Trans-dimensional geoacoustic inversion

Jan Dettmer, Stan E. Dosso, and Charles W. Holland

Citation: The Journal of the Acoustical Society of America **128**, 3393 (2010); doi: 10.1121/1.3500674
View online: https://doi.org/10.1121/1.3500674
View Table of Contents: http://asa.scitation.org/toc/jas/128/6
Published by the Acoustical Society of America

Articles you may be interested in

Trans-dimensional matched-field geoacoustic inversion with hierarchical error models and interacting Markov chains
The Journal of the Acoustical Society of America **132**, 2239 (2012); 10.1121/1.4746016

Model selection and Bayesian inference for high-resolution seabed reflection inversion
The Journal of the Acoustical Society of America **125**, 706 (2009); 10.1121/1.3056553

Quantifying uncertainty in geoacoustic inversion. I. A fast Gibbs sampler approach
The Journal of the Acoustical Society of America **111**, 129 (2002); 10.1121/1.1419086

Sequential trans-dimensional Monte Carlo for range-dependent geoacoustic inversion
The Journal of the Acoustical Society of America **129**, 1794 (2011); 10.1121/1.3557052

Uncertainty estimation in simultaneous Bayesian tracking and environmental inversion
The Journal of the Acoustical Society of America **124**, 82 (2008); 10.1121/1.2918244

Parallel tempering for strongly nonlinear geoacoustic inversion
The Journal of the Acoustical Society of America **132**, 3030 (2012); 10.1121/1.4757639

Trans-dimensional geoacoustic inversion

Jan Dettmer[a)] and Stan E. Dosso
School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, V8W 3P6, Canada

Charles W. Holland
Applied Research Laboratory, The Pennsylvania State University, State College, Pennsylvania 16804-0030

(Received 7 July 2010; revised 13 September 2010; accepted 15 September 2010)

This paper develops a general trans-dimensional Bayesian methodology for geoacoustic inversion. Trans-dimensional inverse problems are a generalization of fixed-dimensional inversion that includes the number and type of model parameters as unknowns in the problem. By extending the inversion state space to multiple subspaces of different dimensions, the posterior probability density quantifies the state of knowledge regarding inversion parameters, including effects due to limited knowledge about appropriate parametrization of the environment and error processes. The inversion is implemented here using a reversible-jump Markov chain Monte Carlo algorithm and the seabed is parametrized using a partition model. Unknown data errors are addressed by including a data-error model. Jumps between dimensions are implemented with a birth–death methodology that allows transitions between dimensions by adding or removing interfaces while maintaining detailed balance in the Markov chain. Trans-dimensional inversion results in an inherently parsimonious solution while partition modeling provides a naturally self-regularizing algorithm based on data information content, not on subjective regularization functions. Together, this results in environmental estimates that quantify appropriate seabed structure as supported by the data, allowing sharp discontinuities while approximating smooth transitions where needed. This approach applies generally to geoacoustic inversion and is illustrated here with seabed reflection-coefficient data.
© *2010 Acoustical Society of America*. [DOI: 10.1121/1.3500674]

PACS number(s): 43.30.Pc, 43.30.Ma, 43.60.Pt [AIT] Pages: 3393–3405

I. INTRODUCTION

Knowledge of geoacoustic properties and associated uncertainties is important for a variety of acoustic and sonar applications. Inferring geoacoustic parameters requires the assumption of a model describing the observed physical system including the physical theory, its appropriate parameterization, and a statistical representation for the data-error processes. In the past, Bayesian inference has been applied widely to geoacoustic inverse problems.[1–8] However, model selection and comparison, a fundamental aspect of Bayesian inference, has seen only limited applications in acoustics.[2,9–12] In addition, ambiguity and subjectiveness in the choice of model causes parameter uncertainties that have been ignored in geoacoustic inversion. The choice of model parametrization strongly influences parameter uncertainty estimates, with under-parametrized models generally under-estimating uncertainties while over-parametrized models over-fit the data and over-estimate uncertainties.[11] In addition, since the model is an approximation of the actual environment, the ambiguities resulting from this approximation cause parameter uncertainty that should be accounted for by integrating over the range of applicable parametrizations. In addition, parameter estimates can appear biased if an inappropriate parametrization is chosen.

This paper develops a trans-dimensional formulation of the geoacoustic inverse problem, where the number of parameters (environmental, data-error model, etc.) is itself an unknown in the problem. This results in a trans-dimensional posterior probability density (PPD) that intrinsically addresses model selection and accounts for parameter uncertainty due to the choice of model parametrization by integrating over possible choices rather than picking a single model. Trans-dimensional inference was introduced by the landmark papers of Green[13,14] and has since been applied to several problems in geophysics[15–18] and other fields.[19]

To sample from trans-dimensional distributions, Green[13] introduced a generalization of the Metropolis–Hastings (MH) algorithm[20,21] termed the reversible-jump Markov chain Monte Carlo (rjMCMC) sampler that allows the Markov chain to transition between dimensions of the state space (i.e., the model parameter space) while maintaining detailed balance[22] of the chain. The rjMCMC formulation is based on an extended acceptance rule similar to MH acceptance and can be applied to a wide range of problems and dimension transitions.[14,19] The rjMCMC methodology used in this paper is based on a partition modeling[19] approach and trans-dimensional jumps of the birth–death[15] form, which allow for a straightforward implementation. Partition modeling is a popular form of parametrization in biological and medical applications[19] where spatially variable parameters are common. This paper applies a partition model to the layering structure of the seabed sediment by describing the sediment as an interval over a certain depth with layer interface locations determined by the data. Partition models have been shown[17] to give results similar to regularized inversions when ensemble inference is carried out on the posterior. However, common

[a)]Author to whom correspondence should be addressed. Electronic mail: jand@uvic.ca

regularizations[23] used in geophysical inverse problems put subjective constraints on the solution, such as requiring the smoothest environmental model. With such regularization terms, the parameter-uncertainty estimates are difficult to interpret. Here, the partition model, together with a trans-dimensional approach, results in a naturally parsimonious self-regularization that is driven by the data. Results combine the ability to resolve sharp discontinuities as well as approximate smooth transitions (such as gradients) of arbitrary shape. The integrated "map" of interfaces shows increased probability where the data support structure.[24]

The trans-dimensional Bayesian inference approach is applied here to geoacoustic inversion of seabed reflection-coefficient data measured in the Straits of Sicily, Mediterranean Sea. The results highlight the strength of the approach, recovering both sharp discontinuities and smooth transitions at different depths, as supported by core measurements.

II. TRANS-DIMENSIONAL BAYESIAN INFERENCE

A. Bayesian inference theory

Let $\mathbf{d} \in \mathbb{R}^N$ be a random variable of N observed data containing information about a physical system. Further, let $\mathcal{I}$ denote the model specifying a particular choice of physical theory, model parameterization, and error statistics to explain that system. Let $\mathbf{m} \in \mathbb{R}^M$ be a random variable of the M free parameters representing a realization of model $\mathcal{I}$. Bayes' rule can then be written

$$P(\mathbf{m}|\mathbf{d},\mathcal{I}) = \frac{P(\mathbf{d}|\mathbf{m},\mathcal{I})P(\mathbf{m}|\mathcal{I})}{P(\mathbf{d}|\mathcal{I})}, \tag{1}$$

where the conditional probability $P(\mathbf{m}|\mathbf{d},\mathcal{I})$ represents the PPD of the unknown model parameters given the observed data, prior information, and choice of model $\mathcal{I}$. The conditional probability $P(\mathbf{d}|\mathbf{m},\mathcal{I})$ describes the data-error statistics. Since data errors include both measurement and theory errors (which cannot generally be separated), the specific form of this distribution is often not known. To interpret Eq. (1) quantitatively, some particular form that describes the data-error statistics reasonably well must be assumed for this distribution. In practice, mathematically simple distributions, such as multivariate Gaussian distributions, are commonly used; the validity of such assumptions should be checked *a posteriori* using statistical tests.[6,25] The general multivariate Gaussian distribution for real data is given by

$$P(\mathbf{d}|\mathbf{m},\mathcal{I}) = \frac{1}{(2\pi)^{N/2}|\mathbf{C}_d|^{1/2}} \times \exp\left(-\frac{1}{2}(\mathbf{d}-\mathbf{d}(\mathbf{m}))^\top \mathbf{C}_d^{-1}(\mathbf{d}-\mathbf{d}(\mathbf{m}))\right), \tag{2}$$

where $\mathbf{C}_d$ is the data covariance matrix and $\mathbf{d}(\mathbf{m})$ is the modeled data. The covariance matrix $\mathbf{C}_d$ is often unknown since the source of errors may be poorly understood. In some cases, data-error statistics can be parameterized (e.g., as variances or as a covariance matrix based on an assumed form such as an auto-regressive moving average[26]) and included in the inversion, either implicitly[27] or explicitly as unknown parameters with assigned priors.[16,27,28] Data-error covariance matrices can also be estimated non-parametrically from data residuals.[25] In Bayesian inversion, $P(\mathbf{d}|\mathbf{m},\mathcal{I})$ is interpreted as the likelihood function $\mathcal{L}(\mathbf{m},\mathcal{I})$ of $\mathbf{m}$ for fixed (observed) data $\mathbf{d}$. Note that given a Gaussian data-error distribution, the likelihood function is not Gaussian distributed for nonlinear inverse problems. The term $P(\mathbf{m}|\mathcal{I})$ in Eq. (1) represents the model parameter prior distribution and can contain any information about model parameters that is available *a priori* and can be expressed probabilistically. The conditional probability $P(\mathbf{d}|\mathcal{I})$ in Eq. (1) is commonly referred to as the *normalizing constant* or *evidence* of $\mathcal{I}$. It describes how likely a particular choice of model $\mathcal{I}$ is given the observed data and prior information. Since the evidence $P(\mathbf{d}|\mathcal{I})$ normalizes Eq. (1), it can be written

$$\mathcal{Z}(\mathcal{I}) = P(\mathbf{d}|\mathcal{I}) = \int_{\mathcal{M}} P(\mathbf{d}|\mathbf{m},\mathcal{I})P(\mathbf{m}|\mathcal{I})d\mathbf{m}, \tag{3}$$

where the integration is over the state space $\mathcal{M}$. Solving Eq. (3) for nonlinear geoacoustic inference is a challenging problem and can be addressed using thermodynamic integration,[29] annealed importance sampling,[30] or nested sampling.[10,31] Computing the evidence is one approach to model selection that provides a quantitative measure for the *natural parsimony*[22] intrinsic to Bayes' rule. This natural parsimony is the key to model selection which is a fundamental component of estimating parameter uncertainties, as illustrated in Fig. 1. Let $\mathbf{d}^{\text{obs}}$ be an observed data vector given as one point in the data space $\mathbf{D}$. Further, let $\mathcal{I}_1$ be one model parameterization (hypothesis) and $\mathcal{I}_2$ be a second model parameterization that is more complex than $\mathcal{I}_1$. Since model $\mathcal{I}_1$ has a limited number of parameters, it can only access a certain subspace $\mathbf{D}_1$ of the total data space $\mathbf{D}$. Since $\mathcal{I}_2$ is a more complex model with more parameters, it can access a larger subspace $\mathbf{D}_2$ which includes $\mathbf{D}_1$. As long as $\mathbf{d}^{\text{obs}}$ falls into a region of the data space that can be accessed by both models, Bayes' theorem [Eq. (1)] generally favors the simpler model $\mathcal{I}_1$ [i.e., $P(\mathbf{d}|\mathcal{I})$ is larger for $\mathcal{I}_1$].

