

Multichannel Deconvolution of Seismic Signals Using Statistical MCMC Methods

Idan Ram, Israel Cohen, *Senior Member, IEEE*, and Shalom Raz

Abstract—In this paper, we propose two multichannel blind deconvolution algorithms for the restoration of two-dimensional (2D) seismic data. Both algorithms are based on a 2D reflectivity prior model, and use iterative multichannel deconvolution procedures which deconvolve the seismic data, while taking into account the spatial dependency between neighboring traces. The first algorithm employs in each step a modified maximum posterior mode (MPM) algorithm which estimates a reflectivity column from the corresponding observed trace using the estimate of the preceding reflectivity column. The second algorithm takes into account estimates of both the preceding and subsequent columns in the estimation process. Both algorithms are applied to synthetic and real data and demonstrate better results compared to those obtained by a single-channel deconvolution method. Expectedly, the second algorithm which utilizes more information in the estimation process of each reflectivity column is shown to produce better results than the first algorithm.

Index Terms—Markov Chain Monte Carlo, maximum posterior mode method, multichannel deconvolution, reflectivity estimation, seismic signals.

I. INTRODUCTION

REFLLECTION seismology is a common method in oil and natural gas exploration, in which a picture of the subsurface sedimentary layers of the earth is generated from surface measurements. Seismic data is obtained by transmitting an acoustic wave into the ground and measuring the reflected energy resulting from impedance discontinuities. The seismic pulse (wavelet) is time-varying, however here we make the usual assumption that it is approximately time-invariant for the received section of the seismic data. Therefore, the observed seismic data can be modeled as a convolution between a two-dimensional (2D) reflectivity section and the wavelet, which has been further degraded by additive noise. Deconvolution is used to minimize the effect of the wavelet and produce an increased resolution estimate of the reflectivity, where closely spaced reflectors can be identified.

Many methods utilize the fact that the wavelet is a one-dimensional (1D) vertical signal and break the multichannel

deconvolution problem into independent vertical 1D deconvolution problems. A 1D reflectivity column appears in the vertical direction as a sparse spike train where each spike (reflector) corresponds to a boundary between two adjacent homogenous layers. Mendel *et al.* [1], [2] use an autoregressive moving average model for the wavelet and model the reflectivity as a Bernoulli-Gaussian (BG) process [3], [4]. For each sample of the reflectivity sequence, a Bernoulli variable characterizes the presence or absence of a reflector, and the amplitude of the reflector follows a Gaussian distribution given the Bernoulli variable is nonzero. They use second order statistics methods to estimate the wavelet and recover the reflectivity by maximum likelihood estimation. The maximum likelihood criterion is maximized using the single most likely replacement (SMLR) algorithm [2], which improves the likelihood by iteratively choosing a reflectivity sequence that varies at each iteration by only one sample. Kaaresen and Tøft [5] introduced an algorithm which alternately estimates a finite impulse response wavelet and a Bernoulli-Gaussian reflectivity. The wavelet is estimated using a least-squares fit and the reflectivity is recovered using the iterated window maximization algorithm [6]. This algorithm is similar to the SMLR, but produces better results since it updates many samples at each step instead of only one. Cheng, Chen, and Li [7] simultaneously estimate a Bernoulli-Gaussian reflectivity and a moving average wavelet using a Bayesian framework in which prior information is imposed on the seismic wavelet, BG reflectivity parameters and the noise variance. These parameters along with the reflectivity sequence are estimated using a Markov chain Monte Carlo (MCMC) method called a Gibbs sampler [8], [9]. Rosec *et al.* [10] use a moving average wavelet and model the reflectivity sequence as a mixture of Gaussian distributions [11]. They propose two parameter estimation methods. The first method performs maximum likelihood estimation and use the stochastic expectation maximization (SEM) algorithm [12], [13] to maximize the likelihood criterion. The second method performs a Bayesian estimation resembling the method of Cheng *et al.* [7]. The estimated parameters are employed by the maximum posterior mode (MPM) algorithm [14], which uses realizations of the reflectivity simulated by a Gibbs sampler to estimate the reflectivity.

Application of 1D restoration methods to 2D seismic data is clearly suboptimal, as it does not take into account the correlation between neighboring columns of the seismic data (traces), which stems from the presumed continuous and roughly horizontal structure of the earth layers. Idier and Goussard [15] proposed two versions of a multichannel deconvolution method which takes into account the stratification of the layers.

Manuscript received June 25, 2009; accepted December 16, 2009. Date of publication February 17, 2010; date of current version April 14, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Alfred Hanssen.

The authors are with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Technion City, Haifa 32000, Israel (e-mail: idanram@tx.technion.ac.il; icohen@ee.technion.ac.il; raz@ee.technion.ac.il).

Digital Object Identifier 10.1109/TSP.2010.2042485

The two versions are based on two 2D reflectivity models: Markov–Bernoulli–Gaussian (MBG) I and II. Each model is composed of a Markov–Bernoulli random field (MBRF) [16], which controls the geometrical characteristics of the reflectivity, and an amplitude field, defined conditionally to the MBRF. The deconvolution is carried out using a suboptimal maximum *a posteriori* (MAP) estimator, which iteratively recovers the columns of the reflectivity section. Each reflectivity column is estimated from the corresponding observed trace and the estimate of the previous reflectivity column, using an SMLR-type method. Kaaresen and Tøft [5] also suggest a multichannel version of their blind deconvolution algorithm, which accounts for the dependencies across the traces. However, this method encourages spatial continuity of the estimated reflectors using an optimization criterion which penalizes nonsparse and non-continuous configurations. Heimer, Cohen, and Vassiliou [17], [18] introduced a multichannel blind deconvolution method which combines the algorithm of Kaaresen and Tøft with dynamic programming [19], [20] to find continuous paths of reflectors across the channels of the reflectivity section. However, layer discontinuities are not taken into account by this method. Heimer and Cohen [21] also proposed a multichannel blind deconvolution algorithm which is based on the MBG I reflectivity model. They first define a set of reflectivity states and legal transitions between configurations of neighboring reflectivity columns. Then they apply the Viterbi algorithm [22] for finding the most likely sequences of reflectors that are connected across the reflectivity section by legal transitions.

In this paper, we propose two multichannel blind deconvolution algorithms. Both algorithms are based on the MBG I reflectivity model and iteratively deconvolve the seismic data, while taking into account the spatial dependency between neighboring traces. The first algorithm employs in each step a modified version of the maximum posterior mode (MPM) algorithm which estimates the current reflectivity column from the corresponding observed trace and the estimate of the preceding reflectivity column. The modified MPM algorithm is a two step procedure. First, it employs a Gibbs sampler to simulate realizations of the MBRF and amplitude variables by iteratively sampling from their conditional distributions, which depend on the estimate of the preceding reflectivity column. Then, a decision step takes place in which the MBRF and amplitude variables are estimated from their realizations. The second algorithm is an extension of the first. It takes into account the dependency between each reflectivity column and both the preceding and *subsequent* neighbors, in the deconvolution process. It employs in each step a further modified maximum posterior mode algorithm which simultaneously estimates both the current and subsequent reflectivity columns. These columns are determined from the corresponding observed traces and the estimate of the preceding reflectivity column. Again, the estimation is carried out in two steps. First, a Gibbs sampler is employed to simulate realizations of the MBRF and amplitude variables corresponding to the current and subsequent reflectivity columns, by iteratively sampling from their conditional distributions. Then, the MBRF and amplitude variables are determined from their realizations in a decision step. Out of the two obtained estimates, only the estimate of the current reflectivity column is kept. The estimate

of the subsequent column is discarded, as this column will be determined from estimates of both its preceding and subsequent neighbors in the next step.

Both multichannel deconvolution algorithms are applied to synthetic and real data, and demonstrate better results compared to those obtained by the single-channel deconvolution method of Rosencorn *et al.* [10]. The second algorithm which utilizes more information in the estimation process of each reflectivity column is shown to produce better results than the first algorithm.

The paper is organized as follows: In Section II we formulate the multichannel blind deconvolution problem. Then we describe the MBG I reflectivity model and the parameter estimation method. In Section III and IV, we introduce the first and second proposed algorithms, respectively. In Section V we present simulation and real data results demonstrating the performance of both proposed algorithms compared to a single-channel deconvolution. We summarize the paper in Section VI.

II. PROBLEM FORMULATION AND REFLECTIVITY MODEL

A. Problem Formulation

Multichannel blind seismic deconvolution aims at restoring a 2D reflectivity section and an unknown seismic wavelet from a 2D observed seismic data. The seismic wavelet $\mathbf{h} = [h(1), \dots, h(N_h)]^T$ is a 1D vertical vector of length N_h , which is assumed to be invariant in both horizontal and vertical directions. The reflectivity section \mathbf{R} is a matrix of size $N_r \times J$ and the 2D seismic data \mathbf{Y} is a matrix of size $N_y \times J$, where $N_y = N_h + N_r - 1$. \mathbf{Y} can be modeled as the following noise-corrupted convolution product:

$$\mathbf{Y} = \mathbf{h} * \mathbf{R} + \mathbf{W} \quad (1)$$

where \mathbf{W} is a matrix of size $N_y \times J$ which denotes an additive white Gaussian noise independent of \mathbf{R} with zero mean and variance σ_w^2 .

We use the MBG I reflectivity model so that the stratification of the layers of the Earth will be taken into account in the deconvolution process. Since the deconvolution problem is blind, i.e., \mathbf{h} , σ_w^2 , and the MBG I model parameters are unknown, a suitable estimation method needs to be derived. We next describe the MBG I reflectivity model and subsequently propose a method for estimating the missing parameters.

B. Prior Model

The Markov–Bernoulli–Gaussian I reflectivity model [15] is a 2D extension of the 1D Bernoulli–Gaussian representation. It is composed of a Markov–Bernoulli random field, which controls the geometrical characteristics of the reflectivity, and an amplitude field, defined conditionally to the MBRF. The MBRF comprises two types of binary variables: location variables and transition variables. The location variables, set in a $N_r \times J$ matrix \mathbf{Q} , indicate the position of layer boundaries. Let $q_{k,j}$ denote the location variable in the (k, j) position of \mathbf{Q} . Then $q_{k,j}$ is set to one if a reflector exists in the (k, j) position of \mathbf{R} , and is set to zero otherwise. The transition variables, set in three $N_r \times J - 1$ matrices \mathbf{T}^+ , \mathbf{T}^- , and \mathbf{T}^\setminus , determine whether adjacent location variables belong to the same layer boundary or not. Let $t_{k,j}^+$, $t_{k,j}^-$, $t_{k,j}^\setminus$ denote the transition variables in the (k, j)

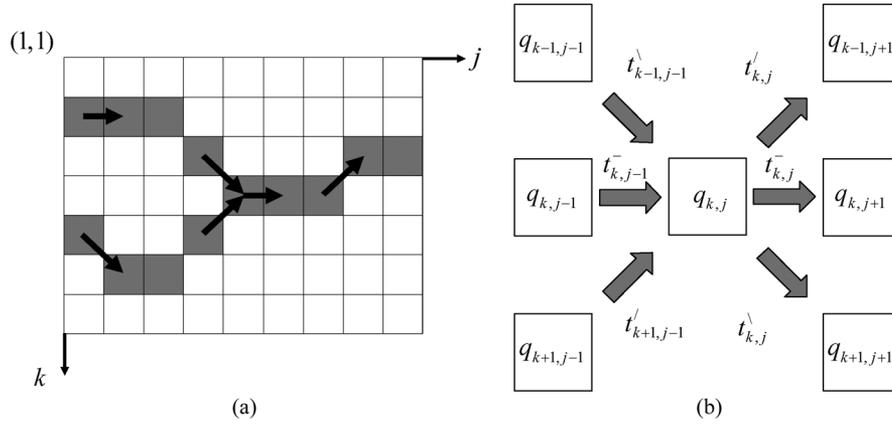


Fig. 1. Location and transition variables: (a) Layer boundaries representation. (b) Location variable $q_{k,j}$ and other location and transition variables affected by it.

positions of \mathbf{T}' , \mathbf{T}^- , and \mathbf{T}^\setminus , respectively. Then $t_{k,j}'$ is set to one if $q_{k,j}$ and $q_{k-1,j+1}$ belong to the same layer boundary and to zero otherwise. Similarly $t_{k,j}^-$ and $t_{k,j}^\setminus$ are set to one if $q_{k,j}$ belongs to the same layer boundary as $q_{k,j+1}$ and $q_{k+1,j+1}$, respectively, and to zero otherwise. Therefore, $t_{k,j}'$, $t_{k,j}^-$ and $t_{k,j}^\setminus$ determine whether layer boundaries whose orientation is diagonally ascending, horizontal, and diagonally descending, respectively, exist in the (k, j) position of \mathbf{Q} . Fig. 1(a) shows a representation of layer boundaries and their orientation in several locations using location and transitions variables. Gray squares denote the presence of a layer boundary and arrows facing upward, rightward and downward correspond to positions in which $t_{k,j}'$, $t_{k,j}^-$ and $t_{k,j}^\setminus$, respectively, are set to one. Fig. 1(b) shows the location variable $q_{k,j}$, all the location variables which may be on the same boundary with it, and the transition variables between them.

Let $p(\cdot)$ denote a probability distribution function. Then the MBRF has the following properties.

