

Markov Modeling for Bayesian Restoration of Two-Dimensional Layered Structures

Jérôme Idier and Yves Goussard

Abstract— Bayesian estimation of two-dimensional stratified structures is described. The major point addressed here is the derivation of a statistical prior model that adequately describes such layered media. This problem is of interest in application domains such as seismic exploration, nondestructive testing, and medical imaging, where the data are generally processed in one dimension only. In order to take local interactions into account, a Markovian description is used. The model is derived so as to fulfill a set of constraints which summarize physical and geometrical characteristics of the problem as well as practical requirements. The approach adopted is reminiscent of the work of Pickard. The resulting class of Markov random fields presents a unilateral structure on a nonrectangular lattice and a hierarchical organization which involves a *line process*. In addition, it is shown to be an extension of one-dimensional models already used in the application domains previously mentioned. After an investigation of the properties of the model, its practical interest is demonstrated by an application to seismic deconvolution. Simulation results show significant improvements with respect to the usual one-dimensional methods.

Index Terms— Markov random fields, unilateral processes, Pickard random fields, Bayesian estimation, modeling of stratified media.

I. INTRODUCTION

IN RECENT YEARS, the use of Markov random fields (MRF's) has raised considerable attention in the area of image processing [1]–[4]. This interest stems from the fact that MRF's provide a general framework for specifying local interactions, and can adequately model the inhomogeneous structures encountered in many real world images. Several Bayesian methods using MRF's as priors have been proposed for performing such tasks as reconstruction, restoration, deconvolution, and segmentation of images.

Specification of adequate priors is also important in other application fields such as geophysics, nondestructive testing, and medical ultrasonic imaging. In these domains, a two-dimensional (2-D) signal which often represents a stratified structure has to be estimated. Accounting for the stratification presents important difficulties and, to our knowledge, no fully satisfactory answer has been given to this problem yet. Existing estimation methods fall into two categories. The ones

which rely on empirical 2-D treatments such as [5] or [6], and the ones in which the 2-D stratified field is considered as a juxtaposition of one-dimensional (1-D) signals perpendicular to the general orientation of the layers. Methods of the first kind present the disadvantage of questionable theoretical properties. Among methods of the second kind, Mendel *et al.* [7], [8], followed by others [9]–[11], proposed and investigated the properties of a Bayesian approach to seismic deconvolution based upon a Bernoulli–Gaussian (B–G) description of 1-D signals; interesting results were obtained. However, in this approach, the 1-D signals are processed independently from one another and these methods do not account for one of the major characteristic of a 2-D stratified field, i.e., continuity of the layers or, equivalently, correlation between the 1-D signals. This illustrates the inherent limits of 1-D methods in 2-D problems and underlines the need for an adequate and theoretically sound prior model of 2-D layered structures.

As the continuity of the layers is essentially a local characteristic, Markov representations appear to be good candidates for specifying this 2-D prior. The simplest MRF's, e.g., homogeneous first- or second-order MRF's, cannot provide an adequate answer since they do not capture an essential property of 2-D stratified fields, i.e., the presence of discontinuities at the layer boundaries. To account for this characteristic, it seems natural to introduce a *line process* as suggested by Geman and Geman [2]. However, direct application of the techniques described in [2] is not appropriate since the model should also fulfill some additional statistical properties connected to the general orientation of the layers, their thickness distribution, etc. Therefore, derivation of a specific MRF seems necessary to model 2-D stratified fields adequately.

The goal of the paper is to propose such a model, to investigate its theoretical properties, and to demonstrate its practicality through simulations and application to seismic deconvolution. This model is made up of two parts which are organized in a hierarchical manner. The higher level is rather general as it controls the geometrical characteristics of the 2-D field, i.e., the location and shape of the layers. It will be referred to as the *layer boundary model*. Its salient characteristics are the presence of hidden *transition variables* which allow an explicit modeling of the boundaries while preserving a simple structure with a small neighborhood size, and a unilateral structure on a nonrectangular lattice which simplifies simulation of the field and implementation of the estimation method. The lower level of the model specifies the distribution of attributes of the layers, such as the value of physical quantities, textures, etc., and is therefore more

Manuscript received November 22, 1990; revised April 14, 1992. This work was presented in part at the IEEE International Conference on Acoustics, Speech, and Signal Processing, Albuquerque, NM, 1990.

J. Idier is with the Laboratoire des Signaux & Systèmes, Ecole Supérieure d'Electricité, Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, France.

Y. Goussard is with Ecole Polytechnique, Biomedical Engineering Institute, P.O. Box 6079, Station A, Montreal, PQ, H3C 3A7, Canada.

IEEE Log Number 9209456.

0018-9448/93\$03.00 © 1993 IEEE

© 1993 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

application-specific. The resulting MRF constitutes a 2-D extension of the 1-D B-G model already used with some success for seismic deconvolution.

Section II is dedicated to the derivation of the layer boundary model and to the investigation of its theoretical properties. A precise statement of the problem is given in Section II-A, and Section II-B contains some background material on Markov-type random fields which will be useful later. Hidden transition variables are introduced in Section II-C, so as to build a compound unilateral MRF (UMRF). In Section II-D, it is shown that the UMRF can be characterized by a local probability measure on simple cells of a nonrectangular lattice. This measure is further specified in Section II-E so as to fulfill global constraints; this yields a parsimonious parametrization of the probability measure. Section II-F provides a physical interpretation of the parameters, and additional properties of the UMRF are investigated in Section II-G.

