Simultaneous Wavelet Estimation and Deconvolution of Reflection Seismic Signals

Qiansheng Cheng, Rong Chen, and Ta-Hsin Li, *Member, IEEE*

***Abstract*—In this paper, the problem of simultaneous wavelet estimation and deconvolution is investigated with a Bayesian approach under the assumption that the reflectivity obeys a Bernoulli–Gaussian distribution. Unknown quantities, including the seismic wavelet, the reflection sequence, and the statistical parameters of reflection sequence and noise are all treated as realizations of random variables endowed with suitable prior distributions. Instead of deterministic procedures that can be quite computationally burdensome, a simple Monte Carlo method, called Gibbs sampler, is employed to produce random samples iteratively from the joint posterior distribution of the unknowns. Modifications are made in the Gibbs sampler to overcome the ambiguity problems inherent in seismic deconvolution. Simple averages of the random samples are used to approximate the minimum mean-squared error (MMSE) estimates of the unknowns. Numerical examples are given to demonstrate the performance of the method.**

I. Introduction

In reflection seismology experiments (e.g., [16]), a test signal (or seismic source) is sent to probe the layered earth and the reflected signal (seismic trace) is recorded by a geophone or hydrophone at the surface. Geophysical structure of the earth is then investigated through an analysis of the reflectivity from deep layers of the earth. The true reflectivity, however, is not directly accessible; instead, the recorded signal is a smeared version of the reflectivity, caused by reverberations due to the surface layers. One of the important objectives in seismic data processing is to undo the effects of the degradation in order to recover the true reflectivity. This usually requires a certain *deconvolution* technique, since the received signal can be regarded as a convolution of the reflection sequence with an unknown seismic wavelet. As a common practice, the seismic wavelet is often modeled as a deterministic finite moving-average filter, while the reflectivity is assumed to be a sequence of random variables with scattered nonzero values (e.g., [16]). In most of the available deconvolution techniques, the reflection sequence is further assumed to be *white* so that the random variables in the sequence are mutually independent (e.g., [14], [15]).

Manuscript received August 16, 1994; revised July 25, 1995. This work was supported in part by the NSF under Grant DMS-9301193 and by the ONR under Grant N00014-92-C-0019.

Q. Cheng is with the Institute of Mathematics, Peking University, Beijing, China.

R. Chen and T.-H. Li are with the Statistics Department, Texas A&M University, College Station, TX 77843 USA.

Publisher Item Identifier S 0196-2892(96)00997-7.

To successfully recover the reflection sequence, it is crucial for the deconvolution algorithms to incorporate as much prior knowledge about the reflectivity as possible. Wiggins [18], for example, proposed a criterion that describes the "simplicity of appearance" of the reflection sequence, resulting in the well-known minimum entropy method (see also [2] and [4]). Probabilistic models for the reflection sequence are also helpful, among which the *Bernoulli–Gaussian white sequence model* has been proved very successful (e.g., [10], [14], [15], [17]).

In the Bernoulli–Gaussian white sequence model, the occurrence of nonzero values in the reflection sequence is governed by a Bernoulli law while the magnitude of nonzero values by an independent Gaussian distribution (e.g., [14], [15]). With suitable choice of parameters, the Bernoulli–Gaussian model can imitate the appearance of reflection sequences characterized by the sparseness of nonzero values. The problem, then, becomes that of simultaneous estimation of the seismic wavelet, the statistical parameters, and the reflection sequence. Since ambient noise is always present in the recorded signal, the statistical parameters to be estimated also include those that describe the probability distribution of the noise.

To estimate these unknown quantities, maximum likelihood (ML) approach has been shown to yield satisfactory results (e.g., [14], [15]); yet, to obtain the ML estimates numerically involves highly complicated iterative algorithms and detection schemes that can be quite computationally expensive (e.g., [9], [11], [14], [15]). Besides, the convergence of these algorithms to desired solutions is not always guaranteed, depending on the choice of initial guess.

Under the Bernoulli–Gaussian white sequence model, the present paper deals with the deconvolution problem from a *Bayesian* point of view: The unknown quantities are treated as random variables with suitable prior distribution and inferences about the unknowns obtained on the basis of their posterior distribution. This approach is found to be able to incorporate nonminimum-phase wavelet and even colored reflection sequences into a unified framework and make effective use of prior information about the unknown quantities.

Instead of directly computing the Bayesian estimates, we employ a Monte Carlo procedure—the *Gibbs sampler*—to iteratively generate random samples from the joint posterior distribution of the unknowns. Bayesian estimators, such as the minimum mean-squared error (MMSE) estimator and the maximum *a posteriori* (MAP) estimator, of the unknown quantities can be easily approximated using the random samples. For instance, to find the MMSE estimate of the reflection sequence, one simply calculates the average of the corresponding components in the sample, and thus avoids the burdensome computation of the conditional expectation of the reflection sequence given the recorded signal.

The Gibbs sampler has been applied with success to image restoration problems when the degradation filter and the statistical parameters of the signal and noise are known *a priori* (e.g., [7]). Extensions to the cases of *unknown* degradation filters and statistical parameters have also been developed recently by Chen and Li [1] and Liu and Chen [13] for the deconvolution of discrete-valued input signals (see also [12]). Our experiment in this paper on simulated data further proves the suitability of the Gibbs sampler in seismic deconvolution.

The rest of the paper is organized as follows. Section II provides the mathematical formulation of the problem. Section III introduces the Bayesian approach and the prior distributions to be used. Then the joint posterior and conditional posterior distributions of the unknowns to be used in the Gibbs sampler are derived. Section IV briefly reviews the Gibbs sampler in general and details its implementation in our particular problem. Some simulation results are presented in Section V.