1) Separability property:

$$p\left(t_{k,j}', t_{k,j}^-, t_{k,j}^\setminus\right) = p\left(t_{k,j}'\right) p\left(t_{k,j}^-\right) p\left(t_{k,j}^\setminus\right).$$

- 2) The j th columns of \mathbf{Q} , \mathbf{T}' , \mathbf{T}^- and \mathbf{T}^\setminus , denoted \mathbf{q}_j , \mathbf{t}_j' , \mathbf{t}_j^- and \mathbf{t}_j^\setminus , respectively, are white and Bernoulli distributed marginally from the rest of the field.
- 3) The characteristic parameters of the Bernoulli distributions are given by

$$\begin{aligned} \lambda &= p(q_{k,j} = 1), & \mu' &= p\left(t_{k,j}' = 1\right), \\ \mu^- &= p\left(t_{k,j}^- = 1\right), & \mu^\setminus &= p\left(t_{k,j}^\setminus = 1\right). \end{aligned}$$

- 4) Horizontal symmetry: $p(q_{k,j}, t_{k,j}', t_{k,j}^-, t_{k,j}^\setminus) = p(q_{k,j}, t_{k+1,j-1}^-, t_{k,j-1}^-, t_{k-1,j-1}^\setminus)$.
- 5) Isolated transition variables cannot be set to one: $p(t_{k,j}' = 0, t_{k,j}^- = 0, t_{k,j}^\setminus = 0 | q_{k,j} = 0) = 1$.
- 6) Discontinuities along layer boundaries are possible

$$p\left(q_{k,j} = 1 | t_{k+1,j-1}^- = 0, t_{k,j-1}^- = 0, t_{k-1,j-1}^\setminus = 0\right) = \varepsilon.$$

7) λ is related to $\{\mu', \mu^-, \mu^\setminus, \varepsilon\}$ according to: $\lambda = 1 - (1 - \mu')(1 - \mu^-)(1 - \mu^\setminus)(1 - \varepsilon)$.

We now turn to the amplitude field \mathbf{R} . The MBG I model assumes that the amplitudes of the reflectors are independent in the vertical direction and that marginally from the rest of the field, the amplitude of each reflector is normally distributed with zero mean and variance equal to σ_r^2 . The conditional probability of the amplitude field of the reflectivity $p(\mathbf{R} | \mathbf{T}', \mathbf{T}^-, \mathbf{T}^\setminus, \mathbf{Q})$ is assumed to have a first-order Markov chain structure, and each reflector is assumed to be correlated only with reflectors located on the same boundary. Let \mathbf{r}_j denote the j th column of \mathbf{R} , and let $r_{k,j}$ denote the k th reflector in \mathbf{r}_j . Then the correlation between $r_{k,j}$ and reflectors in previous columns depends on the local geometry of the layers and is described through $p(r_{k,j} | q_{k,j}, \mathbf{t}_{j-1}', \mathbf{t}_{j-1}^-, \mathbf{t}_{j-1}^\setminus, \mathbf{r}_{j-1})$. Let $t_{k+1,j-1}'$ (respectively, $t_{k,j-1}^-$, $t_{k-1,j-1}^\setminus$) be set to one, then we will further refer to the reflector $r_{k,j}$ as a successor of $r_{k+1,j-1}$ (respectively, $r_{k,j-1}$, $r_{k-1,j-1}$) and symmetrically $r_{k+1,j-1}$ (respectively, $r_{k,j-1}$, $r_{k-1,j-1}$) will be referred to as a predecessor of $r_{k,j}$. The conditional probabilities $p(r_{k,j} | q_{k,j}, \mathbf{t}_{j-1}', \mathbf{t}_{j-1}^-, \mathbf{t}_{j-1}^\setminus, \mathbf{r}_{j-1})$ can be separated into four cases which depend on the existence and uniqueness of successors and predecessors.

- 1) If $q_{k,j} = 0$ then there is no reflector at position (k, j) , and $r_{k,j} = 0$.
- 2) If $q_{k,j} = 1$, and if $r_{k,j}$ is the unique successor of a unique predecessor $r_{k+dk,j-1}$ ($-1 \leq dk \leq 1$), then $r_{k,j}$ is sampled from a first-order autoregressive (AR) process, conditionally to $r_{k+dk,j-1}$. This case corresponds to interactions along a single layer boundary. Let $a \in [0, 1]$ control the degree of correlation between reflector amplitudes along the same boundary and let $w_r \sim N[0, (1 - a^2)\sigma_r^2]$, then the AR process is defined by

$$r_{k,j} = ar_{k+dk,j-1} + w_r. \tag{2}$$

- 3) If $q_{k,j} = 1$ and if $r_{k,j}$ has no predecessor, then $r_{k,j}$ is sampled from the basic Gaussian distribution $N(0, \sigma_r^2)$.
- 4) If $q_{k,j} = 1$ and if $r_{k,j}$ has more than one predecessor, or symmetrically when $r_{k,j}$ is not a unique successor, then $r_{k,j}$ is sampled from the basic Gaussian distribution $N(0, \sigma_r^2)$.

Before the deconvolution can be performed, the parameters of the 2D reflectivity model described above need to be estimated from the data, along with the seismic wavelet and the noise variance. We next describe the parameter estimation method.

C. Parameter Estimation

Our goal is to estimate the parameters $\theta = (\mathbf{h}, \lambda, \sigma_r, \sigma_w)$ and the MBG I parameters $\theta_{MBG} = (a, \mu', \mu^-, \mu^\lambda, \varepsilon)$ from the observed seismic data \mathbf{Y} . The parameters θ can be estimated using the stochastic expectation maximization algorithm of Rosec *et al.* [10]. Let \mathbf{y}_j denote the j th trace of the seismic data. Then we apply the SEM algorithm to each of the traces \mathbf{y}_j and obtain from each trace an estimate $\hat{\theta}_j$. Since the parameters θ are assumed common to all the seismic traces, the final estimate $\hat{\theta} = (\hat{\mathbf{h}}, \hat{\lambda}, \hat{\sigma}_r, \hat{\sigma}_w)$ is obtained by averaging the estimates $\hat{\theta}_j$. We now proceed to estimate the parameters θ_{MBG} by employing the parameters $\hat{\theta}$ to deconvolve each of the traces \mathbf{y}_j using the maximum posterior mode algorithm of Rosec *et al.* [10]. Subsequently, we remove all the isolated reflectors from the obtained reflectivity estimate and term the result $\hat{\mathbf{R}}_{TBT}^c$. Letting I' , I^- and I^λ denote the number of positions in $\hat{\mathbf{R}}_{TBT}^c$ in which the orientation of the layer boundaries is ascending, horizontal and descending, respectively, we propose the following estimators:

$$\begin{aligned}\hat{\mu}' &= \frac{I'}{(J-1)N_r} \\ \hat{\mu}^- &= \frac{I^-}{(J-1)N_r} \\ \hat{\mu}^\lambda &= \frac{I^\lambda}{(J-1)N_r} \\ \hat{\varepsilon} &= 1 - (1 - \hat{\lambda}) / (1 - \hat{\mu}') / (1 - \hat{\mu}^-) / (1 - \hat{\mu}^\lambda). \quad (3)\end{aligned}$$

For the estimation of a , we use a heuristic estimator, which calculates the average attenuation ratio between neighboring reflectors. Let I_l be the number of layer boundaries in $\hat{\mathbf{R}}_{TBT}^c$, and let $\mathbf{l}_m = [l_m(1), \dots, l_m(L_m)]^T$ be the reflectors of the m th boundary arranged in a vector of length L_m . Note that when $\hat{\mathbf{R}}_{TBT}^c$ contains layer boundaries which split or merge, each section of a boundary before and after the node where the splitting or merging occurs, is treated as a separate boundary layer. Then

$$\hat{a} = \frac{1}{I_l} \sum_{m=1}^{I_l} \left[\frac{1}{L_m - 1} \sum_{k=1}^{L_m - 1} \min \left(\frac{l_m(k+1)}{l_m(k)}, \frac{l_m(k)}{l_m(k+1)} \right) \right]. \quad (4)$$

Once all the missing parameters are known, multichannel deconvolution can be performed. We next describe the first proposed multichannel deconvolution procedure.

III. RECURSIVE CAUSAL MULTICHANNEL BLIND DECONVOLUTION

In this section, we propose a multichannel blind deconvolution algorithm, which iteratively deconvolves the seismic data, while taking into account the spatial dependency between neighboring traces. The proposed method is based on the MBG I reflectivity prior model and employs the parameter estimation method proposed in the previous section. We next describe the deconvolution scheme of the proposed algorithm.

A. Deconvolution Scheme

The MAP estimator of the matrices $\{\mathbf{T}', \mathbf{T}^-, \mathbf{T}^\lambda, \mathbf{Q}\}$, comprising the MBRF, and the amplitude field \mathbf{R} is

$$(\hat{\mathbf{T}}', \hat{\mathbf{T}}^-, \hat{\mathbf{T}}^\lambda, \hat{\mathbf{Q}}, \hat{\mathbf{R}}) = \arg \max_{\mathbf{T}', \mathbf{T}^-, \mathbf{T}^\lambda, \mathbf{Q}, \mathbf{R}} p(\mathbf{T}', \mathbf{T}^-, \mathbf{T}^\lambda, \mathbf{Q}, \mathbf{R} | \mathbf{Y}). \quad (5)$$

Obtaining the exact MAP solution is very difficult, even when the efficient Viterbi algorithm is used, because of the large dimension of the state-space of $\{\mathbf{T}', \mathbf{T}^-, \mathbf{T}^\lambda, \mathbf{Q}\}$. However, Idier and Goussard [15] showed that the a posteriori likelihood $p(\mathbf{T}', \mathbf{T}^-, \mathbf{T}^\lambda, \mathbf{Q}, \mathbf{R} | \mathbf{Y})$ can be expressed as

$$p(\mathbf{T}', \mathbf{T}^-, \mathbf{T}^\lambda, \mathbf{Q}, \mathbf{R} | \mathbf{Y}) \propto p(\mathbf{q}_1, \mathbf{r}_1, \mathbf{y}_1) \times \prod_{j=2}^J p(\mathbf{t}_{j-1}, \mathbf{q}_j, \mathbf{r}_j, \mathbf{y}_j | \mathbf{q}_{j-1}, \mathbf{r}_{j-1}) \quad (6)$$

where $\mathbf{t}_{j-1} = \{\mathbf{t}_{j-1}', \mathbf{t}_{j-1}^-, \mathbf{t}_{j-1}^\lambda\}$. This formula led them to propose the following suboptimal iterative maximization procedure:

$$1) \text{ First column : } (\hat{\mathbf{r}}_1, \hat{\mathbf{q}}_1) = \arg \max_{\mathbf{r}_1, \mathbf{q}_1} p(\mathbf{q}_1, \mathbf{r}_1, \mathbf{y}_1) \quad (7)$$

$$2) \text{ For } j \in [2, J] : (\hat{\mathbf{r}}_j, \hat{\mathbf{q}}_j, \hat{\mathbf{t}}_{j-1}) = \arg \max_{\mathbf{r}_j, \mathbf{q}_j, \mathbf{t}_{j-1}} p(\mathbf{r}_j, \mathbf{q}_j, \mathbf{t}_{j-1}, \mathbf{y}_j | \hat{\mathbf{q}}_{j-1}, \hat{\mathbf{r}}_{j-1}) \quad (8)$$

where SMLR-type algorithms were used for the optimization of the partial criteria (7) and (8). In the first step \mathbf{r}_1 and \mathbf{q}_1 are determined from the first observed trace \mathbf{y}_1 . In each following step, the reflectivity column \mathbf{r}_j , $j \in [2, J]$ and corresponding hidden binary vectors \mathbf{t}_{j-1} , \mathbf{q}_j are determined from the current observed trace \mathbf{y}_j and the estimates of \mathbf{r}_{j-1} and \mathbf{q}_{j-1} , obtained in the previous step. This maximization procedure is suboptimal since for $j > 1$ each partial criterion is maximized only with respect to \mathbf{r}_j , \mathbf{q}_j , \mathbf{t}_{j-1} and all the previously estimated quantities remain unchanged. Also, the determination of \mathbf{r}_j is based on observations only up to \mathbf{y}_j and subsequent columns of the observed data, which are very informative about \mathbf{r}_j , are not taken into account in its estimation. On the other hand, this method is much simpler than global maximization of $p(\mathbf{T}', \mathbf{T}^-, \mathbf{T}^\lambda, \mathbf{Q}, \mathbf{R} | \mathbf{Y})$, and does take into account the dependency between neighboring reflectivity columns, unlike single-channel deconvolution methods.

