldsymbol{\Psi}_{\text{inv}} \mathbf{p}. \tag{3}$$

The energy conservation constraint plays no role in most linear inverse problems. The mode filtering problem is special in this respect. The left-hand side of Eq. (3) is the total energy in the measured wavefield. In the presence of noisy measurements, estimation of this quantity can be improved if the energy in the noise field $\mathbf{e}^H \mathbf{e}$ can be independently estimated, which is usually the case.

Ideally one would like to find an inverse operator $\mathbf{\Psi}_{\text{inv}}$ that: (1) is unbiased $\mathbf{\Psi}_{\text{inv}}\mathbf{\Psi} = \mathbf{I}$; (2) satisfies the CRLB condition; and (3) satisfies the energy conservation condition, Eq. (3). For the special case of uncorrelated errors, all with the same variance σ_e^2 (so $\mathbf{K}_{ee} = \sigma_e^2\mathbf{I}$), all three conditions can be satisfied by choosing

$$\mathbf{\Psi}_{\text{inv}} = \mathbf{\Psi}^T. \tag{4}$$

It should be emphasized, however, that problems for which all three conditions can be satisfied are exceptional. More generally, some compromise must be reached between the often competing desires to satisfy all three conditions. In the following, we shall refer to $\mathbf{\Psi}_{\text{inv}}$ defined in Eq. (4) as the DP (for direct projection) mode filter.

Consider now the more realistic case for which $\mathbf{K}_{ee}$ is not simply a constant times the identity matrix. In that case, the DP filter can still be used. In fact, the DP filter still has two desirable properties: it is an unbiased estimator and it satisfies the energy conservation constraint. (But it does not satisfy the CRLB condition.) An alternative to the DP mode filter is

$$\mathbf{\Psi}_{\text{inv}} = (\mathbf{\Psi}^T\mathbf{K}_{ee}^{-1}\mathbf{\Psi})^{-1}\mathbf{\Psi}^T\mathbf{K}_{ee}^{-1}. \tag{5}$$

This filter is known by many names, including best linear unbiased estimator[3] (BLUE). The BLUE mode filter also has two desirable properties: it is an unbiased estimator and it satisfies the CRLB condition. (But it does not satisfy the energy conservation constraint.) Several alternatives to the DP and BLUE mode filters have been proposed,[1,4–8] including generalizations of the BLUE filter that account for *a priori* knowledge of both the statistics of measurement errors and the statistics of modal amplitudes.[1,8] Those filters will not be further discussed, except to note that they are not constrained to conserve energy. Finally, consider a mode filter composed of a weighted sum of the BLUE and DP mode filters,

$$\mathbf{\Psi}_{\text{inv}} = \gamma(\mathbf{\Psi}^T\mathbf{K}_{ee}^{-1}\mathbf{\Psi})^{-1}\mathbf{\Psi}^T\mathbf{K}_{ee}^{-1} + (1-\gamma)\mathbf{\Psi}^T, \tag{6}$$

where $0 \leq \gamma \leq 1$ is an adjustable parameter. We shall refer to this filter as the blended mode filter. This filter has the following properties: (1) it is unbiased for any choice of γ; (2) it satisfies the CLRB condition when $\gamma = 1$; and (3) it satisfies the energy conservation constraint when $\gamma = 0$. The blended filter will be used below to illustrate the trade-off between the competing desires to keep the uncertainty of modal amplitude estimates small (which favors γ near 1) and to satisfy the energy conservation constraint (which favors γ near 0).

3. Deficient receiving arrays

In most applications, the dominant source of uncertainty in $\mathbf{p}$ is missing measurements, often at many depths, rather than noisy measurements. In this section, we describe three different methods that might, under different circumstances, be used to deal with missing measurements, and we discuss the related issue of how the energy conservation constraint can be applied in each case. Also, simulations are presented that focus on elucidating the importance of satisfying, approximately at least, the energy conservation constraint when mode filtering is performed.