II. Formulation of the Problem

Given a seismic trace, the observed data, $\{x_1, \cdots, x_n\}$, can be modeled as a convolution of the unknown seismic wavelet $\{\phi_i\}$ with the unknown reflection sequence $\{u_t\}$, namely

$$x_t = \sum_{i=0}^{q} \phi_i \, u_{t-i} + \epsilon_t \tag{1}$$

where the ϵ_t stand for the additive measurement error and noise. Assume that the ϵ_t are mutually independent and have a common Gaussian distribution $N(0, \sigma^2)$ with mean zero and a certain unknown variance σ^2.

It is clear that some information about the reflection sequence must be provided in order to distinguish it from the seismic wavelet. The Bernoulli–Gaussian white sequence model ([11], [14], [15]) is one of the well-established models in this regard. The model assumes that $\{u_t\}$ is a white sequence whose common probability distribution is a mixture of Bernoulli and Gaussian distributions. In other words, the distribution of u_t takes the form

$$\eta \, \delta(u_t = 0) + (1 - \eta) \, N(0, r^2) \tag{2}$$

where $\eta \in (0, 1)$ is the probability that $u_t = 0$. According to this model, the nonzero values of u_t follow a Gaussian distribution with mean zero and a certain variance r^2. Both η and r^2 are unknown in practice, and hence, to be estimated along with the reflection sequence from the observed data.

The seismic wavelet $\{\phi_i\}$ is a broad-band sequence often modeled as being deterministic with finite length (e.g., [8], [14], [15]). The order q of the seismic wavelet is assumed to be available in the sequel. The causality assumption of $\{\phi_i\}$, i.e., the assumption that $\phi_i = 0$ for any $i < 0$, is not crucial since a noncausal wavelet of finite length [17] can be easily transformed into a causal one by index substitution (both i and t). The scale ambiguity in $\{\phi_i\}$, however, is crucial to the existence of unique solution. In fact, since multiplying the ϕ_i by any nonzero constant does not change the problem if the reflectivity is re-scaled accordingly, the deconvolution problem cannot be uniquely solved without imposing extra conditions that eliminate the ambiguity. To this end, we assume that the wavelet is normalized with $\phi_0 = 1$, although other conditions, such as the total power of the seismic wavelet, also help to achieve the same objective.

The ultimate goal of simultaneous wavelet estimation and deconvolution is to estimate $\{\phi_i\}$, η, r^2, σ^2, as well as $\{u_1, \cdots, u_n\}$, solely on the basis of $\{x_1, \cdots, x_n\}$. Clearly, there are more unknowns to be estimated than the observed data. This partially explains the difficulties that arise in the maximum likelihood method in seeking to maximize a function of such enormous dimension. The Bayesian approach, on the other hand, helps to ease the problem, especially when combined with Gibbs sampling.

III. The Bayesian Approach

To deal with the aforementioned problem of simultaneous wavelet estimation and deconvolution, the Bayesian approach assumes that the unknown quantities are realizations of random variables governed by certain *prior* probability distributions. This assumption is particularly reasonable for the seismic wavelet since the reverberations in the surface layers of the earth are indeed random in nature [14]. For other parameters, the imposed prior distributions simply reflect the degree of uncertainty of belief about the values of the parameters in a mathematical way that can be easily coped with by the Bayes theorem.

In the Bayesian paradigm, inferences about the unknowns are made by investigating their joint *posterior* probability distribution that reflects the gain of knowledge from the observed data toward the unknowns. Widely-used Bayesian point estimates include the conditional expectation, known as the minimum mean-squared error (MMSE) estimator, and the mode of the posterior distribution, known as the maximum *a posteriori* (MAP) estimator; the former minimizes the mean-squared error criterion while the latter intends to provide the maximum certainty about the parameter in a way similar to the maximum likelihood principle.

A. Prior Distributions

A unique feature of Bayesian approach rests on its ability to incorporate prior information about unknown parameters in the form of *prior probability distribution.* A suitable choice of priors is essential not only to the accommodation of prior information but to the insurance of simplicity of the resulting posterior distributions as well. Roughly speaking, the priors should contain predetermined parameters that reflect the degree of uncertainty toward the unknowns. If no prior information is available, flat or "nearly" flat priors should be used. The prior distributions should also be simple enough so as to minimize the computational burden of the posterior distributions. Conjugate priors are commonly used for this purpose.

A *conjugate* family of prior distributions is closed under sampling, i.e., if the prior distribution is in the family, the posterior distribution given a sample will also be in the family, regardless of the observed values in the sample and the sample size. Hence, if a conjugate prior is used, the posterior distribution will have the same form with updated information reflected by a change in the parameters of the distribution. For example, the beta distribution is a conjugate family for samples from a Bernoulli distribution—when (binary) data are observed from a Bernoulli(θ) distribution and θ follows a Beta(a, b) prior, the posterior distribution of θ given the data will always be a beta distribution. If fact, the posterior distribution is Beta($a + X$, $b + Y$), where X and Y are the numbers of 0's and 1's, respectively, in the data set.

For most commonly-used distribution families, a conjugate prior also ensures that the resulting MMSE estimator is consistent (e.g., [3], p. 335). The MMSE estimator may be biased if the mean of the prior does not coincide with the parameter. However, as the sample size grows, the effect of the prior distribution fades away and the bias eventually vanishes.

Based on these principles, the following priors are employed in our problem of deconvolution and all of them are conjugate priors. Other choices are possible but may not lead to computationally tractable posteriors.