Here we use a similar iterative maximization procedure, which employs MCMC methods for the optimization of its partial criteria. We first rewrite (7) and (8) as

$$1) \text{ First column : } (\hat{\mathbf{r}}_1, \hat{\mathbf{q}}_1) = \arg \max_{\mathbf{r}_1, \mathbf{q}_1} p(\mathbf{r}_1, \mathbf{q}_1 | \mathbf{y}_1) \quad (9)$$

$$2) \text{ For } j \in [2, J] : (\hat{\mathbf{r}}_j, \hat{\mathbf{q}}_j, \hat{\mathbf{t}}_{j-1}) = \arg \max_{\mathbf{r}_j, \mathbf{q}_j, \mathbf{t}_{j-1}} p(\mathbf{r}_j, \mathbf{q}_j, \mathbf{t}_{j-1} | \mathbf{y}_j, \hat{\mathbf{r}}_{j-1}, \hat{\mathbf{q}}_{j-1}). \quad (10)$$

The first partial criterion can be optimized using the maximum posterior mode algorithm presented by Rosec *et al.* [10]. Finding an optimal solution for the maximization problem (10) is very hard, since it requires examination of all the possible configurations of \mathbf{q}_j , \mathbf{t}_{j-1} , whose number ranges from 2^{N_r} to 8^{N_r} .

Therefore, we apply instead a modified version of the MPM algorithm. This algorithm estimates the vectors \mathbf{r}_j , \mathbf{q}_j , \mathbf{t}_{j-1} from realizations simulated by a Gibbs sampler, described next.

B. Gibbs Sampler

The Gibbs sampler generates samples of \mathbf{r}_j , \mathbf{q}_j , \mathbf{t}_{j-1} from the joint distribution $p(\mathbf{r}_j, \mathbf{q}_j, \mathbf{t}_{j-1} | \mathbf{y}_j, \hat{\mathbf{r}}_{j-1}, \hat{\mathbf{q}}_{j-1})$. Let $\mathbf{v}_{-k,j}$ denote a vector \mathbf{v}_j without its k th sample, i.e., $\mathbf{v}_{-k,j} = [v_{1,j}, \dots, v_{k-1,j}, v_{k+1,j}, \dots, v_{N_r,j}]^T$. Also, let $B(\alpha)$ denote a Bernoulli distribution with parameter α . Then instead of sampling directly from the joint distribution, the Gibbs sampler iteratively samples from the conditional distributions:

- $p(t'_{k,j-1} | \mathbf{r}_j, \mathbf{q}_j, \mathbf{t}'_{-k,j-1}, \mathbf{t}'_{j-1}, \mathbf{t}'_{j-1}, \hat{\mathbf{r}}_{j-1}, \hat{\mathbf{q}}_{j-1}) \sim B(\mu'_{k,j-1})$
- $p(t^-_{k,j-1} | \mathbf{r}_j, \mathbf{q}_j, \mathbf{t}'_{j-1}, \mathbf{t}^-_{-k,j-1}, \mathbf{t}'_{j-1}, \hat{\mathbf{r}}_{j-1}, \hat{\mathbf{q}}_{j-1}) \sim B(\mu^-_{k,j-1})$
- $p(t^\backslash_{k,j-1} | \mathbf{r}_j, \mathbf{q}_j, \mathbf{t}'_{j-1}, \mathbf{t}^-_{j-1}, \mathbf{t}^\backslash_{-k,j-1}, \hat{\mathbf{r}}_{j-1}, \hat{\mathbf{q}}_{j-1}) \sim B(\mu^\backslash_{k,j-1})$
- $p(r_{k,j}, q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \hat{\mathbf{r}}_{j-1}) \sim \lambda_{k,j}^b N(m_{k,j}^b, V_{k,j}^b) + (1 - \lambda_{k,j}^b) \delta(r_{k,j})$

where the derivation of $\mu'_{k,j-1}$, $\mu^-_{k,j-1}$, $\mu^\backslash_{k,j-1}$, $\lambda_{k,j}^b$, $m_{k,j}^b$ and $V_{k,j}^b$ can be found in subsections I and II of the Appendix.

For the simulation of the vectors \mathbf{r}_j , \mathbf{q}_j , and \mathbf{t}_{j-1} , the Gibbs sampler follows these steps iteratively:

- 1) Initialization: choice of $\mathbf{q}_j^{(0)}$, $\mathbf{r}_j^{(0)}$ and $\mathbf{t}_{j-1}^{(0)}$.
- 2) For $i = 1, \dots, I$

For $k = 1, \dots, N_r$

- compute $\mu'_{k,j-1}$ using (36) and simulate

$$t'^{(i)}_{k,j-1} \sim B(\mu'_{k,j-1})$$

- compute $\mu^-_{k,j-1}$ using (37) and simulate

$$t'^{(i)-}_{k,j-1} \sim B(\mu^-_{k,j-1})$$

- compute $\mu^\backslash_{k,j-1}$ using (38) and simulate

$$t'^{(i)\backslash}_{k,j-1} \sim B(\mu^\backslash_{k,j-1})$$

- compute $\lambda_{k,j}^b$ using (26) and simulate

$$q_{k,j}^{(i)} \sim B(\lambda_{k,j}^b)$$

- simulate $r_{k,j}^{(i)} \sim N(m_{k,j}^b, V_{k,j}^b)$ if $q_{k,j}^{(i)} = 1$, otherwise $r_{k,j}^{(i)} = 0$.

C. MPM Algorithm

We estimate each column \mathbf{r}_j , $1 < j \leq J$ using a modified version of the MPM algorithm. This algorithm employs the Gibbs sampler described above to generate realizations of \mathbf{r}_j , \mathbf{q}_j and \mathbf{t}_{j-1} drawn from $p(\mathbf{r}_j, \mathbf{q}_j, \mathbf{t}_{j-1} | \mathbf{y}_j, \hat{\mathbf{r}}_{j-1}, \hat{\mathbf{q}}_{j-1})$. The Gibbs sampler performs I_0 iterations until it reaches a steady-state period. The samples $(\mathbf{r}_j^{(i)}, \mathbf{q}_j^{(i)}, \mathbf{t}_{j-1}^{(i)})$ produced in the following

iterations $I_0 < i \leq I$ are used to first estimate each of $q_{k,j}$, $t'_{k,j-1}$, $t'^{-}_{k,j-1}$, $t'^{\backslash}_{k,j-1}$, and then determine $r_{k,j}$ conditionally to the estimate of $q_{k,j}$. The modified MPM algorithm follows these steps iteratively.

- 1) For $i = 1, \dots, I$ simulate $(\mathbf{r}_j^{(i)}, \mathbf{q}_j^{(i)}, \mathbf{t}_{j-1}^{(i)})$ using the Gibbs sampler.
- 2) For $k = 1, \dots, N_r$
 - detection step:

$$\hat{t}'_{k,j-1} = \begin{cases} 1, & \text{if } \frac{1}{I-I_0} \sum_{i=I_0+1}^I t'^{(i)}_{k,j-1} > 0.5 \\ 0, & \text{otherwise,} \end{cases}$$

$$\hat{t}'^-_{k,j-1} = \begin{cases} 1, & \text{if } \frac{1}{I-I_0} \sum_{i=I_0+1}^I t'^{(i)-}_{k,j-1} > 0.5 \\ 0, & \text{otherwise,} \end{cases}$$

$$\hat{t}'^\backslash_{k,j-1} = \begin{cases} 1, & \text{if } \frac{1}{I-I_0} \sum_{i=I_0+1}^I t'^{(i)\backslash}_{k,j-1} > 0.5 \\ 0, & \text{otherwise,} \end{cases}$$

$$\hat{q}_{k,j} = \begin{cases} 1, & \text{if } \frac{1}{I-I_0} \sum_{i=I_0+1}^I q_{k,j}^{(i)} > 0.5 \\ 0, & \text{otherwise} \end{cases}$$

- estimation step

$$\hat{r}_{k,j} = \begin{cases} \frac{\sum_{i=I_0+1}^I q_{k,j}^{(i)} r_{k,j}^{(i)}}{\sum_{i=I_0+1}^I q_{k,j}^{(i)}(k)}, & \text{if } \hat{q}_{k,j} = 1 \\ 0, & \text{otherwise} \end{cases}$$

- 3) $\hat{\mathbf{r}}_j = [r_{1,j}, \dots, r_{N_r,j}]^T$.

IV. RECURSIVE NONCAUSAL MULTICHANNEL BLIND DECONVOLUTION

The algorithm proposed in this section is an extended version of the first proposed algorithm, and uses the same reflectivity model and parameter estimation method. However, it takes into account the dependency between each reflectivity column and both the preceding and subsequent neighbors, in the deconvolution process. We next describe the deconvolution scheme of this algorithm.

A. Deconvolution Scheme

Our goal is to improve the performance of the first proposed algorithm, by taking into account information from both preceding and subsequent traces in the deconvolution process of each trace. More specifically, we wish to utilize estimates of both \mathbf{r}_{j-1} and \mathbf{r}_{j+1} , in the estimation process of \mathbf{r}_j . However, an estimate of \mathbf{r}_{j+1} is not available from previous steps in the j th step of the algorithm, in which \mathbf{r}_j is estimated. Therefore, instead of estimating only \mathbf{r}_j in the j th step, we simultaneously estimate both \mathbf{r}_j and \mathbf{r}_{j+1} , conditionally to the estimate of \mathbf{r}_{j-1} , obtained in the previous step. This way, the dependency between \mathbf{r}_j and \mathbf{r}_{j+1} is taken into account when the former is estimated. In each step besides the last, only the estimate of \mathbf{r}_j is kept out of the two obtained estimates. The estimate of \mathbf{r}_{j+1} is discarded,

as this column will be reestimated in the next step. It is kept only in the last step, as the estimate of the last column in the 2D reflectivity section \mathbf{r}_J . We note that extending the algorithm to utilize more than one subsequent reflectivity column in the estimation process of \mathbf{r}_j is straightforward. However, simulation results showed that only a small improvement in the performance is gained when more than one reflectivity column is estimated along with \mathbf{r}_j , at the cost of a larger computational burden. We next define the following vectors:

$$\begin{aligned}\bar{\mathbf{y}}_j &= [\mathbf{y}_j^T, \mathbf{y}_{j+1}^T]^T, \quad \bar{\mathbf{r}}_j = [\mathbf{r}_j^T, \mathbf{r}_{j+1}^T]^T, \quad \bar{\mathbf{q}}_j = [\mathbf{q}_j^T, \mathbf{q}_{j+1}^T]^T, \\ \bar{\mathbf{t}}_{j-1}^{\prime} &= \left[\left(\mathbf{t}_{j-1}^{\prime} \right)^T, \left(\mathbf{t}_j^{\prime} \right)^T \right]^T, \quad \bar{\mathbf{t}}_{j-1}^- = \left[\left(\mathbf{t}_{j-1}^- \right)^T, \left(\mathbf{t}_j^- \right)^T \right]^T, \\ \bar{\mathbf{t}}_{j-1}^{\setminus} &= \left[\left(\mathbf{t}_{j-1}^{\setminus} \right)^T, \left(\mathbf{t}_j^{\setminus} \right)^T \right]^T, \quad j=1, \dots, J-1\end{aligned}$$

where $\mathbf{t}_0^{\prime} = \mathbf{t}_0^- = \mathbf{t}_0^{\setminus} = \mathbf{0}_{N_r \times 1}$ and $\mathbf{0}_{N_r \times 1}$ denotes a vector of N_r zeros. Using these concatenated vectors in the deconvolution process allows us to simultaneously estimate the amplitude, location and transition variables associated with both \mathbf{r}_j and \mathbf{r}_{j+1} . The deconvolution is carried out iteratively, using the following maximization procedure:

$$1) \text{ First column : } (\hat{\mathbf{r}}_1, \hat{\mathbf{q}}_1, \hat{\mathbf{t}}_0) = \arg \max_{\mathbf{r}_1, \mathbf{q}_1, \mathbf{t}_0} p(\bar{\mathbf{r}}_1, \bar{\mathbf{q}}_1, \bar{\mathbf{t}}_0 | \bar{\mathbf{y}}_1) \quad (11)$$

$$2) j \in [2, J-1] : (\hat{\mathbf{r}}_j, \hat{\mathbf{q}}_j, \hat{\mathbf{t}}_{j-1}) = \arg \max_{\mathbf{r}_j, \mathbf{q}_j, \mathbf{t}_{j-1}} p(\bar{\mathbf{r}}_j, \bar{\mathbf{q}}_j, \bar{\mathbf{t}}_{j-1} | \bar{\mathbf{y}}_j, \hat{\mathbf{r}}_{j-1}, \hat{\mathbf{q}}_{j-1}) \quad (12)$$

where $\bar{\mathbf{t}}_{j-1} = \{\bar{\mathbf{t}}_{j-1}^{\prime}, \bar{\mathbf{t}}_{j-1}^-, \bar{\mathbf{t}}_{j-1}^{\setminus}\}$.

Similarly to the deconvolution scheme of the first proposed algorithm, a single partial criterion is optimized in each step. Direct optimization of these partial criteria is practically impossible, since it requires examination of all the possible configurations of $\bar{\mathbf{q}}_j, \bar{\mathbf{t}}_{j-1}$, whose number ranges from 2^{2N_r} to 8^{2N_r} . Instead, we apply a further modified version of the MPM algorithm to these partial criteria. This MPM algorithm estimates in the first step $\bar{\mathbf{t}}_0, \bar{\mathbf{q}}_1$ and $\bar{\mathbf{r}}_1$ from $\bar{\mathbf{y}}_1$, and estimates in the j th step, $j \in [2, J-1]$, $\bar{\mathbf{t}}_{j-1}, \bar{\mathbf{q}}_j$, and $\bar{\mathbf{r}}_j$ from $\bar{\mathbf{y}}_j, \hat{\mathbf{r}}_{j-1}$ and $\hat{\mathbf{q}}_{j-1}$. The first N_r samples of the estimates of $\bar{\mathbf{r}}_j, \bar{\mathbf{q}}_j, \bar{\mathbf{t}}_{j-1}^{\prime}, \bar{\mathbf{t}}_{j-1}^-$ and $\bar{\mathbf{t}}_{j-1}^{\setminus}$, $j \in [1, J-1]$ are kept as the desired estimates $\hat{\mathbf{r}}_j, \hat{\mathbf{q}}_j, \hat{\mathbf{t}}_{j-1}$. In the $J-1$ th step the last N_r samples of the estimates are also kept as $\hat{\mathbf{r}}_J, \hat{\mathbf{q}}_J, \hat{\mathbf{t}}_{J-1}^{\prime}, \hat{\mathbf{t}}_{J-1}^-, \hat{\mathbf{t}}_{J-1}^{\setminus}$.