In Section III, an application to Bayesian multichannel seismic deconvolution is presented. The problem is briefly described in Section III-A, and the main characteristics of usual 1-D approaches are outlined in Section III-B. In Section III-C, the 2-D Markov model studied in Section II is completed so as to incorporate amplitude information. The resulting model is used as priors in a *maximum a posteriori* (MAP) deconvolution procedure described in Section III-D. Section III-E contains simulation results which illustrate the performances and the practicality of the approach.

Finally, conclusions and perspectives for future work are presented in Section IV.

II. CONSTRUCTION OF THE LAYER BOUNDARY MODEL

This part of the model accounts for the geometric characteristic of the 2-D stratified field. It stands for the shape of the boundaries, regardless of the content of the layers, since the latter induces no definite information on the former. The resulting binary structure is reminiscent of *line processes* used for image restoration and segmentation [2], [4], though it fulfills several specific properties such as being a UMRF, whereas general line processes may not even be MRF's [4].

Very important to the derivation of the model is the information about the general direction of the layers. Indeed, the whole construction relies on the fact that the most layers are only slightly bent. Rigorously speaking, we mean that all but a few must only bend into less than ninety degrees from an average direction (without loss of generality, it is assumed horizontal in the sequel). In practice, for a wide range of application domains, such an assumption is merely valid as an understood consequence of stratification. On the other hand, the incorporation of discontinuities in Section II-C-3 allows scarce abruptly bent layers or vertical boundary lines.

A. Problem Statement and Basic Constraints

Assume that the 2-D layered field is sampled on a finite rectangular grid $\Lambda^\circ = \{(i, j): 1 \leq i \leq I; 1 \leq j \leq J\}$. Integer variables i and j index the vertical and horizontal directions respectively (see Fig. 1). To each site $(i, j) \in \Lambda^\circ$, one associates a binary 0-1 *location variable* Q_{ij} that

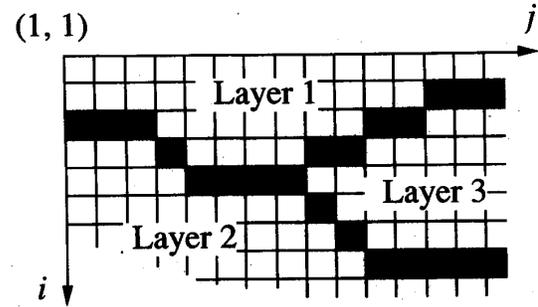


Fig. 1. Discrete binary representation of layer boundaries: Black squares denote the presence of a layer boundary on the corresponding site.

indicates whether a boundary is absent or present on the corresponding site. It shall remain implicit that vertical and horizontal sampling periods have been chosen small enough so that the boundaries can be represented as unbroken chains of nonzero location variables.

In the statistical approach adopted here, $\mathbf{Q} = \{Q_{ij}: (i, j) \in \Lambda^\circ\}$ is considered as a random field, and specifying a model for \mathbf{Q} is equivalent to defining a probability measure $P(\mathbf{Q} = \mathbf{q})$ for every possible realization \mathbf{q} of \mathbf{Q} . In order to account for the specific characteristics of layered media, three structural constraints are introduced on P :

- 1) Each binary column should follow a Bernoulli distribution.
- 2) P should be insensitive to reverting horizontal or vertical indexation.
- 3) P should incorporate local spatial dependencies to ensure lateral continuity between layer boundary sites.

According to constraint 1), each column \mathbf{Q}_j should contain independent identically distributed (i.i.d.) Bernoulli random variables Q_{ij} :

$$P(\mathbf{Q}_j = \mathbf{q}_j) = \prod_{i=1}^I P(Q_{ij} = q_{ij}) = \lambda^{I_j} (1 - \lambda)^{I - I_j}$$

where I_j is the number of ones in \mathbf{q}_j , standing for layer boundary sites, and λ controls the sparsity of layer boundaries along the vertical direction. Constraint 1) indicates that the model should yield a true extension of 1-D Bernoulli models which were fruitful in the area of seismic deconvolution [8]. Such an extension motivated part of the study. Note that a Bernoulli model for boundary sites corresponds to a geometrical thickness distribution of the layers, which also arises from other 1-D seismic models [12].

Constraint 2) is introduced in order to preserve physical symmetries intrinsic to the layered structure. More precisely, the indexation order of the sites is arbitrary, i.e., it is not connected to any physical orientation of the layers. Consequently, the model should not be sensitive to the indexation order.

Constraint 3) simply reflects the basic requirement of correlation between adjacent columns so as to favor roughly horizontal and unbroken boundaries.

In addition to structural constraints 1)–3), a fourth practical constraint is introduced so as to cope with the numerical requirements of many applications:

- 4) P should have a simple enough structure so as to yield numerically tractable estimation.

B. Markovian Modeling

Markovian modeling is specifically suited to account for local spatial dependencies as stated in 3). Numbers of Markov-type random fields have been introduced during the past two decades, and an exhaustive overview can be found in [13]. The most general class of Markovian processes on finite lattices is formed by Markov random fields (MRF's), from which almost every other family derives as particular case or subclass. In the following, we only give the general definition of MRF's on finite lattices and present developments directly connected with our problem. References are given for a more extensive comprehension of Markovian modeling.

1) *Markov Random Fields*: Henceforth, whenever unambiguous, probability distribution $P(X = x | Y = y)$ will be noted $P(x | y)$, where X (resp. Y) and x (resp. y) denote any random variable and one of its realizations.