1) Assume that η has a prior beta distribution $\pi(\eta) \sim$ Beta(a, b), where $a > 0$ and $b > 0$ are predetermined parameters that control the shape of the distribution.
2) Assume that r^2 is a known constant. Another possibility is to regard r^2 as a random variable with an inverted chi-square distribution so that $\mu\gamma/r^2 \sim \chi^2(\mu)$ for some $\gamma > 0$ and $\mu > 0$. For simplicity, we shall not pursue this option in the present paper. It should be noted that in practice r^2 may vary with time so that a time-dependent model for r^2 is more realistic. For instance, r^2 may obey a gain law such as exponential decay with some unknown parameters. If r^2 is allowed to change with time, the problem becomes more complicated, though still solvable. The difficulty comes from the fact that the conditional posterior distribution of the parameters in r^2 may not have a simple form and hence more sophisticated sampling procedures such as the rejection method have to be employed to generate random samples for these parameters. At this stage of development, stationarity of the data record is assumed.
3) Given η and r^2, the reflection sequence $\{u_t\}$ is assumed to be i.i.d. (independent and identically distributed) with a common distribution of the form (2). For convenience, it is also assumed that $u_t = 0$ with probability one for $t = -q+1, \cdots, -1, 0$.
4) With the restriction $\phi_0 = 1$, assume that the rest parameters $\Phi := [\phi_1, \cdots, \phi_q]^T$ in the seismic wavelet obey the multivariate Gaussian law of the form $\pi(\Phi) \sim N_q(\Phi_0, \Sigma_0)$ for some predetermined mean vector Φ_0 and covariance matrix Σ_0.
5) The noise variance σ^2 is assumed to have an inverted chi-square prior distribution $\pi(\sigma^2) \sim \chi^{-2}(\nu, \lambda)$, i.e., $\nu\lambda/\sigma^2 \sim \chi^2(\nu)$, for some known parameters $\lambda > 0$ and $\nu > 0$.

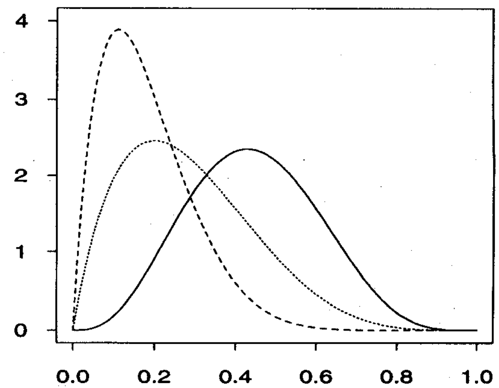


Fig. 1. Prior probability density of η, the Beta(a, b) distribution: solid line for $a = 4$ and $b = 5$, dotted line for $a = 2$ and $b = 5$, and dashed line for $a = 2$ and $b = 9$.

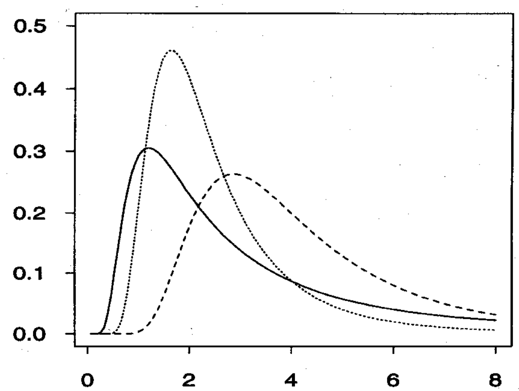


Fig. 2. Prior probability density of σ^2, the $\chi^{-2}(\nu, \lambda)$ distribution: solid line for $\nu = 3$ and $\lambda = 2$, dotted line for $\nu = 9$ and $\lambda = 2$, and dashed line for $\nu = 9$ and $\lambda = 3.5$.

Typical graphs of the beta and inverted chi-square distributions with various parameters can be found in Figs. 1 and 2, respectively. These graphs help to understand the impact of each parameter on the shape and center of the priors. Since the degree of uncertainty about the unknowns is controlled by these parameters, one is always advised to experiment with a few combinations of the parameter values until satisfactory results can be obtained.

B. Posterior Distributions

Equipped with these priors, one can easily derive the *conditional posterior distributions* for a later use in the Gibbs sampler. To this end, it is crucial to note that under the assumption that $\{\epsilon_t\}$ is Gaussian white noise the joint distribution of $X := \{x_1, \cdots, x_n\}$, Φ, $U := \{u_{1-q}, \cdots, u_n\}$, σ^2, and η takes the form of

$$\begin{aligned}\pi(X, \Phi, U, \sigma^2, \eta) &= \pi(X|\Phi, U, \sigma^2, \eta)\pi(\Phi)\pi(U|\eta)\pi(\sigma^2)\pi(\eta)\\ &= \frac{1}{(2\pi\sigma^2)^{n/2}}\end{aligned}$$

$$\cdot \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=1}^{n}\left(x_t - \sum_{t=1}^{q}\phi_i u_{t-1}\right)^2\right\}$$
$$\cdot\ \pi(\Phi)\pi(U|\eta)\pi(\sigma^2)\pi(\eta). \tag{3}$$

This expression leads to the following conditional posterior distributions.

1) Let $U_t := [u_{t-1}, \cdots, u_{t-q}]^T$; then

$$\pi(\Phi|\,\text{rest}) \sim N_q(\Phi_*, \Sigma_*) \tag{4}$$

where "rest" $:= (X, U, \sigma^2, \eta)$,

$$\Sigma_*^{-1} := \frac{1}{\sigma^2}\sum_{t=1}^{n} U_t U_t^T + \Sigma_0^{-1}$$

and

$$\Phi_* := \Sigma_*\left[\frac{1}{\sigma^2}\sum_{t=1}^{n}(x_t - u_t)U_t + \Sigma_0\Phi_0\right].$$

2) Let $s^2 := \sum_{t=1}^{n}(x_t - \sum_{i=0}^{q}\phi_i u_{t-i})^2$; then

$$\pi(\sigma^2|\,\text{rest}) \sim \chi^{-2}(\nu_*, \lambda_*)$$

where "rest" $:= (X, \Phi, U, \eta)$, $\nu_* := \nu + n$, and $\lambda_* := \nu\lambda + s^2$.