The MPM algorithm employs two different Gibbs samplers in the estimation process of \mathbf{r}_1 and $\mathbf{r}_j, j \in [2, J-1]$. We describe these Gibbs samplers next.

B. Gibbs Samplers

In the estimation process of the first reflectivity column, we employ a Gibbs sampler to generate samples of $\bar{\mathbf{r}}_1, \bar{\mathbf{q}}_1, \bar{\mathbf{t}}_0$ from the joint distribution $p(\bar{\mathbf{r}}_1, \bar{\mathbf{q}}_1, \bar{\mathbf{t}}_0 | \bar{\mathbf{y}}_1)$. Instead of sampling directly from this joint distribution, the Gibbs sampler iteratively samples from the conditional distributions of $\bar{r}_{k,1}, \bar{q}_{k,1}, \bar{t}_{k,0}^{\prime}, \bar{t}_{k,0}^-$ and $\bar{t}_{k,0}^{\setminus}$. The first N_r samples of $\bar{\mathbf{t}}_0^{\prime}, \bar{\mathbf{t}}_0^-, \bar{\mathbf{t}}_0^{\setminus}$ equal zero and the first N_r samples of $\bar{\mathbf{r}}_1, \bar{\mathbf{q}}_1$ are sampled from:

- $p(r_{k,j}, q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j) \sim \lambda_{k,j}^f N(m_{k,j}^f, V_{k,j}^f) + (1 - \lambda_{k,j}^f) \delta(r_{k,j})$ where the derivation of $\lambda_{k,j}^f, m_{k,j}^f$ and $V_{k,j}^f$ can be found in subsection III of the Appendix. The last N_r samples of $\bar{\mathbf{r}}_1, \bar{\mathbf{q}}_1, \bar{\mathbf{t}}_0^{\prime}, \bar{\mathbf{t}}_0^-, \bar{\mathbf{t}}_0^{\setminus}$ are sampled from:
 - $p(\mathbf{t}_{k,j}^{\prime} | \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_{-k,j}^{\prime}, \mathbf{t}_j^{\prime}, \mathbf{r}_j, \mathbf{q}_j) \sim B(\mu_{k,j}^{\prime})$
 - $p(\mathbf{t}_{k,j}^- | \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j^{\prime}, \mathbf{t}_{-k,j}^-, \mathbf{t}_j^{\setminus}, \mathbf{r}_j, \mathbf{q}_j) \sim B(\mu_{k,j}^-)$
 - $p(\mathbf{t}_{k,j}^{\setminus} | \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j^{\prime}, \mathbf{t}_j^-, \mathbf{t}_{-k,j}^{\setminus}, \mathbf{r}_j, \mathbf{q}_j) \sim B(\mu_{k,j}^{\setminus})$
 - $p(r_{k,j+1}, q_{k,j+1} | \mathbf{y}_{j+1}, \mathbf{r}_{-k,j+1}, \mathbf{q}_{-k,j+1}, \mathbf{t}_j, \mathbf{r}_j) \sim \lambda_{k,j+1}^b N(m_{k,j+1}^b, V_{k,j+1}^b) + (1 - \lambda_{k,j+1}^b) \delta(r_{k,j+1})$
- For the simulation of the vectors $\bar{\mathbf{r}}_1, \bar{\mathbf{q}}_1$ and $\bar{\mathbf{t}}_0$, the Gibbs sampler follows these steps iteratively:

- 1) Initialization: choice of $\bar{\mathbf{r}}_1^{(0)}, \bar{\mathbf{q}}_1^{(0)}, \bar{\mathbf{t}}_0^{(0)}$.
- 2) For $i = 1, \dots, I$

For $k = 1, \dots, N_r$

- compute $\lambda_{k,j}^f$ using (54) and simulate

$$\bar{q}_{k,j}^{(i)} \sim B(\lambda_{k,j}^f)$$

- if $\bar{q}_{k,j} = 1$ simulate $\bar{r}_{k,j}^{(i)} \sim N(m_{k,j}^f, V_{k,j}^f)$, otherwise $\bar{r}_{k,j}^{(i)} = 0$.

For $k = N_r + 1, \dots, 2N_r$

- compute $\mu_{k-N_r,j}^{\prime}$ using (36) and simulate

$$\bar{t}_{k,j-1}^{\prime(i)} \sim B(\mu_{k-N_r,j}^{\prime})$$

- compute $\mu_{k-N_r,j}^-$ using (37) and simulate

$$\bar{t}_{k,j-1}^{- (i)} \sim B(\mu_{k-N_r,j}^-)$$

- compute $\mu_{k-N_r,j}^{\setminus}$ using (38) and simulate

$$\bar{t}_{k,j-1}^{\setminus (i)} \sim B(\mu_{k-N_r,j}^{\setminus})$$

- compute $\lambda_{k-N_r,j+1}^b$ using (26) and simulate

$$\bar{q}_{k,j}^{(i)} \sim B(\lambda_{k-N_r,j+1}^b)$$

- simulate $\bar{r}_{k,j}^{(i)} \sim N(m_{k-N_r,j+1}^b, V_{k-N_r,j+1}^b)$ if $\bar{q}_{k,j}^{(i)} = 1$, otherwise $\bar{r}_{k,j}^{(i)} = 0$.

Similarly, in the estimation process of the j th reflectivity column, $j \in [2, J-1]$, we employ a different Gibbs sampler to generate samples of $\bar{\mathbf{r}}_j, \bar{\mathbf{q}}_j, \bar{\mathbf{t}}_{j-1}$ from the joint distribution $p(\bar{\mathbf{r}}_j, \bar{\mathbf{q}}_j, \bar{\mathbf{t}}_{j-1} | \bar{\mathbf{y}}_j, \hat{\mathbf{r}}_{j-1}, \hat{\mathbf{q}}_{j-1})$. This Gibbs sampler iteratively samples from the conditional distributions of $\bar{r}_{k,j}, \bar{q}_{k,j}, \bar{t}_{k,j-1}^{\prime}, \bar{t}_{k,j-1}^-$ and $\bar{t}_{k,j-1}^{\setminus}$, where the first N_r samples of $\bar{\mathbf{r}}_j, \bar{\mathbf{q}}_j, \bar{\mathbf{t}}_{j-1}^{\prime}, \bar{\mathbf{t}}_{j-1}^-, \bar{\mathbf{t}}_{j-1}^{\setminus}$ are sampled from:

- $p(\mathbf{t}_{k,j-1}^{\prime} | \mathbf{r}_j, \mathbf{q}_j, \mathbf{t}_{-k,j-1}^{\prime}, \mathbf{t}_{j-1}^{\prime}, \hat{\mathbf{r}}_{j-1}, \hat{\mathbf{q}}_{j-1}) \sim B(\mu_{k,j-1}^{\prime})$
- $p(\mathbf{t}_{k,j-1}^- | \mathbf{r}_j, \mathbf{q}_j, \mathbf{t}_{j-1}^{\prime}, \mathbf{t}_{-k,j-1}^-, \mathbf{t}_{j-1}^{\setminus}, \hat{\mathbf{r}}_{j-1}, \hat{\mathbf{q}}_{j-1}) \sim B(\mu_{k,j-1}^-)$

- $p(t_{k,j-1}^{\setminus} | \mathbf{r}_j, \mathbf{q}_j, \mathbf{t}_{j-1}^{\setminus}, \mathbf{t}_{j-1}^{\setminus}, \mathbf{t}_{-k,j-1}^{\setminus}, \hat{\mathbf{r}}_{j-1}, \hat{\mathbf{q}}_{j-1}) \sim B(\mu_{k,j-1}^{\setminus})$
- $p(r_{k,j}, q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \hat{\mathbf{r}}_{j-1}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j) \sim \lambda_{k,j}^m N(m_{k,j}^m, V_{k,j}^m) + (1 - \lambda_{k,j}^m) \delta(r_{k,j})$

and the derivation of $\lambda_{k,j}^m$, $m_{k,j}^m$, and $V_{k,j}^m$ can be found in subsection IV of the Appendix. The last N_r samples of $\bar{\mathbf{r}}_j$, $\bar{\mathbf{q}}_j$, $\bar{\mathbf{t}}_{j-1}^{\setminus}$, $\bar{\mathbf{t}}_{j-1}^{\setminus}$, $\bar{\mathbf{t}}_{j-1}^{\setminus}$ are sampled from the same conditional distributions as the last N_r samples of $\bar{\mathbf{r}}_1$, $\bar{\mathbf{q}}_1$, $\bar{\mathbf{t}}_0^{\setminus}$, $\bar{\mathbf{t}}_0^{\setminus}$, $\bar{\mathbf{t}}_0^{\setminus}$.

For the simulation of the vectors $\bar{\mathbf{r}}_j$, $\bar{\mathbf{q}}_j$ and $\bar{\mathbf{t}}_{j-1}$, the Gibbs sampler follows these steps iteratively:

- 1) Initialization: choice of $\bar{\mathbf{r}}_j^{(0)}$, $\bar{\mathbf{q}}_j^{(0)}$, $\bar{\mathbf{t}}_{j-1}^{(0)}$.
- 2) For $i = 1, \dots, I$

For $k = 1, \dots, N_r$

- compute $\mu_{k,j-1}^{\setminus}$ using (36) and simulate

$$\bar{t}_{k,j-1}^{\setminus(i)} \sim B(\mu_{k,j-1}^{\setminus})$$

- compute $\mu_{k,j-1}^{\bar{}}$ using (37) and simulate

$$\bar{t}_{k,j-1}^{\bar{(i)}} \sim B(\mu_{k,j-1}^{\bar{}})$$

- compute $\mu_{k,j-1}^{\setminus}$ using (38) and simulate

$$\bar{t}_{k,j-1}^{\setminus(i)} \sim B(\mu_{k,j-1}^{\setminus})$$

- compute $\lambda_{k,j}^m$ using (64) and simulate

$$\bar{q}_{k,j}^{(i)} \sim B(\lambda_{k,j}^m)$$

- simulate $\bar{r}_{k,j}^{(i)} \sim N(m_{k,j}^m, V_{k,j}^m)$ if $\bar{q}_{k,j} = 1$, otherwise $\bar{r}_{k,j}^{(i)} = 0$.

For $k = N_r + 1, \dots, 2N_r$ follow the same steps as the Gibbs sampler described above.

C. MPM Algorithm

The second proposed algorithm employs a further modified MPM algorithm. This MPM algorithm estimates $\bar{\mathbf{r}}_1$, $\bar{\mathbf{q}}_1$, $\bar{\mathbf{t}}_0$ in the first step, and estimates $\bar{\mathbf{r}}_j$, $\bar{\mathbf{q}}_j$, $\bar{\mathbf{t}}_{j-1}$, $j \in [2, J-1]$ in the following steps. It uses the two versions of the Gibbs sampler described in the previous subsection to generate realizations of $\bar{\mathbf{r}}_j$, $\bar{\mathbf{q}}_j$, and $\bar{\mathbf{t}}_{j-1}$, where the first I_0 iterations are considered a learning period. Only the samples $(\bar{\mathbf{r}}_j^{(i)}, \bar{\mathbf{q}}_j^{(i)}, \bar{\mathbf{t}}_{j-1}^{(i)})$ produced in the subsequent steady state period $I_0 < i \leq I$ are used to first estimate each of $\bar{q}_{k,j}$, $\bar{t}_{k,j-1}^{\setminus}$, $\bar{t}_{k,j-1}^{\bar{}}$, $\bar{t}_{k,j-1}^{\setminus}$ and then determine $\bar{r}_{k,j}^{(i)}$ conditionally to the estimate of $\bar{q}_{k,j}$.

The further modified MPM algorithm follows these steps iteratively.