At short range and low frequency, when mode coupling is either very weak or can be approximately described deterministically, an acoustic propagation model prediction can be used to fill in the missing measurements. One must then also estimate the uncertainty of the model prediction and fill the matrix $\mathbf{K}_{ee}$ in a way consistent with the estimated uncertainty. The total energy can then be computed as $\mathbf{p}^H\mathbf{p}$ where $\mathbf{p}$ is filled with a combination of measurements and model predictions.

The situation in which mode coupling is not describable deterministically is much more common. In the presence of even weak mode coupling, due, for example, to internal waves in the deep ocean, elements of $\mathbf{p}$ can usually be accurately described as zero-mean complex Gaussian random variables. In other words, if one computes an ensemble of simulations of the wavefield at a fixed range, with each member of the ensemble corresponding to a different realization of the internal wave field, then the resulting ensemble of complex pressures at each depth is, to a good approximation, described by a zero-mean Gaussian distribution. Simulations

in deep ocean conditions similar to those presented below reveal that this behavior is quite robust. Generally, both the variance σ_p^2 and the vertical correlation length z_c of the complex pressure distributions are slowly varying functions of depth. Both quantities can be estimated from simulations or measurements when the latter are available.

Under these circumstances the following strategy for dealing with deficient receiving arrays can be used. At each depth z_i at which no measurement is available, p_i is set equal to zero, the mean of the complex Gaussian distribution at that depth, the diagonal element of the covariance matrix $(\mathbf{K}_{ee})_{ii}$ is set equal to the variance σ_p^2 of the distribution, and the off-diagonal elements $(\mathbf{K}_{ee})_{ij} = (\mathbf{K}_{ee})_{ji}$ are set equal to $\sigma_p^2 \exp(-|z_i - z_j|/z_c)$. At depths at which measurements are available, p_i is the measured pressure and $(\mathbf{K}_{ee})_{ij}$ is set equal to $\sigma_e^2 \delta_{ij}$, where σ_e^2 is estimated from the noise. Typically $\sigma_p^2 \gg \sigma_e^2$. For simplicity, we assume here that the variance of the pressure distribution σ_p^2 and the variance of measurement error σ_e^2 are both independent of depth, and that measurement errors are uncorrelated; these assumptions can be relaxed. Provided Δz, the depth interval between adjacent measurements, is sufficiently small that $\boldsymbol{\psi}_n^T \boldsymbol{\psi}_m = \delta_{nm}$ to good approximation, the idea of filling an expanded $\mathbf{p}$ vector with zeros between measurements is not relevant.

Filling the vector $\mathbf{p}$ with zeros at the depths where measurements are missing suggests the following strategy for imposing the energy conservation constraint: estimate the energy in the field using only the measured p-values. Application of the energy conservation constraint (3) will then lead to a set of modal amplitude estimates $\hat{\mathbf{d}}$ whose total energy is less than that of the true modal amplitudes. We shall refer to this form of the energy conservation constraint as the strong form of the constraint. Note that the strong form of the energy conservation constraint is consistent with the exact result, Eq. (3), inasmuch as, in the limit of no missing receiving array elements, the two are equal. But the strong form of the constraint is just that—a somewhat stronger assumption than the exact (limiting) constraint. While the strong form of the energy conservation constraint is not exact, it remains useful in that it prevents the inclusion of spurious energy in modal amplitude estimates. This will be illustrated below.

