

# Hypothesis testing for geoacoustic environmental models using likelihood ratio

Christoph F. Mecklenbräuer<sup>a)</sup>  
Ruhr-University of Bochum, D-44780 Bochum, Germany

Peter Gerstoft<sup>b)</sup>  
Marine Physical Laboratory, Scripps Institution of Oceanography, University of California at San Diego,  
La Jolla, California 92093-0704

Johann F. Böhme<sup>c)</sup> and Pei-Jung Chung<sup>c)</sup>  
Signal Theory Group, Department of Electrical Engineering, Ruhr-University of Bochum, D-44780 Bochum,  
Germany

(Received 2 June 1997; revised 30 October 1998; accepted 5 November 1998)

A generalized likelihood ratio test is developed for testing acoustic environmental models with an application to parameter inversion using an acoustic propagation code. The signal-to-noise ratio in acoustic measurements proves to limit the details on geoacoustic environments that can be determined by matched field processing methods. A hypothesis test serves in Monte Carlo simulations as a tool to determine minimal signal levels for the bottom parameter inversion. The term ‘‘hierarchy of models’’ is used for denoting a sequence of models in which each particular model contains all previous ones. For determining the model order and its parameters, a combined parameter estimation and multiple sequential test is proposed. Given the observed data, how many parameters should be included in the model? The last question is important for the order selection in hierarchies of models with an increasing number of parameters. Multiple sequential hypotheses testing provides a procedure to determine the model order in a statistically justified way. Monte Carlo simulations show the behavior of the test for selecting a model order as a function of the signal-to-noise (SNR) ratio. The test is applied to broadband data measured using a vertical array near the island of Elba in the Mediterranean Sea and compared with Akaike’s Information Criterion. © 1999 Acoustical Society of America. [S0001-4966(99)03502-X]

PACS numbers: 43.60.Pt, 43.30.Pc [JCB]

## INTRODUCTION

Previous literature on geoacoustic inversion focused primarily on the parameter estimation problem.<sup>1–4</sup> The *structure* of the acoustic propagation model (i.e., its parametrization) was assumed to be known *a priori*. Which parameters are relevant and should be included into the model as *unknowns* is usually decided on intuitive physical grounds— independently of the observed experimental data. However, both background noise and fluctuations in the quantities severely limit the *observable details* of an acoustic environment.

Various approaches to structural model identification are available in the signal processing literature.<sup>5–7</sup> Ljung<sup>7</sup> gives a good discussion of the subject. Most problems considered so far assume that the true model structure (although unknown in detail) is embedded in a hierarchy. This hierarchy is constructed from model structures of increasing complexity. Good examples of such model structures are the familiar autoregressive moving average (ARMA) models.

There seems no general optimum way to *build* such a hierarchy for geoacoustic environmental models, since the

acoustic parameters (e.g., sound speed and attenuation profiles) do not have a natural *order*: Obviously, there is no specific first, second, third,..., parameter. As a consequence, building such a hierarchy is subjective and the acoustician stays responsible for the design according to specific needs. However, indications of the ‘‘relative importance’’ of individual parameters can be numerically obtained.

In this paper we deal with statistical hypothesis tests for acoustic environments based on observed data and a replica on a vertical array of sensors. It is an extension of the work in Refs. 8 and 9. The replica is generated using environmental model parameters (sound speed profile, attenuations, and densities) and an acoustic propagation model. The array output is modeled as a superposition of a stationary noise process and the signal of interest. Both the parameter estimation<sup>10</sup> and testing are performed by analysis of data in the frequency domain using a finite Fourier transform.

The purpose of this hypothesis testing is to verify estimated parameters of environmental models that were obtained previously by the inversion of observed data.<sup>11</sup> The inversion was based on data from a calibrated vertical array and carried out using information at multiple frequencies from a single broadband source. A range-dependent adiabatic normal mode code<sup>12</sup> was used as the forward model. The global optimization was implemented by a directed Monte Carlo search based on genetic algorithms (GA) and

<sup>a)</sup>Present address: Siemens AG, A-1101 Vienna, Austria. Electronic mail: christoph.mecklenbraeuer@siemens.at

<sup>b)</sup>Electronic mail: gerstoft@mpl.ucsd.edu

<sup>c)</sup>Electronic mail: {boehme,pjc}@sth.ruhr-uni-bochum.de

the Bartlett objective function.<sup>10</sup> The inversion included the estimation of all important forward model parameters, which can be divided into geometric, geoacoustic, and ocean sound-speed variables. From analyzing the *a posteriori* parameter distributions of the GA, it is known that not all parameters can be equally well estimated.<sup>10</sup>

In order not to bias the test, a different dataset is used for testing and estimation, but the datasets are based on the same acquisition. The Fourier-transformed data is split in two subsets for estimation and testing in a comb-like fashion. In this way, the whole frequency range is available to both stages of processing—adjacent frequency bins are merely spaced farther apart.

Global search necessitates a huge number of forward solutions to be computed, vastly limiting the number of frequencies used in the estimation procedure. The test can verify an obtained model against a set of alternatives by incorporating data from more frequencies than were used to estimate the associated environmental parameters. A set of alternative environmental models is readily available from the inversion itself and given by the final populations of the GA.

The definition of objective functions for environmental parameter estimation, and the choice of test statistics in hypothesis testing using multifrequency data is still under discussion.<sup>13,14</sup> Exploiting the asymptotic Gaussianity of data in the frequency domain allows defining approximate likelihood functions, which can be maximized for the parameter estimation, and can be used for hypothesis tests based on likelihood ratios.<sup>13,14</sup> The proposed likelihood ratio test (LRT) is based on multifrequency data and exploits the asymptotic Gaussianity of short-time Fourier-transformed measurement data.

This test is related to a solution to the detection problem in passive sonar, seismics, and radar applications using a multiple sequential  $F$  test that is based on a frequency–domain regression.<sup>15–17</sup> The estimated signal-to-noise ratio (SNR) for the observed data in the frequency domain turns out to be the basic quantity from which the LRT is constructed. This is very appealing from a physical point of view. In this paper, we define the SNR in the frequency domain via the quotient of eigenvalues of the spectral density matrix (SDM) of the sensor outputs. In Sec. IID we give a more detailed discussion of these quotients in signal space.

The LRT compares the geometric means over the frequency of the estimated noise spectral levels under the hypothesis and alternatives. In the case of only one single source, this is related to the sum of Bartlett powers in dB.

Monte Carlo simulations are performed for a numerical analysis of the proposed algorithm. The behavior of the combined inversion and hypotheses testing is computed over the SNR.

Finally, the test is applied to broadband data measured using a vertical array near the island of Elba in the Mediterranean Sea.<sup>11</sup> It is assumed that optimum parameter estimates for a given environmental model have already been found by using a global optimization approach.<sup>10</sup>

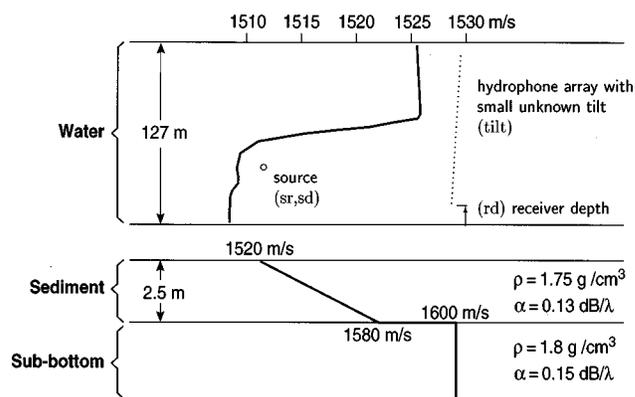


FIG. 1. Range-independent baseline model for acoustic propagation in the shallow ocean north of Elba island; see Ref. 19 for further details.

## I. PROPAGATION AND DATA MODEL

We consider a single wide-band source and the usual linear model in frequency domain for the output of an array of  $N$  sensors with spatially uncorrelated additive noise,

$$\mathbf{x}(\omega) = \mathbf{d}(\omega; \mathbf{m})s(\omega) + \mathbf{u}(\omega). \quad (1)$$

In this paper, vectors are denoted by lowercase boldface, matrices in uppercase boldface, and  $*$  is the Hermitian transpose operation. Let  $\mathbf{x}$  be the Fourier transform of the output of the vertical sensor array. Correspondingly the source signal is denoted by  $s$  and the noise by  $\mathbf{u}$ . The transfer function  $\mathbf{d}(\omega; \mathbf{m}) \in \mathbb{C}^N$  is calculated by the SNAP forward model code, which calculates the Green’s function  $G$  for the ocean acoustic frequency–domain wave equation.<sup>12,18</sup>

The source/receivers geometry and geoacoustic environment for the baseline model are shown in Fig. 1 for a range independent environment and in Fig. 2 for a range dependent environment. The source signal  $s$  and the transfer function  $\mathbf{d}(\omega; \mathbf{m})$  are assumed to be deterministic. The additive noise in time domain is stochastic, stationary, and zero mean.