A MRF is a stochastic process $X = \{X_n, 1 \leq n \leq N\}$ whose labelling corresponds to an arbitrary enumeration of a finite set of sites $\Lambda = \{s_n, 1 \leq n \leq N\}$. Random variables X_1, \dots, X_N are not necessarily sampled from a common state space. In the general theory as found in Kinderman and Snell [14], Besag [15], or Geman and Geman [2], Λ does not need to be a subset of \mathbb{Z}^2 , nor hold any other structural prerequisite.

A neighborhood system $\mathcal{N} \triangleq \{N_n, 1 \leq n \leq N\}$ is defined as follows:

$$\forall n, N_n \subset \Lambda \text{ and } s_n \notin N_n;$$

$$\forall m, \forall n, s_n \in N_m \Leftrightarrow s_m \in N_n.$$

Elements of N_n are called the *neighbors* of site s_n . Sometimes, this designation appears to be compatible with some preexistent spatial distribution of the sites, but such a correspondence is only optional. X is a MRF with respect to the neighborhood system \mathcal{N} , if and only if, for any site s_n and for any joint realizations $\mathbf{x} = (x_1, \dots, x_N)$ such that $P(\mathbf{x}) > 0$,

$$\begin{aligned} P(x_n | x_1, \dots, x_{n-1}, x_{n+1}, \dots, x_N) \\ = P(x_n | x_m, s_m \in N_n). \end{aligned} \quad (1a)$$

Rigorously, a minimality condition must be added on each N_n , otherwise \mathcal{N} would not be unique:

Minimality Condition: For any subset R of N_n such that $R \neq N_n$,

$$P(x_n | x_m, s_m \in R) \neq P(x_n | x_m, s_m \in N_n) \quad (1b)$$

for at least one \mathbf{x} , otherwise R could be substituted for N_n as the set of neighbors of site s_n .

Note that the neighborhood system \mathcal{N} defines the graph (Λ, \mathcal{N}) of the neighborhood relation, so that neighbors are pairs of mutually related sites. Any set of sites which either

consists of a single site or else in which every pair of distinct sites are neighbors is called a *clique*.

Characteristic property (1) provides an easy way to check whether a random field is a MRF and, if it is, to determine its neighborhood system. On the other hand, (1) does not guarantee that specification of local conditional probability measures $P(x_n | x_m, s_m \in N_n)$ yields a consistent joint probability measure. Indeed, this crucial problem of inferring global statistical properties from local conditional ones has been studied in detail by Besag [15]. It appears that severe restrictions must be imposed on the functional forms of the conditional probability distributions $P(x_n | x_m, s_m \in N_n)$ to yield a mathematically consistent joint probability measure. In this respect, an important result has been established by Hammersley and Clifford (see [15]). The Hammersley-Clifford theorem states that every MRF for which $P(\mathbf{x}) > 0$ for every realization \mathbf{x} (positivity condition), is equivalent to a Gibbs random field (GRF). Since Gibbs distributions can be explicitly expressed using local functions (namely, Gibbs potentials), the Markov-Gibbs equivalence provides the most general way of specifying local conditional probability measures which are consistent with a valid MRF structure, under the positivity condition. Further discussion and proofs can be found in [15].

2) *Unilateral MRF's*: In spite of its fruitfulness, the Markov-Gibbs equivalence should not be regarded as the only way of defining a MRF from local characteristics, as some perfectly valid MRF's may not fulfill the positivity condition. Pickard [16] proposed an alternative approach based upon a *unilateral* specification of local probability distributions. To begin with, note that any arbitrary enumeration of the sites in Λ allows one to factor the joint probability measure of X according to

$$P(\mathbf{x}) = P(x_1) \prod_{n=2}^N p(x_n | x_1, \dots, x_{n-1}).$$

If, for some enumeration, each conditional probability $P(x_n | x_1, \dots, x_{n-1})$ only depends on a restricted subset P_n of *predecessors* of current site s_n , then it is straightforward to show that X is a MRF with restricted sets of neighbors and with *maximal* cliques $P_n \cup \{s_n\}$ (*maximal* cliques are those which are not subsets of any strictly larger ones). MRF's which can be defined in this manner are called *unilateral MRF's* (UMRF's). They form a subclass of general MRF's, which can be thought as the most natural extension of Markov chains to multidimensional lattices [16]. Unlike general MRF's, UMRF's yield a recursive representation of their probability distribution:

$$P(\mathbf{x}) = P(x_1) \prod_{n=2}^N P(x_n | x_m, s_m \in P_n).$$

This formulation allows easy simulations and does not require the positivity condition. Moreover, it has been shown by Goutsias [17] that, under certain conditions which guarantee translation invariance, the only isotropic GRF's on rectangular lattices are a special case of UMRF's, namely the *Curious*

Lattice Process introduced by Pickard in [18] and renamed Pickard random field (PRF).

On the other hand, unilaterality requires a sequential enumeration of the sites, which may not arise naturally. For instance, in order to perform model-based recursive image processing, UMRF's have often been related to lexicographic enumerations on 2-D rectangular lattices or to the idea of quarter-plane causality borrowed from 2-D recursive filtering. The former gave rise to nonsymmetric half-plane models [3], [13] and the latter yielded *Markov mesh* processes [19]–[21]. As mentioned by Besag [1] and Geman and Geman [2], such models can be considered as useful approximations of more complex MRF's, even if the resulting neighborhoods are irregular. Unilateral approximation of GRF's is specifically addressed by Goutsias in [22].

However, unilaterality does not necessarily involve such enumeration schemes. The approach taken in this paper provides an original alternative example. In Section II-C, hidden variables forming a *line process* (in Geman and Geman's terminology) are introduced *between* the sites of the initial rectangular grid Λ° . Then the resulting compound lattice Λ , which is not rectangular any more, provides an enumeration of the sites which is natural given the problem statement. Hence, a unilateral approach can be adopted in conformity with the work of Pickard [16].