A more common approach to the treatment of missing receiving array elements is to remove the missing elements of the vectors $\mathbf{p}$ and $\mathbf{e}$, and the corresponding rows of $\boldsymbol{\Psi}$ from the system (2) to be solved. Like zero filling of $\mathbf{p}$, truncation of the system (2) suggests consideration of the strong form of the energy conservation constraint. This will be discussed in more detail below. Let $\boldsymbol{\Psi}_t$ denote the truncated modal matrix. The desirable properties of $\boldsymbol{\Psi}$ do not carry over to $\boldsymbol{\Psi}_t$; in particular, $\boldsymbol{\Psi}_t^T \boldsymbol{\Psi}_t \neq \mathbf{I}$. Thus, much of the rationale for using the DP mode filter is lost when the system is truncated. In contrast, the BLUE mode filter (5) with $\boldsymbol{\Psi}$ replaced by $\boldsymbol{\Psi}_t$ retains its desirable properties: it remains an unbiased estimator that satisfies the CRLB condition. Note also that the truncated BLUE filter can be thought of as a limiting case of the zero-filled nontruncated BLUE filter in the limit of infinite σ_p^2. In other words, the truncated BLUE filter assumes that the uncertainty of p at depths where measurements are missing is unbounded. Like the zero-filled BLUE mode filter, the truncated BLUE mode filter is not constrained to conserve energy. Also, truncation often results in the matrix $\boldsymbol{\Psi}_t^T \mathbf{K}_{ee,t}^{-1} \boldsymbol{\Psi}_t$ being poorly conditioned. A commonly used way to overcome this problem is to employ diagonal loading,

$$\boldsymbol{\Psi}_{\text{inv}} = (\boldsymbol{\Psi}_t^T \mathbf{K}_{ee,t}^{-1} \boldsymbol{\Psi}_t + \beta \mathbf{I})^{-1} \boldsymbol{\Psi}_t^T \mathbf{K}_{ee,t}^{-1}. \tag{7}$$

This filter is unbiased only in the limit $\beta = 0$. Complementing the *ad hoc* arguments presented here, Eq. (7) can be derived as a damped least-squares solution to Eq. (2) that minimizes the cost function $(\mathbf{p}_t - \boldsymbol{\Psi}_t \mathbf{d})^H \mathbf{K}_{ee,t}^{-1}(\mathbf{p}_t - \boldsymbol{\Psi}_t \mathbf{d}) + \beta \mathbf{d}^H \mathbf{d}$ (see, e.g., Ref. 9). Also, Eq. (7) can be thought of as a special case of the MAP filter discussed in Buck *et al.*,[1] where $\beta^{-1}\mathbf{I}$ is an *a priori* estimate of the modal covariance matrix $\mathbf{K}_{dd}$.

We now test, using simulations, some of the approaches to mode filtering that we have described. Figure 1 shows four mode-processed transient wavefields produced by an axial source with $f_0 = 75$ Hz and $\Delta f = 18.75$ Hz (full width at half amplitude) at 1000 km range in the same deep ocean environment. (The environment and source function used in these simulations are identical to those used in Ref. 10. Mode processing was done at each frequency after Fourier trans-
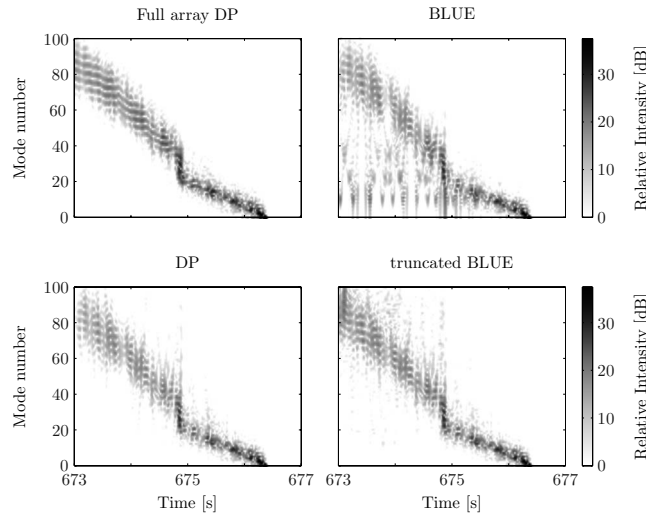
Fig. 1. Mode-processed transient wavefields (f_0=75 Hz, Δf=18.75 Hz) in a deep ocean environment at 1000 km range. Upper left panel: correct mode-processed field constructed using a dense water-column-spanning receiving array and the DP mode filter. Lower left panel: mode-processed field constructed using a deficient receiving array and the DP mode filter with zero filling. Upper right panel: mode-processed field constructed using a deficient receiving array and the BLUE mode filter with zero filling. Lower right panel: mode-processed field constructed using a deficient receiving array and the BLUE mode filter with truncated modes and small diagonal loading.