B. Trans-dimensional formulation

Another approach to using natural parsimony in inversion is to formulate the inverse problem as trans-dimensional,

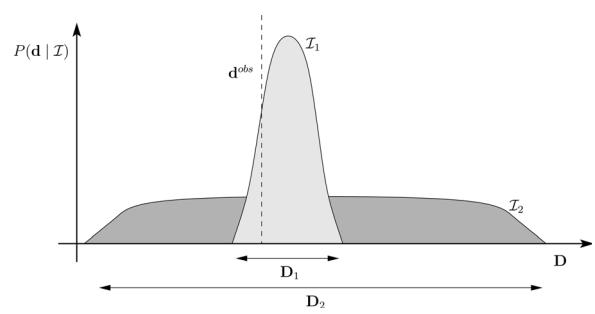

FIG. 1. The concept of natural parsimony as given by Bayes theorem [after MacKay (Ref. 22)].

meaning that the number of unknowns in the problem is itself an unknown that is integrated over, resulting in a posterior that spans multiple spaces of different dimensions. A crucial advantage of treating a problem as trans-dimensional is that the posterior includes the uncertainty due to the limited knowledge of the model. Green[13] shows that Eq. (1) can be written as a Bayesian hierarchical model to include a parameter k,

$$P(k,\mathbf{m}_k|\mathbf{d}) = \frac{P(k)P(\mathbf{d}|k,\mathbf{m}_k)P(\mathbf{m}_k|k)}{\sum_{k'\in\mathcal{K}}\int_{\mathcal{M}} P(k')P(\mathbf{d}|k',\mathbf{m}'_{k'})P(\mathbf{m}'_{k'}|k')d\mathbf{m}'_{k'}}, \tag{4}$$

where $k \in \mathcal{K}$ ($\mathcal{K}$ is a countable set) can be interpreted as indexing possible choices of models (as commonly used in model selection using evidence or point estimates thereof) or, in a trans-dimensional sense, where k is a variable dimension parameter of a single model that spans several multidimensional spaces. In Eq. (4) $P(k)$ is the prior over the $\mathcal{K}$ models to be considered. The state variables of the inversion for the trans-dimensional case are then given by (k, $\mathbf{m}_k$) and are of dimension M_k. The state space is trans-dimensional and given by the union of all fixed dimension spaces in $\mathcal{K}$, i.e., $\bigcup_{k\in\mathcal{K}}(\{k\}\times\mathbb{R}^{M_k})$. A Markov chain that samples this state space can be defined which converges to the trans-dimensional joint posterior $P(k,\mathbf{m}_k|\mathbf{d})$ (see Sec. II D). The denominator in Eq. (4) can be interpreted as the *total evidence* over all models in $\mathcal{K}$.

C. Partition modeling

Partition modeling has been widely applied in biological and medical applications[19] where spatially variable parameters and data are common. Bodin and Sambridge[17] apply a partition modeling approach to two-dimensional seismic tomography to obtain a self-parametrizing algorithm that gives results similar to regularized inversions while using substantially fewer cells and maintaining resolution of sharp discontinuities. However, partition modeling can be applied in a Monte Carlo sampling framework and requires fewer discretizing parameters since the data introduce a self-smoothing through marginalization of the resulting PPD estimate, removing unrealistic high-frequency spatial structure from the result.

A partition modeling approach is applied here to parametrize the positions of sediment interfaces in the environmental model for the sediment extending from the seafloor to a maximum depth of interest. Environmental models in geoacoustic inversion most commonly treat the sediment structure as a stack of layers with each layer assigned a thickness parameter. Perturbing this type of parametrization leads to changes for the whole model below the layer being considered. In MCMC inversion, this can lead to high rejection rates/poor mixing in the algorithm since a relatively small perturbation to one layer thickness parameter can result in relatively large changes to the model. Perturbing one interface in a partition model only changes the neighboring two layers and leaves the rest of the model invariant; this is consistent with the idea that relatively small steps are of interest and should improve acceptance in the Markov chain. These perturbations are generally small compared to the distance to the nearest interface neighbor to avoid crossing of interfaces. Mixing can also be improved by occasionally allowing large changes to the partition model, which can be achieved by removing a randomly chosen interface and placing it at a new (random) depth.

The PPD estimate from an MCMC sampling process then contains many partition models and inference can be carried out with ensemble averages over the structure. The results are similar to the regularized inversions, but without the problematic requirement of specifying subjective regularization terms. In essence, the environmental model is regularized by the data which determine wherein the partition structure is supported. The amount of structure allowed can vary with depth and does not depend on a subjective regularization term that applies to the whole depth interval of interest. For example, in this framework, gradients in sound speed may be approximated by the ensemble structure of several interfaces spread over some parts of the partition model while other areas can have homogeneous structure. Sharp discontinuities can be supported by ensembles of interfaces being confined to a tightly limited depth range.

The number of sediment interfaces required for this parametrization can also be determined by the data and is addressed here by treating the problem as trans-dimensional, with the model index k in Eq. (4) indicating the number of interfaces in the partition.

D. Reversible-jump MCMC

A Markov chain that samples the posterior in Eq. (4) in a trans-dimensional sense must be able to undergo dimension changes (*jumps*) in the state space while not violating the requirement of detailed balance. Green[13] accomplishes this by introducing a transformation

$$(k,\mathbf{m}_k,\mathbf{u}) \rightleftharpoons (k',\mathbf{m}'_{k'},\mathbf{u}'), \tag{5}$$

where (k, $\mathbf{m}$) is the current state of the chain and $(k',\mathbf{m}'_{k'})$ is a proposed state in a state space of possibly different dimension or type. The vectors $\mathbf{u}$ and $\mathbf{u}'$ are random variables accounting for potential dimension changes so that the dimensions δ match: $\delta(k) + \delta(\mathbf{m}) + \delta(\mathbf{u}) = \delta(k') + \delta(\mathbf{m}') + \delta(\mathbf{u}')$. A jump between states can then be defined by a transformation h as

$$\begin{aligned}(k',\mathbf{m}'_{k'},\mathbf{u}') &= h(k,\mathbf{m}_k,\mathbf{u}),\\ (k,\mathbf{m}_k,\mathbf{u}) &= h^{-1}(k',\mathbf{m}'_{k'},\mathbf{u}').\end{aligned} \tag{6}$$

The function h and its inverse must exist and be differentiable (h is a diffeomorphism). Otherwise, the choice is arbitrary which allows for powerful and general choices in jumps including state spaces spanning multiple subspaces of varying dimension as well as different types of state spaces (i.e., the underlying physical theory and type of parameters can change). In the fixed-dimension case ($k = k'$), the MH algorithm can be used to sample from the posterior by

implementing the well known acceptance criterion for a step from the current model $\mathbf{m}_k$ to a proposed model $\mathbf{m}'_k$, as

$$\alpha = \min\left[1, \frac{P(\mathbf{m}'_{k'})P(\mathbf{d}|\mathbf{m}'_{k'})}{P(\mathbf{m}_k)P(\mathbf{d}|\mathbf{m}_k)}\frac{Q(\mathbf{m}_k|\mathbf{m}'_{k'})}{Q(\mathbf{m}'_{k'}|\mathbf{m}_k)}\right], \tag{7}$$

where the product consists of, from left to right, the prior ratio, the likelihood ratio (containing the information introduced by the data), and the proposal ratio (the proposal distribution Q is used to propose a perturbation to the parameters that can be accepted or rejected according to α). In practice, Eq. (7) is implemented by drawing a random number $u \in [0, 1]$ and a perturbation is accepted if $\alpha > u$. Green[13,14] shows that the rjMCMC can be implemented in a manner similar to the MH algorithm by generalizing the Metropolis acceptance [Eq. (7)] to give the acceptance for a step from the current model $(k, \mathbf{m}_k)$ to a proposed model $(k', \mathbf{m}'_{k'})$ as

$$\alpha = \min\left[1, \frac{P(k', \mathbf{m}'_{k'})}{P(k, \mathbf{m}_k)}\frac{P(\mathbf{d}|k', \mathbf{m}'_{k'})}{P(\mathbf{d}|k, \mathbf{m}_k)}\frac{Q(k, \mathbf{m}_k|k', \mathbf{m}'_{k'})}{Q(k', \mathbf{m}'_{k'}|k, \mathbf{m}_k)}|\mathbf{J}|\right], \tag{8}$$

where $|\mathbf{J}|$ is the determinant of the Jacobian for the diffeomorphism h from $(k, \mathbf{m}_k)$ to $(k', \mathbf{m}'_{k'})$,

$$|\mathbf{J}| = \left|\frac{\partial h(k, \mathbf{m}_k, \mathbf{u})}{\partial(k, \mathbf{m}_k, \mathbf{u})}\right|. \tag{9}$$

Note that in many fixed-dimensional MCMC applications, the proposal ratio cancels when symmetric proposals are used and the proposal steps are inherently reversible. In addition, commonly used uniform priors result in unit prior ratios. The classic MH acceptance [Eq. (7)] is then given simply by the ratio of the likelihood values. The value of the likelihood function generally increases (for nested models) with additional parameters and will therefore favor models with large numbers of parameters. In the trans-dimensional formulation [Eq. (8)] this property of the likelihood ratio is opposed by the other factors (proposal and prior ratios) when a jump is proposed to higher dimension, since the volume of the state space increases. This behavior is an intrinsic property of trans-dimensional Bayesian inversions, which provides natural parsimony and constitutes a general and powerful approach for treating data inference. In addition, sampling over dimensions effectively integrates over the (subjective) choices of models and quantifies the lack of knowledge about the parametrization as parameter uncertainty, resulting in more realistic uncertainty estimates. Further, the different dimensions are sampled proportional to this support by the observed data, resulting in an efficient approach to sampling a trans-dimensional space. Hence, trans-dimensional sampling can be more efficient than sampling a sequence of individual state spaces. Green[14] also suggests that trans-dimensional chains can result in better mixing of the chain for some cases. A disadvantage is that rjMCMC has no memory of the various state spaces it has visited which can negatively affect the performance of jumping dimensions.