C. Introducing Transition Variables T_{ij}

Definition of Q requires the specification of the joint probability measure $P(Q = q)$, for any $I \times J$ binary matrix q . Usual values of I and J (256 at least) obviously preclude any global specification of $P(Q = q)$. Adequate factoring must be sought instead, in order to define the random field Q from tractable short-range interactions which fulfill the constraints specified in Section II-A. As mentioned in the Introduction, Geman and Geman [2] showed that hidden variables such as the *line process* can effectively model local interactions which occur at discontinuities. Here, it is not appropriate to directly use the models proposed in [2], as constraints 1) and 4) could hardly be fulfilled. However, in order to indicate whether two adjacent nonzero location variables belong to the same layer boundary, it is possible to introduce another kind of hidden binary *transition* variables T_{ij} placed between location variables Q_{ij} : T_{ij} is set to 1 when the two adjacent location variables belong to the same boundary, and to 0, otherwise. Since the field is stratified in the horizontal direction, it is sufficient to define transition sites between pairs of location sites which are either diagonally or horizontally adjacent (see Fig. 2).

Let Λ^+ , Λ^- , and Λ^\setminus denote the sets of diagonally ascending, horizontal, and diagonally descending transition sites. The set of transition variables $T = \{T_{ij}^+\}_{\Lambda^+} \cup \{T_{ij}^-\}_{\Lambda^-} \cup \{T_{ij}^\setminus\}_{\Lambda^\setminus}$ constitutes a random field on the nonrectangular lattice $\Lambda^+ \cup \Lambda^- \cup \Lambda^\setminus$, and it will be referred to as the transition field.

The major interest of these hidden transition variables is the possibility of specifying interactions between location variables in a local manner. On the other hand, characterization of the compound field $\{T, Q\}$ on $\Lambda = \Lambda^\circ \cup \Lambda^+ \cup \Lambda^- \cup \Lambda^\setminus$ is required in order to fully define the layer boundary model.

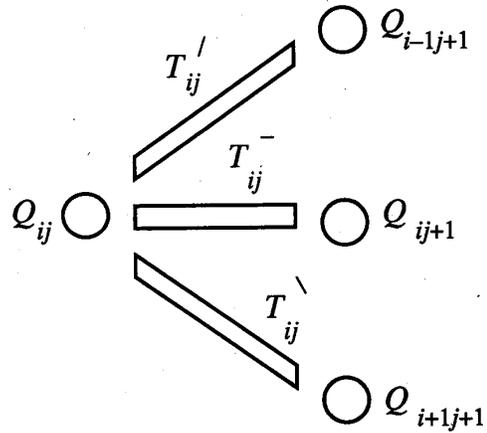


Fig. 2. Introducing transition variables T_{ij} : Transition variables are placed between pairs of location sites which are either diagonally or horizontally adjacent. They are meant to indicate whether two adjacent nonzero location variables belong to the same layer boundary.

D. Markov Random Field $\{T, Q\}$

1) *Unilateral Characterization*: We now turn to the unilateral joint characterization of $\{T, Q\}$ on lattice Λ instead of Q on Λ° alone. In [16], Pickard starts the derivation of the so-called *Curious Lattice Process* by suggesting that it is possible to construct unilateral processes from probability measures on generic cliques. On an $I \times J$ rectangular lattice, Pickard considers a measure τ on the basic cell $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ and introduces the probability distribution P of a unilateral process X in the following factored form

$$p(x) = \tau(x_{1,1}) \prod_{i=2}^I \tau(x_{i1} | x_{i-1,1}) \prod_{j=2}^J \tau(x_{1j} | x_{1j-1}) \cdot \prod_{i=2}^I \prod_{j=2}^J \tau(x_{ij} | x_{i-1j}, x_{ij-1}, x_{i-1j-1}).$$

Then Pickard investigates the structural conditions to be imposed on measure τ so that the restriction of P to any clique $(X_{ij}, X_{i-1j}, X_{ij-1}, X_{i-1j-1})$ is invariant and identical to τ .

Similarly, we are looking for a factored form of $P(t, q)$, which corresponds to a unilateral enumeration of the sites and in which each factor can be derived from a measure on basic cells. This can be achieved from three mere assumptions:

Assumption H1): Columns of $\{T, Q\}$, i.e., $\{q_1, t_1, q_2, \dots, t_{j-1}, q_j\}$, form a *first-order* vector Markov chain, so that $P(t, q)$ factors according to

$$P(t, q) = P(q_1) \prod_{j=1}^{J-1} P(t_j | q_j) P(q_{j+1} | t_j). \quad (2)$$

Higher order chains could be considered in order to incorporate more complex local interactions, such as a smooth curvature on boundaries. This raises the question of the trade-off between model accuracy and complexity. The approach taken in the paper remains valid for higher order vector Markov chains, as we have checked that second-order models

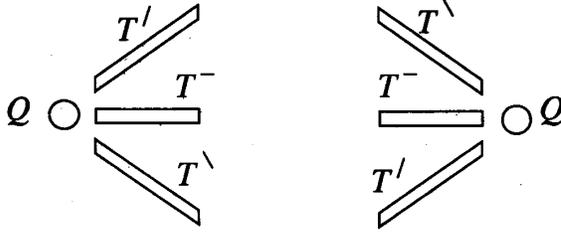


Fig. 3. Generic cells on lattice Λ : The unilateral joint characterization of $\{T, Q\}$ is based on the decomposition of $P(t, q)$ as a product of local probabilities on these cells.

which meet Constraints 1)–3) do exist. But while there is no theoretical obstacle (but no theoretical stake either) in developing higher-order models, the computational burden quickly becomes very heavy. For instance, 32 parameters are needed to specify second-order models, as opposed to 4 for first-order ones (see Section II-G). Hence we feel that the potential performance improvements are outbalanced by the additional computational complexity, and we will not pursue further in this direction.