3) For any fixed $j \in \{1, \cdots, n\}$, it is not difficult to show from (3) that

$$\pi(u_j|\,\text{rest}) \propto$$
$$\left[\eta\delta(u_j = 0) + (1-\eta)\frac{1}{\sqrt{2\pi r^2}}\exp\left\{\frac{-\left(\frac{u_j}{r}\right)^2}{2}\right\}\right]$$
$$\times \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=j}^{j+q}\left(x_t - \sum_{i=0}^{q}\phi_i u_{t-i}\right)^2\right\}.$$

Let $\epsilon'_t := x_t - \sum' \phi_i u_{t-i}$ where the sum $\sum'$ is over $i = 0, 1, \cdots, q$ such that $i \neq t - j$; then, after some straightforward manipulations, one obtains

$$\pi(u_j|\,\text{rest}) \sim \eta_j\,\delta(u_j = 0) + (1-\eta_j)\,N(u_j^*, r_j^2)$$

where

$$\eta_j := \frac{\eta}{\eta + (1-\eta)\dfrac{r_j}{r}\exp\left\{\dfrac{\left(\dfrac{u_j^*}{r_j}\right)^2}{2}\right\}},$$

$$u_j^* := \frac{r^2\sum_{i=0}^{q}\epsilon'_{j+i}\phi_i}{r^2\sum_{i=0}^{q}\phi_i^2 + \sigma^2},$$

and

$$r_j^2 := \frac{r^2\sigma^2}{r^2\sum_{i=0}^{q}\phi_i^2 + \sigma^2}.$$

It is worthwhile to note that using standard routines one can easily generate random samples from these conditional posterior distributions. This turns out to be crucial to successful application of the Gibbs sampler to the deconvolution problem. We also remark that the proposed approach can be easily extended to handle colored Gaussian noise (e.g., a MA(q_ε) process). In this case, the probability density function $\pi(X|\Phi, U, \sigma^2, \eta)$ in (3) needs slight modification (though still an exponential function) in order to incorporate the correlation structure of the noise. The other calculations are similar. For simplicity, however, we use the Gaussian white noise model in developing our algorithms.

IV. The Gibbs Sampler

Because of its high complexity, direct inference from the joint posterior distribution of the unknown quantities in our problem is extremely difficult, if not impossible. The Gibbs sampler serves as a means to circumvent the difficulties.

Rather than pinpoint the optimal solutions, as many numerical optimization algorithms do, the Gibbs sampler generates a Markov chain of random samples whose equilibrium distribution coincides with the desired joint posterior distribution of the unknowns. It is mostly helpful when the joint posterior distribution is complicated while the conditional posterior distributions are simple. This is indeed the case in our problem.

A. A Generic Example

To be more specific, let us consider a generic example of making inference about three unknown variables/vectors z_1, z_2, and z_3 from the data vector x. Suppose the joint posterior distribution of (z_1, z_2, z_3) given x is known to be $p(z_1, z_2, z_3|\,x)$. The objective of Gibbs sampling is to generate random samples of (z_1, z_2, z_3) according to the distribution $p(z_1, z_2, z_3|\,x)$. A commonly-used procedure of Gibbs sampling proceeds iteratively as follows: Starting with a trivial set of initial values $(z_{1,0}, z_{2,0}, z_{3,0})$ and $m = 0$:

1) Step 1. Generate $z_{1,m+1}$ from $p(z_1|\,z_2 = z_{2,m}, z_3 = z_{3,m}, x)$;
2) Step 2. Generate $z_{2,m+1}$ from $p(z_2|\,z_1 = z_{1,m+1}, z_3 = z_{3,m}, x)$;
3) Step 3. Generate $z_{3,m+1}$ from $p(z_3|\,z_1 = z_{1,m+1}, z_2 = z_{2,m+1}, x)$;
4) Step 4. Set $m = m + 1$ and go to Step 1.

In so doing, the Gibbs sampler produces a sample Markov chain $\{(z_{1,m}, z_{2,m}, z_{3,m}), m = 0, 1, \cdots\}$ whose limiting (or equilibrium) distribution as $m \to \infty$ can be shown to agree with the desired posterior distribution $p(z_1, z_2, z_3|\,x)$ under some regularity conditions. For details, see [5] and [7].

To ensure the convergence in practice, the first N samples in sequence $\{(z_{1,m}, z_{2,m}, z_{3,m}), m = 0, 1, \cdots\}$ are usually discarded and the remaining ones, $(z_{1,N+m}, z_{2,N+m}, z_{3,N+m})$ for $m = 1, \cdots, M$, can be regarded as independent samples governed by the distribution $p(z_1, z_2, z_3|x)$. To eliminate possible correlations between the samples, one can also save one sample from every k samples to obtain $(z_{1,N+mk}, z_{2,N+mk}, z_{3,N+mk})$ for $m = 1, \cdots, M$. From the