- 1) For $i = 1, \dots, I$ simulate $(\bar{\mathbf{t}}_{j-1}^{(i)}, \bar{\mathbf{q}}_j^{(i)}, \bar{\mathbf{r}}_j^{(i)})$ using the Gibbs samplers.
- 2) Use the same detection step as in Section III-C, where $r_{k,j}$, $q_{k,j}$, $t_{k,j-1}^{\setminus}$, $t_{k,j-1}^{\bar{}}$ and $t_{k,j-1}^{\setminus}$ are replaced by $\bar{r}_{k,j}$, $\bar{q}_{k,j}$, $\bar{t}_{k,j-1}^{\setminus}$, $\bar{t}_{k,j-1}^{\bar{}}$ and $\bar{t}_{k,j-1}^{\setminus}$, respectively.

$$3) \hat{\mathbf{r}}_j = [\hat{r}_{1,j}, \dots, \hat{r}_{N_r,j}]^T.$$

V. EXPERIMENTAL RESULTS

A. Synthetic Data

We generated a 2D reflectivity section of size 76×100 , shown in Fig. 2(a), using the MBG I model. We then convolved it with a 25 samples long Ricker wavelet and added white Gaussian noise, with signal-to-noise ratios (SNRs) of 0 and 5 dB, where the SNR is defined as

$$\text{SNR} = 10 \log_{10} \left(\frac{\lambda \sigma_r^2 E_h}{\sigma_w^2} \right). \quad (13)$$

We created 20 realizations for each SNR, two of them with SNRs of 0 and 5 dB are shown in Fig. 2(b) and (c), respectively. We then used the proposed parameter estimation method to find the missing parameters corresponding to each of the data sets. The total number of iterations of the SEM algorithm was set to 4000, where the first 3000 iterations served as the burn-in period after which the algorithm reaches a steady state. The true wavelet, along with wavelets estimated for the realizations in Fig. 2(b) and (c) are shown in Fig. 2(d) and (e), respectively. The true parameters are shown in Table I, along with the means and standard deviations (in brackets) of the parameters estimated from the realizations with the different SNRs.

The true and estimated parameters were employed by the deconvolution schemes of the two proposed multichannel algorithms, and the single-channel MPM algorithm of Rosenc *et al.* [10]. The reflectivity sections recovered by single-channel deconvolution and those obtained with the true parameters were used for comparison reasons. The total number of iterations I of the MPM algorithms of Rosenc *et al.* and the first and second proposed algorithms was set to 8000, 8000, and 16000, and the corresponding burn-in period I_0 was set to 4000, 4000, and 8000 iterations, respectively. The average processing times of a data set of size 100×100 on Pentium Core 2 Duo E8400, by Matlab implementations of the single-channel and the first and second proposed algorithms, were 8.86, 9.17, and 47.69 min, respectively. Note that each reflectivity column estimated by the single-channel and multichannel deconvolution algorithms had gone through a postprocessing procedure. Whenever this procedure found two or three successive reflectors, or two reflectors separated by one sample, it replaced them by their center of mass. We will hereafter refer to the first and second proposed algorithms as MC-I and MC-II, respectively. The results of single-channel deconvolution and the MC-I and MC-II algorithms, obtained with the true and estimated parameters for the seismic data with SNR of 5 dB, depicted in Fig. 2(c), are shown in Fig. 3. The results obtained with the estimated parameters for the seismic data with SNR of 0 dB depicted in Fig. 2(b), are shown in Fig. 4. Visual comparison between these results shows improved performance of both the MC-I and MC-II algorithms over the performance of the single-channel deconvolution algorithm. For both SNR levels the estimates of MC-I and MC-II are more continuous, contain less false detections and are generally closer to the true reflectivity than the single-channel deconvolution results. It can also be seen that, as one would expect, better

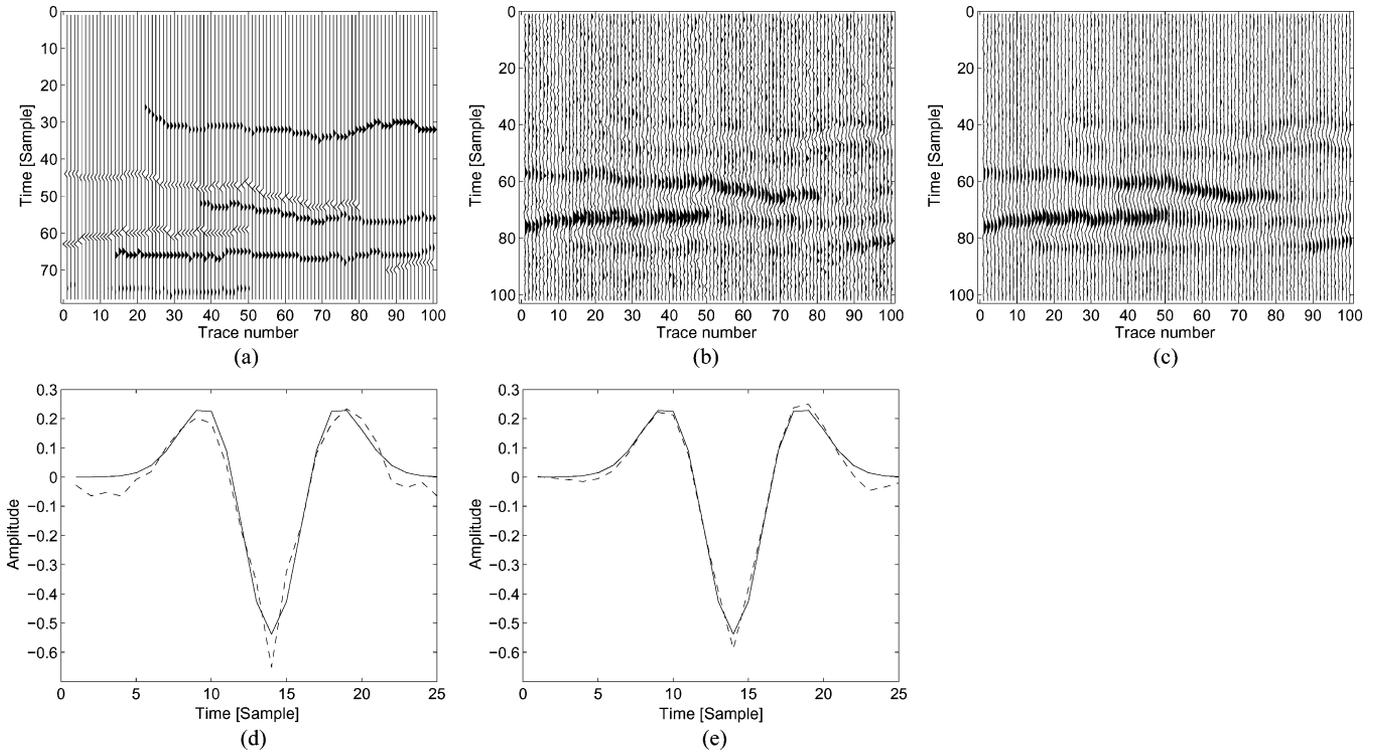


Fig. 2. Synthetic reflectivity, wavelet and data sets: (a) Synthetic 2D reflectivity section. (b) 2D seismic data (SNR = 0 dB). (c) 2D seismic data (SNR = 5 dB). (d) True wavelet (solid) and its estimate (dashed) for SNR = 0 dB. (e) True wavelet (solid) and its estimate (dashed) for SNR = 5 dB.

TABLE I
SYNTHETIC 2D EXAMPLE: TRUE AND ESTIMATED PARAMETERS

	True	Estimated (0 dB)	Estimated (5 dB)
λ	0.0489	0.0479 (0.0033)	0.0553 (0.0013)
σ_r	1	1.2851 (0.0603)	1.152 (0.0249)
σ_w (0 dB)	0.2211	0.1742 (0.0035)	-
σ_w (5 dB)	0.1243	-	0.1014 (0.0012)
a	0.999	0.7752 (0.0202)	0.7899 (0.0152)
μ^+	0.008	0.0058 (0.0007)	0.0076 (0.0008)
μ^-	0.033	0.01 (0.0011)	0.0228 (0.0017)
μ^{\setminus}	0.008	0.0061 (0.0008)	0.0086 (0.0007)
ε	0.0005	0.0266 (0.0034)	0.0174 (0.0019)

results are obtained with the true parameters than with the estimated ones.

In order to quantify the performances of the MC-I and MC-II algorithms and compare them to each other and to the performance of the single-channel algorithm, we used the four loss functions suggested by Kaaresen in [23]. Let \mathbf{r} be a 1D reflectivity sequence and $\hat{\mathbf{r}}$ be its estimate, and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be the L_1 and L_2 norms, respectively. Also, let $N^{\text{miss}} = \#\{n : \hat{r}(n) = 0, r(n) \neq 0\}$ and $N^{\text{false}} = \#\{n : \hat{r}(n) \neq 0, r(n) = 0\}$ denote the number of missed and false detections in $\hat{\mathbf{r}}$, respectively. Then the loss functions are

$$\begin{aligned}
 L^{\text{miss+false}} &= \|\hat{\mathbf{r}} - \mathbf{r}\|_1 + N^{\text{miss}} + N^{\text{false}} \\
 L^{\text{miss}} &= \|\hat{\mathbf{r}} - \mathbf{r}\|_1 + N^{\text{miss}} \\
 L^{\text{false}} &= \|\hat{\mathbf{r}} - \mathbf{r}\|_1 + N^{\text{false}} \\
 L^{\text{SSQ}} &= \|\hat{\mathbf{r}} - \mathbf{r}\|_2.
 \end{aligned} \tag{14}$$

Kaaresen also suggested to make the loss functions more realistic, by regarding estimated reflectors that were close to their true positions as partially correct. Therefore we added three loss functions which treated reflectors in $\hat{\mathbf{r}}$ with an offset of one sample from their true location as if they were set in their true locations, with half their amplitude. For these reflectors a penalty of 0.5 was added to both the missed and false detection measures. The new loss functions are

$$\begin{aligned}
 L_2^{\text{miss+false}} &= D + N_2^{\text{miss}} + N_2^{\text{false}} \\
 L_2^{\text{miss}} &= D + N_2^{\text{miss}} \\
 L_2^{\text{false}} &= D + N_2^{\text{false}}
 \end{aligned} \tag{15}$$

where $D = \sum_n |\hat{r}(n) + (1/2)\hat{r}(n-1) + (1/2)\hat{r}(n+1) - r(n)|$ is a difference measure, $N_2^{\text{miss}} = N^{\text{miss}} - (1/2) \times \#\{n : r(n) \neq 0, \hat{r}(n) = 0, \hat{r}(n-1) \neq 0 \text{ or } \hat{r}(n+1) \neq 0\}$ is a missed detection measure, and $N_2^{\text{false}} = N^{\text{false}} - (1/2) \times \#\{n : \hat{r}(n) \neq 0, r(n) = 0, r(n-1) \neq 0 \text{ or } r(n+1) \neq 0\}$ is a false detection measure. Since we are dealing with 2D reflectivity signals, we calculated the loss functions for their column stack forms. We also normalized L^{SSQ} by the L_2 norm of the column stack form of the true reflectivity and normalized the rest of the loss functions by the number of reflectors contained in the true reflectivity. The means and standard deviations of the loss functions calculated for the estimates obtained by single-channel deconvolution and the MC-I and MC-II algorithms are shown in percents in Tables II and III. The values displayed in Table II correspond to the results obtained with the true and estimated parameters, for the seismic data with SNR of 5 dB. Similarly,

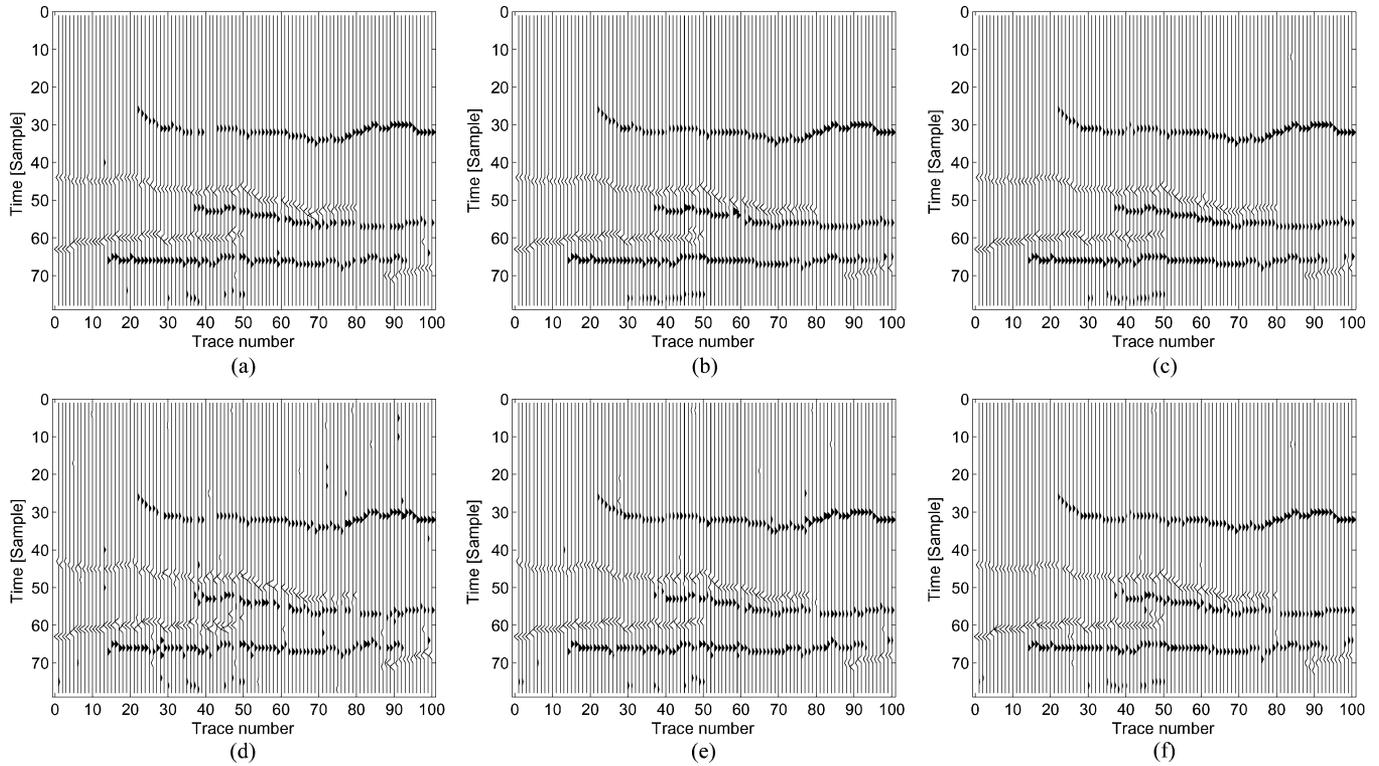


Fig. 3. Synthetic 2D data deconvolution results obtained with the true parameters (TP) and the estimated parameters (EP) for SNR of 5 dB. (a) Single-channel deconvolution results (TP). (b) MC-I results (TP). (c) MC-II results (TP). (d) Single-channel deconvolution results (EP). (e) MC-I results (EP). (f) MC-II results (EP).