E. Birth–death rjMCMC

A potential challenge in implementing Eq. (8) is computing the Jacobian of the diffeomorphism. However, by choosing the proposal distributions for dimension jumps appropriately, analytic expressions for $|\mathbf{J}|$ can be found (e.g., Refs. 14, 32). The most commonly applied form of rjMCMC is the birth–death scheme,[15] in which case the determinant of the Jacobian is unity. The scheme must be carefully implemented to ensure unbiased sampling and equal probability of birth and death moves to maintain detailed balance. In this scheme, a birth move is defined to be the introduction of new parameters. For example, in a partition model a new layer can be created by introducing a new interface at a random position in the seabed model, increasing the model indexing parameter k by 1. Alternatively, a death move randomly picks and deletes an interface, reducing k by 1. This paper implements the birth–death approach as follows. The geoacoustic model is represented as $(k, \mathbf{m}_k) = (k, \mathbf{z}, \gamma, \boldsymbol{\sigma})$, where k is the number of interfaces in the partition and the partition is the interval of the sediment that is of interest, extending from $z = z_0 = 0$ to $z = z^{\max}$. The sediment interval partitioning is denoted $\mathbf{z}$ and contains the depths of the k sediment interfaces, while γ contains the geoacoustic properties of all partitions. Here, $\gamma = (\mathbf{c}, \boldsymbol{\rho}, \boldsymbol{\alpha})$, where $\mathbf{c}$ is a vector of sound velocities, $\boldsymbol{\rho}$ a vector of densities, and $\boldsymbol{\alpha}$ a vector of attenuations. The data-error model is denoted σ. The dimension of a sub-state space for model index k is $M_k = 4k + 3 + N_\sigma$, where k is the number of interfaces and N_σ is the dimension of the data-error model (when sampling over the data-error model). The birth–death rjMCMC algorithm is implemented such that birth, death, and perturbation moves are proposed with probabilities that ensure equal probability of birth and death acceptance so that parameter k is determined by the data and not biased toward too many or too few interfaces. In detail, the moves are as follows:

(1) Birth: Parameter $k' = k + 1$ and a new interface position is drawn from a uniform prior over $[z_0, z_{\max}]$. The new interface is inserted into the model under the constraint of a minimum required layer thickness. The γ values that go with the layer that was divided by the new interface are copied into both new layers, and one of the two new layers is randomly picked and the γ parameters of that layer are randomly perturbed with a Gaussian proposal distribution.
(2) Death: Parameter $k' = k - 1$ and an interface is randomly chosen from the current interfaces and deleted from the partition. The γ parameters of the new layer are the average of the two layers that were combined by the death move.
(3) Perturb: $k' = k$ and an interface is randomly picked and perturbed using a Gaussian proposal distribution.

All three move types are assigned a probability of 1/6. In addition, moves that leave the partition invariant are applied with probability of 1/2 (so that all move-type probabilities sum to unity). The choice of these probabilities is arbitrary and can be adjusted according to the problem. In addition, other moves are possible and can be advantageous to improve chain mixing for certain problems. Since k is an integer variable bounded by $k^{\min}$ and $k^{\max}$, care must be

taken to sample uniformly over k. While $k^{\min} < k < k^{\max}$, equal probabilities between the three move types above assures this. However, the probability needs to be changed to preclude deaths when $k = k^{\min}$ (e.g., 1/4 for perturb and 1/4 for birth). Likewise, the probability must preclude births when $k = k^{\max}$ (e.g., 1/4 for perturb and 1/4 for death). During these moves, only γ and the data-error model parameters $\boldsymbol{\sigma}$ are perturbed using one-dimensional Gaussian proposal distributions. Each move perturbs all parameters, one at a time, in random sequence. Once a model is proposed, acceptance is computed according to Eq. (8). If the proposed model is accepted, the algorithm moves to the new state, otherwise it stays at the current point. Care must be taken when computing the acceptance rates as discussed in Sec. II G. This approach effectively separates the partition model from the acoustic parameters (γ) in the sediment layers which is useful in computing the correct acceptance probability $\alpha((k, \mathbf{m}_k), (k', \mathbf{m}'_{k'}))$ (see Sec. II G).

F. Prior information

The prior information used in this paper consists of bounded uniform distributions with the range of parameter values chosen to represent physically reasonable values and set wide enough that the data predominantly determine the posterior. Since $(k, \mathbf{m}_k) = (k, \mathbf{z}, \gamma, \boldsymbol{\sigma})$, the prior probability $P(k)P(\mathbf{m}_k|k) = P(k)P(\mathbf{z}|k)P(\gamma|k)P(\boldsymbol{\sigma}|k)$. In particular, the prior distribution for the model index parameter is given by (assuming equal preference for all models)

$$P(k) = \begin{cases} (k^{\max} - k^{\min})^{-1} & \text{if } k^{\min} \le k \le k^{\max} \\ 0 & \text{else.} \end{cases} \tag{10}$$

The prior distribution for γ is given by

$$P(\gamma|k) = \begin{cases} \prod_{i=1}^{3(k+1)} (\gamma_i^{\max} - \gamma_i^{\min})^{-1} & \text{if } \gamma_i^{\min} \le \gamma_i \le \gamma_i^{\max}, \\ 0 & \text{else,} \end{cases} \tag{11}$$

and the prior distribution for the data-error model is given by

$$P(\boldsymbol{\sigma}|k) = \begin{cases} \prod_{i=1}^{N_\sigma} (\sigma_i^{\max} - \sigma_i^{\min})^{-1} & \text{if } \sigma_i^{\min} \le \sigma_i \le \sigma_i^{\max}, \\ 0 & \text{else,} \end{cases} \tag{12}$$

where superscripts min and max indicate the lower and upper bounds applied. To derive the prior probability of the partitioning, it is convenient to temporarily assume an underlying grid of K grid points (that will later cancel out) on which the different combinations of k interfaces can be arranged and assigned equal probability.[19,24] The prior distribution for $\mathbf{z}$ is then

$$P(\mathbf{z}|k) = \left[\frac{K!}{k!(K-k)!}\right]^{-1}. \tag{13}$$

G. Prior and proposal ratios for birth and death

To compute the rjMCMC acceptance [Eq. (8)], the proposal distributions for all move types must be defined in addition to the prior information (Sec. II F) and the likelihood function [Eq. (2)]. The proposal distributions for moves that do not jump between dimensions are chosen to be Gaussian distributions (which are symmetric proposals) and hence do not affect acceptance. Therefore, the proposal ratio needs to be computed only for birth and death moves. Further, the proposal for $\mathbf{z}$ is considered uniform over $z_{\min}$ to $z_{\max}$ and birth and death moves have the same probability. The proposal ratio can be separated into several components[19]

$$\frac{Q(k, \mathbf{m}_k|k', \mathbf{m}'_{k'})}{Q(k', \mathbf{m}'_{k'}|k, \mathbf{m}_k)} = \frac{Q(k|k', \mathbf{m}'_{k'})}{Q(k'|k, \mathbf{m}_k)} \frac{Q(\mathbf{z}|k', \mathbf{m}'_{k'})}{Q(\mathbf{z}'|k, \mathbf{m}_k)} \frac{Q(\gamma|k', \mathbf{m}'_{k'})}{Q(\gamma'|k, \mathbf{m}_k)} \frac{Q(\sigma|k', \mathbf{m}'_{k'})}{Q(\sigma'|k, \mathbf{m}_k)}. \tag{14}$$

To obtain simplified proposal ratios for the birth and death moves consider the probabilities of each move and their reverse. For a birth move, the probabilities associated with creating a new interface are

$$Q(\mathbf{z}'|k, \mathbf{m}_k) = \frac{1}{K-k},$$
$$Q(\gamma'|k, \mathbf{m}_k) = \frac{1}{\left[(2\pi)^3|\hat{\mathbf{C}}|\right]^{1/2}} \times \exp\left[-\frac{1}{2}(\hat{\gamma}' - \hat{\gamma})^\top \hat{\mathbf{C}}^{-1}(\hat{\gamma}' - \hat{\gamma})\right], \tag{15}$$

where $\hat{\mathbf{C}}$ is the 3×3 covariance matrix of the Gaussian proposal (the dimension is due to γ containing three acoustic parameters for each layer). In Eq. (15), $\hat{\gamma}$ is a 1×3 vector with the acoustic parameters of the layer that is split by the birth, and $\hat{\gamma}'$ is a 1×3 vector with the proposed values. The probabilities of the reverse move are

$$Q(\mathbf{z}|k', \mathbf{m}'_{k'}) = \frac{1}{k+1},$$
$$Q(\gamma|k', \mathbf{m}_k) = 1. \tag{16}$$

Substituting Eqs. (15) and (16) in the proposal ratio gives the proposal ratio for a birth move

$$\left[\frac{Q(k, \mathbf{m}_k|k', \mathbf{m}'_{k'})}{Q(k', \mathbf{m}'_{k'}|k, \mathbf{m}_k)}\right]_{\text{birth}} = \frac{K-k}{k+1}[(2\pi)^3|\hat{\mathbf{C}}|]^{1/2} \exp\left[\frac{1}{2}(\hat{\gamma}' - \hat{\gamma})^\top \hat{\mathbf{C}}^{-1}(\hat{\gamma}' - \hat{\gamma})\right]. \tag{17}$$

Similarly, the proposal ratio for a death move becomes

$$\left[\frac{Q(k, \mathbf{m}_k|k', \mathbf{m}'_{k'})}{Q(k', \mathbf{m}'_{k'}|k, \mathbf{m}_k)}\right]_{\text{death}} = \frac{k}{K-k+1}[(2\pi)^3|\hat{\mathbf{C}}|]^{-1/2} \times \exp\left[-\frac{1}{2}(\hat{\gamma}' - \hat{\gamma})^\top \hat{\mathbf{C}}^{-1}(\hat{\gamma}' - \hat{\gamma})\right]. \tag{18}$$

The prior ratios for birth and death moves follow from Eqs. (10) to (13),

$$\left[\frac{P(k',\mathbf{m}'_{k'})}{P(k,\mathbf{m}_k)}\right]_{\text{birth}} = \frac{k+1}{K-k}\left(\prod_{i=1}^{3}\Delta\hat{\gamma}_i\right)^{-1},$$
$$\left[\frac{P(k',\mathbf{m}'_{k'})}{P(k,\mathbf{m}_k)}\right]_{\text{death}} = \frac{K-k+1}{k}\prod_{i=1}^{3}\Delta\hat{\gamma}_i, \tag{19}$$

where $\Delta\hat{\gamma}_i$ is the width of the prior distribution for parameter $\hat{\gamma}_i$.