The data  $\mathbf{x}(\omega)$  at the receiving array are asymptotically complex normal distributed with mean  $\mathbf{d}(\omega; \mathbf{m})s(\omega)$  and covariance  $\mathbf{C}_u(\omega) = \nu(\omega)\mathbf{I}$ , where  $\nu(\omega) > 0$  denotes the power spectral density of the noise. The vector-valued transfer function  $\mathbf{d}(\omega; \mathbf{m})$  depends nonlinearly on the environmental parameters, which are summarized in a vector  $\mathbf{m}$ . In this paper, the dimension of  $\mathbf{m}$  may vary with the structure of the model. The *true model* corresponding to the data is characterized by the  $r$ -dimensional parameter vector  $\mathbf{m}^*$ .

We use the word “model” for the geoacoustic parameters of an ocean acoustic wave guide with given structure. Let us denote the set of all models by  $\mathcal{M}$ . The set of all replica vectors  $\mathbf{d}(\omega; \mathbf{m})$  that can be calculated by SNAP for all models is denoted by  $\mathcal{D}$ . The models are indexed by the parameter vector  $\mathbf{m}$ , which is an element of the associated parameter set  $\mathcal{M} \subset \mathbb{R}^r$ . We assume  $\mathbf{m}^* \in \mathcal{M}$ . The *mapping* between the parameter set  $\mathcal{M}$  and the replicas  $\mathcal{D}$  is termed a *model structure*. In Sec. III the sets  $\mathcal{M}$  and  $\mathcal{D}$  will be given a hierarchical buildup.

## II. LIKELIHOOD RATIO TEST

Considering the model (1) at a discrete set of frequencies  $\omega_j$  ( $j=1,\dots,J$ ), it can be shown that the log-likelihood function is given by

$$\mathcal{L}(\mathbf{m}) = -\frac{1}{J} \sum_{j=1}^J \log(1 - B(\omega_j; \mathbf{m})), \quad (2a)$$

$$B(\omega; \mathbf{m}) = \frac{\mathbf{d}^*(\omega; \mathbf{m}) \hat{\mathbf{C}}_x(\omega) \mathbf{d}(\omega; \mathbf{m})}{|\mathbf{d}(\omega; \mathbf{m})|^2 \text{tr} \hat{\mathbf{C}}_x(\omega)}, \quad (2b)$$

where  $\hat{\mathbf{C}}_x(\omega)$  denotes a nonparametric estimate of the spectral density matrix (SDM) of the data.<sup>20</sup> We use the sample mean

$$\hat{\mathbf{C}}_x(\omega) = \frac{1}{KL} \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} \mathbf{x}_{k,l}(\omega) \mathbf{x}_{k,l}^*(\omega) \quad (3)$$

of  $KL$  Fourier-transformed data snapshots  $\mathbf{x}_{k,l}(\omega)$ . The observed data  $\mathbf{x}(t)$  in time domain  $t=0,1,\dots,KT-1$  is divided into  $K$  snapshots of duration  $T$  each. These are Fourier transformed using  $L$  orthogonal windows  $\nu_i^{(l)}$  depending on snapshot duration  $T$  and selected analysis bandwidth  $W$ . Here, we used discrete prolate spheroidal sequences as data tapers.<sup>21–23,14</sup> The bandwidth of resolution  $0 < W < \frac{1}{2}$  is selected such that (3) is nonsingular. The number  $L$  of data tapers in (3) is essentially equal to the time–bandwidth product:  $L = \lceil 2WT - 1 \rceil$ , where  $\lceil x \rceil$  is the largest integer smaller than or equal to  $x$ . The Fourier-transformed snapshots are calculated for  $k=0,\dots,K-1$  and  $l=0,\dots,L-1$  by

$$\mathbf{x}_{k,l}(\omega) = \sum_{j=0}^{T-1} \nu_j^{(l)} \mathbf{x}(t+kT) e^{-j\omega t}.$$

The maximum-likelihood (ML) estimate for the parameter vector  $\mathbf{m}$  of  $\mathcal{M}$  is defined as

$$\hat{\mathbf{m}} = \arg \max_{\mathbf{m} \in \mathcal{M}} \mathcal{L}(\mathbf{m}). \quad (4)$$

Note, that the log-likelihood function can be easily interpreted as the *averaging of the Bartlett power over frequency bins in dB*. This approach of averaging in log units was empirically found to have optimal side lobe suppression.<sup>24</sup>

After estimates have been obtained in this way, the resulting models can be validated using hypotheses tests. We start with an introductory example for which the likelihood ratio test is known to be optimal in Sec. II A. However, this simple setting does not apply here. Proceeding to Sec. II B, we formulate the test problem that is appropriate for validation. It turns out that a straightforward implementation of the LRT is not feasible and a possible solution is presented in Sec. II C.

### A. Simple hypothesis and alternative

We test which one of two models  $\mathcal{M}_0 = \{\mathbf{m}_0\}$ ,  $\mathcal{M}_1 = \{\mathbf{m}_1\}$  has generated the measurement data  $\mathbf{x}(\omega)$ . In this test problem, the hypothesis and alternative are both of the *simple type*, i.e., they do not contain unknown parameters. The Neyman–Pearson Lemma applies, dictating that the most powerful test is based on the likelihood ratio.<sup>7</sup> The

hypothesis  $H_0$  states “ $\mathcal{M}_0$  is true” and the alternative  $H_1$  reads “ $\mathcal{M}_1$  is true.” The LRT of  $H_0$  against  $H_1$  constructs the test statistic  $T = \mathcal{L}(\mathbf{m}_0) - \mathcal{L}(\mathbf{m}_1)$  and compares its value with a critical value  $t_\alpha$ , which depends on the chosen level  $\alpha$  of the test.

### B. Validating a specified model against a set of alternatives

A more interesting setting is to validate a particular selected model against a set of competing models with the same parametrization, deciding whether  $\mathbf{m}_0 \in \mathcal{M}_0 = \{\mathbf{m}^*\}$  or  $\mathbf{m}_0 \in \mathcal{M}_1 = \mathcal{M} \setminus \mathcal{M}_0$ . The model  $\mathcal{M}_0$  should be compared with all competing models inside the alternative  $\mathcal{M}_1$ , which contains all models except for  $\mathbf{m}_0$ . The test uses the quantity  $t(\mathbf{x}) = \mathcal{L}(\mathbf{m}_0) - \max_{\mathbf{m}_1 \in \mathcal{M}_1} \mathcal{L}(\mathbf{m}_1)$ . We are dealing with a *composite* alternative, and, therefore, optimality of the LRT cannot be guaranteed. In terms of the Bartlett Power, the test statistic is written as

$$t(\mathbf{x}) = \min_{\mathbf{m}_1 \in \mathcal{M}_1} \frac{1}{J} \sum_{j=1}^J \log \frac{1 - B(\omega_j; \mathbf{m}_1)}{1 - B(\omega_j; \mathbf{m}_0)}. \quad (5)$$

The LRT now compares this quantity with a predetermined threshold value  $t_\alpha$  that depends on the level  $\alpha$  of the test and the distribution of the test statistic (5).

In general, there exists a set  $\mathcal{K}_\alpha \subset \mathcal{M}_1$  of parameter vectors for which  $H_0$  is rejected. This  $\alpha$ -critical region can be viewed as a confidence region for  $\mathbf{m}$ : parameter vectors  $\mathbf{m} \in \mathcal{K}_\alpha$  cannot be rejected against  $H_0$ . By sampling  $\mathcal{M}_1$  and repeating the test for each individual  $\mathbf{m} \in \mathcal{M}_1$ , numerical approximations to  $\mathcal{K}_\alpha$  are obtained.

Unfortunately, this direct approach is not feasible, due to the incomplete knowledge about the distribution of the test statistic. The fraction inside (5) is the ratio of two  $\chi^2$ -distributed random variables, but they are not independent. This problem arises because we are testing a particular  $\mathbf{m}_0 \in \mathbb{R}^r$  against all alternatives with the same parametrization: they are of the same order  $r$ .