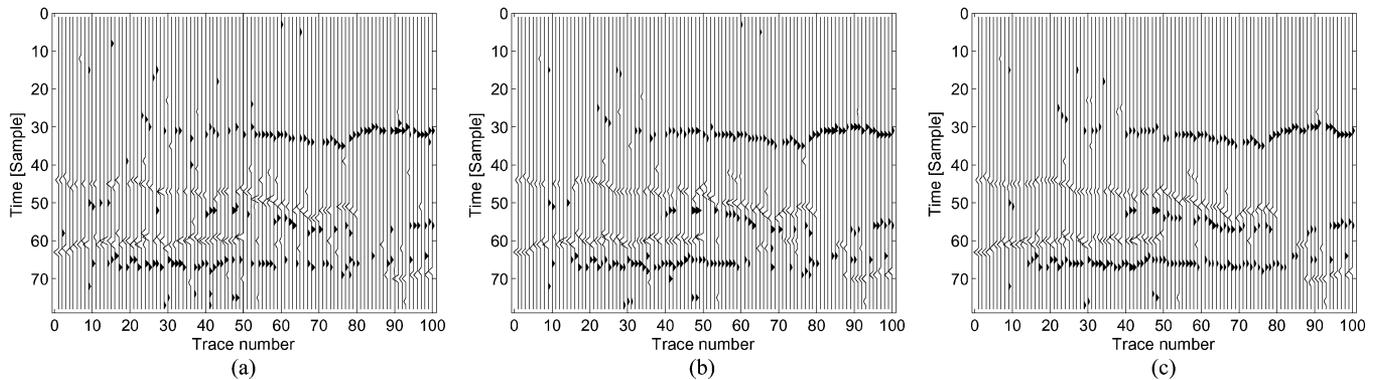


Fig. 4. Synthetic 2D data deconvolution results. (a) Single-channel deconvolution results for SNR = 0 dB. (b) MC-I results for SNR = 0 dB. (c) MC-II results for SNR = 0 dB.

TABLE II
COMPARISON BETWEEN THE QUALITY OF RESTORATION OF THE SINGLE-CHANNEL DECONVOLUTION (SC), MC-I, AND MC-II ALGORITHMS, OBTAINED WITH THE TRUE AND ESTIMATED PARAMETERS, FOR SNR OF 5 dB

SNR=5 dB	Estimated parameters			True parameters		
	SC	MC-I	MC-II	SC	MC-I	MC-II
$L^{miss+false}$	112.36 (10.89)	81.15 (13.61)	69.31 (10.42)	59 (5.81)	46.24 (8.74)	45.35 (5.43)
L^{miss}	88.55 (7.95)	64.33 (10.02)	54.81 (7.67)	50.98 (4.8)	38.53 (6.58)	37.47 (3.88)
L^{false}	79.76 (8.38)	57.67 (10.11)	49.35 (7.92)	37.96 (4.01)	31.24 (6.42)	30.85 (4.29)
L^{SSQ}	67.8 (4.32)	54.01 (5.68)	47.16 (5.31)	44.13 (3.57)	35.24 (5.9)	34.67 (3.73)
$L_2^{miss+false}$	76.38 (6.39)	54.13 (8.24)	46.48 (6.17)	44.45 (4.22)	33.7 (6.48)	31.54 (3.03)
L_2^{miss}	59.78 (4.46)	42.95 (5.87)	36.98 (4.34)	39.34 (3.56)	28.75 (4.82)	26.83 (2.05)
L_2^{false}	51.14 (4.82)	36.54 (5.8)	31.73 (4.54)	26.39 (2.55)	21.7 (4.48)	20.54 (2.41)

the values in Table III correspond to the results obtained for the data with SNR of 0 dB.

It can be seen that for both SNR levels, and for all the loss functions, the mean values calculated for the estimates of the

TABLE III
COMPARISON BETWEEN THE QUALITY OF RESTORATION OF THE SINGLE-CHANNEL DECONVOLUTION (SC), MC-I, AND MC-II ALGORITHMS, OBTAINED WITH THE TRUE AND ESTIMATED PARAMETERS, FOR SNR OF 0 dB

SNR=0 dB	Estimated parameters			True parameters		
	SC	MC-I	MC-II	SC	MC-I	MC-II
$L^{miss+false}$	207.13 (13.76)	185.57 (18.01)	165.26 (16.67)	130.42 (6.61)	104.48 (8.09)	102.15 (9.01)
L^{miss}	167.94 (9.51)	148.91 (13.54)	132.2 (12.05)	118.6 (5.18)	91.46 (6.46)	87.75 (7.04)
L^{false}	145.55 (11.82)	131.9 (14.08)	117.96 (13.57)	77.56 (5.19)	65.62 (5.97)	65.94 (6.83)
L^{SSQ}	101.35 (4.46)	94.1 (6.23)	86.59 (6.06)	74.18 (3.35)	63.11 (4.16)	62.48 (4.26)
$L_2^{miss+false}$	146.65 (8.99)	123.08 (11.64)	106.65 (9.72)	105.67 (3.66)	76.52 (5.27)	70.63 (4.43)
L_2^{miss}	118.4 (5.69)	98.16 (8.46)	84.93 (6.49)	98.31 (3.05)	68.89 (4.41)	62.36 (3.54)
L_2^{false}	96.05 (8.06)	81.17 (8.93)	70.7 (7.99)	57.27 (2.52)	43.07 (3.38)	40.58 (2.96)

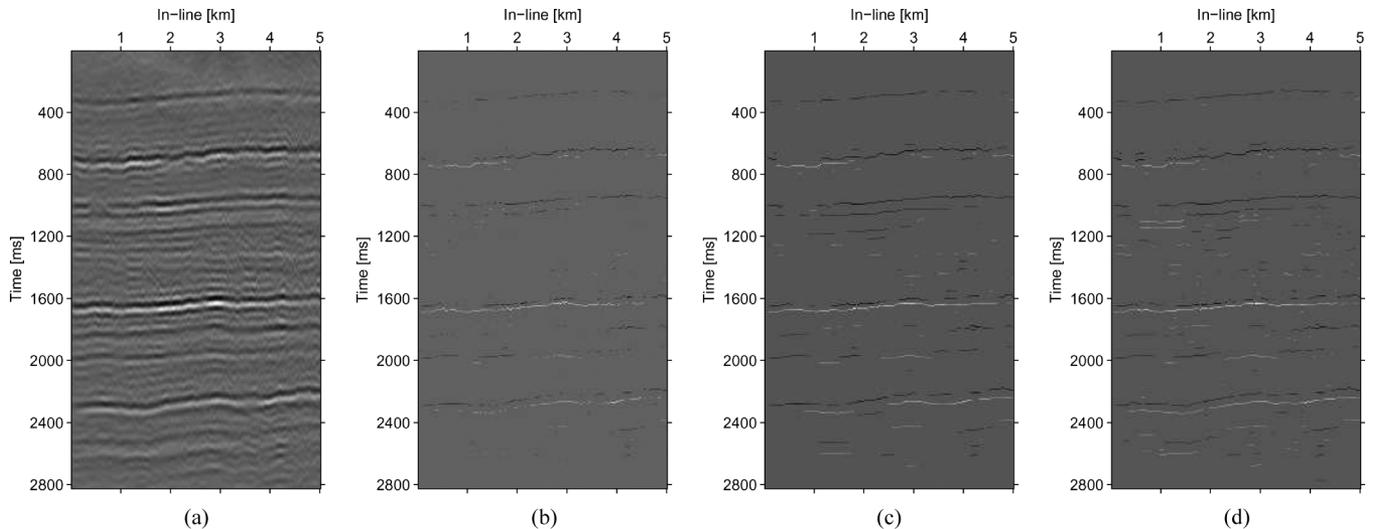


Fig. 5. Real data deconvolution results: (a) Real seismic data. (b) Single-channel deconvolution result. (c) MC-I results. (d) MC-II results.

MC-I and MC-II algorithms are smaller than the respective mean values calculated for the estimates of the single-channel deconvolution algorithm. This implies that both the MC-I and MC-II algorithms produce better results than the single-channel algorithm. It can also be seen that for both SNR levels the MC-II algorithm outperforms the MC-I algorithm, and that the improvement is getting smaller as the SNR increases. Not surprisingly, lower mean values of the loss functions are measured for all the higher SNR estimates, meaning that all the algorithms performed better when the noise level was low. Also, lower mean values of the loss functions are obtained with the true parameters than with the estimated ones, however the difference between the mean values obtained in these two cases is getting smaller as the SNR increases. Finally, MC-II seems to be more robust to model parameter inaccuracies than MC-I. Although MC-II performs only slightly better than MC-I when the model parameters are known, it produces significant improvement in the case of blind deconvolution.

B. Real Data

We applied the proposed parameter estimation method to real seismic data from a small land 3D survey in North America (courtesy of GeoEnergy Inc., TX) of size 400×200 , shown in Fig. 5(a). Three-dimensional denoising was applied to the data [24], which was subsequently decimated in both time and space. The time interval is 8 ms and the in-line trace spacing

is 25 m. The estimated wavelet is shown in Fig. 6, and the estimated parameters are presented in Table IV. Similarly to the case of the synthetic data, the estimated parameters were employed by the deconvolution schemes of the MC-I and MC-II algorithms, and the MPM algorithm of Rosec *et al.* The reflectivity sections obtained by single-channel deconvolution, MC-I and MC-II are shown in Fig. 5(b), (c), and (d), respectively. Comparing these reflectivity sections, it can be seen that the estimates obtained by MC-I and MC-II contain layer boundaries which are more continuous and smooth than the ones obtained by the single-channel deconvolution. These algorithms also manage to detect parts of the layers that the single-channel deconvolution missed. It can also be seen that the estimates produced by the MC-I and MC-II algorithms are quite close, however the latter managed to recover parts of the layer boundaries missed by both the single-channel and first proposed algorithms. Note that since the true reflectivity section is unknown, the loss functions (14) and (15) cannot be used to assess the performance of the proposed algorithms on real data.

VI. CONCLUSION

We have proposed two multichannel blind deconvolution algorithms. Both algorithms, which take into account the spatial dependency between neighboring traces in the deconvolution process, produce visually superior deconvolution results, compared to a single-channel deconvolution algorithm, for synthetic and real data. The second algorithm uses more information from

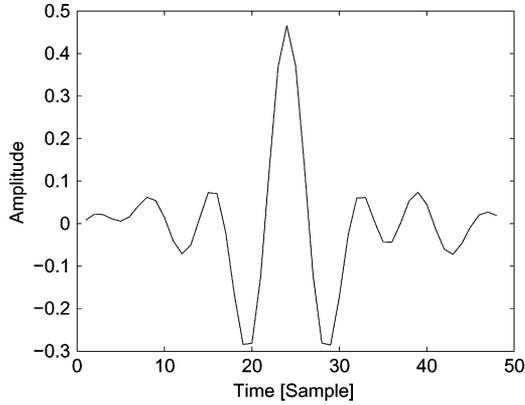


Fig. 6. Real data estimated seismic wavelet.

TABLE IV
REAL DATA EXAMPLE: PARAMETERS ESTIMATED FOR THE REAL DATA

λ	σ_r	σ_w	a	μ^+	μ^-	μ^\backslash	ε
0.0385	3.945	0.701	0.901	0.0015	0.011	0.0009	0.0254

neighboring traces in the deconvolution process of each trace, and therefore performs better than the first proposed algorithm, on synthetic and real data. Qualitative assessment of synthetic data deconvolution results shows improved performance of both proposed algorithms compared to the single-channel algorithm. It also shows that the second proposed algorithm improves on the first, but this improvement is getting smaller as the SNR increases.

One topic for future research is developing new versions of the two proposed algorithms, based on the MBG II model [15]. This model uses a different amplitude field than the MBG I model, which may lead to different quality of the deconvolution results. The performance of the new algorithms can be assessed and compared to that of the original ones. Another topic for future research is the extension of the second proposed algorithm to handle 3D input data. In this case the recovered reflectivity is a 3D signal and a 3D estimation window can be used so that neighboring reflectivity columns from 8 directions will be taken into account in the estimation process of each reflectivity column.