forming from time to frequency, and prior to transforming back to the time domain.) The upper left panel of Fig. 1 was produced by mode processing a dense ($\Delta z \approx 27$ m) set of pressure time histories spanning the entire 5 km deep water column using the DP mode filter. That mode-processed field can be thought of as the correct field. The other three panels of Fig. 1 show the result of mode processing the same wavefield, but measured using a deficient receiving array, with three different mode filters. The deficient receiving array included only those measurements in the full array in the depth intervals 350–1750, 2150–2850, and 3570–4270 m. The three deficient array mode-processed fields shown in Fig. 1 correspond to the DP mode filter with zero filling, the BLUE mode filter with zero filling, and the BLUE mode filter with truncated modes and diagonal loading. In the implementation of the BLUE mode filter with zero-filling that was used to construct Figs. 1 and 2(a), it was assumed that σ_p^2, σ_e^2, and z_c are constants, with σ_p^2/σ_e^2=1000 and z_c=150 m. (Note that $\boldsymbol{\Psi}_{\text{inv}}$ for the DP mode filter does not depend on these parameters, but the corresponding $\mathbf{K}_{\hat{d}\hat{d}}$ does.) The loading coefficient β=0.032 was used to produce the mode-processed field shown in Fig. 1 corresponding to the BLUE mode filter with truncated modes (7). Some diagonal loading was essential because the matrix $\boldsymbol{\Psi}_t^T \mathbf{K}_{ee,t}^{-1} \boldsymbol{\Psi}_t$ is very poorly conditioned (condition number approximately 10^{16} at
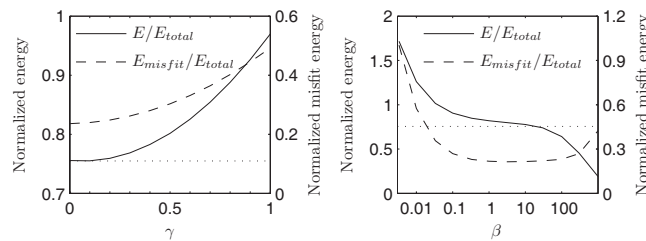

Fig. 2. Normalized energy and normalized misfit energy: (a) as a function of γ for the blended mode filter with zero filling (6); and (b) as a function of β for the truncated BLUE mode filter with diagonal loading (7). Dotted horizontal lines at E/E_{total}=0.75 correspond to the energy value associated with the strong form of the energy conservation constraint.

75 Hz). With $\beta=0.032$, the condition number of the matrix $\mathbf{\Psi}_t^T\mathbf{K}_{ee,t}^{-1}\mathbf{\Psi}_t+\beta\mathbf{I}$ is approximately 3 $\times 10^4$ at 75 Hz. A visual examination of Fig. 1 reveals that the DP mode filter with zero filling does a good job of reproducing the correct mode-processed field, with the caveat that some energy appears to be missing. That filter is constrained to satisfy the strong form of the energy conservation constraint. In contrast, both implementations of the BLUE mode filter—which are energetically unconstrained—appear to predict spurious energy that is not seen in the correct mode-processed field.