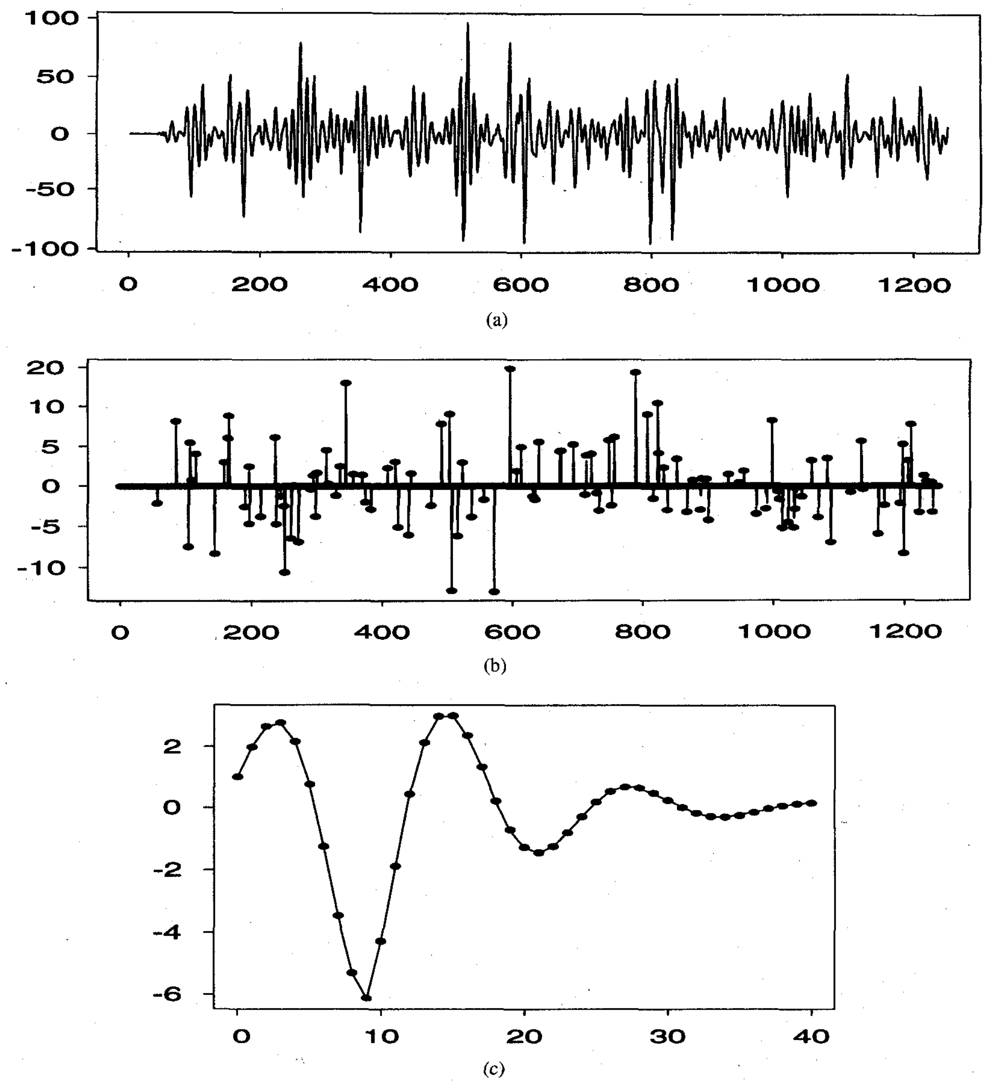


Fig. 3. (a) Observed time series with SNR = 26 dB. (b) MMSE estimates of reflectivity from the data in (a), where the dots stand for true values and the solid lines for estimates. (c) MMSE estimates of the seismic wavelet in (5) from the data in (a), where true values of the wavelet are indicated by dots and their estimates by solid curve.

saved samples, inference can be made about the unknowns. For instance, the MMSE estimate of z_1, namely $E(z_1|x)$, can be approximated by averaging $z_{1,N+m}$ (or $z_{1,N+mk}$) for $m = 1, \cdots, M$.

B. Remarks on Implementation

In our problem of deconvolution, the Gibbs sampler is implemented by using the conditional posterior distribution in Section III. To accelerate the convergence, the reflection sequence is "visited" five times more often than the other parameters due to the high correlation between successive samples of the reflection sequence.

It is also worth pointing out that even with the constraint $\phi_0 = 1$ there are still possible ambiguities in the deconvolution problem, of which the shift ambiguity is especially prominent in some cases. For example, if the leading coefficients $\phi_0, \cdots, \phi_k$, (for some $k < q$ with $\phi_k \approx 1$), are relatively small in the seismic wavelet, the Gibbs sampler may not be able to effectively distinguish the time-shifted model

$$y_t = \sum_{j=0}^{q} \phi'_j x'_{t-j} + \varepsilon_t$$

from the true model in (1), where $\phi'_j := \phi_{j+k}$ for $j = 0, 1 \cdots, q$ (assuming $\phi_j = 0$ for $j > q$) and $x'_t := x_{t-k}$ for $t = 0, 1 \cdots, q$. In such cases, the posterior distribution of the unknowns becomes approximately a *mixture* of several distributions—each corresponding to a possible time delay—that play the same role as local equilibrium in optimization problems. As a result, the convergence speed of the Gibbs sampler may be considerably decreased.