Assumption H2): Each conditional probability in (2) further factors into the following product of local probabilities

$$P(t_j | q_j) = \prod_{i=1}^I P(t'_{ij}, t^-_{ij}, t^\backslash_{ij} | q_{ij}), \quad (3a)$$

$$P(q_{j+1} | t_j) = \prod_{i=1}^I P(q_{ij+1} | t'_{i+1j}, t^-_{ij}, t^\backslash_{i-1j}), \quad (3b)$$

and the initial term $P(q_1)$ is factored as

$$P(q_1) = \prod_{i=1}^I P(q_{i1}). \quad (4)$$

At horizontal grid boundaries, some transition variables are not defined and adjustments are required for (3) to be correct. Here a free boundary assumption is made, which means that all undefined variables are simply omitted.

Equation (3) corresponds to intuitive short-range interactions between *transition* variables and *location* variables. It actually confers a crucial role to the two generic cells depicted in Fig. 3. Each cell comprises four sites, one from each of the elementary lattices Λ^o , Λ' , Λ^- , and Λ^\backslash . It is shown in Section II-D-2 that these cells are the *maximal* cliques of the MRF $\{T, Q\}$.

Assumption H3): We choose to make all probability measures appearing in (3) and (4) derive from an arbitrary unique probability measure $\tau(q, t', t^-, t^\backslash)$ defined on both generic cells:

$$P(t'_{ij}, t^-_{ij}, t^\backslash_{ij} | q_{ij}) \triangleq \tau(t'_{ij}, t^-_{ij}, t^\backslash_{ij} | q_{ij}), \quad (5a)$$

$$P(q_{ij+1} | t'_{i+1j}, t^-_{ij}, t^\backslash_{i-1j}) \triangleq \tau(q_{ij+1} | t'_{i+1j}, t^-_{ij}, t^\backslash_{i-1j}), \quad (5b)$$

$$P(q_{i1}) \triangleq \tau(q_{i1}). \quad (5c)$$

In (5) and henceforth, the same generic symbol τ indifferently denotes all conditional or marginal probabilities

deducible from $\tau(q, t', t^-, t^\backslash)$ by projection (i.e., application of Bayes' theorem) or summation. For instance, we shall write

$$\tau(t^- | q) = \frac{\tau(q, t^-)}{\tau(q)} = \frac{\sum_{t', t^\backslash} \tau(q, t', t^-, t^\backslash)}{\sum_{t', t^-, t^\backslash} \tau(q, t', t^-, t^\backslash)},$$

where the sums extend over all the possible realizations of the corresponding variables. As in [16], H3) imposes an invariant functional form to the local conditional probability measures which appear in (3) and (4). Defining a single measure τ on both generic cells simply accounts for the symmetry requirement of constraint 2).

Using (2)–(5), we obtain the following factored form of the joint probability measure of $\{T, Q\}$:

$$P(t, q) = \prod_{i=1}^I \tau(q_{i1}) \prod_{j=1}^{J-1} \left\{ \tau(t^-_{1j}, t^\backslash_{1j} | q_{1j}) \tau(t'_{1j}, t^-_{1j} | q_{1j}) \cdot \tau(q_{1j+1} | t'_{2j}, t^-_{1j}) \tau(q_{1j+1} | t^-_{1j}, t^\backslash_{1-1j}) \cdot \prod_{i=2}^{I-1} \tau(t'_{ij}, t^-_{ij}, t^\backslash_{ij} | q_{ij}) \tau(q_{ij+1} | t'_{i+1j}, t^-_{ij}, t^\backslash_{i-1j}) \right\}. \quad (6a)$$

Equation (6a) can be rewritten in a simpler form, in which the free boundary assumption remains implicit:

$$P(t, q) = \prod_{j=1}^J \sum_{i=1}^I \tau(t'_{ij}, t^-_{ij}, t^\backslash_{ij} | q_{ij}) \cdot \tau(q_{ij} | t'_{i+1j-1}, t^-_{ij-1}, t^\backslash_{i-1j-1}). \quad (6b)$$

It should be stressed that for any valid probability measure τ , (6) defines a true MRF with a unilateral structure. The form of its neighborhood system is examined in the next subsection.

A remark should be made about the enumeration of the sites associated with the unilateral structure of the field $\{T, Q\}$. Whereas the *Curious Lattice Process* is defined on a rectangular lattice with a lexicographic enumeration of the sites (all other orderings available for *Markov mesh* models are also possible), the compound UMRF $\{T, Q\}$ is built on the more complex lattice Λ in which the sites are ordered column by column. The order is arbitrary within each column, but the transition variables must be associated three by three ($T'_{ij}, T^-_{ij}, T^\backslash_{ij}$ are kept together).