With expressions for prior and proposal ratios for the birth and death moves, the acceptance for these moves can now be computed as

$$\alpha_{\text{birth}} = \min\left[1, \left(\prod_{i=1}^{3}\Delta\hat{\gamma}_i\right)^{-1}[(2\pi)^3|\hat{\mathbf{C}}|]^{1/2} \times \exp\left[\frac{1}{2}(\hat{\gamma}'-\hat{\gamma})^{\top}\hat{\mathbf{C}}^{-1}(\hat{\gamma}'-\hat{\gamma})\right]\frac{\mathcal{L}(k',\mathbf{m}'_{k'})}{\mathcal{L}(k,\mathbf{m}_k)}\right],$$
$$\alpha_{\text{death}} = \min\left[1, \left(\prod_{i=1}^{3}\Delta\hat{\gamma}_i\right)^{-1}[(2\pi)^3|\hat{\mathbf{C}}|]^{-1/2} \times \exp\left[-\frac{1}{2}(\hat{\gamma}'-\hat{\gamma})^{\top}\hat{\mathbf{C}}^{-1}(\hat{\gamma}'-\hat{\gamma})\right]\frac{\mathcal{L}(k',\mathbf{m}'_{k'})}{\mathcal{L}(k,\mathbf{m}_k)}\right]. \tag{20}$$

Note that K (hypothetical grid dimension) has canceled out of these equations. The determinant of the Jacobian of the diffeomorphism is unity in this case and therefore does not appear in the acceptance expressions.

III. INVERSION EXAMPLES

This section illustrates how the trans-dimensional approach can be applied to practical geoacoustic inverse problems, using the particular example of seabed reflection-coefficient inversion. The forward model for the inversion is a plane wave reflection-coefficient computation[33] based on homogeneous sediment layers using a fluid (no shear) approximation. Trans-dimensional inversions were carried out allowing 1–25 interfaces ($k^{\min} = 1$ and $k^{\max} = 25$) to allow for a large range of possible partition models. The range of k corresponds to allowing models with more than 100 free parameters; however, the highest number of parameters sampled in these examples was 48. Bounds for the uniform prior distributions were set to 1450–1700 m/s for sound velocity, 1.2–2.0 g/cm^3 for density, and 0.001–1.0 dB/m/kHz for attenuation. The partition reached from 0 to 4.0 m below the seafloor which equals the extent of the time windowing for the seabed response for a sound velocity of 1600 m/s. A constraint for minimum layer thickness was applied, allowing layers >0.05-m thick (avoiding layers <1/20 of the shortest wavelengths for efficiency). An auto load-balancing algorithm[34] was implemented for parallel computers to allow large numbers of independent chains, providing a thorough search of the state space while speeding up convergence. In all case, 168 chains were run in parallel until >1000 models were collected in each chain. To avoid highly dependent samples in the PPD, chain thinning[35] was applied and the length of the chain between thinning steps was randomly chosen between 500 and 1000. Hence, the length of each chain was of the order of ~750 k moves and the final PPD is based on $>10^8$ moves. To ensure independence from the starting points, the first 1/3 of each chain was discarded as burn-in. This ensured sampling from a stationary chain, which was confirmed by examining the likelihood values of the sampled models (see Fig. 2), which is randomly distributed after the burn-in with no systematic change. The acceptance rate of the chains was around 20%–30% for both simulation and measured data inversions, suggesting that the proposal distributions were scaled reasonably well for this problem.[22]

A. Simulation

To examine the algorithm and gain insight as to how complex seabed environments are handled, a simulation based on core data is considered first. A simulated layered-

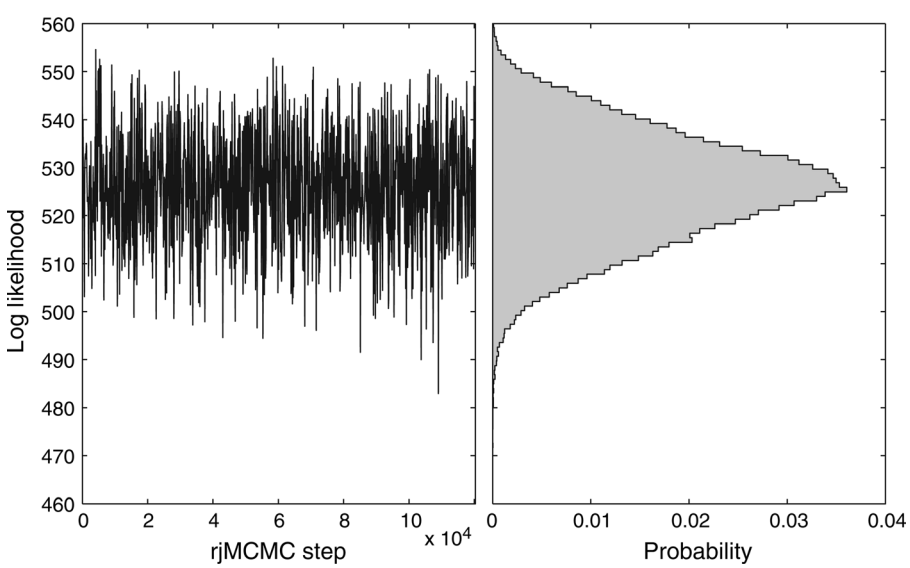

FIG. 2. Log-likelihood as a function of rjMCMC move step and as a histogram for the measured data inversion. For display purposes only every 100th move is plotted in the left panel.

seabed environment was constructed from three cores that were taken close to the experiment site (see Sec. III B). Since the density measurements were not as closely spaced as the velocity measurements on some of the cores, densities were interpolated onto the same depth spacing as velocities. Attenuation, which was not measured in the cores, was assumed to be constant at 0.1 dB/m/kHz over the entire depth. The simulated seabed consisted of 120 layers of varying thickness to a maximum depth of 3.8 m. The environment was chosen to be complex with many layers so that the trans-dimensional inversion is unable to resolve all layers but must approximate the underlying environment with a simpler model. Simulated reflection-coefficient data were obtained (Fig. 3) by computing the reflection-coefficient as a function of angle and frequency over the same ranges used for the measured data (31 angles from 20° to 83° and eight frequency bands from 300 to 1600 Hz with 1/15 fractional bandwidth and Gaussian averaging). Finally, random Gaussian-distributed data errors with a standard deviation of 0.015 were added to the simulated data.

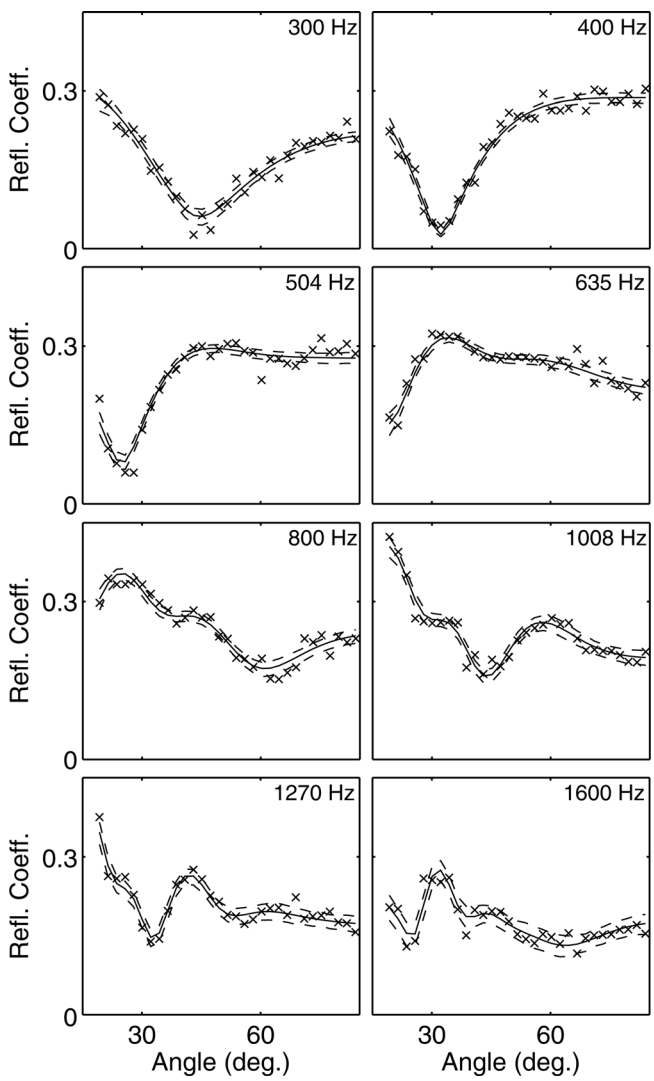

FIG. 3. Simulated reflection-coefficient data (crosses), mean fit (solid line; computed from PPD), and 95% highest probability density (HPD) intervals (dashed), at frequencies indicated.

The algorithm was initiated at a random point in the trans-dimensional state space and a burn-in phase was carried out where models were discarded until the chain appeared to be stationary. Convergence of the algorithm was judged by the convergence of the mean ensemble-profiles as well as by the running histogram of the model index parameter k. While convergence cannot be guaranteed, once both indicators do not change significantly over a large number of rjMCMC moves, the algorithm can be considered to have converged with reasonable confidence.

Figure 4 shows the model index parameter histogram and indicates that the naturally parsimonious inversion approximates the true environment (120 layers) by integrating over models with five to ten interfaces. The highest probability is for models with seven layers. Figure 5 shows the marginal profile distributions[12] for the simulation, which are an ensemble estimate of the seabed structure based on the trans-dimensional PPD and constitute the main result of the Bayesian inference carried out in this paper. The panels show the velocity, density, and attenuation marginal profiles over depth as well as the probability of positions of layer interfaces in the partition. The profile marginal distributions are individually normalized at each depth for display purposes. The probability distribution of sediment interfaces is a useful measure of where the data support structure in the environmental model. The figure also shows the "true" underlying seabed. Several interesting observations can be made by comparing the complex, 120-layer underlying seabed model and the probability distribution recovered by the inversion. The inversion results capture the seabed structure by a combination of sharp discontinuities as well as gradients which are smoothly approximated through the ensemble of many partition models in the PPD. Hence, the inversion is effectively self-regularized by the data, a feature intrinsic to the trans-dimensional partition model.

