EFFICIENCY OF TRANS-DIMENSIONAL BAYESIAN INFERENCE FOR GEOACOUSTIC INVERSION

Stan E. Dosso
School of Earth and Ocean Sciences, University of Victoria, Victoria B.C. Canada V8W 3P6

1 Overview

This paper considers trans-dimensional (trans-D) Bayesian inference applied to a representative geoacoustic inverse problem of estimating seabed parameters from acoustic reflectivity measurements. Trans-D methods include model selection in inversion by sampling the posterior probability density (PPD) over models with differing numbers of parameters (dimensions) [1-3]. The approach is applied here to samplie over seabed geoacoustic models with a variable number of layers [2,3], providing seabed profile estimates with uncertainties that include the uncertainty in the model parameterization. However, trans-D sampling can be computationally intensive. Sampling efficiency is largely determined by the proposal schemes applied to generate perturbed values for existing parameters and for new parameters assigned to layers added to the model. Perturbations of existing parameters are considered in a principal-component (PC) space based on an eigenvector decomposition of the unit-lag parameter covariance matrix (computed from the history of sampled models, a diminishing adaptation) [3]. The relative efficiency of proposing new-layer parameters from the prior versus a Gaussian distribution focused near existing values (the common approch) is examined [3]. Parallel tempering [4], which employs a sequence of interacting samplers with successively-relaxed likelihoods, is also applied to increase the acceptance rate of new layers. The relative efficiency of various proposal schemes is compared through repeated inversions with a pragmatic convergence criterion [3].

2 Theory

Consider measured data $\mathbf{d}$ and a set of possible models $\{\mathbf{m}_k\}$ indexed by $k \in K$ (where K is a countable set); e.g., k can indicate the number of seabed interfaces in a geoacoustic model. Bayes' theorem for a hierarchical model including hyper-parameter k can be written [1]

$$P(k,\mathbf{m}_k \,|\, \mathbf{d}) = \frac{P(k)P(\mathbf{m}_k \,|\, k)P(\mathbf{d} \,|\, k,\mathbf{m}_k)}{\sum_{k' \in K} \int P(k')P(\mathbf{m}'_{k'} \,|\, k')P(\mathbf{d} \,|\, k',\mathbf{m}'_{k'})\, d\mathbf{m}'_{k'}}. \quad (1)$$

In Eq. (1), $P(k)$ and $P(\mathbf{m}_k|k)$ represent prior probability densities, and $P(\mathbf{d}|k,\mathbf{m}_k)$ is the conditional probability of $\mathbf{d}$ given k and $\mathbf{m}_k$, which, for (fixed) measured data is interpreted as the likelihood $L(k,\mathbf{m}_k)$. The PPD $P(k,\mathbf{m}_k|\mathbf{d})$ is defined over the trans-D model space defined by the union of all fixed-dimensional spaces spanned by K. A Markov chain which samples the trans-D PPD can be obtained via reversible-jump Markov-Chain Monte Carlo (rjMCMC) sampling [1] in which proposed state $(k',\mathbf{m}'_{k'})$ is generated from an existing state $(k,\mathbf{m}_k)$ via a proposal distribution $Q(k',\mathbf{m}'_{k'}|k,\mathbf{m}_k)$ and accepted with the Metropolis-Hastings-Green acceptance probability

$$A(k',\mathbf{m}'_{k'} \,|\, k,\mathbf{m}_k) = \min\left[1, \frac{Q(k,\mathbf{m}_k \,|\, k',\mathbf{m}'_{k'})}{Q(k',\mathbf{m}'_{k'} \,|\, k,\mathbf{m}_k)} \frac{P(k')P(\mathbf{m}'_{k'} \,|\, k')}{P(k)P(\mathbf{m}_k \,|\, k)} \frac{L^{1/T}(k',\mathbf{m}'_{k'})}{L^{1/T}(k,\mathbf{m}_k)} \,|\, \mathbf{J} \,|\right]. \quad (2)$$

In Eq. (2), T is referred to as the sampling temperature which relaxes the likelihood as used in parallel tempering (discussed below) and $|\mathbf{J}|$ is the determinant of the Jacobian matrix for transitions from $(k,\mathbf{m}_k)$ to $(k',\mathbf{m}'_{k'})$. For transitions which perturb the parameters of the present state but do not change model dimension (i.e., $k' = k$) it follows that $|\mathbf{J}| = 1$. The birth/death rjMCMC scheme [1,2] changes k by 1 (e.g., adds/deletes a single seabed interface) such that only the parameters added/deleted are changed, with proposed parameters in a birth step depending only on the parameters in the current state; in this case it also follows that $|\mathbf{J}| = 1$.