APPENDIX I

DERIVATION OF THE PARAMETERS $\lambda_{k,j}^b$, $m_{k,j}^b$ AND $V_{k,j}^b$

Our goal is to derive the BG distribution $p(r_{k,j}, q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1})$. We start by factoring this distribution as

$$\begin{aligned} & p(r_{k,j}, q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}) \\ & \propto p(\mathbf{y}_j | \mathbf{r}_j) p(r_{k,j} | q_{k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}) \\ & \quad \times p(q_{k,j} | \mathbf{t}_{k+1,j-1}^\backslash, \mathbf{t}_{k,j-1}^-, \mathbf{t}_{k-1,j-1}^\backslash). \end{aligned} \quad (16)$$

Noting that

$$p(q_{k,j} | \mathbf{t}_{k+1,j-1}^\backslash, \mathbf{t}_{k,j-1}^-, \mathbf{t}_{k-1,j-1}^\backslash) \sim B(\lambda_{k,j}^t) \quad (17)$$

with $\lambda_{k,j}^t = \varepsilon^{(1-\mathbf{t}_{k+1,j-1}^\backslash)(1-\mathbf{t}_{k,j-1}^-)(1-\mathbf{t}_{k-1,j-1}^\backslash)}$, and defining

$$p(r_{k,j} | q_{k,j} = 1, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}) \sim N(m_{k,j}, \sigma_{k,j}^2) \quad (18)$$

we get, after some algebraic manipulations

$$\begin{aligned} & p(r_{k,j}, q_{k,j} = 1 | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}) \\ & \propto \lambda_{k,j}^t \frac{\sqrt{V_{k,j}^b}}{\sigma_{k,j}} \exp \left[-\frac{m_{k,j}^2}{2\sigma_{k,j}^2} + \frac{(m_{k,j}^b)^2}{2V_{k,j}^b} \right] \frac{1}{\sqrt{2\pi V_{k,j}^b}} \\ & \quad \times \exp \left[-\frac{(r_{k,j} - m_{k,j}^b)^2}{2V_{k,j}^b} \right] \end{aligned} \quad (19)$$

with

$$V_{k,j}^b = \left(\frac{1}{\sigma_{k,j}^2} + \frac{E_h}{\sigma_w^2} \right)^{-1}, \quad m_{k,j}^b = V_{k,j}^b \left(\frac{m_{k,j}}{\sigma_{k,j}^2} + m_w \right) \quad (20)$$

$$m_w = \frac{1}{\sigma_w^2} \sum_{i=1}^{N_h} h(i) \left[y_{k+i-1,j} - \sum_{\substack{s=1 \\ s \neq i}}^{N_h} h(s) r_{i+k-s,j} \right] \quad (21)$$

$$E_h = \sum_{i=1}^{N_h} h^2(i). \quad (22)$$

Similarly, we get that

$$p(r_{k,j}, q_{k,j} = 0 | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}) \propto (1 - \lambda_{k,j}^t) \delta(r_{k,j}). \quad (23)$$

Finally, from (19) and (23) we get that

$$p(q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}) \sim B(\lambda_{k,j}^b) \quad (24)$$

and

$$\begin{aligned} & p(r_{k,j}, q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}) \\ & \sim \lambda_{k,j}^b N(m_{k,j}^b, V_{k,j}^b) + (1 - \lambda_{k,j}^b) \delta(r_{k,j}) \end{aligned} \quad (25)$$

where

$$\begin{aligned} & \lambda_{k,j}^b = p(q_{k,j} = 1 | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}) \\ & = \left\{ 1 + \frac{1 - \lambda_{k,j}^t}{\lambda_{k,j}^t} \frac{\sigma_{k,j}}{\sqrt{V_{k,j}^b}} \exp \left[-\frac{(m_{k,j}^b)^2}{2V_{k,j}^b} + \frac{m_{k,j}^2}{2\sigma_{k,j}^2} \right] \right\}^{-1}. \end{aligned} \quad (26)$$

APPENDIX II

DERIVATION OF THE PARAMETERS $\mu_{k,j-1}^\backslash$, $\mu_{k,j-1}^-$ AND $\mu_{k,j-1}^\backslash$

We will now derive the conditional probabilities for the three types of transition variables. Let $S_{k,j} = \{-1 \leq dk \leq 1 : q_{k+dk,j} = 1\}$, then we start with

$p(t'_{k,j-1} | \mathbf{r}_j, \mathbf{q}_j, \mathbf{t}'_{-k,j-1}, \mathbf{t}'_{j-1}, \mathbf{r}_{j-1}, \mathbf{q}_{j-1})$, which can be expressed by

$$\begin{aligned} & p(t'_{k,j-1} | \mathbf{r}_j, \mathbf{q}_j, \mathbf{t}'_{-k,j-1}, \mathbf{t}'_{j-1}, \mathbf{r}_{j-1}, \mathbf{q}_{j-1}) \\ & \propto p(q_{k-1,j} | t'_{k,j-1}, t'_{k-1,j-1}, t'_{k-2,j-1}) \\ & \quad \times p(t'_{k,j-1}, t'_{k,j-1}, t'_{k,j-1} | q_{k,j-1}) \\ & \quad \times \prod_{dk \in S_{k,j}} p(r_{k+dk,j} | q_{k+dk,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}). \end{aligned} \quad (27)$$

Let $\mathbf{u}'_{j-1} = \{\mathbf{t}'_{-k,j-1}, \mathbf{t}'_{j-1}, \mathbf{r}_{j-1}\}$, then we first note that $p(t'_{k,j-1} = 1 | \mathbf{r}_j, \mathbf{q}_j, \mathbf{u}'_{j-1}, \mathbf{q}_{j-1}) \neq 0$ only when $q_{k,j-1} = 1$ and $q_{k-1,j} = 1$. Next, let

$$\mathbf{x}'_{j-1} = \{\mathbf{q}_{-(k-1),j}, \mathbf{u}'_{j-1}, \mathbf{q}_{-k,j-1}\} \quad (28)$$

$$\rho'_{1,dk} = p(r_{k+dk,j} | q_{k+dk,j}, t'_{k,j-1} = 1, \mathbf{u}'_{j-1}) \quad (29)$$

$$\rho'_{0,dk} = p(r_{k+dk,j} | q_{k+dk,j}, t'_{k,j-1} = 0, \mathbf{u}'_{j-1}) \quad (30)$$

$$\eta'_1 = p(t'_{k,j-1} = 1, t'_{k,j-1}, t'_{k,j-1} | q_{k,j-1} = 1) \quad (31)$$

$$\begin{aligned} \eta'_0 &= p(q_{k-1,j} = 1 | t'_{k,j-1} = 0, t'_{k-1,j-1}, t'_{k-2,j-1}) \\ & \quad \times p(t'_{k,j-1} = 0, t'_{k,j-1}, t'_{k,j-1} | q_{k,j-1} = 1). \end{aligned} \quad (32)$$

Then using (27), (29), and (31), we get

$$\begin{aligned} p(t'_{k,j-1} = 1 | \mathbf{r}_j, q_{k-1,j} = 1, \mathbf{x}'_{j-1}, q_{k,j-1} = 1) \\ \propto \eta'_1 \prod_{dk \in S_{k,j}} \rho'_{1,dk} \end{aligned} \quad (33)$$

and using (27), (30), and (32) we get

$$\begin{aligned} p(t'_{k,j-1} = 0 | \mathbf{r}_j, q_{k-1,j} = 1, \mathbf{x}'_{j-1}, q_{k,j-1} = 1) \\ \propto \eta'_0 \prod_{dk \in S_{k,j}} \rho'_{0,dk}. \end{aligned} \quad (34)$$

Now, from (33) and (34), we get that

$$p(t'_{k,j-1} | \mathbf{r}_j, \mathbf{q}_j, \mathbf{u}'_{j-1}, \mathbf{q}_{j-1}) \sim B(\mu'_{k,j-1}) \quad (35)$$

with

$$\begin{aligned} \mu'_{k,j-1} &= p(t'_{k,j-1} = 1 | \mathbf{r}_j, \mathbf{q}_j, \mathbf{u}'_{j-1}, \mathbf{q}_{j-1}) \\ &= \begin{cases} \left(1 + \frac{\eta'_0}{\eta'_1} \prod_{dk \in S_{k,j}} \frac{\rho'_{0,dk}}{\rho'_{1,dk}}\right)^{-1}, & \text{if } q_{k-1,j} = q_{k,j-1} = 1. \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (36)$$

Similarly, it can be shown that

$$\begin{aligned} \mu_{k,j-1} &= p(t_{k,j-1} = 1 | \mathbf{r}_j, \mathbf{q}_j, \mathbf{u}_{j-1}, \mathbf{q}_{j-1}) \\ &= \begin{cases} \left(1 + \frac{\eta_0}{\eta_1} \prod_{dk \in S_{k,j}} \frac{\rho_{0,dk}}{\rho_{1,dk}}\right)^{-1} & \text{if } q_{k,j} = q_{k,j-1} = 1 \\ 0 & \text{else} \end{cases} \end{aligned} \quad (37)$$

and

$$\begin{aligned} \mu_{k,j-1} &= p(t_{k,j-1} = 1 | \mathbf{r}_j, \mathbf{q}_j, \mathbf{u}_{j-1}, \mathbf{q}_{j-1}) \\ &= \begin{cases} \left(1 + \frac{\eta_0}{\eta_1} \prod_{dk \in S_{k,j}} \frac{\rho_{0,dk}}{\rho_{1,dk}}\right)^{-1} & \text{if } q_{k+1,j} = q_{k,j-1} = 1. \\ 0 & \text{else} \end{cases} \end{aligned} \quad (38)$$

APPENDIX III

DERIVATION OF THE PARAMETERS $\lambda_{k,j}^f$, $m_{k,j}^f$ AND $V_{k,j}^f$

Our goal is to derive the BG distribution $p(r_{k,j}, q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j)$. Let $S_{k,j+1} = \{-1 \leq dk \leq 1 : q_{k+dk,j+1} = 1\}$, then this distribution can be rewritten as

$$\begin{aligned} & p(r_{k,j}, q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j) \\ & \propto p(\mathbf{y}_j | \mathbf{r}_j) p(t'_{k,j}, t'_{k,j}, t'_{k,j} | q_{k,j}) p(r_{k,j} | q_{k,j}) p(q_{k,j}) \\ & \quad \times \prod_{dk \in S_{k,j+1}} p(r_{k+dk,j+1} | q_{k+dk,j+1}, \mathbf{t}_j, \mathbf{r}_j, \mathbf{q}_j). \end{aligned} \quad (39)$$

Now, defining

$$g_1 = p(t'_{k,j}, t'_{k,j}, t'_{k,j} | q_{k,j} = 1) p(q_{k,j} = 1) \quad (40)$$

$$\tilde{\sigma}_r^2 = (1 - a^2) \sigma_r^2 \quad (41)$$

$$\chi_{dk} = \begin{cases} 1 & \text{if } r_{k+dk,j+1} \text{ and } r_{k,j} \text{ are correlated} \\ 0 & \text{otherwise} \end{cases} \quad (42)$$

$$\mathbf{f}_j = \{\mathbf{t}_j, \mathbf{r}_j, \mathbf{q}_{-k,j}\} \quad (43)$$

and

$$\rho_{1,dk} = p(r_{k+dk,j+1} | q_{k+dk,j+1} = 1, \mathbf{f}_j, q_{k,j} = 1, \chi_{dk} = 0) \quad (44)$$

we get that

$$\begin{aligned} & p(r_{k,j}, q_{k,j} = 1 | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j) \\ & \propto \frac{\sqrt{V_{k,j}^f} g_1}{\sigma_r} \exp\left[-\frac{(m_{k,j}^f)^2}{2V_{k,j}^f}\right] \\ & \quad \times \frac{1}{\sqrt{2\pi V_{k,j}^f}} \exp\left[-\frac{(r_{k,j} - m_{k,j}^f)^2}{2V_{k,j}^f}\right] \\ & \quad \times \prod_{dk \in S_{k,j+1}} (\rho_{1,dk})^{1 - \chi_{dk}} \\ & \quad \times \left\{ \frac{1}{\sqrt{2\pi \tilde{\sigma}_r^2}} \exp\left[-\frac{r_{k+dk,j+1}^2}{2\tilde{\sigma}_r^2}\right] \right\}^{\chi_{dk}} \end{aligned} \quad (45)$$

with

$$V_{k,j}^f = \left[\frac{1}{\sigma_r^2} + \frac{E_h}{\sigma_w^2} + \sum_{dk \in S_{k,j+1}} \chi_{dk} \frac{a^2}{\tilde{\sigma}_r^2} \right]^{-1} \quad (47)$$

$$m_{k,j}^f = V_{k,j}^f \left[\sum_{dk \in S_{k,j+1}} \chi_{dk} \frac{a r_{k+dk,j+1}}{\tilde{\sigma}_r^2} + m_w \right] \quad (48)$$

(37) and m_w, E_h as defined in (21) and (22), respectively.