2) **Neighborhood Structure:** From (6), it is straightforward to check that the UMRF $\{T, Q\}$ fulfills MRF characteristic property (1). In Appendix A, it is shown that the sets of neighbors of interior sites (i, j) in Λ^o , Λ' , Λ^- , and Λ^\backslash take the respective forms (see also Fig. 4)

$$\begin{aligned} N_{ij}^o &= \{T'_{ij}, T^-_{ij}, T^\backslash_{ij}, T'_{i+1j-1}, T^-_{i-1j}, T^\backslash_{i-1j-1}\}, \\ N_{ij}^{\prime} &= \{Q_{ij}, T^-_{ij}, T^\backslash_{ij}, Q_{i-1j+1}, T^-_{i-1j}, T^\backslash_{i-2j}\}, \\ N_{ij}^- &= \{Q_{ij}, T'_{ij}, T^\backslash_{ij}, Q_{ij+1}, T'_{i+1j}, T^\backslash_{i-1j}\}, \\ N_{ij}^\backslash &= \{Q_{ij}, T'_{ij}, T^-_{ij}, Q_{i+1j+1}, T'_{i+2j}, T^-_{i+1j}\}, \end{aligned} \quad (7)$$

and that the corresponding bilateral conditional probabilities can, respectively, be expressed as (8) (see (8) at the bottom of the page) where $\bar{b} = 1 - b$ for any 0-1 binary variable b .

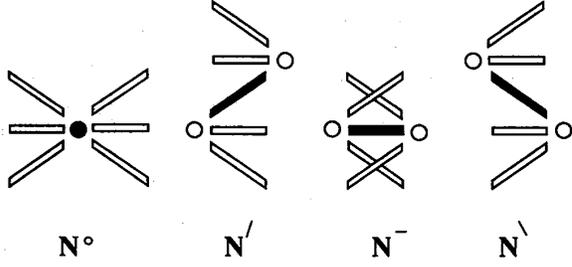


Fig. 4. Sets of neighbors: Graphic representation of the sets of neighbors N_{ij}^o , N_{ij}' , N_{ij}^- , and N_{ij}^\setminus defined by (7). The central site is shaded in black, and its neighbors are not shaded.

Moreover, it is easy to check from (7) or from Fig. 4 that the generic cells depicted in Fig. 3 are indeed the *maximal cliques* of MRF $\{T, Q\}$.

E. Defining the Measure τ

Introduction of hidden variables T_{ij} has allowed the construction of a UMRF $\{T, Q\}$ for any probability measure τ on generic cells. Now τ should be defined so as to incorporate desired features into this unilateral scheme, in accordance with constraints 1)–4). The practical constraint of numerical tractability (constraint 4)) cannot be satisfactorily studied at this point, and it will be dealt with more specifically in Section III when the problem of multichannel seismic deconvolution is addressed. However, it should already be stressed that the unilateral structure of $\{T, Q\}$ is a valuable property with respect to constraint 4). Therefore, this subsection is dedicated to the characterization of τ in accordance with structural constraints 1)–3).

As shown in the sequel, fulfillment of constraints 1) and 2) is a consequence of the invariance of the restriction of probability measure $P(t, q)$ to maximal cliques of the UMRF $\{T, Q\}$. By means of a theorem, we first prove that this invariance property holds, if and only if measure τ presents a separable form. Then, constraints 1) and 2) are fulfilled as consequences of the separability property. Finally additional constraints are imposed on the separable form of τ so as to fulfill constraint 3).

1) *Invariance of the Restriction of P to Maximal Cliques:* In

Section II-D, the joint probability measure P of $\{T, Q\}$ was derived from the probability measure τ on basic cells, according to (6). Inversely, from (6), it is possible to establish the expression of the restriction of P to any maximal clique $\{Q, T', T^-, T^\setminus\}$, by summation over the possible combinations of all the other variables on Λ . Now, a natural question arises: is this measure independent from the position of the selected clique and, if so, is it identical to the generic measure τ ? The answer is given by the following theorem.

Theorem 1: The restriction of P to maximal cliques of UMRF $\{T, Q\}$ is invariant, if and only if measure τ fulfills the following separability condition: for any triplet of transition variables (t', t^-, t^\setminus) ,

$$\tau(t', t^-, t^\setminus) = \tau(t')\tau(t^-)\tau(t^\setminus). \quad (9)$$

If measure τ fulfills condition (9), it is said *separable*.

Proof: See Appendix B.

Remark: Theorem 1 clearly indicates that, even though the UMRF $\{T, Q\}$ is completely defined from a single measure τ , not all possible forms of τ yield the invariance of P on maximal cliques.

2) *Sufficient Conditions for Fulfillment of Constraints 1) and 2):* Fulfillment of constraints 1) and 2) is a consequence of the separability condition introduced above. More precisely, we have the following two corollaries of Theorem 1.

Corollary 1 (Constraint 1): If measure τ is separable, then each column Q_j of Λ^o follows a Bernoulli distribution whose characteristic parameter is

$$\lambda = \tau(q = 1). \quad (10)$$

Furthermore each column T_j' , T_j^- , and T_j^\setminus defined on elementary lattices Λ' , Λ^- , and Λ^\setminus , respectively, follows a Bernoulli distribution. Their respective characteristic parameters are given by

$$\begin{cases} \mu' = \tau(t' = 1) \\ \mu^- = \tau(t^- = 1) \\ \mu^\setminus = \tau(t^\setminus = 1) \end{cases} \quad (11)$$

Proof: See Appendix B.