### C. Sequential LRT using three steps

The difficulties described in the previous section disappear when we test a *smaller* against a *bigger* model.<sup>7</sup> They are avoided by testing the models in a three step sequential procedure. We will use the following hypotheses  $H_i$  and alternatives  $A_i$  ( $i=1,2,3$ ):

$H_1$ : no signal in the data,  $A_1$ :  $\mathbf{d}_0 \in \mathcal{D}_0$  or  $\mathbf{d}_1 \in \mathcal{D}_1$  generated the data,

$H_2$ : replica vector  $\mathbf{d}_0 \in \mathcal{D}_0$  generated the data,

$A_2$ : the data cannot be adequately modeled by  $\mathbf{d}_0$ ,

$H_3$ : replica vector  $\mathbf{d}_1 \in \mathcal{D}_1$  generated the data,

$A_3$ : the data cannot be adequately modeled by  $\mathbf{d}_1$ .

We have omitted the dependency on  $\omega$  in notation for all quantities,  $\mathbf{d}_0 = \mathbf{d}(\omega; \mathbf{m}_0)$ , and  $\mathbf{d}_1 = \mathbf{d}(\omega; \mathbf{m}_1)$ .

Some comment is necessary on  $A_1$ : physically, we know that it is not possible for two *differing* geoacoustic environments to be correct at the same time. So in  $A_1$ , the logical “or” can be substituted by an “exclusive or” opera-

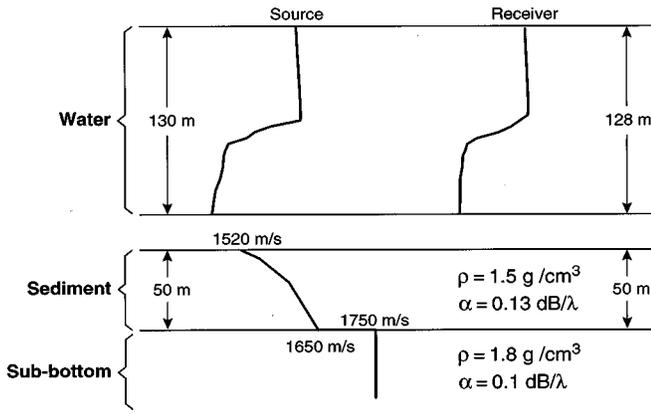


FIG. 2. Range-dependent geoacoustic baseline model with differing properties at the source and receiver locations.

tion. For the algorithm, however, it is more suitable to implement the “inclusive or” which is easier to formulate and makes no logical difference in this case.

In mathematical terms, the above  $H_i$ ,  $A_i$  are formulated as

$$H_1: \mathbf{x} = \mathbf{u},$$

$$A_1: \mathbf{x} = (\mathbf{d}_0 \quad \mathbf{d}_1) \begin{pmatrix} s_0 \\ s_1 \end{pmatrix} + \mathbf{u}, \quad \text{with } |s_0|^2 + |s_1|^2 \neq 0;$$

$$H_2: \mathbf{x} = \mathbf{d}_0 s_0 + \mathbf{u}, \quad \text{with arbitrary } s_0 \neq 0$$

$$A_2: \mathbf{x} = (\mathbf{d}_0 \quad \mathbf{d}_1) \begin{pmatrix} s_0 \\ s_1 \end{pmatrix} + \mathbf{u}, \quad \text{with } s_1 \neq 0 \quad \text{and} \quad s_0 \neq 0;$$

$$H_3: \mathbf{x} = \mathbf{d}_1 s_1 + \mathbf{u}, \quad \text{with arbitrary } s_1 \neq 0$$

$$A_3: \mathbf{x} = (\mathbf{d}_0 \quad \mathbf{d}_1) \begin{pmatrix} s_0 \\ s_1 \end{pmatrix} + \mathbf{u}, \quad \text{with } s_0 \neq 0 \quad \text{and} \quad s_1 \neq 0.$$

*Step 1:* First, we test  $H_1$  against the alternative  $A_1$ , and if the hypothesis is rejected, we conclude that *the data contain a signal, and at least one of the models will be correct*; we proceed with Step 2. If the hypothesis is accepted the test stops at this point. In this case, we have an identifiability problem due to lack of signal power.

*Step 2:* We test the hypothesis  $H_2$  with arbitrary  $s_0 \neq 0$  against the alternative  $A_2$ . If the hypothesis is accepted, we conclude that  $H_2$  is true and the test stops here. If, on the other hand,  $H_2$  is rejected, we go to Step 3.

*Step 3:* This is a cross-check: we test the hypothesis  $H_3$  with arbitrary  $s_1 \neq 0$  against the alternative  $A_3$ .

In each step, the test statistics  $t_1(\mathbf{x})$ ,  $t_2(\mathbf{x})$ , and  $t_3(\mathbf{x})$  can be put into the form

$$t_i(\mathbf{x}) = -\frac{1}{J} \sum_{j=1}^J \log \left( 1 + \frac{n_1}{n_2} V_i(\omega_j) \right), \quad (6)$$

with

$$V_i(\omega) = \frac{n_2}{n_1} \frac{\text{tr}[(\mathbf{P}_3(\omega) - \mathbf{P}_{i-1}(\omega)) \hat{\mathbf{C}}_x(\omega)]}{\text{tr}[(\mathbf{I} - \mathbf{P}_3(\omega)) \hat{\mathbf{C}}_x(\omega)]}, \quad (7)$$

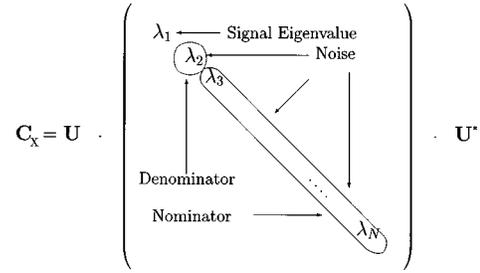


FIG. 3. An interpretation of the test statistic and spectral density matrix eigenvalues.

where the projection matrices  $\mathbf{P}_{i-1}(\omega)$  ( $i=1,2,3$ ) are associated with the signal subspaces under  $H_i$  in each step of the test; see Sec. 1 of the Appendix.

$$\mathbf{P}_0(\omega) = \mathbf{0},$$

$$\mathbf{P}_1(\omega) = \mathbf{d}_0 \mathbf{d}_0^* / |\mathbf{d}_0|^2,$$

$$\mathbf{P}_2(\omega) = \mathbf{d}_1 \mathbf{d}_1^* / |\mathbf{d}_1|^2, \quad (8)$$

$$\mathbf{P}_3(\omega) = \mathbf{Q} \mathbf{Q}^*, \quad \text{where } \mathbf{Q} = \text{orth}(\mathbf{d}_0, \mathbf{d}_1).$$

The integers  $n_1, n_2 \in \mathbb{N}$  are explained in Sec. II E. The projection matrix  $\mathbf{P}_3$  is associated with the alternatives: it has rank two. The signal space under the alternatives is spanned by both  $\mathbf{d}_0$  and  $\mathbf{d}_1$ . If the geoacoustic model is identifiable, then it is asserted that the vectors  $\mathbf{d}_0$  and  $\mathbf{d}_1$  are linearly independent. Matrix  $\mathbf{Q}$  contains two columns being a unitary basis of the two-dimensional space spanned by  $\mathbf{d}_0$  and  $\mathbf{d}_1$ .

This test strategy ensures that the difference of the projection matrices in the denominator of (6) is a projection operator that is orthogonal to  $\mathbf{I} - \mathbf{P}_3$ . This makes  $V_i(\omega)$  an  $F_{n_1, n_2}$ -distributed random variable.

#### D. Interpretation in signal space

The test statistics can be easily interpreted in the signal/noise space analogy. The following discussion is for *Step 2* of the test, but the reasoning for the third step is completely equivalent. Figure 3 describes graphically what eigenvalues of the spectral density  $\mathbf{C}_x$  are grouped together in the numerator and denominator of the  $F$  variable (7) if the hypothesis  $H_2$  is true. In this discussion, we replace  $\hat{\mathbf{C}}_x(\omega_j)$  in (6) by its expectation  $\mathbf{C}_x(\omega_j)$  to make the analogy more clear. The frequency dependence of the quantities will be suppressed in notation for convenience.