Figure 2(a) shows normalized energy and normalized misfit energy as a function of γ for the blended mode filter (6) with zero filling. The same deficient array that was used to produce the deficient array mode-processed fields shown in Fig. 1 was used to produce Fig. 2. Consider first the normalized energy, which is defined as the ratio of energy E contained in a mode-processed wavefield to E_{total}, the total energy that is recorded on the water-column spanning array (corresponding to the upper left panel of Fig. 1). For a broadband signal, energy can be computed as the integral over frequency of $\mathbf{p}^H\mathbf{p}$ or $\hat{\mathbf{d}}^H\hat{\mathbf{d}}$. Clearly, E_{total} is an upper bound on the energy recorded on a deficient receiving array; for the conditions used to produce Fig. 2, only 75% of E_{total} is captured by the receiving array. The blended filter was applied with zero filling, so the strong form of the energy conservation constraint is satisfied in the $\gamma=0$ limit. (Recall that in that limit the blended filter reduces to the DP filter.) Because the strong form of the energy conservation constraint is satisfied in that limiting case, the normalized energy in that case is 0.75. In Fig. 2(a) normalized energy for the blended mode filter is seen to increase with increasing γ. It is not clear, however, from the normalized energy alone whether the energy that is added as γ increases is (a) "true" energy that improves the quality of the mode-processed field, or (b) spurious energy that degrades the quality of the mode-processed field. To address this question we compute the misfit energy.

The misfit energy E_{misfit}, defined as $(\hat{\mathbf{d}}-\mathbf{d})^H(\hat{\mathbf{d}}-\mathbf{d})$ (or, for a broadband signal, the integral over frequency of this quantity), is a good measure of the total error in a mode-processed wavefield. To compute E_{misfit} the vector $\mathbf{d}$ was computed using the wavefield measured on water-column-spanning receiving array. We define the normalized misfit energy as the ratio $E_{\text{misfit}}/E_{\text{total}}$. Consider the normalized misfit energy vs γ curve in Fig. 2(a). That curve reveals that the cause of the increase in the energy in the mode-processed wavefield as γ increases is inclusion of an increasing amount of spurious energy. Based on the misfit energy, we conclude that the DP mode filter ($\gamma=0$) clearly outperforms the BLUE mode filter ($\gamma=1$). Note that the monotonic increase in the misfit energy with increasing γ that is seen in Fig. 2(a) is exactly opposite the trend expected based on the uncertainty estimates (in the matrix $\mathbf{K}_{\hat{d}\hat{d}}$) described above; recall that the BLUE mode filter satisfies the minimum variance (CRLB) condition, while the DP filter does not. We infer from this observation that $\mathbf{K}_{\hat{d}\hat{d}}$ defined above is a meaningful measure of the uncertainty of estimates of modal amplitudes only when the $\mathbf{\Psi}_{\text{inv}}$ used in that expression is, to a good approximation, both unbiased and energy conserving. The poor performance of the BLUE mode filter is linked to the fact that it is not constrained, even approximately, to conserve energy.

Consider now Fig. 2(b). That figure shows normalized energy and normalized misfit energy for the truncated BLUE mode filter with diagonal loading (7) as a function of the loading coefficient β. $\beta\approx 10^{-3}$ is sufficiently large to stabilize the inverse; for β-values significantly larger than this value, this filter should be thought of as a parametric damped weighted least-squares filter rather than a BLUE filter. Figure 2(b) shows that the misfit energy has a broad minimum for β-values between approximately 1 and 20, and that the strong form of the energy constraint, $E/E_{\text{total}}=0.75$, is satisfied at $\beta=26$. In other words, within the one-parameter family of mode filters defined by Eq. (7), good performance is achieved by choosing the parameter β in such a way as to satisfy the strong form of the energy conservation constraint. For this choice of β ($\beta=26$), the corresponding mode-processed wavefield looks almost identical to the DP-based mode-processed wavefield shown in Fig. 1. While the observation that choosing β in Eq. (7) to satisfy an energy conservation constraint leads to good filter performance is somewhat encouraging, it should be emphasized that, for most one-parameter families of filters, this procedure cannot be expected to help;

a simple scaling of the magnitude of $\mathbf{\Psi}_{\text{inv}}$ is generally not helpful, for instance. Stated somewhat differently, energy conservation is a necessary condition, but not a sufficient condition, for good mode filter performance.