$$\begin{aligned} P(q_{ij} | N_{ij}^o) &= \left[1 + \frac{\tau(\bar{q}_{ij}, t'_{ij}, t^-_{ij}, t^\setminus_{ij})\tau(\bar{q}_{ij}, t'_{i+1j-1}, t^-_{i+1j-1}, t^\setminus_{i+1j-1})\tau(q_{ij})}{\tau(q_{ij}, t'_{ij}, t^-_{ij}, t^\setminus_{ij})\tau(q_{ij}, t'_{i+1j-1}, t^-_{i+1j-1}, t^\setminus_{i+1j-1})\tau(\bar{q}_{ij})} \right]^{-1} \\ P(t'_{ij} | N_{ij}') &= \left[1 + \frac{\tau(q_{ij}, \bar{t}'_{ij}, t^-_{ij}, t^\setminus_{ij})\tau(q_{i-1j+1}, \bar{t}'_{i-1j}, t^-_{i-1j}, t^\setminus_{i-1j})\tau(t'_{ij}, t^-_{i-1j}, t^\setminus_{i-2j})}{\tau(q_{ij}, t'_{ij}, t^-_{ij}, t^\setminus_{ij})\tau(q_{i-1j+1}, t'_{i-1j}, t^-_{i-1j}, t^\setminus_{i-2j})\tau(\bar{t}'_{ij}, t^-_{i-1j}, t^\setminus_{i-2j})} \right]^{-1} \\ P(t^-_{ij} | N_{ij}^-) &= \left[1 + \frac{\tau(q_{ij}, t'_{ij}, \bar{t}^-_{ij}, t^\setminus_{ij})\tau(q_{ij+1}, t'_{i+1j}, \bar{t}^-_{i+1j}, t^\setminus_{i+1j})\tau(t'_{ij}, t^-_{ij}, t^\setminus_{i-1j})}{\tau(q_{ij}, t'_{ij}, t^-_{ij}, t^\setminus_{ij})\tau(q_{ij+1}, t'_{i+1j}, t^-_{i+1j}, t^\setminus_{i+1j})\tau(\bar{t}^-_{i+1j}, t^-_{ij}, t^\setminus_{i-1j})} \right]^{-1} \\ P(t^\setminus_{ij} | N_{ij}^\setminus) &= \left[1 + \frac{\tau(q_{ij}, t'_{ij}, t^-_{ij}, \bar{t}^\setminus_{ij})\tau(q_{i+1j+1}, t'_{i+2j}, t^-_{i+1j}, \bar{t}^\setminus_{i+1j})\tau(t'_{i+2j}, t^-_{i+1j}, t^\setminus_{ij})}{\tau(q_{ij}, t'_{ij}, t^-_{ij}, t^\setminus_{ij})\tau(q_{i+1j+1}, t'_{i+2j}, t^-_{i+1j}, t^\setminus_{ij})\tau(t'_{i+2j}, t^-_{i+1j}, \bar{t}^\setminus_{ij})} \right]^{-1} \end{aligned} \quad (8)$$

Corollary 2 (Constraint 2): If measure τ is separable, the joint probability distribution of $\{T, Q\}$ is left unchanged by simultaneous reversion of horizontal and vertical indices. In addition, if τ presents a horizontal symmetry, i.e.,

$$\tau(q, t', t^-, t^\backslash) = \tau(q, t^\backslash, t^-, t').$$

then $P(t, q)$ is left unchanged by independent reversion of horizontal or vertical indices.

Proof: If τ is separable, a new functional form of the joint probability distribution of $\{T, Q\}$ is straightforward to derive from (6). Under implicit free boundary assumption, it can be expressed as

$$P(t, q) = \prod_{j=1}^J \prod_{i=1}^I \frac{\tau(q_{ij}, t'_{ij}, t^-_{ij}, t^\backslash_{ij}) \tau(q_{ij}, t'_{i+1j-1}, t^-_{i+1j-1}, t^\backslash_{i+1j-1})}{\tau(q_{ij}) \tau(t'_{ij}) \tau(t^-_{ij}) \tau(t^\backslash_{ij})} \quad (12)$$

It is readily seen on (12) that the conclusions of Corollary 2 hold.

In order to facilitate further derivations, it is convenient to get an exact form of (12), where the free boundary assumption becomes explicit. First, let us define *left* and *right* cliques as follows:

$$\begin{aligned} C_{ij}^l &\triangleq \{Q_{ij}, T'_{i+1j-1}, T^-_{ij-1}, T^\backslash_{i-1j-1}\} \cap \{T, Q\}, \\ C_{ij}^r &\triangleq \{Q_{ij}, T'_{ij}, T^-_{ij}, T^\backslash_{ij}\} \cap \{T, Q\}. \end{aligned} \quad (13)$$

It is now easy to get an exact expression of (12) of the form of (14) (see (14) at the bottom of the page).

3) *Spatial Interactions (Constraint 3):* Since measure τ rules the interactions between neighboring *location* and *transition* variables, further constraints must be added on τ to make these interactions compatible with the physical properties of layered media. In this respect, we impose two basic rules of dependency:

$$\tau(t' = 0, t^- = 0, t^\backslash = 0 | q = 0) = 1, \quad (15a)$$

$$\tau(q = 1 | t' = 0, t^- = 0, t^\backslash = 0) = \epsilon \quad (\epsilon \ll 1). \quad (15b)$$

Equation (15a) precludes any isolated transition variable to be set: transition variables can be set only if they link two location variables which are set to one, according to Section II-C. Note that (15a) breaks the positivity condition, which does not generate any difficulty with the unilateral approach adopted here.

Equation (15b) is somewhat similar to (15a). In order to obtain a perfect symmetry between (15a) and (15b), ϵ would

have to be set to zero that would forbid discontinuities along layer boundaries. Actually, such a constraint would be too restrictive, since local breaks may appear in a layered structure (for instance, geological faults in seismic exploration). Moreover, it is generally desirable to keep a degree of freedom when modeling *real-world* structures, so as to deal with unaccounted phenomena. Therefore, ϵ is set to a value small with respect to 1.