In the equation described in Fig. 3,  $\mathbf{U}$  denotes the unitary eigenvector matrix of  $\mathbf{C}_x$ . The dominant eigenvalue (EV) corresponding to the signal eigenvector  $\mathbf{d}_0$  is denoted by  $\lambda_1$ . The noise EVs are given by  $\lambda_2, \dots, \lambda_N$ . The projection matrix  $\mathbf{P}_3$  projects onto a two-dimensional subspace of  $\mathbb{C}^N$ . The dominant eigenvector of  $\mathbf{C}_x$  and one of the noise eigenvectors are a basis of this subspace. By resorting the eigenvectors in  $\mathbf{U}$ , we can always achieve that  $\mathbf{P}_3$  is associated with the first two eigenvectors—and thus with  $\lambda_1, \lambda_2$  in the following sense:

$$\text{tr}[\mathbf{P}_3 \mathbf{C}_x] = \lambda_1 + \lambda_2 = |\mathbf{d}_0|^2 |s|^2 + 2\nu.$$

On the other hand,  $\mathbf{P}_1$  is associated with the dominant signal EV  $\lambda_1$  alone,

$$\text{tr}[\mathbf{P}_1 \mathbf{C}_x] = \lambda_1 = |\mathbf{d}_0|^2 |s|^2 + \nu.$$

And the projection matrix of the difference projector  $\mathbf{P}_3 - \mathbf{P}_1$  is associated with the single noise EV  $\lambda_2$  only. The nominator in (6) is associated with all the other noise EVs,

$$\text{tr}[(\mathbf{I} - \mathbf{P}_3) \mathbf{C}_x] = \lambda_3 + \dots + \lambda_N = (N - 2)\nu.$$

Thus, we see that the signal power is canceled out of the  $F$  variable (7) if the hypothesis  $H_1$  is true:

$$\frac{n_1}{n_2} V_2 \approx \frac{\text{tr}[(\mathbf{P}_3 - \mathbf{P}_1) \mathbf{C}_x]}{\text{tr}[(\mathbf{I} - \mathbf{P}_3) \mathbf{C}_x]} = \frac{1}{N - 2}.$$

In this way, the test statistic becomes a pivot, i.e., independent of the unknown parameters.<sup>25</sup> Somewhat loosely stated, the test statistic compares one selected noise EV to the arithmetic mean of all the other noise EV. If this ratio exceeds some predetermined threshold, we conclude that the selected replica vector  $\mathbf{d}_0$  is not the only dominant eigenvector of  $\mathbf{C}_x$ , and the hypothesis  $H_2$  is rejected. If the alternative  $A_2$  is true,  $V_2$  depends on  $|s|^2/\nu$ .

In the nonasymptotic case,  $V_i$  in (7) can be interpreted as the increase of the signal-to-noise ratio, if the model is enlarged by the alternative  $A_i$ .

### E. Degrees of freedom

The degrees of freedom (DOF)  $n_1, n_2$  of the  $F_{n_1, n_2}$ -distributed random variable are given by<sup>13</sup>

$$\text{in Step 1: } n_1 = 4KL, \quad n_2 = KL(2N - 4),$$

$$\text{in Steps 2 and 3: } n_1 = 2KL, \quad n_2 = KL(2N - 4),$$

if  $\mathbf{m}_1$  is not estimated from the same data where the test statistic is based upon. The factor  $KL$  stems from Eq. (3). This was ensured by using a different and larger set of frequency bins  $\omega_j$  for the test than for the inversion.

The ratio  $n_1/n_2 = 1/(N - 2)$  is equal to the ratio of the signal- and noise-subspace dimensions. Asymptotically, i.e., for a large observation time,  $V_i(\omega_j)$  and  $V_i(\omega_k)$  are independent if  $\omega_j \neq \omega_k$ .

We must be more careful in the case when  $\mathbf{m}_1$  is itself dependent on the data. This is the case if  $\mathbf{m}_1$  is estimated from the measurement data. In this case, the DOF are given by the more complicated formulas:

$$\text{in Step 1, } n_1 = KL(r_1 + r_2 + 4),$$

$$n_2 = KL(2N - (r_1 + r_2 + 4));$$

$$\text{in Steps 2 and 3, } n_1 = KL(r_1 + r_2 + 2),$$

$$n_2 = KL(2N - (r_1 + r_2 + 4)),$$

where  $r_1, r_2$  denote the number of environmental parameters (i.e., the dimension of  $\mathbf{m}_1, \mathbf{m}_2$ ).<sup>13</sup> This can be circumvented easily by splitting the data into two disjoint sets for the purpose of estimation and test. We can either straightforwardly split in the time domain or exploit asymptotic independence in the frequency domain for adjacent frequency bins. In the first case, weak effects of nonstationarity in the background

noise could lead to trouble. In the latter case, we recommend an interleaved scheme: a comb-like separation of estimator and test frequencies. If such an approach is feasible then the DOF do not depend on  $r_1$  or  $r_2$ .

Alternatively, the DOF might be obtained from the Fischer Information Matrix of the unknown parameters  $\mathbf{m}$ . The ‘‘nonlinear DOF’’  $r_p$  are strongly related to the numerical rank of the Fischer Information Matrix.<sup>26</sup>

### F. Calculation of thresholds

The test statistics from (6) are easily interpreted as the arithmetic mean of a sample of independent identically distributed random variables  $T_j = \log(1 + (n_1/n_2)V_i(\omega_j))$  whose probability density and cumulants can be evaluated in closed form (see Sec. 2 of the Appendix), e.g., for mean and variance, we obtain

$$\begin{aligned} \mu_T &= \Psi\left(\frac{n_1 + n_2}{2}\right) - \Psi\left(\frac{n_2}{2}\right) \approx \log\left(1 + \frac{n_1}{n_2}\right), \\ \sigma_T^2 &= \Psi'\left(\frac{n_2}{2}\right) - \Psi'\left(\frac{n_1 + n_2}{2}\right) \approx \frac{2n_1}{n_2(n_1 + n_2)}. \end{aligned} \quad (9)$$

Here,  $\Psi(x) = d \log \Gamma(x)/dx$  and  $\Psi'(x) = d\Psi(x)/dx$  are polygamma functions.<sup>27</sup> The  $\alpha$ -critical value  $t_\alpha$  can be derived from a normal approximation for the distribution  $F_T(t)$  of (6) for a large number of frequencies  $J$  using the inverse error function<sup>27</sup>

$$t_\alpha = \mu_T + \frac{\sigma_T \sqrt{2}}{\sqrt{J}} \text{erf}^{-1}(2\alpha - 1). \quad (10)$$

A more accurate value for  $t_\alpha$  can be obtained via the Cornish–Fisher expansion of the inverse cumulative distribution function of the test statistic<sup>28</sup> or bootstrapping.<sup>25</sup> Experimentally, it was shown that the performance of the test does not improve using these more elaborate approximations.

### III. MODEL STRUCTURE IDENTIFICATION

The replica-vector set  $\mathcal{D}$  is now given an additional hierarchical buildup. We define an increasing sequence  $\mathcal{D}_1, \mathcal{D}_2, \dots$ , of subsets, such that

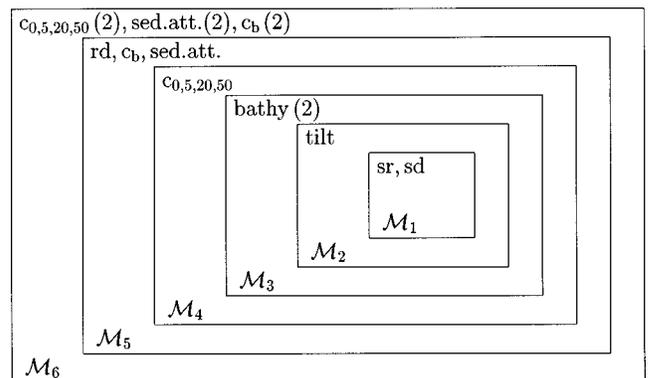


FIG. 4. The box-like structure for geoacoustic models  $\mathcal{M}_1, \dots, \mathcal{M}_6$  used for the inversion of real data. The meaning of the parameter name abbreviations is given in Table I.

TABLE I. The models used for the inversion (Ref. 9) of real data. BP is Bartlett power in dB from (13); the maximum is zero.  $r$  is the number of parameters over which optimization was performed. The notation (2) denotes that a value of this parameter was estimated for both the source and receiver range. Inversion parameters: source range, source depth, bathymetry, array tilt, velocity profile, receiver depth of deepest hydrophone in array, bulk velocity in sediment, sediment attenuation.