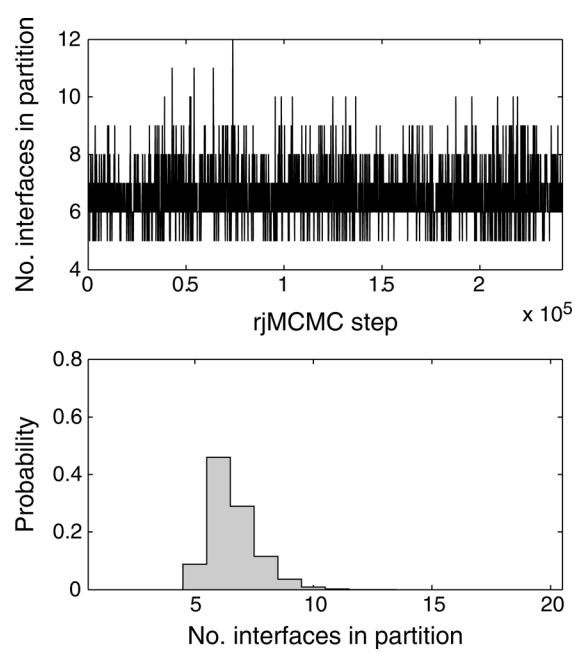

FIG. 4. Number of interfaces k in the partition model.

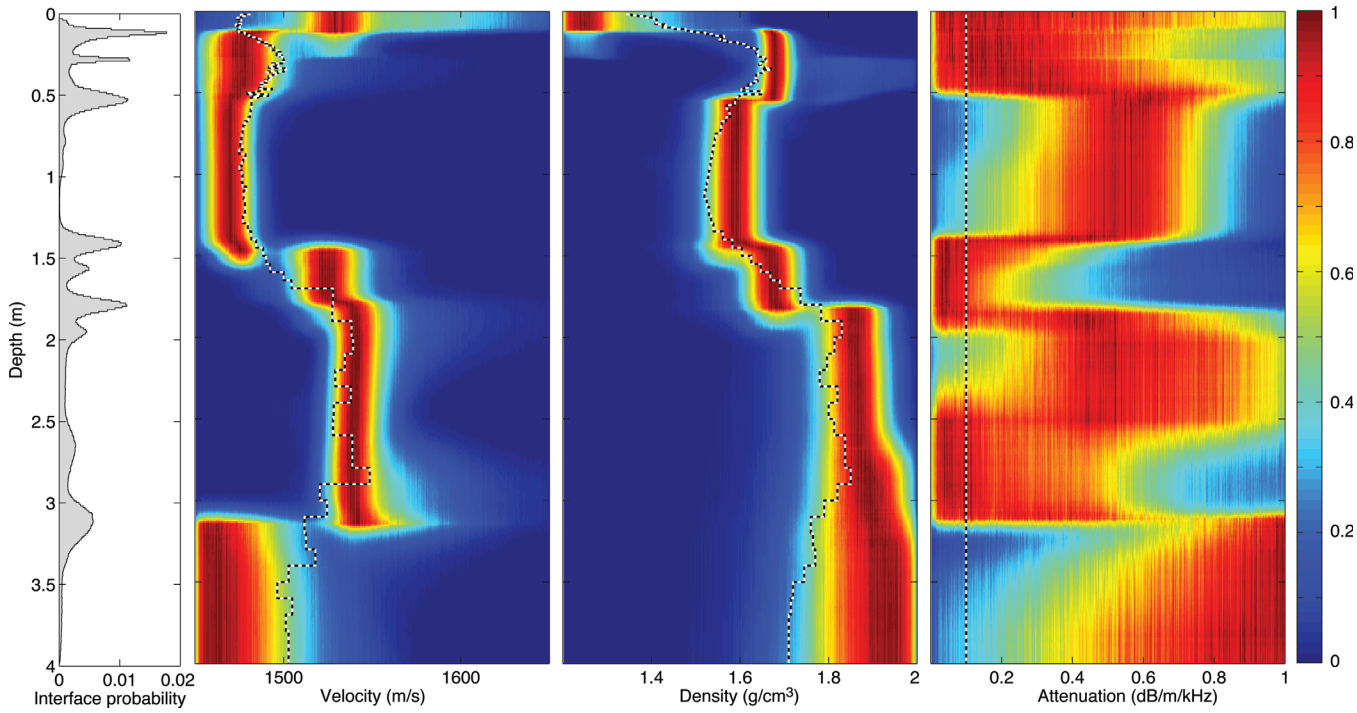

FIG. 5. (Color online) Trans-dimensional profile marginal distributions for simulation. Plot boundaries correspond to prior bounds. Solid lines indicate the simulated seabed model.

Figure 3 also shows the 95% HPD credibility intervals for the data misfit that are calculated from the data predicted for all models in the PPD (the *predicted-data sample*).

Figure 5 shows that attenuation (not a primary focus of this work) is largely unresolved in this example. This is likely due to the limited frequency bandwidth chosen for this example, since sensitivity of the reflection-coefficient data to attenuation would be expected to increase within a given layer as the ratio of layer thickness to wavelength increases. Overall, this simulation illustrates the advantages of the trans-dimensional inversion in capturing a wide range of environmental structures (gradients and discontinuities in different parts of the same environmental model) in a general way.

B. Experiment and data

Seismo-acoustic reflection data were collected April 22, 1998, during the SCARAB98 experiment at 36°30'39"N, 14°32'59"E on the Malta Plateau, Mediterranean Sea. The acoustic data were generated with an electro-mechanical impulsive source (EG&G model 265 Uniboom) with a short pulse length (< 1 ms) and a broad bandwidth (0.2–10 kHz). The data were recorded at a single receiver of a vertical line array of 16 Benthos AQ-4 hydrophones. The source depth was 0.35 m, the hydrophone depth was 112 m, and the water depth was 153 m. These data have been previously considered in Ref. 6, where a more detailed account of the experiment can be found.

Reflection-coefficient data as a function of grazing angle and frequency are computed from time-windowed direct and bottom-reflected arrivals using the method of Holland[36] and are shown in Fig. 6. In this case, the bottom response is time windowed to approximately 3.75-m depth below the seafloor. The data are averaged into eight frequency bands from 300 to 1600 Hz using a Gaussian frequency average[37,38] with a fractional bandwidth of 1/15 (this bandwidth was found to retain structure in the reflection-coefficient data while reducing noise). The data are interpolated onto a uniform spacing in angle; points with a signal to noise ratio of less than 6 dB were excluded. Further, interpolated data that fall into recording gaps (due to experiment design) are excluded from the inversion. This results in approximately 30 data at each frequency. Previous studies have shown that spherical wave-front effects are negligible and that plane-wave modeling is sufficient for this experiment geometry and sub-bottom depth.[38]

C. Posterior parameter inference

The MCMC sampling convergence was judged in the same way as in the simulation. Figure 7 shows a histogram of the model indexing parameter which indicates that five interfaces have the highest probability. However, significant probability exists for layering structure from three to seven layers. This uncertainty in the number of interfaces supported by the data translates into an uncertainty for geoacoustic parameters and the position of interfaces that cannot be assessed in a fixed-dimensional approach. However, the trans-dimensional approach intrinsically accounts for this source of uncertainty, resulting in more realistic uncertainty estimates based on the information content of the data. In combination with a partitioning of the seabed, this result is a general approach that can be recognized as regularizing by the data.

Figure 8 shows the maximum *a posteriori* (MAP) model (the model that maximizes the posterior) from the trans-dimensional PPD and compares it to three cores taken at the experiment site. The MAP model shows seven interfaces in the partition and indicates several strong discontinuities that

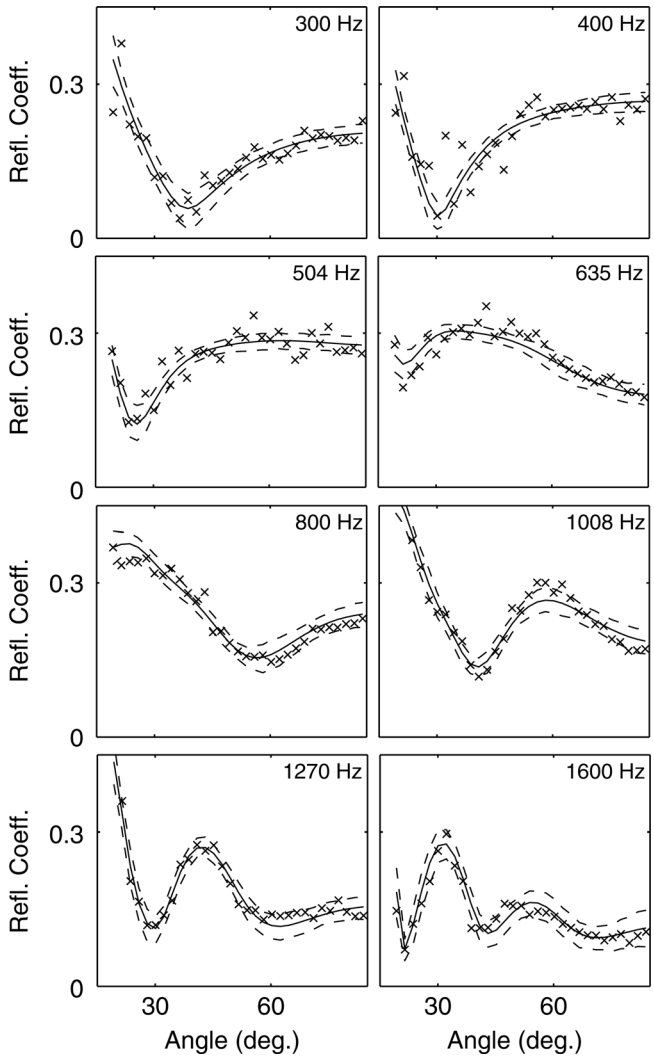

FIG. 6. Observed reflection-coefficient data (crosses), mean fit (solid line; computed from PPD), and 95% HPD intervals (dashed) at indicated frequencies.