The rjMCMC algorithm draws dependent samples from the trans-D PPD such that representative parameter estimates and uncertainties (e.g., marginal probability densities) can be computed. However, the approach can be computationally intensive, with efficiency largely determined by the effectiveness of the proposal schemes employed for perturbing existing parameters in fixed dimensions and for adding new parameters in birth steps. Several approaches have been considered to improve sampling efficiency (described in detail in [3]). These include: (1) Parallel tempering [4], in which a sequence of interacting Markov chains are run at a range of temperatures $T \geq 1$ to successively relax the likelihood and improve the acceptance rate of birth and death steps. In this case either only the $T = 1$ chain is retained as parameter samples or all chains are retained and importance reweighting is applied to remove the sampling bias due to the non-unity sampling temperatures. (2) Drawing parameter perturbations in a PC parameter space determined by an eigenvector decomposition of the unit-lag covariance matrix estimated adaptively from the sampling history, effectively de-correlating parameters in the sampling. (3) Drawing new parameters in birth steps from their prior rather than from Gaussian distribution focussed near existing parameter values. Some of these approaches are considered in terms of sampling efficiency in this paper.

3 Results

Sampling approaches are considered here for the geoacoustic inverse problem of estimating profiles of seabed properties (sound speed, density, attenuation) from angle- and frequency-dependent measurements of acoustic reflection coefficients [2,3]. Figure 1 shows an example of the results