To overcome this problem, we employ the following *constrained* Gibbs sampler as reported in [1]. Since in many cases the seismic wavelet has a unique maximum in absolute value, one can impose a constraint so that the largest $|\phi_j|$ appears at

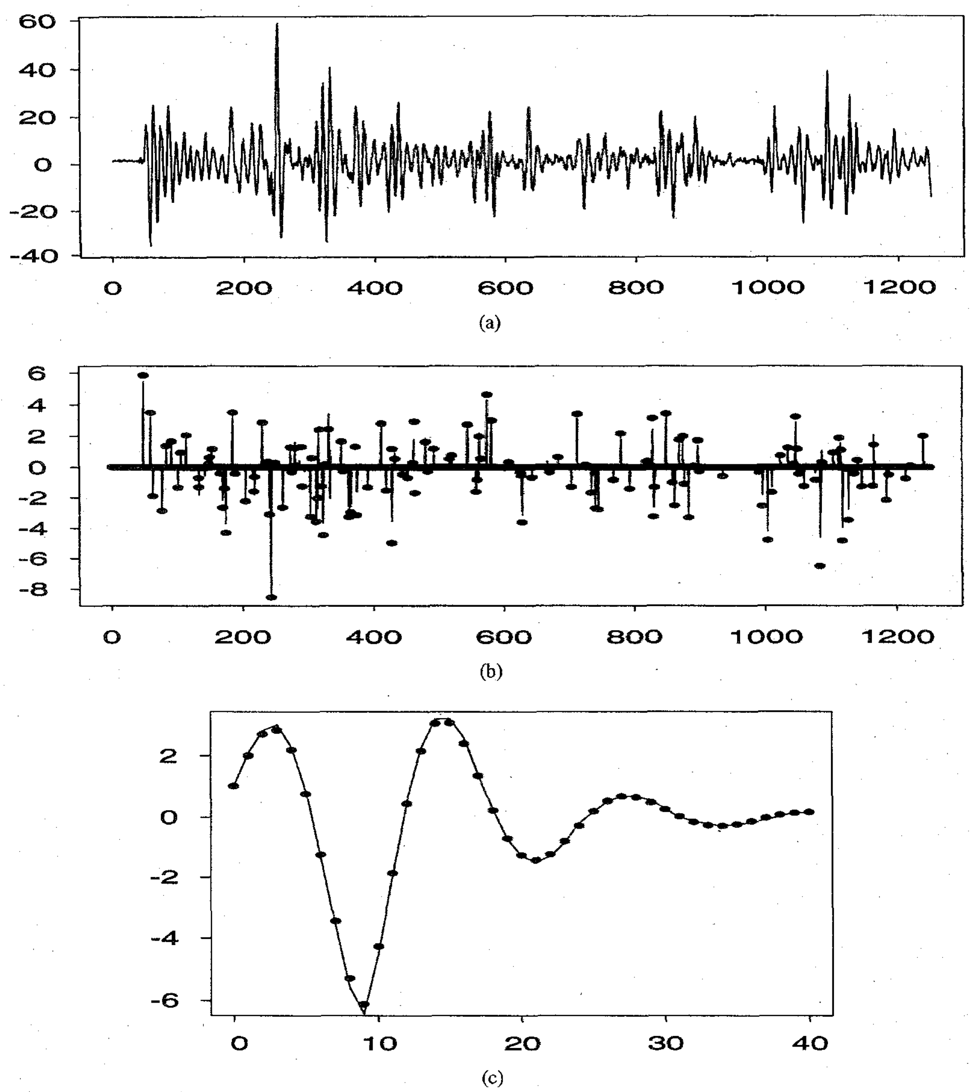


Fig. 4. (a) Observed time series with SNR = 18.6 dB. (b) MMSE estimates of reflectivity from the data in (a), where the dots stand for true values and the solid lines for estimates. (c) MMSE estimates of the seismic wavelet in (5) from the data in (a), where true values of the wavelet are indicated by dots and their estimates by solid curve.

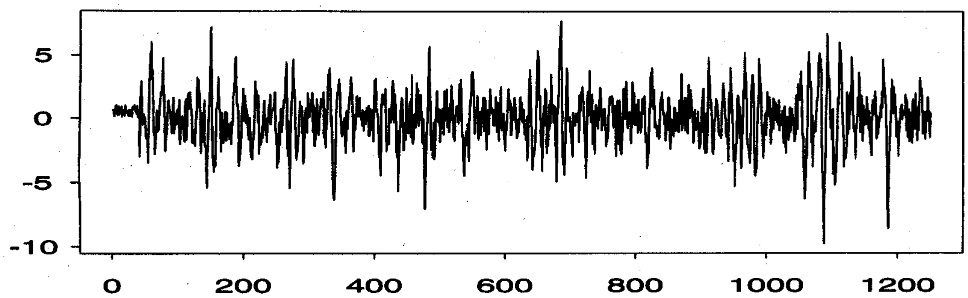
a known location. For instance, assuming that $|\phi_k| > |\phi_j|$ for some $1 \leq k \leq q$ and all $j \neq k$ forces the largest wavelet coefficient to appear at lag k. With this restriction, the prior of Φ becomes $\pi(\Phi) \propto N_q(\Phi_0, \Sigma_0)\, I(|\phi_k| > |\phi_j|, j = 1, \cdots, q$ and $j \neq k)$. It is easy to show that the conditional posterior distribution of Φ given the rest unknowns remains the same as in (4) except for an extra factor $I(|\phi_k| > |\phi_j|, j = 1, \cdots, q$ and $j \neq k)$. An easy way to sample from this constrained Gaussian distribution is to use the *rejection method,* i.e., when a sample is drawn from the unconstrained distribution (4), it is checked to see if the constraint is satisfied; if not, the sample is rejected and a new sample is generated. This procedure continues until a sample that satisfies the constraint is obtained. One can also restart the Gibbs sampler when it takes too long (e.g., 1000 runs) for the rejection method to generate a desired sample. In this case, it seems plausible to start with the last rejected Φ and shift it forward or backward (with vacancies filled by zeros) so that the resulting wavelet satisfies the constraint.