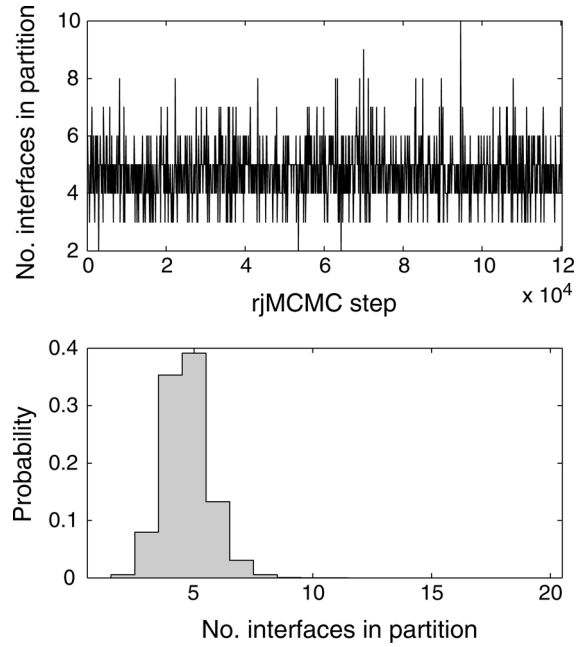

FIG. 7. The number of interfaces k in the partition model.

do not appear to be physically reasonable. This type of structure is commonly observed in over-parametrized models, where too much structure is introduced to fit (or over-fit) the data, resulting in under-constrained solutions. In linear or linearized inverse theory, these problems are often addressed by regularizing with some subjective, global regularization term that constrains the form of the solution (e.g., the smoothest model that fits the data to within a pre-determined limit). The global nature of regularizations can cause problems when some parts of the model are required to be smooth while other parts should allow discontinuities. This is commonly solved by breaking the regularization at some location chosen by the user. Subjective regularization terms are problematic with regard to parameter uncertainty estimates, since the parameter uncertainties are strongly influenced by the regularization and may not be representative of the environment or the data information content. It is evident that MAP models are of limited value in trans-dimensional inversions and posterior inference should rather be based on ensembles and marginalization.

Figure 9 shows profile marginal distributions for the measured data. In this case, there are four dominant peaks for layer interfaces, indicating four significant reflectors. However, other areas of increased probability exist that do not show strong reflectors in the profile marginals. These areas exhibit velocity and density gradients that are captured through the ensemble and regularization qualities of the trans-dimensional partition model. Velocity and density profile marginals are compared to measurements from three cores which were taken at the experiment site. High quality 1.3 m gravity cores were collected using a thin-walled core barrel (to minimize sediment disturbance) with a sphincter device designed to minimize end effects caused by typical core closing mechanisms. The cores were maintained in a vertical position until density (gamma-ray attenuation) and sound speed (time of flight at 200 kHz) were measured using a Geotek Multi-Sensor Core Logger (MSCL). The system was calibrated and the error bars shown in Fig. 9 represent the errors from the measurement system only. Core-sampling errors (compression or extension of sediments from core-barrel impact, removal, and handling) are not included. There is some evidence that the core-sampling uncertainties are of the same order or smaller than the measurement errors for fine-grained materials that likely dominated the upper 1.5 m at this site.

In order to sample deeper than 1.5 m, a 4-m piston corer (with thicker barrel walls and conventional core catcher) was employed. While the uncertainties for the these data are much larger than those associated with the 1.3 m gravity cores, the plot shows the same uncertainties which should be viewed as an extreme lower bound. These uncertainties represent a lower bound because the measurement errors are higher (the time of flight measurements were conducted by hand rather than with the MSCL and density was measured using dried weight) and mostly because the sampling errors are believed to be much higher. The sampling errors include effects of the thicker walls and effects of the piston on compression/extension of the sediments, but almost certainly the largest factor was the placement and storage of the core in a horizontal

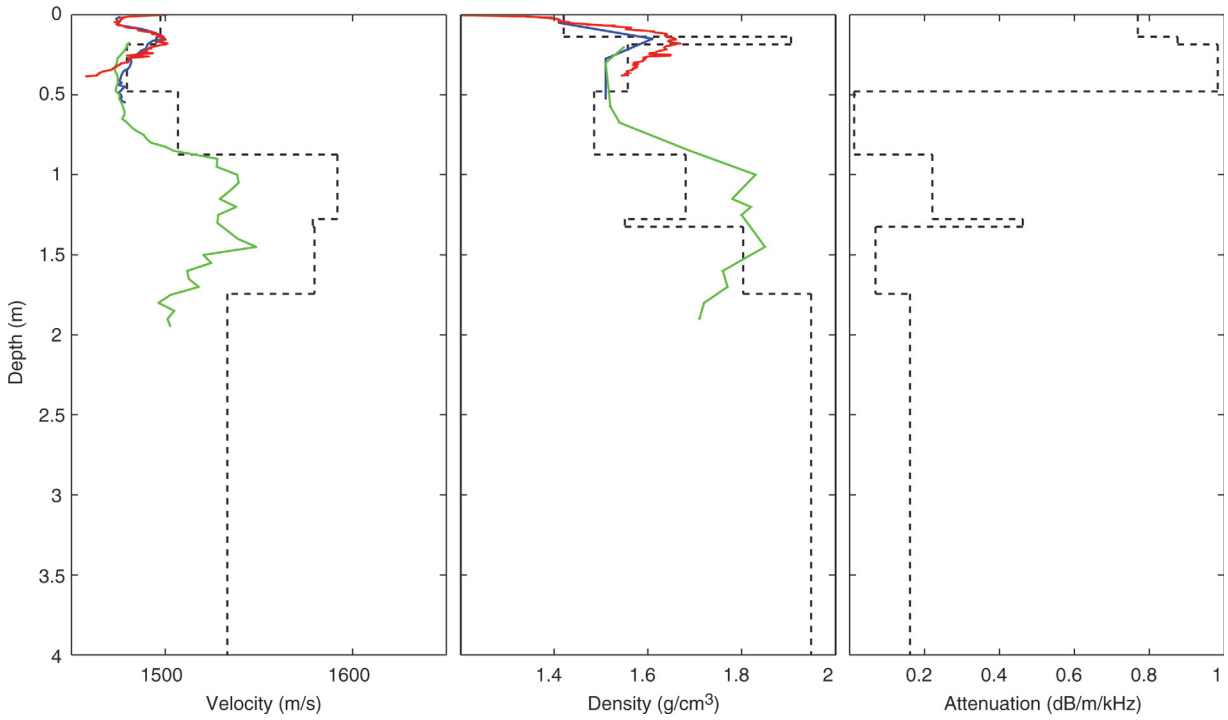

FIG. 8. (Color online) MAP model (dashed) of the trans-dimensional PPD and three cores taken at the site (solid lines).

position before measurements. The effect of turning the core on its side would tend to have the largest effects on granular (as opposed to cohesive) sediments, tending to fluidize the sediment matrix and almost certainly reducing the bulk modulus (hence sound speed). Since cohesive sediment dominated the upper 1.7 m, the largest bias errors would be expected to occur in the sediments below 1.7 m. Core-density measurements introduced two further sampling uncertainties: First splitting the entire core lengthwise and then scooping out by hand small samples for weighing. The magnitude of errors introduced by these two procedures is not known, but both would tend to reduce porosity and increase density.

The agreement between core and inversion results is remarkable over the first 1.7 m. The inversion results match core values in both velocity and density and capture the shape of the profiles indicated by the cores. This is an indication of

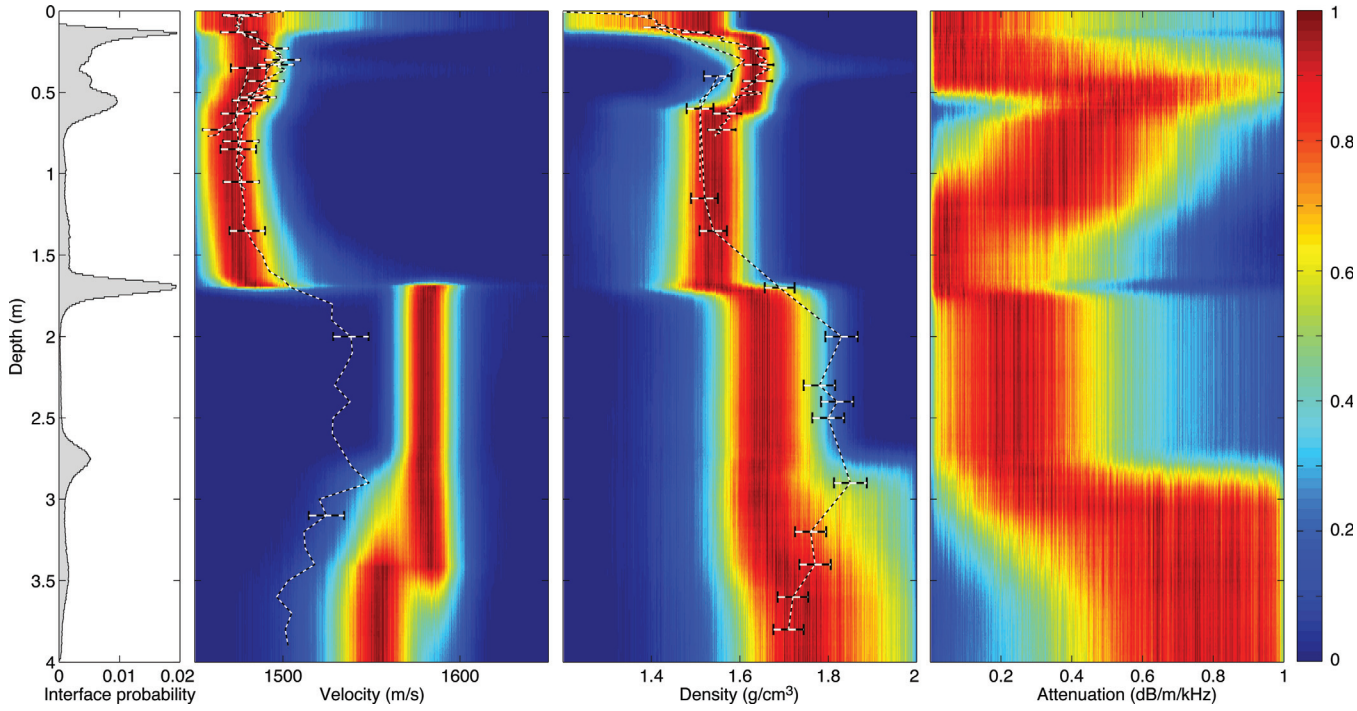

FIG. 9. (Color online) Trans-dimensional profile marginal distributions. Plot boundaries correspond to prior bounds. The solid lines indicate three cores taken at the experiment site with error estimates due to measurement errors.

the data regularization intrinsic to the method and suggests that the trans-dimensional approach can successfully approximate gradients by homogeneous layers where needed without over-fitting the data. To extract this information, it is crucial to rely on ensemble properties of the trans-dimensional PPD rather than point estimates such as the MAP model.