Model	$r$	BP	Inversion parameters $\mathbf{m}$
$\mathcal{M}_1$	2	-0.41	sr, sd
$\mathcal{M}_2$	4	-0.28	sr, sd, bathy.(2)
$\mathcal{M}_3$	5	-0.25	sr, sd, bathy.(2), tilt
$\mathcal{M}_4$	9	-0.23	sr, sd, bathy.(2), tilt, $c_{0,5,20,50}$
$\mathcal{M}_5$	12	-0.23	sr, sd, bathy.(2), tilt, $c_{0,5,20,50}$ , rd, $c_b$ , sed. att.
$\mathcal{M}_6$	18	-0.22	sr, sd, bathy.(2), tilt, rd, $c_{0,5,20,50}$ (2), sed. att. (2), $c_b$ (2)

$$\mathcal{D}_1 \subset \mathcal{D}_2 \subset \dots \subset \mathcal{D}_p \subset \dots \subset \mathcal{D} \subset \mathbb{C}^N.$$

The model subsets  $\mathcal{D}_p$  are associated with parameter sets  $\mathcal{M}_p$  because each replica vector  $\mathbf{d}(\cdot; \mathbf{m}_p) \in \mathcal{D}_p$  is indexed by a parameter vector  $\mathbf{m}_p \in \mathcal{M}_p \subset \mathbb{R}^{r_p}$ . The dimension  $r_p \in \mathbb{N}$  of the parameter set  $\mathcal{M}_p$  increases monotonically with  $p$ .

For two arbitrary models, we say ‘‘model structure  $\mathcal{M}_p$  is included in  $\mathcal{M}_q$ ’’ iff  $\mathcal{D}_p \subset \mathcal{D}_q$ . We write  $\mathcal{M}_p < \mathcal{M}_q$ . We can think of the *smaller* model structure  $\mathcal{M}_p$  being generated from  $\mathcal{M}_q$  by *freezing* some elements of  $\mathbf{m}_2$  to constant nominal values of a *baseline model*.

As a specific example, consider Fig. 4. The ‘‘Russian-doll’’-like figure gives an intuitive visualization of the problem. The meaning of the parameter name abbreviations is given in Table I. Model  $\mathcal{M}_1$  only includes the two source coordinates (sr, sd) as unknown parameters; all other physical quantities are given from the reference model, Fig. 1. The next model additionally includes the tilt of the receiving (nominally vertical) array (tilt is defined as horizontal displacement between the bottom and top hydrophone in the array). Model  $\mathcal{M}_3$  is enlarged to include two parameters for bathymetry: the ocean depth at the source and receiver location.

The order of the parameters  $\mathbf{m}_p = (m_1, \dots, m_p) \in \mathcal{M}_p$  in the  $p$ th model structure  $\mathcal{M}_p$  seems to be completely arbitrary at first. However, the acoustician has an intuitive idea about

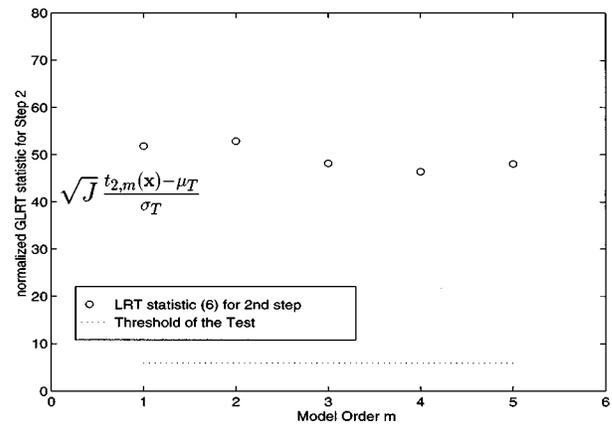


FIG. 5. The standardized test statistic  $t_{2,m}$  for each stage in the model hierarchy as applied to the real data. The horizontal index  $m$  denotes the test of hypothesis  $H_{2,m} = \mathcal{M}_{p-1}$  against the alternative  $A_{2,m} = \mathcal{M}_{p-1}$  or  $\mathcal{M}_p$ . For details see Table I. The dashed horizontal line shows the threshold value  $t_\alpha$  for a false-alarm rate of  $\alpha = 5\%$ .

the ‘‘relative importance’’ of the individual parameters. In this way, the question arises as to how to design the hierarchy.

The rigorous approach consists in fixing one parameter of the highest-order model  $\mathcal{M}_p$  to the nominal value of the baseline model. After this has been done for all parameters in turn, the very set of  $P - 1$  parameters is selected, which gives the *least degradation of the maximum-likelihood value*. This procedure is repeated for  $\mathcal{M}_{p-1}, \mathcal{M}_{p-2}, \dots$ , and so on. This approach is computationally prohibitive for a practical numbers of inversion parameters.

The experienced acoustician can get help from two sources: first, he can calculate the Fischer information matrix for the parameters at their *nominal values*: this gives an indication of variances in the single parameter estimates and their mutual correlations. This could be called the *local* approach and should be repeated for several model vectors. Finally, the *a posteriori* distributions of Genetic Algorithms<sup>10,14</sup> also indicate spreading of the estimates: this is a practical *global* approach for hierarchy construction.

TABLE II. Models used for simulation (Ref. 28).  $r$  is the number of parameters over which optimization was performed. Parameters: source range, source depth, bathymetry, array tilt. Unspecified parameters are given in Fig. 1. Model  $\mathcal{M}$  generated the simulated datasets.

Model	$r$	Inversion parameters $\mathbf{m}$	Search ranges	Fixed values
$\mathcal{M}_1$	2	sr, sd	5500 m < sr < 5700 m 75 m < sd < 85 m	tilt = 0 m bathy = 127 m
$\mathcal{M}_2$	3	sr, sd, tilt	5500 m < sr < 5700 m 75 m < sd < 85 m -3 m < tilt < 3 m	bathy = 127 m
$\mathcal{M}_3$	4	sr, sd, tilt, bathy	5500 m < sr < 5700 m -3 m < tilt < 3 m 125 m < bathy < 133 m	...
$\mathcal{M}$	...	...	...	sr = 5600 m, sd = 80 m tilt = 0.3 m, bathy = 130.5 m

## A. Identification algorithm using LRT

We can now formulate Algorithm 1 for the structure identification problem. It returns the estimated model order  $\hat{p}$  and the corresponding parameter  $\hat{\mathbf{m}}_p$ . We need hypotheses and alternatives analogous to Sec. II B for each model order  $m$ : they are denoted by  $H_{i,m}$ ,  $A_{i,m}$ , respectively. In particu-

lar,  $\mathbf{d}_0 \leftarrow \mathbf{d}(\cdot; \hat{\mathbf{m}}_p)$  and  $\mathbf{d}_1 \leftarrow \mathbf{d}(\cdot; \hat{\mathbf{m}}_{p+1})$  are substituted in Eq. (8). The algorithm starts with Step 1 for  $\mathcal{M}_1$ : this decides whether there is a significant input signal. This step is only needed once, whereas Steps 2 and 3 are repeated for all model orders.

Algorithm 1: model identification using a LRT.

```

 $\hat{\mathbf{m}}_1 := \arg \max_{\mathbf{m}_1 \in \mathcal{M}_1} \mathcal{L}(\mathbf{m}_1)$ 
if  $H_{1,1}$  cannot be rejected against  $A_{1,1}$  then
  stop “identifiability problem: we don’t know what  $p$  is”
main loop:
for  $p := 1, 2, \dots$  do
   $\hat{\mathbf{m}}_{p+1} := \arg \max_{\mathbf{m}_{p+1} \in \mathcal{M}_{p+1}} \mathcal{L}(\mathbf{m}_{p+1})$ 
  if  $H_{2,p}$  cannot be rejected against  $A_{2,p}$  then
     $\hat{p} := p$ , stop “this is a conservative selection”
  else
    if  $H_{3,p}$  cannot be rejected against  $A_{3,p}$  then
       $\hat{p} := p+1$ , stop
    else
       $\hat{p} \geq p+1$ , continue with the loop over  $p$ 
end of for loop
output  $\hat{p}, \hat{\mathbf{m}}_p$ 

```

## B. Akaike’s information criterion

Alternatively, we can adopt Akaike’s Information Criterion (AIC) approach to the order selection problem for acoustic models; cf. Ref. 7. In the present application, the AIC selects the model order  $\hat{m} \in \mathbb{N}$ , which minimizes the criterion:

$$\text{AIC}[m] = -\mathcal{L}(\hat{\mathbf{m}}_p) + \frac{r_p}{N}. \quad (11)$$

From this simple form of  $\text{AIC}[m]$ , we can directly calculate the required increase in likelihood to equalize the cost of additional parameters. The AIC prefers  $\mathcal{M}_{p+1}$  over  $\mathcal{M}_p$  iff

$$\mathcal{L}(\hat{\mathbf{m}}_{p+1}) - \mathcal{L}(\hat{\mathbf{m}}_p) > \frac{1}{N} (r_{p+1} - r_p). \quad (12)$$

## IV. EXAMPLES

The examples are all based on the environment from the North Elba sea trial;<sup>19</sup> see Figs. 1 and 2. In Sec. IV A the examples are based on synthetic data and in Secs. IV B and IV C on real data.