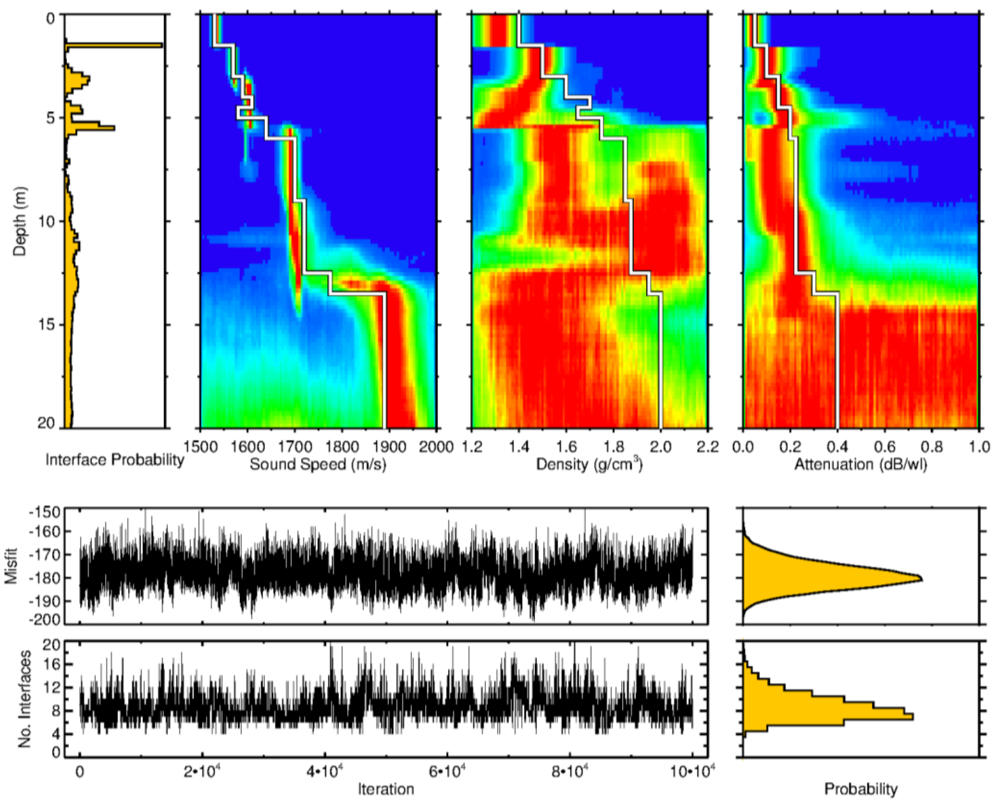

Figure 1: Example result of geoacoustics inversion results for (simulated) reflection coefficient data. The upper row, from left to right, shows marginal probability profiles of interface occurrence and of geoacoustic properties sound speed, density, and attenuation (solid lines indicate true values; hot colours indicate high probabilities). The middle and lower rows show the sampling history of the rjMCMC algorithm and marginal distributions for the misfit (negative log likelihood function) and the number of interfaces (the true number is 9).

of the trans-D geoacoustic inversion algorithm applied to (simulated) reflection coefficient data. To consider the efficiency of various approaches to sampling, a set of 10 inversions were run for each of four proposal schemes (PSs) with an automated convergence criterion based on examining the differences between parallel samplers [3]. PS1 consisted of PC parameter perturbations, geoacoustic parameters drawn from the prior in birth moves, and parallel tempering with 5 chains and $T_i = 1.25^{i-1}$. PS2 is the same as PS1 except that importance re-weighting is applied such that marginal profiles are computed using all parallel-tempering chains. PS3 is the same as PS1 except that parameter perturbations are drawn from a Gaussian distribution centred at current values with standard deviation equal to 1/10 of the prior-bound width. Finally, PS4 is the same as PS1 except that geoacoustic parameters in birth moves are drawn from a Gaussian distribution with standard deviation 1/30 of the prior-bound width and the number of parallel tempering chains is increased to 10 to help offset the resulting decrease in birth/death acceptance.

Figure 2 shows that proposal schemes PS1 and PS2 are similar in efficiency, indicating that retaining all samples in parallel tempering and applying importance re-weighting does not provide a significant benefit in sampling efficiency. However, both PS1 and PS2 are much more efficient than PS3 and PS4. This indicates that drawing new parameters in birth steps from the prior is more efficient than the commonly-used approach of applying a focused (Gaussian) proposal density. Mathematical reasons for this are discussed in [3], but are beyond the scope of this proceedings paper.

References

[1] P. J. Green, Trans-dimensional Markov chain Monte Carlo in highly structured stochastic systems. *Oxford Statistical Sciences Series* (Oxford: Oxford University Press), 179–98, 2003.

[2] J. Dettmer, S. E. Dosso, and C. W. Holland, Trans-dimensional geoacoustic inversion, *J. Acoust. Soc. Am.* 128:3393–405, 2010.

[3] S. E. Dosso, J. Dettmer, G. A. W. Steininger, and C. W. Holland, Efficient trans-dimensional inversion for geoacoustic parameter estimation. *Inverse Problems*, 30:1-29, 2014.

[4] S. E. Dosso, C. W. Holland, and M. Sambridge, Parallel tempering in strongly nonlinear geoacoustic inversion, *J. Acoust. Soc. Am*. 132:3030–3040, 2012.

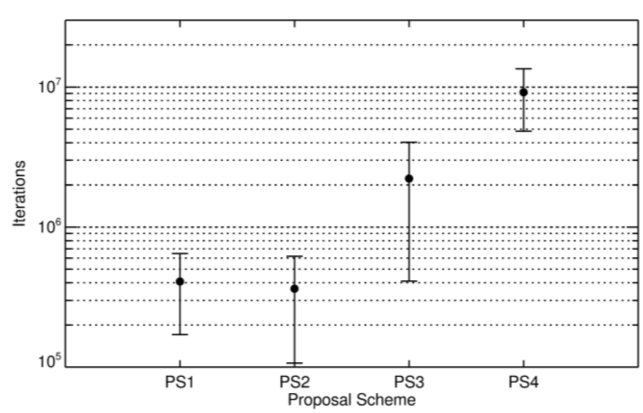

Figure 2: Comparison of the average number of iterations to convergence (with one standard-deviation error bars) for various proposal schemes (PSs) as described in the text.