The inversion results show a strong velocity discontinuity at 1.7-m depth and a weak gradient preceding this discontinuity. This feature is also supported by the piston core which extends to this depth; however, the core estimates a velocity approximately 40 m/s lower than what the inversion results suggest. A likely explanation for the difference between the core and inversion sound velocity is that the core measurements are biased low. This conclusion is supported because the reflection coefficient is very sensitive to the velocity structure and the simulation (Fig. 5) indicates that the inversion estimates sound velocities well. In addition, the velocities below 1.7-m depth suggest sediment with a significant silt or sand component, which would tend to bias the velocities low in the core measurements as discussed above. Hence, although no conclusive statement can be made about the actual sediment velocities from 1.7 to 4 m, a likely explanation is that the core measurements are biased low. An additional discontinuity is evident at ~3.35 m. Here, the inversion results indicate a rather smooth transition from high velocities (~1590 m/s) to velocities ~25 m/s lower. The transition is evident from 2.8 to 3.5 m, matching the pattern of transition in the core very well.

Densities appear to be well resolved in the uppermost 1.5 m of the sediment, closely matching core estimates. Below 2.5-m depth, the data appear to be less sensitive to changes in density, although reasonable density values are recovered even at depth.

Overall, it is evident that this approach provides a useful method to capture both smooth spatial changes as well as sharp discontinuities in the environment, providing the capability of describing a large variety of environments without explicit specification. In addition, no artificial/subjective regularization terms need to be introduced that can potentially contaminate the parameter uncertainty estimates.

Data errors are addressed by including a data-error model in the inversion. In this case, the error model consists of one parameter, the standard deviation of the data, which is assumed to be independent of frequency. Figure 10 shows the marginal distribution of the standard deviation from the trans-dimensional PPD. Data errors are closely tied to the environmental parametrization and a poor choice of environmental parametrization can result in complicated and strongly correlated data errors. In practice, independent knowledge of data errors are often not readily available, since the effects due to parametrization, random processes in the environment, and other potential error sources cannot be separated. In such cases, data residuals (difference between observed and predicted data) are used to estimate the data errors. For an appropriate choice of environmental parametrization, data residuals should be largely uncorrelated (assuming no other source of correlated errors such as modal noise) and Gaussian distributed (as supported by the central limit theorem), since any significant structure/correlations in the residuals would indicate that something is missing in the computation of the predicted data. Hence, it is important to examine the data residuals *a posteriori* and validate that the assumption of a random Gaussian error distribution is reasonably fulfilled.

Posterior error validation can be carried out by applying quantitative and qualitative tests[11,12] to the residuals resulting from the MAP model. However, whether an individual model can be considered representative of the PPD is questionable for nonlinear inverse problems (where multiple modes and other complicated PPD structure commonly exist). This is particularly true in trans-dimensional approaches where the interpretation is best based on ensemble properties of the PPD.[15,24,39] Hence, data residuals are examined here based on ensembles. In particular, a large random subset of the PPD is used to compute a large sample of data residual sets (the *residual sample*), which are then used to examine the assumptions about data errors. Figure 6 shows the mean fit of data predicted for the PPD ensemble to the observed data. In addition, 95% HPD credibility intervals are shown. The figure shows that the PPD predictions fit the observed data well and that the 95% HPD bounds include most data points, indicating a good fit.

Figure 11 shows the mean autocorrelation function of the *residual sample* with 95% HPD bounds for the various frequencies. A narrow center peak of the autocorrelation function indicates residuals with short correlation lengths and supports the assumption of random errors. Figure 11 shows that, although data at 635 and 800 Hz have center

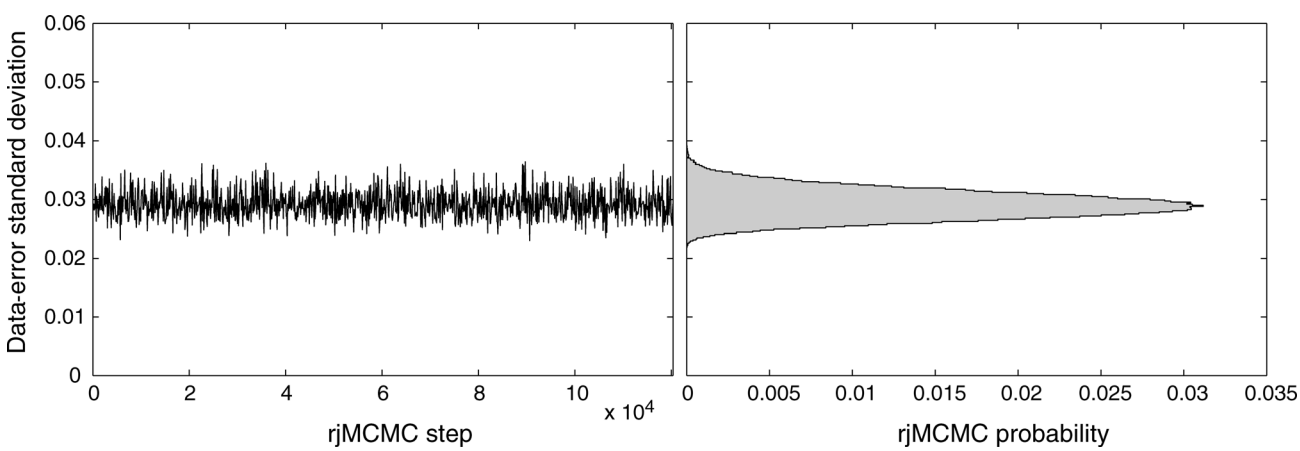

FIG. 10. Data-error standard deviation.

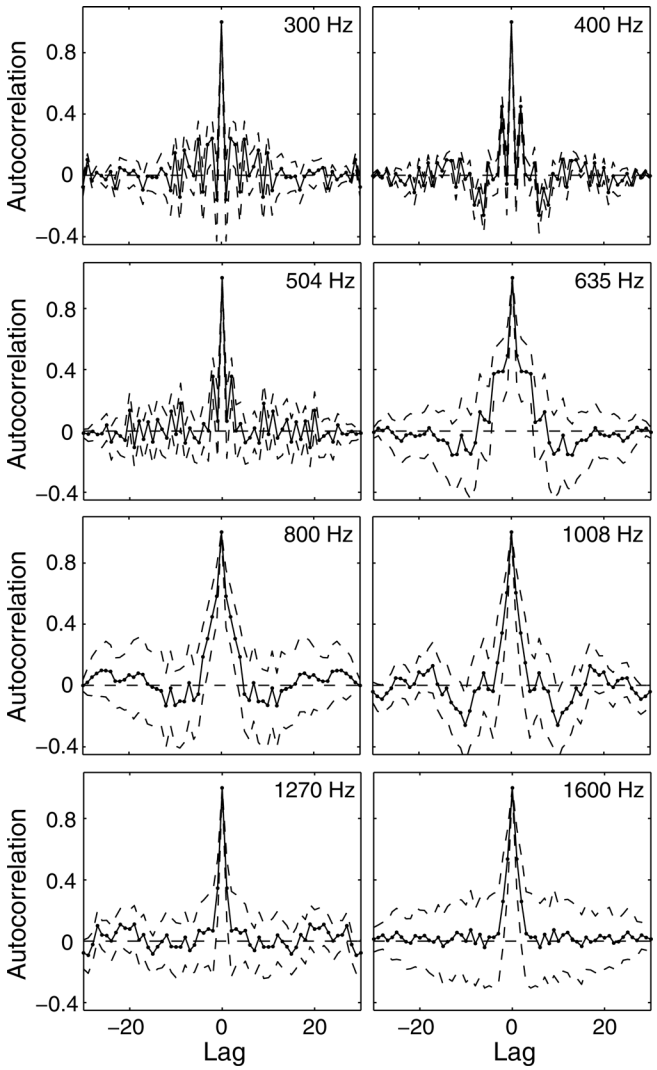

FIG. 11. Mean data residual autocorrelation function (solid line) and 95% highest probability bounds (dashed lines) from PPD.

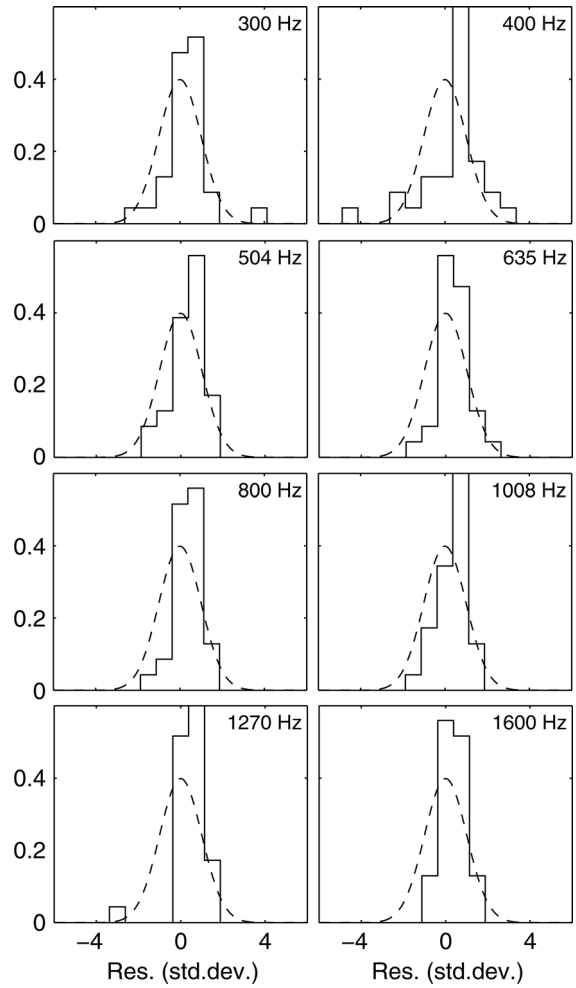

FIG. 12. Histograms of mean data residuals (solid) compared to Gaussian distributions.

peaks a few points wide, all other frequencies do not show obvious strong correlation structure, indicating the assumption of uncorrelated errors appears reasonably supported.