The experimental data were collected near the Island of Elba in the Mediterranean Sea during a sea trial in October 1993.<sup>11,4</sup> A vertical array of  $N=48$  sensors approximately spanning the whole water column of 128-m ocean depth was used. The signal source was located at a range of 5600 m and at depth 80 m below the surface. The transmitted signal was centered at 170 Hz and modulated with a pseudorandom noise sequence giving a total bandwidth of approximately 20 Hz. A record of one minute ( $6 \times 10^4$  samples at a sampling frequency of 1 kHz) of time samples was used. The SDM  $\hat{\mathbf{C}}_x(\omega)$  was estimated from  $K=15$  nonoverlapping windowed

Fourier-transformed snapshots of 4-s duration each. The 4000 samples per snapshot were zero padded to enable a FFT of length  $2^{12}$ .  $L=4$  data windows with bandwidth parameter  $W=6.25 \times 10^{-4}$  were used in (3), making  $KL=60$ . Adjacent frequency bins are separated by  $\Delta f = \frac{125}{512} \text{ Hz} \approx 0.244 \text{ Hz}$ .

## A. Simulation of sequential test

First the algorithm is used to determine the correct model order as a function of the SNR. Simulated datasets are generated as follows: The SNAP code calculates the replica vector for the range-independent environment  $\mathcal{M}$  given in Table II. Uncorrelated Gaussian noise is added to the replica for obtaining a desired SNR. For each SNR, we conduct 50 independent random experiments: each experiment is carried out with a new choice of signal and corrupted by noise that is independent of the other realizations. The data were generated by a selected model in the largest model structure considered and corrupted by additive noise at a prescribed SNR. In each experiment, we calculated the ML estimates and applied the proposed algorithm for the first model structures by global optimization of (2a) using a genetic algorithm, analogous to Ref. 10. For each SNR, this is repeated 50 times in order to make a statistical interpretation of the result. For each model order and SNR, 22 500 forward models were calculated during optimization.

The nonoptimized parameters in each model  $\mathcal{M}_p$  are given fixed values from the baseline model, Fig. 1. The baseline model is different from the *true* model, although its parameters are close guesses to the true values based on geological archives. If the fixed parameters were *true* it is clear that we would always favor a low-order model.

Monte Carlo estimates for the probabilities of the test

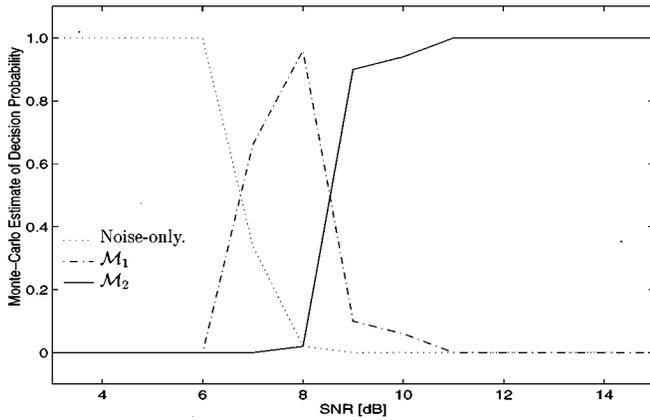


FIG. 6. The simulation of LRT  $\mathcal{M}_1$  against  $\mathcal{M}_2$  for various SNR levels.

decisions as a function of the SNR are given in Figs. 6 and 7. Typical threshold effects can be clearly observed in the figures: at a characteristic SNR value, the test decides with high probability in favor of the correct model order. In specific intervals of SNR, *downgraded* models are preferred to a full model containing all parameters. Each downgraded lower-order model has its own range in the SNR in which it is dominant, indicating identification problems for higher-order models at the corresponding SNR levels. Thus, in Fig. 6 for a SNR lower than 6 dB, there is not enough signal in the data to determine even the  $\mathcal{M}_1$  model, and above 10 dB there is enough information to determine at least the  $\mathcal{M}_2$  model. Around 8 dB, the model structure  $\mathcal{M}_1$  is best. In Fig. 7 we also include the model  $\mathcal{M}_3$ . It is seen that above 10-dB model  $\mathcal{M}_3$  is favored, and  $\mathcal{M}_2$  is only favored in a narrow region around 9 dB.

The corresponding results for the AIC criterion (12) are shown in Fig. 8. If the improvement in the maximum likelihood between  $\mathcal{M}_p$  and  $\mathcal{M}_{p+1}$  exceeds  $1/N = \frac{1}{48} \approx 2.083 \times 10^{-2}$  per additional parameter,  $\mathcal{M}_{p+1}$  is preferred. It is seen that this approach is not as conservative as the LRT, but AIC does not guarantee a false-alarm rate  $\alpha$ . We accept model  $\mathcal{M}_3$  down to a level of 6 dB and for a low SNR we accept model  $\mathcal{M}_2$ .

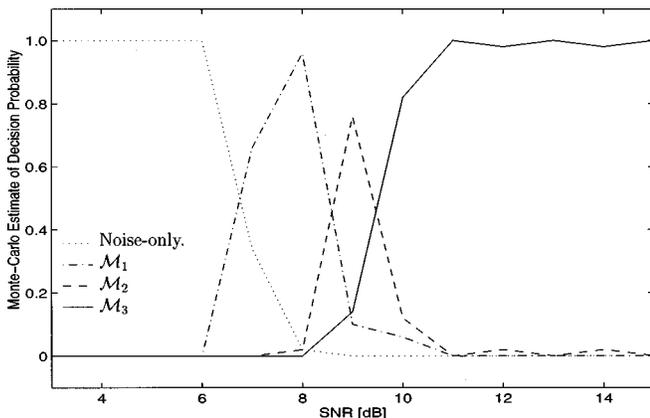


FIG. 7. The simulation of multiple sequential LRT for the first three models.

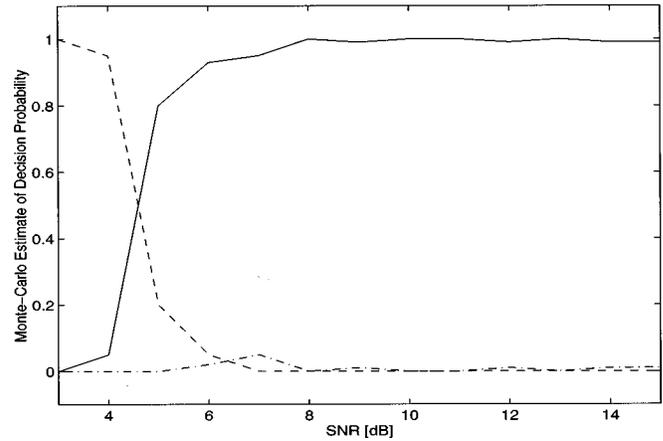


FIG. 8. The AIC simulation result based on 100 Monte Carlo runs.

## B. Application to North Elba data

For a set of six range-dependent ocean acoustic models,<sup>11</sup> the environmental parameters were estimated by a globally convergent stochastic algorithm using a frequency-domain measurement data at three distinct frequencies (167.72, 169.19, and 172.61 Hz). The search range used for each parameter is indicated in Table III. This rather slim broadband approach was selected for computational reasons, due to the large number of forward models that had to be calculated during the optimization: 20 000 forward models are calculated for each model order. In a second step, the obtained results were validated by the LRT using measurement data at the much broader frequency range 159.91–179.93 Hz, incorporating the measurement at 42–3 = 39 distinct frequency bins. From the selected frequencies  $\omega_j/2\pi = \Delta f(2j+653)$ , for  $j=1, \dots, 42$ , we discarded the three frequency bins mentioned before that were used for parameter inversion. The SNR in the data varies between about 33 dB at the center frequency and 22 dB at the ends of the range—which is quite high.