Similarly, let

$$g_0 = p\left(t'_{k,j}, t^-_{k,j}, t^\backslash_{k,j} | q_{k,j} = 0\right) p(q_{k,j} = 0) \quad (49)$$

and

$$\rho_{0,dk} = p(r_{k+dk,j+1} | q_{k+dk,j+1} = 1, \mathbf{f}_j, q_{k,j} = 0) \quad (50)$$

then

$$p(r_{k,j}, q_{k,j} = 0 | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j) \propto g_0 \delta(r_{k,j}) \prod_{dk \in S_{k,j+1}} \rho_{0,dk}. \quad (51)$$

Finally, from (46) and (51) we get that

$$p(q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j) \sim B\left(\lambda_{k,j}^f\right) \quad (52)$$

and

$$p(r_{k,j}, q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j) \sim \lambda_{k,j}^f N\left(m_{k,j}^f, V_{k,j}^f\right) + \left(1 - \lambda_{k,j}^f\right) \delta(r_{k,j}) \quad (53)$$

where

$$\begin{aligned} \lambda_{k,j}^f &= p(q_{k,j} = 1 | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j) \\ &= \left\{ 1 + \frac{\sigma_r g_0}{\sqrt{V_{k,j}^f} g_1} \prod_{dk \in S_{k,j+1}} \frac{\rho_{0,dk} (2\pi \tilde{\sigma}_r^2)^{\frac{\chi dk}{2}}}{(\rho_{1,dk})^{1-\chi dk}} \right. \\ &\quad \left. \times \exp\left[-\frac{(m_{k,j}^f)^2}{2V_{k,j}^f} + \sum_{dk \in S_{k,j+1}} \chi dk \frac{r_{k+dk,j+1}^2}{2\tilde{\sigma}_r^2}\right] \right\}^{-1}. \end{aligned} \quad (54)$$

APPENDIX IV

DERIVATION OF THE PARAMETERS $\lambda_{k,j}^m$, $m_{k,j}^m$, AND $V_{k,j}^m$

The BG distribution $p(r_{k,j}, q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j)$ can be rewritten as

$$\begin{aligned} &p(r_{k,j}, q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j) \\ &\propto p(\mathbf{y}_j | \mathbf{r}_j) p\left(t'_{k,j}, t^-_{k,j}, t^\backslash_{k,j} | q_{k,j}\right) p(r_{k,j} | q_{k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}) \\ &\quad \times p\left(q_{k,j} | t'_{k+1,j-1}, t^-_{k,j-1}, t^\backslash_{k-1,j-1}\right) \\ &\quad \times \prod_{dk \in S_{k,j+1}} p(r_{k+dk,j+1} | q_{k+dk,j+1}, \mathbf{t}_j, \mathbf{r}_j). \end{aligned} \quad (55)$$

Let

$$\begin{aligned} c_1 &= p\left(t'_{k,j}, t^-_{k,j}, t^\backslash_{k,j} | q_{k,j} = 1\right) \\ &\quad \times p\left(q_{k,j} = 1 | t'_{k+1,j-1}, t^-_{k,j-1}, t^\backslash_{k-1,j-1}\right) \end{aligned} \quad (56)$$

then using (18), (41), (42), and (44) we get that

$$p(r_{k,j}, q_{k,j} = 1 | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j) \propto \frac{\sqrt{V_{k,j}^m} c_1}{\sigma_{k,j}} \exp\left[-\frac{m_{k,j}^m}{2\sigma_{k,j}^2} + \frac{(m_{k,j}^m)^2}{2V_{k,j}^m}\right] \frac{1}{\sqrt{2\pi V_{k,j}^m}}$$

$$\begin{aligned} &\times \exp\left[-\frac{(r_{k,j} - m_{k,j}^m)^2}{2V_{k,j}^m}\right] \prod_{dk \in S_{k,j+1}} (\rho_{1,dk})^{1-\chi dk} \\ &\quad \times \left\{ \frac{1}{\sqrt{2\pi \tilde{\sigma}_r^2}} \exp\left[-\frac{r_{k+dk,j+1}^2}{2\tilde{\sigma}_r^2}\right] \right\}^{\chi dk} \end{aligned} \quad (57)$$

with

$$V_{k,j}^m = \left[\frac{1}{\sigma_{k,j}^2} + \frac{E_h}{\sigma_w^2} \sum_{dk \in S_{k,j+1}} \chi dk \frac{a^2}{\tilde{\sigma}_r^2} \right]^{-1} \quad (58)$$

$$m_{k,j}^m = V_{k,j}^m \left[\sum_{dk \in S_{k,j+1}} \chi dk \frac{ar_{k+dk,j+1}}{\tilde{\sigma}_r^2} + \frac{m_{k,j}}{\sigma_{k,j}^2} + m_w \right] \quad (59)$$

and m_w, E_h as defined in (21) and (22), respectively.

Similarly, let

$$\begin{aligned} c_0 &= p\left(t'_{k,j}, t^-_{k,j}, t^\backslash_{k,j} | q_{k,j} = 0\right) \\ &\quad \times p\left(q_{k,j} = 0 | t'_{k+1,j-1}, t^-_{k,j-1}, t^\backslash_{k-1,j-1}\right) \end{aligned} \quad (60)$$

then using (50) we get that

$$p(r_{k,j}, q_{k,j} = 0 | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j) \propto c_0 \delta(r_{k,j}) \prod_{dk \in S_{k,j+1}} \rho_{0,dk}. \quad (61)$$

Finally, from (57) and (61) we get that

$$p(q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j) \sim B\left(\lambda_{k,j}^m\right) \quad (62)$$

and

$$p(r_{k,j}, q_{k,j} | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j) \sim \lambda_{k,j}^m N\left(m_{k,j}^m, V_{k,j}^m\right) + \left(1 - \lambda_{k,j}^m\right) \delta(r_{k,j}) \quad (63)$$

where

$$\begin{aligned} \lambda_{k,j}^m &= p(q_{k,j} = 1 | \mathbf{y}_j, \mathbf{r}_{-k,j}, \mathbf{q}_{-k,j}, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}, \mathbf{r}_{j+1}, \mathbf{q}_{j+1}, \mathbf{t}_j) \\ &= \left\{ 1 + \frac{\sigma_{k,j} c_0}{\sqrt{V_{k,j}^m} c_1} \exp\left[-\frac{(m_{k,j}^m)^2}{2V_{k,j}^m} + \frac{m_{k,j}^m}{2\sigma_{k,j}^2} \right. \right. \\ &\quad \left. \left. + \sum_{dk \in S_{k,j+1}} \chi dk \frac{r_{k+dk,j+1}^2}{2\tilde{\sigma}_r^2}\right] \right\}^{-1} \\ &\quad \times \prod_{dk \in S_{k,j+1}} \frac{\rho_{0,dk} (2\pi \tilde{\sigma}_r^2)^{\frac{\chi dk}{2}}}{(\rho_{1,dk})^{1-\chi dk}} \left\}^{-1}. \end{aligned} \quad (64)$$

ACKNOWLEDGMENT

The authors thank Dr. A. Vassiliou of GeoEnergy Inc., Houston, TX, for valuable discussions. They also thank the anonymous reviewers for their constructive comments and helpful suggestions.

REFERENCES

- [1] J. Mendel, J. Kormylo, F. Aminzadeh, J. S. Lee, and F. Habibi-Ashrafi, "A novel approach to seismic signal processing and modeling," *Geophys.*, vol. 46, pp. 1398–1414, 1981.
- [2] J. Kormylo and J. Mendel, "Maximum likelihood detection and estimation of Bernoulli-Gaussian processes," *IEEE Trans. Inf. Theory*, vol. 28, no. 3, pp. 482–488, 1982.
- [3] J. M. Mendel, *Optimal Seismic Deconvolution: An Estimation-Based Approach*. New York: Academic, 1983.
- [4] J. Goutsias and J. M. Mendel, "Maximum-likelihood deconvolution: An optimization theory perspective," *Geophys.*, vol. 51, pp. 1206–1220, 1986.
- [5] K. F. Kaareesen and T. Taxt, "Multichannel blind deconvolution of seismic signals," *Geophys.*, vol. 63, no. 6, pp. 2093–2107, 1998.
- [6] K. F. Kaareesen, "Deconvolution of sparse spike trains by iterated window maximization," *IEEE Trans. Signal Process.*, vol. 45, no. 5, pp. 1173–1183, 1997.
- [7] Q. Cheng, R. Chen, and T. H. Li, "Simultaneous wavelet estimation and deconvolution of reflection seismic signals," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 2, pp. 377–384, 1996.
- [8] C. P. Robert, *The Bayesian Choice*. New York: Springer-Verlag, 1994.
- [9] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, 1984.
- [10] O. Rosec, J. Boucher, B. Nsiri, and T. Chonavel, "Blind marine seismic deconvolution using statistical MCMC methods," *IEEE J. Ocean. Eng.*, vol. 28, no. 3, pp. 502–512, 2003.
- [11] M. Lavielle, "A stochastic algorithm for parametric and non-parametric estimation in the case of incomplete data," *Signal Process.*, vol. 42, no. 1, pp. 3–17, 1995.
- [12] G. Celeux and J. Diebolt, "The SEM algorithm: A probabilistic teacher algorithm derived from the EM algorithm for the mixture problem," *Computat. Statist. Quarter.*, vol. 2, no. 1, pp. 73–82, 1985.
- [13] G. Celeux, D. Chauveau, and J. Diebolt, "Stochastic versions of the EM algorithm: An experimental study in the mixture case," *J. Statist. Computat. Simulat.*, vol. 55, no. 4, pp. 287–314, 1996.
- [14] B. Chalmond, "An iterative Gibbsian technique for reconstruction of m -ary images," *Pattern Recogn.*, vol. 22, no. 6, pp. 747–761, Jun. 1989.
- [15] J. Idier and Y. Goussard, "Multichannel seismic deconvolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 5, pp. 961–979, 1993.
- [16] J. Idier and Y. Goussard, "Markov modeling for Bayesian restoration of two-dimensional layered structures," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1356–1373, 1993.
- [17] A. Heimer, I. Cohen, and A. Vassiliou, "Dynamic programming for multichannel blind seismic deconvolution," in *Proc. Soc. Exploration Geophys. Int. Conf., Expo. 77th Annual Meet.*, San Antonio, Sep. 23–28, 2007, pp. 1845–1849.
- [18] A. Heimer and I. Cohen, "Multichannel blind seismic deconvolution using dynamic programming," *Signal Process.*, vol. 88, pp. 1839–1851, Jul. 2008.
- [19] A. A. Amini, T. E. Weymouth, and R. C. Jain, "Using dynamic programming for solving variational problems in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 9, pp. 855–867, 1990.
- [20] M. Buckley and J. Yang, "Regularised shortest-path extraction," *Pattern Recogn. Lett.*, vol. 18, no. 7, pp. 621–629, 1997.
- [21] A. Heimer and I. Cohen, "Multichannel seismic deconvolution using Markov-Bernoulli random field modeling," *IEEE Trans. Geosci. Remote Sens.*, to be published.
- [22] D. G. Forney, "The Viterbi algorithm," *Proc. IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.
- [23] K. F. Kaareesen, "Evaluation and applications of the iterated window maximization method for sparse deconvolution," *IEEE Trans. Signal Process.*, vol. 46, no. 3, pp. 609–624, 1998.
- [24] L. Woog, I. Popovic, and A. Vassiliou, "Applications of adapted wave-form analysis for spectral feature extraction and denoising," in *Proc. Soc. Exploration Geophysicist Int. Conf., Expo., 75th Annu. Meeting*, Houston, TX, Nov. 6–11, 2005, pp. 2181–2184.



Idan Ram received the B.Sc. (*cum laude*) and M.Sc. degrees in electrical engineering in 2004 and 2009, respectively, from the Technion—Israel Institute of Technology, Haifa.

He is currently pursuing the Ph.D. degree in electrical engineering at the Technion. His research interests are image denoising and deconvolution using sparse signal representations.



Israel Cohen (M'01–SM'03) received the B.Sc. (*summa cum laude*), M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, Haifa, in 1990, 1993, and 1998, respectively.

From 1990 to 1998, he was a Research Scientist with RAFAEL Research Laboratories, Haifa, Israel Ministry of Defense. From 1998 to 2001, he was a Postdoctoral Research Associate with the Computer Science Department, Yale University, New Haven, CT. In 2001, he joined the Electrical Engineering Department, Technion, where he is currently an Associate Professor. His research interests are statistical signal processing, analysis and modeling of acoustic signals, speech enhancement, noise estimation, microphone arrays, source localization, blind source separation, system identification, and adaptive filtering. He is a coeditor of the Multichannel Speech Processing section of the "Springer Handbook of Speech Processing" (New York: Springer, 2007), a coauthor of "Noise Reduction in Speech Processing" (New York: Springer, 2009).

Dr. Cohen received the Technion Excellent Lecturer awards in 2005 and 2006, and the Muriel and David Jacknow award for Excellence in Teaching in 2009. He served as an Associate Editor of the IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING and the IEEE SIGNAL PROCESSING LETTERS. He served as Guest Editor of a Special Issue of the EURASIP *Journal on Advances in Signal Processing* on Advances in Multimicrophone Speech Processing and a Special Issue of the EURASIP *Speech Communication Journal* on Speech Enhancement. He was also a co-chair of the 2010 International Workshop on Acoustic Echo and Noise Control.

Shalom Raz, photograph and biography not available at the time of publication.