V. Simulation Results

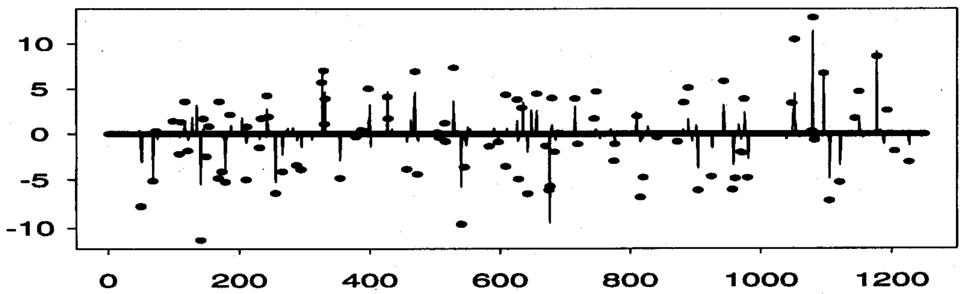
Some simulations are carried out to test the proposed method. The true seismic wavelet $\{\phi_i, i = 1, \cdots, 40\}$ is assumed to be of the form $\phi_i := \phi_i^* / \phi_0^*$, where

$$\phi_i^* := \sin \frac{\pi(i+1)}{6.4} \exp\{-0.12\,|i-9|\}. \quad (5)$$

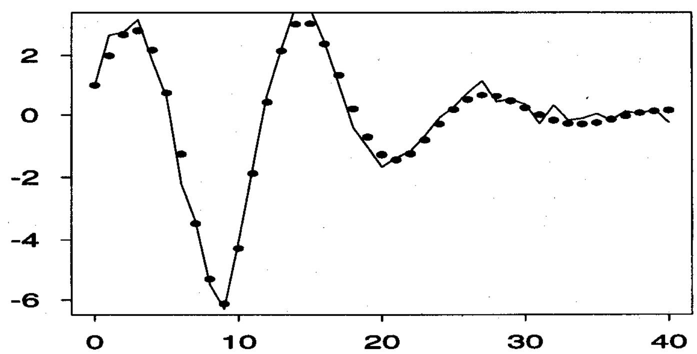
The reflection sequence $\{u_t\}$ is generated from a Bernoulli–Gaussian distribution (2) with $\eta = 0.9$. Three different values of r^2 are considered, namely $r^2 = 25$, 4, and 0.25, corresponding a signal-to-noise ratio (SNR) of 26, 18.6, and 4 dB, respectively. The wavelet estimation and deconvolution results are shown in Figs. 3–5 for the three cases. The observed signal is obtained from (1)



(a)



(b)



(c)

Fig. 5. (a) Observed time series with SNR = 4 dB. (b) MMSE estimates of reflectivity from the data in (a), where the dots stand for true values ($\times$ 10) and the solid lines for estimates. (c) MMSE estimates of the seismic wavelet in (5) from the data in (a), where true values of the wavelet are indicated by dots and their estimates by solid curve.

for $t = 1, \cdots, 2000$, where $\{\epsilon_t\}$ is Gaussian white noise with zero mean and unit variance $\sigma^2 = 1$. Fig. 3(a) shows the observed signals for the case of SNR = 26 dB; Fig. 3(b) is the reflection sequence recovered on the basis of the MMSE criterion with the dots representing the true reflectivity; and Fig. 3(c) shows the MMSE estimate for the seismic wavelet along with the true wavelet. Figures 4 and 5 are organized in a similar way. In each of the cases, the Gibbs sampler is carried out with 4000 iterations and the last 1000 samples are averaged to obtain the MMSE estimates.

Note that checking for convergence in the Gibbs sampler is in general an open problem to be tackled with rigorous methods. Most of the time, the convergence depends on the specific problem to which the Gibbs sampler is applied, although several methods have been proposed recently for checking the convergence in general situations (e.g., [6]). One of the methods proposes to generate multiple sequences and to compare these sequences using some summary statistics in order to determine whether the Gibbs sampler has converged. This, however, requires additional computing time and memory. In our example, experiments show that 4000 iterations are enough for convergence.

In the above simulation, we assume that the location of the largest value of the seismic wavelet is known *a priori* in order to remove the shift ambiguities in the deconvolution problem. The prior distribution for Φ is assumed to be multivariate normal with mean zero and a diagonal covariance matrix with 100 as diagonal elements, namely $\Phi_0 = 0$ and $\Sigma_0 = 100I$. The prior distribution for η is assumed to be Beta(50, 10), i.e., $a = 50$ and $b = 10$. The order q of the seismic wavelet and the variance r^2 of the reflectivity are assumed to be known and the corresponding true values are used in the Gibbs sampler. The prior distribution for σ^2 is assumed to be $\chi^{-2}(2, 0.3)$, i.e., we take $\nu = 2$ and $\lambda = 0.3$. The initial values in the Gibbs sampler are taken to be $\sigma^2 = 1$, $u_t = 0$, for $t = 1, \cdots, 2000$, and $\eta = 0.9$.

As seen from these results, the proposed method is able to recover the reflectivity almost perfectly when the signal-to-noise ratio is high or moderate. For SNR as low as 4 dB, the method still provides satisfactory estimates for most of the reflection coefficients, especially for those with large values, although the performance generally deteriorates with the decrease of SNR.

The computing time of the proposed procedure is minimum due to the fact that all the conditional posterior distributions from which we generate random samples are very simple so that the random number generation can be done rather rapidly. The 5000 iterations in the above example took less than one minute to complete on a Sun SPARC10 workstation. More careful implementation of the method may help to further increase the speed.