The models build a hierarchy of increasing complexity and model  $\mathcal{M}_p$  contains  $\mathcal{M}_{p-1}$  as a special case; see Table I,

TABLE III. GA inversion model with parameter search bounds. Each parameter was discretized into 128 values.

Model parameter	Lower bound	Upper bound
Geometric		
Source range (m)	5300	5900
Source depth (m)	72	82
Array tilt (m)	-3	3
Receiver depth (m)	110	115
Bathymetry-src (m)	127	134
Bathymetry-rcv (m)	127	134
Sediment		
Sound speed, $c_0$ (m/s)	1510	1560
Sound speed, $\delta_1$ (m/s)	0	100
Sound speed, $\delta_2$ (m/s)	0	100
Sound speed, $\delta_3$ (m/s)	0	100
Attenuation (dB/ $\lambda$ )	0	0.4
Bottom		
Sound speed, $\delta_4$ (m/s)	0	200

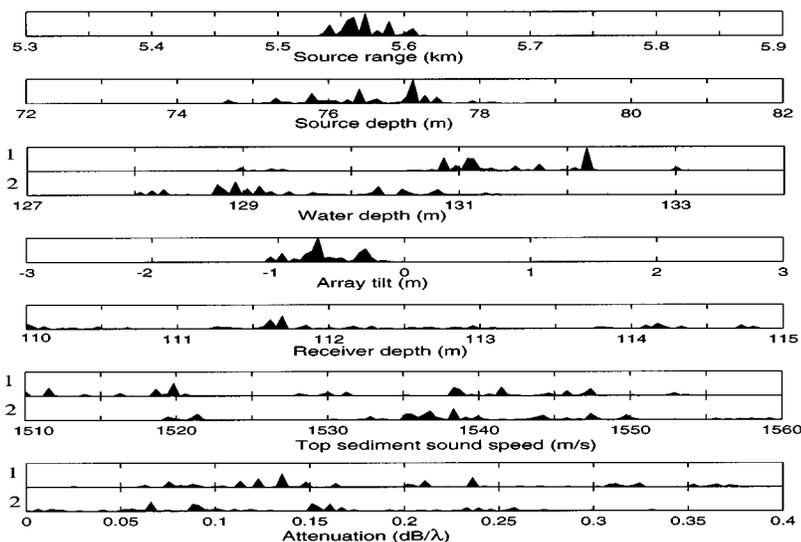


FIG. 9. The *a posteriori* distribution for the most important parameters for model  $\mathcal{M}_6$  using the combined band. 1 and 2 refers to the source and receiver environments, respectively.

where also the Bartlett Power (in dB) for the best estimated models are shown;

$$\text{BP}_p = 10 \log_{10} \sum_j [1 - B(\omega_j; \hat{\mathbf{m}}_p)],$$

with  $j \in \{687, 693, 707\}$ . (13)

Using Eq. (6) with  $H_{2,m}$  as a hypothesis and  $A_{2,m}$  as an alternative, the test statistics in Fig. 5 are computed. We cannot adopt both model structures  $\mathcal{M}_p$  and  $\mathcal{M}_{p+1}$  at the same time, *if they differ in their parameters*. This is what the test decides. If the test statistic  $t_2(\mathbf{x})$  is smaller than the  $\alpha$ -critical value (shown as a solid horizontal line), then  $\mathcal{M}_p$  is rejected. The  $\alpha$ -critical value for the test was found from the normal approximation given by Eq. (10). Figure 5 shows the result of a sequential test for identifying the correct model structure. The calculated LRT test statistics  $t_{2,m}$  for the model hierarchy  $\mathcal{M}_1 < \dots < \mathcal{M}_6$  are shown together with the threshold (dashed horizontal line) for  $\alpha = 5\%$ .

### C. Confidence regions and *posteriori* distributions

During the GA optimization, all obtained samples of the search space are stored and used to produce *a posteriori* probability distributions for the parameters. For a system of  $r$  parameters, the result is an  $r$ -dimensional space. This is difficult to display and therefore only marginal probability distributions are shown. The samples are ordered according to their energy, and the probability distribution is scaled using a Boltzmann distribution.<sup>10</sup> The *a posteriori* distributions are useful for evaluating the convergence of the inverse solution and uniqueness of the solution. While it does not give the model order it clearly indicates which parameters are important for the inversion; see also the comments in Sec. III.

An example of this is shown for the North Elba case.<sup>11</sup> For the  $\mathcal{M}_6$  model the *a posteriori* distributions for the more important parameters are illustrated in Fig. 9. It is seen that, except for the receiver depth, the geometric parameters were well determined. That is, all distributions were compact over the search interval and had well-defined single peaks. The model  $\mathcal{M}_3$  contains all geometric parameters, except for the

receiver depth. By interpretation of the *a posteriori* distributions from  $\mathcal{M}_6$ , we obtain an indication on the order of the parameters that should be incorporated into  $\mathcal{M}_p$  for  $p > 3$ .

In general, it would be expected that the source and receiver depths are about equally important. The source depth is much more important here because it is placed in the thermocline. Further, the search interval for the receiver depth is half of the search interval for the source depth.

### V. CONCLUSIONS

Hypothesis testing is a powerful tool as part of an acoustic inversion procedure. For a given model structure (set of parameters to estimate), the estimated values of parameters found by the optimization can be statistically justified. This is in clear contrast to the usual procedure, where we just arbitrarily select a set of parameters and then optimize the parameter values.

A multiple sequential LRT was applied to the model identification problem and shown suitable for the validation of estimated parameters. Auxiliary hypotheses and alternatives were needed to circumvent problems due to otherwise incomplete knowledge of the test statistics. In each step of the proposed sequential test, the computed LRT test statistic can be interpreted in terms of the estimated incremental SNR if the corresponding hypothesis is true. In the single frequency case, all cumulants of the test statistics are given analytically. In the broadband case, the distribution of the test statistics can be closely approximated by a Gaussian distribution.

The examples illustrate several uses of hypothesis testing. Before the test, an inversion algorithm should have determined the parameter estimates. This is usually done for a set of different model structures. A multiple sequential LRT can then be used to estimate the correct model structure, i.e., how many free parameters should be used in an inversion. It is illustrated that the *a posteriori* distributions obtained from the inversion can be used to order the set of parameters. This helps in designing the structure hierarchy.

How many parameters can be estimated for a given model structure depends on the SNR. The higher the SNR,

the more parameters we can afford to fit. Conversely, for poor SNR, it is best to focus on just the most important parameters. Otherwise, the higher-order parameters are just fitted to the noise (thus increasing the variance of the estimate).

By varying the SNR in Monte Carlo simulations, we can determine a minimum required SNR at the receiver array during a sea trial that will be needed to invert a given set of geoacoustic parameters. This was implemented using both hypothesis testing and Akaike's Information Criterion (AIC). From the simulations, it is observed that the LRT at a false-alarm rate of 5% gives similar results to the AIC, but is slightly more conservative.

Finally, we conclude with some remarks concerning the practical case where the signal space is not strictly rank one. We must distinguish two possible causes for signals with rank higher than one: First, the source might vibrate with higher degrees of freedom, e.g., multipole radiation.<sup>29</sup> A log-likelihood function can be formulated for this case also. Thus, this does not, in principle, present a problem to the LRT. Although the details of the presented algorithm will change. Second, the signal might decorrelate due to random environmental effects, losing coherence between the source and receivers. This case is more difficult. A log-likelihood function is not known to the authors in this case and, consequently, no LRT can be formulated.

## ACKNOWLEDGMENTS

This work was partly performed while CFM was a visiting scientist at SACLANTCEN. The authors wish to thank Dirk Maiwald, Georgios Haralabus, and Finn Jensen for valuable comments, inspirations, and discussions, and SACLANTCEN for providing the experimental data.