Figure 12 shows that histograms for the ensemble average of the residual sample compare favorably to analytic Gaussian distributions at most frequencies. Note that these histograms account for the error residuals for all models in the PPD. Histograms at some frequencies are more peaked near zero; however, this is not considered problematic since the predominant issue for violating the assumption of Gaussian data errors is large outliers, which are not present. These results support the structure and uncertainty estimates in Fig. 9 as reasonable representations of the seabed environment and raise confidence in the trans-dimensional inversion applied here.

IV. SUMMARY AND DISCUSSION

This paper developed a trans-dimensional partition modeling methodology for geoacoustic inverse problems where the number of unknown parameters is itself unknown and shows that under this approach simple environmental parametrizations can capture general seabed structure. This methodology results in an intrinsically parsimonious data self-regularization allowing for smooth transitions as well as sharp discontinuities in the inversion result. The approach is implemented using a reversible-jump MCMC sampler that applies changes in the dimension of the state-space using a birth–death scheme while maintaining detailed balance of the Markov chain, resulting in an unbiased trans-dimensional PPD estimate. The environment was parametrized with a partition model, treating the sediment as a partition over the depth interval of interest with sediment interface positions determined by the data. Data errors were addressed by introducing a simple error model, assuming the data errors to be independent of frequency and uncorrelated within each frequency band. Hence, the data-error model is included in the inversion using the standard deviation as an unknown.

Analysis of the inversion results was based on ensembles and marginalization obtained from the trans-dimensional PPD. In particular, marginal profile distributions were used to estimate the uncertainty of environmental properties such as the velocity- and density-depth profiles. The inversion methodology was applied to both simulated and measured seabed reflection-coefficient data. The simulation was based on an environment generated from three cores taken at an experiment site in the Straits of Sicily. The complicated

structure of the simulated environment (120 layers of varying thickness) was successfully approximated by the trans-dimensional inversion with an environmental model of 5–10 layers. The partition model resulted in the inversion placing interfaces over a range of depths in certain regions that allowed for resolving smooth transitions from ensemble inference applied to the PPD. In addition, sharp discontinuities were resolved in regions where required by the data.

The inversion results for measured data exhibit the parsimonious data self-regularization that is an intrinsic quality of the trans-dimensional partition-modeling methodology, in which the data information content determines the amount and type of structure in the posterior. This feature allows a wide variety of profile shapes to be described with a small number of parameters and without the need to explicitly parametrize in terms of smooth transitions (gradients) or apply subjective global regularization terms. Arbitrary transitions between parameter values, including gradients and sharp discontinuities, are inferred from the data.

Posterior examinations of data residuals based on ensemble properties of the PPD were carried out to ensure that assumptions about data-error statistics were met reasonably well. The results from these examinations increased the confidence in the inversion results.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of an Office of Naval Research (code 322OA) grant. The computational work was carried out on a parallel high-performance computing cluster operated by the authors at the University of Victoria funded by the Natural Sciences and Engineering Research Council of Canada and the Office of Naval Research. The authors thank Dr. Malcolm Sambridge for helpful discussions and two anonymous reviewers for improving the quality of the manuscript.

[1]S. E. Dosso, "Quantifying uncertainty in geoacoustic inversion. I. A fast Gibbs sampler approach," J. Acoust. Soc. Am. **111**, 129–142 (2002).
[2]D. J. Battle, P. Gerstoft, W. S. Hodgkiss, W. A. Kuperman, and P. L. Nielsen, "Bayesian model selection applied to self-noise geoacoustic inversion," J. Acoust. Soc. Am. **116**, 2043–2056 (2004).
[3]Y. Jiang, N. R. Chapman, and H. A. DeFerrari, "Geoacoustic inversion of broadband data by matched beam processing," J. Acoust. Soc. Am. **119**, 3707–3716 (2006).
[4]C.-F. Huang, P. Gerstoft, and W. S. Hodgkiss, "Uncertainty analysis in matched-field geoacoustic inversions," J. Acoust. Soc. Am. **119**, 197–207 (2006).
[5]C.-F. Huang, P. Gerstoft, and W. S. Hodgkiss, "Validation of statistical estimation of transmission loss in the presence of geoacoustic inversion uncertainty," J. Acoust. Soc. Am. **120**, 1932–1941 (2006).
[6]J. Dettmer, S. E. Dosso, and C. W. Holland, "Joint time/frequency-domain inversion of reflection data for seabed geoacoustic profiles," J. Acoust. Soc. Am. **123**, 1306–1317 (2008).
[7]D. Tollefsen, and S. E. Dosso, "Bayesian geoacoustic inversion of ship noise on a horizontal array," J. Acoust. Soc. Am. **124**, 788–795 (2008).
[8]C. Yardim, P. Gerstoft, and W. S. Hodgkiss, "Tracking of geoacoustic parameters using Kalman and particle filters," J. Acoust. Soc. Am. **125**, 746–760 (2009).
[9]C. F. Mecklenbräuker, P. Gerstoft, J. F. Bhme, and P.-J. Chung, "Hypothesis testing for geoacoustic environmental models using likelihood ratio," J. Acoust. Soc. Am. **105**, 1738–1748 (1999).
[10]T. Jasa, and N. Xiang, "Using nested sampling in the analysis of multi-rate sound energy decay in acoustically coupled rooms," AIP Conf. Proc. **803**, 189–196 (2005).
[11]J. Dettmer, S. E. Dosso, and C. W. Holland, "Model selection and Bayesian inference for high resolution seabed reflection inversion," J. Acoust. Soc. Am. **125**, 706–716 (2009).
[12]J. Dettmer, C. W. Holland, and S. E. Dosso, "Analyzing lateral seabed variability with Bayesian inference of seabed reflection inversions," J. Acoust. Soc. Am. **126**, 56–69 (2009).
[13]P. J. Green, "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination," Biometrika 711–732 (1995).
[14]P. J. Green, "Trans-dimensional Markov chain Monte Carlo," in *Highly Structured Stochastic Systems*, Oxford Statistical Science Series (Oxford University Press, Oxford, 2003), pp. 179–198.
[15]A. Malinverno, "Parsimonious Bayesian Markov chain Monte Carlo inversion in a non linear geophysical problem," Geophys. J. Int. **151**, 675–688 (2002).
[16]M. Sambridge, K. Gallagher, A. Jackson, and P. Rickwood, "Trans-dimensional inverse problems, model comparison and the evidence," Geophys. J. Int. **167**, 528–542 (2006).
[17]T. Bodin and M. Sambridge, "Seismic tomography with the reversible jump algorithm," Geophys. J. Int. **178**, 1411–1436 (2009).
[18]N. P. Agostinetti and A. Malinverno, "Receiver function inversion by trans-dimensional Monte Carlo sampling," Geophys. J. Int. **181**, 858–872 (2010).
[19]D. G. T. Denison, C. C. Holmes, B. K. Mallick, and A. F. M. Smith, *Bayesian Methods for Nonlinear Classification and Regression* (Wiley, New York, 2002). pp. 177–207
[20]N. Metropolis, A. Rosenbluth, M. Rosenbluth, and A. T. A. E. Teller, "Equations of state calculations by fast computing machines," J. Chem. Phys. **21**, 1087–1092 (1953).
[21]W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," Biometrika **57**, 97–109 (1970).
[22]D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms* (Cambridge University Press, Cambridge, 2003). pp. 343–386.
[23]R. C. Aster, B. Borchers, and C. H. Thurber, *Parameter Estimation and Inverse Problems*, International Geophysics Series (Elsevier, Amsterdam, 2005). pp. 99–171.
[24]T. Bodin, M. Sambridge, and K. Gallagher, "A self-parametrizing partition model approach to tomographic inverse problems," Inverse Probl. **25**, 055009 (2009).
[25]C. W. Holland, J. Dettmer, and S. E. Dosso, "Remote sensing of sediment density and velocity gradients in the transition layer," J. Acoust. Soc. Am. **118**, 163–177 (2005).
[26]D. C. Montgomery and E. A. Peck, *Introduction to Linear Regression Analysis* (Wiley, New York, 1992), pp. 115–136.
[27]S. E. Dosso and M. J. Wilmut, "Data uncertainty estimation in matched-field geoacoustic inversion," IEEE J. Ocean. Eng. **31**, 470–479 (2005).
[28]A. Malinverno and V. A. Briggs, "Expanded uncertainty quantification in inverse problems: Hierarchical Bayes and empirical Bayes," Geophysics **69**, 1005–1016 (2004).
[29]A. Gelman and X.-L. Meng, "Simulating normalizing constants: From importance sampling to bridge sampling to path sampling," Stat. Sci. **13**, 163–185 (1998).
[30]R. M. Neal, "Annealed importance sampling," Stat. Comp. **11**, 125–139 (2001).
[31]J. Skilling, "Nested sampling for general Bayesian computation," Bayesian Anal. **1**, 833–860 (2006).
[32]P. O. Hopcroft, K. Gallagher, and C. C. Pain, "Inference of past climate from borehole temperature data using Bayesian reversible jump Markov chain Monte Carlo," Geophys. J. Int. **171**, 1430–1439 (2007).
[33]L. M. Brekhovskikh and O. A. Godin, *Acoustics of Layered Media I: Plane and Quasi-Plane Waves*, Wave Phenomena, 1st ed. (Springer, New York, 1990), pp. 203–238.
[34]W. Gropp, E. Lusk, and A. Skjellum, *Using MPI, Portable Parallel Programming With the Message-Passing Interface* (MIT Press, Cambridge, 1999), pp. 171–185.
[35]A. F. M. Smith, "Bayesian computational methods," Philos. Trans. R. Soc. London **337**, 369–386 (1991).
[36]C. W. Holland, "Seabed reflection measurement uncertainty," J. Acoust. Soc. Am. **114**, 1861–1873 (2003).
[37]C. H. Harrison and J. A. Harrison, "A simple relationship between frequency and range averages for broadband sonar," J. Acoust. Soc. Am. **97**, 1314–1317 (1995).
[38]J. Dettmer, S. E. Dosso, and C. W. Holland, "Full wave-field reflection coefficient inversion," J. Acoust. Soc. Am. **122**, 3327–3337 (2007).
[39]A. Malinverno and W. S. Leaney, "Monte-Carlo Bayesian look-ahead inversion of walkaway vertical seismic profiles," Geophys. Prosp. **53**, 689–703 (2005).