VI. Concluding Remarks

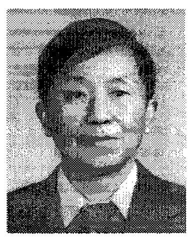
In this paper, we presented a new method for the simultaneous wavelet estimation and deconvolution of reflection seismic signals. The method employs the celebrated Bernoulli–Gaussian white sequence model for the reflectivity. In addition to the reflectivity sequence and the wavelet, statistical parameters in the reflectivity and the noise models are also estimated from the data under a Bayesian framework. The minimum mean-squared error (MMSE) estimators of these unknowns are calculated using a Monte Carlo technique called the Gibbs sampler. Numerical experiments on simulated data have shown that the method is able to provide rather precise estimates even under low signal-to-noise ratios. Further research should extend this method to other reflectivity models (e.g., [14], [15]) and to multichannel seismic signals (e.g., [10]).

Acknowledgment

The authors would like to thank three anonymous referees for their insightful comments and suggestions.

References

[1] R. Chen and T. H. Li, "Blind restoration of linearly degraded discrete signals by Gibbs sampler," *IEEE Trans. Signal Processing,* vol. 43, no. 10, pp. 2410–2413, 1995.
[2] Q. Cheng, "Maximum standardized cumulant deconvolution of non-Gaussian linear processes," *Ann. Statist.,* vol. 18, pp. 1745–1783, 1990.
[3] M. H. DeGroot, *Probability and Statistics,* 2nd ed. Amsterdam: Addison-Wesley, 1986.
[4] D. Donoho, "On minimum entropy deconvolution," in *Applied Time Series Analysis, II,* D. F. Findley, Ed. New York: Academic, 1981.
[5] A. E. Gelfand and A. F. M. Smith, "Sampling-based approaches to calculating marginal densities," *J. Amer. Statist. Assoc.,* vol. 85, pp. 398–409, 1990.
[6] A. Gelman and D. B. Rubin, "Inference from interactive simulation using multiple sequences," *Statist. Sci.,* vol. 7, pp. 457–472, 1992.
[7] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. PAMI-6, pp. 721–741, 1984.
[8] Y. Goussard, G. Demoment, and F. Monfront, "Maximum *a posteriori* detection-estimation of Bernoulli–Gaussian processes," *Mathematics in Signal Processing II,* G. McWhirter, Ed. Oxford: Clarendon, 1990, pp. 121–138.
[9] J. Goutsias and J. M. Mendel, "Maximum likelihood deconvolution: An optimization theory perspective," *Geophysics,* vol. 51, pp. 1206–1220, 1986.
[10] J. Idier and Y. Goussard, "Multichannel seismic deconvolution," *IEEE Trans. Geosci. Remote Sensing,* vol. 31, pp. 961–979, 1993.
[11] J. Kormylo and J. M. Mendel, "Maximum likelihood detection and estimation of Bernoulli–Gaussian processes," *IEEE Trans. Inform. Theory,* vol. IT-28, pp. 482–488, 1982.
[12] T. H. Li, "Blind deconvolution of discrete-valued signals," in *Proc. 27th Asilomar Conf. Signals, Systems, Computers,* Pacific Grove, CA, Nov. 1993, vol. 2, pp. 1240–1244.
[13] J. S. Liu and R. Chen, "Blind deconvolution via sequential imputation," *J. Amer. Statist. Assoc.,* vol. 90, pp. 567–576, 1995.
[14] J. M. Mendel, *Optimal Seismic Deconvolution: An Estimation-Based Approach.* New York: Academic, 1983.
[15] ——, *Maximum-Likelihood Deconvolution: A Journey into Model-Based Signal Processing.* New York: Springer-Verlag, 1990.
[16] E. A. Robinson, "Dynamic predictive deconvolution," *Geophys. Prospecting,* vol. 23, pp. 780–798, 1975.
[17] T.-J. Yu, P. C. Müller, G.-Z. Dai, and C.-F. Lin, "Minimum-variance deconvolution for noncausal wavelets," *IEEE Trans. Geosci. Remote Sensing,* vol. 32, pp. 513–524, 1994.
[18] R. A. Wiggins, "Minimum entropy deconvolution," *Geoexploration,* vol. 16, pp. 21–35, 1978.



Qiansheng Cheng received the B.S. degree in mathematics from Beijing University, Beijing, China, in 1963.

He is now a Professor in the Department of Mathematics, Beijing University, China, and the Vice Director of the Institute of Mathematics, Beijing University. His current research interests include signal processing, time series analysis, system identification, and pattern recognition.

Prof. Cheng is the vice chairman of Chinese Signal Processing Society and has won the 3rd Chinese National Natural Science Award.

Rong Chen received the B.S. degree in mathematics from Beijing University, Beijing, China, in 1985 and the M.S and Ph.D. degrees in statistics from Carnegie Mellon University, Pittsburgh, PA, in 1987 and 1990, respectively.

Since September 1990, he has been an Assistant Professor in the Department of Statistics, Texas A&M University, College Station. His current research interests include nonlinear time series analysis, signal processing and Bayesian analysis.

Dr. Chen is a member of the American Statistical Association and the Institute of Mathematical Statistics.

Ta-Hsin Li (S'89–M'92) received the B.A. degree in mathematics from Beijing Institute of Posts and Telecommunications, Beijing, China, in 1982, the M.S. degree in electrical engineering from Tsinghua University, Beijing, China, in 1984, the M.A. degree in statistics from American University, Washington, DC, in 1989, and the Ph.D. degree in applied mathematics from the University of Maryland at College Park in 1992.

From 1984 to 1987, he was Instructor of Electrical Engineering at Tsinghua University, Beijing, China. He was a Graduate Assistant at American University from 1987 to 1989 and at the Institute for Systems Research, University of Maryland, College Park, from 1989 to 1992. He has been with Texas A&M University at College Station since September of 1992, where he is Assistant Professor of Statistics. His current research interests include time series analysis and statistical signal processing.

Dr. Li is a member of Phi Kappa Phi, the American Statistical Association, and the IEEE Societies of Signal Processing and Information Theory.