## APPENDIX

### 1. Derivation of the likelihood ratios, Eq. (6)

The log likelihoods for the hypotheses  $H_1$ ,  $H_2$ ,  $H_3$  and the alternatives are given by

$$\mathcal{L}_i = -\frac{1}{J} \sum_{j=1}^J \log \text{tr}[(\mathbf{I} - \mathbf{P}_{i-1}(\omega_j))\hat{\mathbf{C}}_x(\omega_j)].$$

The hypotheses are formulated by  $i=1,2,3$  and the alternatives by  $i=4$ . For the cases  $i=2,3$ , this reduces to (2a) up to an irrelevant additive constant; cf. Ref. 13. The LRT calculates differences of these log-likelihood functions. To be more specific: to test  $H_i$  against  $A_i$ , we need to construct

$$\stackrel{\text{Def}}{t_i(\mathbf{x})} := \mathcal{L}_4 - \mathcal{L}_i$$

$$= \frac{1}{J} \sum_{j=1}^J \log \frac{\text{tr}[(\mathbf{I} - \mathbf{P}_{i-1}(\omega_j))\hat{\mathbf{C}}_x(\omega_j)]}{\text{tr}[(\mathbf{I} - \mathbf{P}_3(\omega_j))\hat{\mathbf{C}}_x(\omega_j)]} \quad (i=1,2,3).$$

Inserting  $\mathbf{P}_3(\omega_j) - \mathbf{P}_3(\omega_j)$  into the denominator yields (6).

## 2. Proof of Eq. (9)

Let  $T = \log(1 + (n_1/n_2)F)$  be a random variable, where  $F$  is  $F_{n_1, n_2}$  distributed. Observe that the related random variable  $Z = 1/[1 + (n_1/n_2)F]$  is beta distributed with parameters  $p = n_2/2$  and  $q = n_1/2$ ; cf. Ref. 30. Knowing this, we can express the moment-generating function

$$\stackrel{\text{Def}}{M_T(s)} := \mathbb{E}[e^{Ts}] = \mathbb{E}[e^{-(\log Z)s}] = \mathbb{E}[Z^{-s}] = \frac{B(p-s, q)}{B(p, q)},$$

by means of the beta function. Using the identity  $B(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$ , the cumulants of  $T$  can be expressed by polygamma functions,

$$\begin{aligned} \kappa_r &:= \left. \frac{d^r}{ds^r} \log M_T(s) \right|_{s=0} \\ &= (-1)^r [\Psi^{(r-1)}(p) - \Psi^{(r-1)}(p+q)]. \end{aligned}$$

They are defined as derivatives of the log-gamma function:

$\stackrel{\text{Def}}{\Psi^{(r)}(x)} := (d^r/dx^r) \log \Gamma(x)$ ; see Refs. 25 and 24 for further details and approximations. The mean and variance  $\mu_T, \sigma_T^2$  in (9) are given by  $\kappa_1, \kappa_2$ , respectively. If both DOF  $n_1, n_2$  are even, this reduces to the finite sum

$$\kappa_r = \sum_{k=0}^{n_1/2-1} \frac{1}{(k+n_2/2)^r}.$$

- <sup>1</sup>M. J. Hinich, "Maximum-likelihood estimation of the position of a radiating source in a waveguide," *J. Acoust. Soc. Am.* **66**, 480–483 (1979).
- <sup>2</sup>A. B. Baggeroer, W. A. Kuperman, and P. N. Mikahalevsky, "An overview on matched field methods in ocean acoustics," *IEEE J. Ocean Eng.* **18**, 401–424 (1993).
- <sup>3</sup>A. Tolstoy, *Matched Field Processing for Underwater Acoustics* (World Scientific, Singapore, 1993).
- <sup>4</sup>J. L. Krolik, "The performance of matched-field beamformers with Mediterranean vertical array data," *IEEE Trans. Signal Process.* **44**, 2605–2611 (1996).
- <sup>5</sup>H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control.* **AC-19**, 716–723 (1974).
- <sup>6</sup>J. Rissanen, "A universal prior for integers and estimation by minimum description length," *Ann. Stat.* **11**, 416–431 (1983).
- <sup>7</sup>L. Ljung, *System Identification, Theory for the User*, Information and System Sciences Series (Prentice–Hall, Englewood Cliffs, 1987).
- <sup>8</sup>C. F. Mecklenbräuer and P. Gerstoft, "Hypothesis testing for acoustic environmental models using likelihood ratio," in *3rd European Conference on Underwater Acoustics*, edited by J. Papadakis (Crete U.P., Crete, 1996), pp. 465–470.
- <sup>9</sup>C. F. Mecklenbräuer, P. Gerstoft, P. J. Chung, and J. F. Böhme, "Generalized likelihood ratio test for selecting a geo-acoustic environmental model," *IEEE Proc. ICASSP-97*, 1997, Vol. I, pp. 463–466.
- <sup>10</sup>P. Gerstoft, "Inversion of seismoacoustic data using genetic algorithms and a posteriori distributions," *J. Acoust. Soc. Am.* **95**, 770–782 (1994).
- <sup>11</sup>P. Gerstoft and D. F. Gingras, "Parameter estimation using multifrequency range dependent acoustic data in shallow water," *J. Acoust. Soc. Am.* **99**, 2839–2850 (1996).
- <sup>12</sup>F. B. Jensen and M. C. Ferla, "SNAP: The SACLANTCEN normal-mode acoustic propagation model," SACLANT Undersea Research Centre, SM-121, La Spezia, Italy, 1979.
- <sup>13</sup>J. F. Böhme, "Statistical array signal processing of measured sonar and seismic data," *Proc. SPIE 2563 Advanced Signal Processing Algorithms* (San Diego), July 1995.
- <sup>14</sup>P. Gerstoft and C. F. Mecklenbräuer, "Ocean acoustic inversion with estimation of a posteriori probability distributions," *J. Acoust. Soc. Am.* **104**, 808–819 (1998).

- <sup>15</sup>D. Maiwald and J. F. Böhme, "Multiple testing for seismic data using bootstrap," IEEE Proc. ICASSP-94, 1994, Vol. 6, pp. 89–92.
- <sup>16</sup>H. Messer, "The potential performance gain in using spectral information in passive detection/localization of wideband sources," IEEE Trans. Signal Process. **43**, 2964–2974 (1995).
- <sup>17</sup>A. Ephraty, J. Tabrikian, and H. Messer, "Robust source detection in shallow water based on application of a test for spatial stationarity," IEEE Trans. Signal Process. (submitted).
- <sup>18</sup>F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (AIP Press, Woodbury, 1994).
- <sup>19</sup>D. F. Gingras and P. Gerstoft, "Inversion for geometric and geoacoustic parameters in shallow water: experimental results," J. Acoust. Soc. Am. **97**, 3589–3598 (1995).
- <sup>20</sup>C. F. Mecklenbräuker, D. Maiwald, and J. F. Böhme, "F-test in matched field processing: Identifying multimode propagation," IEEE Proc. ICASSP-95, 1997, Vol. 5, pp. 3123–3126.
- <sup>21</sup>D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty—Part V: The discrete case," Bell Syst. Tech. J. **57**, 1371–1429 (1978).
- <sup>22</sup>D. J. Thomson, "Jackknifing multiple-window spectra," Proc. IEEE ICASSP-94, 1994, Vol. 6, pp. 73–76.
- <sup>23</sup>D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications, Multitaper and Conventional Univariate Techniques* (Cambridge U.P., Cambridge, 1993).
- <sup>24</sup>A. B. Baggeroer, W. A. Kuperman, and H. Schmidt, "Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem," J. Acoust. Soc. Am. **93**, 2798–2808 (1993).
- <sup>25</sup>P. Hall, *The Bootstrap and Edgeworth Expansion* (Springer-Verlag, New York, 1992).
- <sup>26</sup>A. B. Baggeroer and H. Schmidt, "Parameter estimation theory bounds and the accuracy of full field inversions," in *Full Field Inversion Methods in Ocean and Seismo-Acoustics*, edited by O. Diachok, A. Caiti, P. Gerstoft, and H. Schmidt (Kluwer Academic, Dordrecht, 1995), pp. 79–84.
- <sup>27</sup>M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972), 9th ed.
- <sup>28</sup>P.-J. Chung, "Identifikation von geo-akustischen Ausbreitungsmodellen im Flachwasser mittels maximum-likelihood-quotiententests," Diplomarbeit, Ruhr-Universität Bochum, August 1996.
- <sup>29</sup>C. F. Mecklenbräuker, A. Waldhorst, P. Gerstoft, and G. Haralabus, "Matched field processing using multipole expansion," in *4th European Conference on Underwater Acoustics*, edited by A. Alippi and G. B. Cannelli, CNR-IDAC, 1998, pp. 15–20.
- <sup>30</sup>N. L. Johnson and S. Kotz, *Continuous Univariate Distributions—2, Applied Probability and Statistics* (Wiley, New York, 1970).