4. Summary and discussion

In this letter, we have focused on the formulation of mode filtering problem. We have argued that the correctly formulated mode filtering problem involves consideration of the energy conservation constraint, in addition to precision and resolution. Numerical simulations were presented that illustrate the importance of satisfying the energy conservation constraint.

   Although we have not focused on the problem of finding an optimal solution to the mode filtering problem, it is appropriate to make some brief comments relating to that task. The approach that immediately comes to mind is to pose the problem as a constrained optimization problem in which an appropriate quadratic measure of misfit, e.g., $(\mathbf{p}-\mathbf{\Psi d})^H\mathbf{K}_{ee}^{-1}(\mathbf{p}-\mathbf{\Psi d})$, is minimized subject to the constraint that the solution conserve energy. Note that, in general, imposition of the energy conservation constraint precludes satisfying the CLRB condition; the constrained minimum variance will usually be greater than the unconstrained minimum variance. When measurements collected on a deficient array are processed, two approaches to specifying the energy conservation constraint come to mind. One could either: (1) impose the strong form of the constraint that we have discussed (leading to a set of modal amplitude estimates whose total energy is less than that of the true modal amplitudes); or (2) augment the measurements with a propagation model prediction to estimate the total energy. The results shown in Fig. 2 suggest that under conditions similar to those used in our simulations, the former approach is preferable. The constrained optimization problem can be easily formulated by introducing a scalar Lagrange multiplier. The constrained optimization problem does not appear to have an analytic solution, but the numerical method that we have described to constrain the diagonally loaded truncated BLUE filter to conserve energy leads to a solution to this problem. More importantly, the numerical results that we have presented show that nonoptimal mode filters can perform well. A necessary condition for good performance is approximately satisfying the energy conservation constraint.

Acknowledgments

References and links

[1] J. R. Buck, J. C. Preisig, and K. E. Wage, "A unified framework for mode filtering and the maximum a posteriori mode filter," J. Acoust. Soc. Am. **103**, 1813–1824 (1998).

[2] H. L. Van Trees, *Detection, Estimation, and Modulation Theory* (Wiley, New York, 1968).

[3] G. Strang, *Introduction to Applied Mathematics* (Cambridge Press, Wellesley, 1986).

[4] C. T. Tindle, K. M. Guthrie, G. E. J. Bold, M. D. Johns, D. Jones, K. O. Dixon, and T. G. Birdsall, "Measurements of the frequency dependence of normal modes," J. Acoust. Soc. Am. **64**, 1178–1185 (1978).

[5] T. C. Yang, "A method of range and depth estimation by modal decomposition," J. Acoust. Soc. Am. **82**, 1736–1745 (1987).

[6] C. T. Tindle, H. Hobaek, and T. G. Muir, "Normal mode filtering for downslope propagation in a shallow water wedge," J. Acoust. Soc. Am. **81**, 287–294 (1987).

[7] A. G. Voronovich, V. V. Goncharov, A. Yu. Nikol'tsev, and Yu. A. Chepurin, "Comparative analysis of methods for the normal mode decomposition of a sound field in a waveguide: Numerical simulation and full-scale experiment," Sov. Phys. Acoust. **38**, 365–370 (1992).

[8] C.-S. Chiu, C. W. Miller, and J. F. Lynch, "Optimal modal beamforming of bandpass signals using an undersized sparse vertical hydrophone array: Theory and a shallow water experiment," IEEE J. Ocean. Eng. **22**, 522–533 (1997).

[9] C. Wunsch, *The Ocean Circulation Inverse Problem* (Cambridge University Press, Cambridge, 1996).

[10] I. A. Udovydchenkov and M. G. Brown, "Modal group time spreads in weakly range-dependent deep ocean environments," J. Acoust. Soc. Am. **123**, 41–50 (2008).