F. Markov-Bernoulli Random Fields

In the previous section, constraints were imposed on measure τ so as to fulfill the prerequisites stated in Section II-A. The resulting model will be referred to as a Markov-Bernoulli random field (MBRF) due to characteristic property 1). In the course of imposing the constraints, physically meaningful parameters ($\lambda, \mu', \mu^-, \mu^\backslash, \epsilon$) were introduced naturally. We now show that any separable measure τ can be fully characterized by the set of parameters ($\mu', \mu^-, \mu^\backslash, \epsilon$).

Using Bayes' theorem and the separability property, τ can be expressed as

$$\tau(q, t', t^-, t^\backslash) = \tau(q | t', t^-, t^\backslash) \tau(t') \tau(t^-) \tau(t^\backslash). \quad (16)$$

The first term on the right side of (16) can be deduced from (15b) when $t' = t^- = t^\backslash = 0$, or from (15a), otherwise. It is a function of ϵ only. Corollary 1 shows that the three other terms are Bernoulli distributions with respective parameters $\mu', \mu^-,$ and μ^\backslash , hence the result. Then it follows from (12) that $(\mu', \mu^-, \mu^\backslash, \epsilon)$ is a complete parameterization of any MBRF.

Each parameter among $(\mu', \mu^-, \mu^\backslash, \epsilon)$ can be given an intuitive interpretation: $\mu', \mu^-,$ and μ^\backslash are the probabilities of occurrence of upward, horizontal, and downward transition variables; ϵ represents the probability of occurrence of discontinuities along boundaries between layers. The resulting probability of occurrence of boundary sites along a vertical direction is readily given by

$$\begin{aligned} \lambda &= \tau(q = 1) \\ &= 1 - (1 - \mu')(1 - \mu^-)(1 - \mu^\backslash)(1 - \epsilon). \end{aligned}$$

Fig. 5 shows a 51×65 realization of a MBRF, for $\mu' = \mu^\backslash = 0.02, \mu^- = 0.06,$ and $\epsilon = 0.001$. The resulting value of λ is 0.1. This realization was obtained by recursive pseudo-random sampling using the factored form of the joint probability distribution given by (6). For sake of clarity, only the location process is represented in Fig. 5. By construction, this field fulfills constraints 1)–3), and it appears to represent a horizontally layered structure adequately.

Now that MBRF's models have been completely specified as a subclass of MRF's with only four parameters, one might question whether more degrees of freedom should

$$P(t, q) = \frac{\prod_{C_{ij}^l \neq \emptyset} \tau(C_{ij}^l) \prod_{C_{ij}^r \neq \emptyset} \tau(C_{ij}^r)}{\prod_{(i,j) \in \Lambda^0} \tau(q_{ij}) \prod_{(i,j) \in \Lambda'} \tau(t'_{ij}) \prod_{(i,j) \in \Lambda^-} \tau(t^-_{ij}) \prod_{(i,j) \in \Lambda^\backslash} \tau(t^\backslash_{ij})} \quad (14)$$

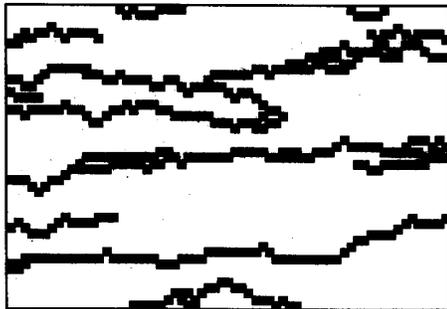


Fig. 5. 51×65 realization of a MBRF, obtained by recursive pseudo-random sampling from the parameters values $\mu^+ = \mu^- = 0.02$, $\mu^* = 0.06$, and $\epsilon = 0.001$. Only the location process is represented. Each column follows a Bernoulli distribution with parameter $\lambda = 1 - (1 - \mu^+)(1 - \mu^-)(1 - \mu^*)(1 - \epsilon) = 0.1$.

have been kept in the model so as to increase its flexibility. This question pertains to a general difficulty in Bayesian estimation, i.e., how to set the trade-off between complexity (and pertinence) of the prior model on one hand and theoretical behavior and practicality of the corresponding estimator on the other hand. In this study, our choices were clearly stated in Constraints 1)–4). However, several possibilities for relaxing these constraints and obtaining a more flexible model can be briefly mentioned. First of all, we would not recommend nonseparable measures τ , because valuable properties such as explicit statistics for columns and single variables would be lost. This seems a high price to pay for an increased flexibility which may be difficult to utilize in practice. As already mentioned, a more interesting possibility is to consider a second-order Markov chain instead of a first-order one in H1). Such an extension preserves the essential properties of MBRF's, including the invariance of the probability measure P on maximal cliques. To a certain extent, it is also possible to make P location dependent and yet controllable, by letting some parameters (e.g., λ) vary between columns, provided that basic properties 1) and 2) remain valid. For instance, any change within $(\mu^+, \mu^-, \mu^*, \epsilon)$ that keeps λ constant still meets 1) and 2); it is also possible to change the value of λ by letting ϵ vary while (μ^+, μ^-, μ^*) are kept constant.

G. Additional Properties of MBRF's

This section deals with some interesting additional characteristics of MBRF's. These properties naturally arise from the particular structure of these fields and were not stated as prerequisites. First, we address the extension of MBRF's to the infinite lattice Λ_∞ associated to $\Lambda^\circ = \mathbb{Z}^2$, which is a problem of theoretical as well as practical interest. The difficulties that one encounters in considering general MRF's on infinite lattices are well described by Derin and Kelly in [13]. Here, we prove in Section II-G-1 that MBRF's can be constructed on Λ_∞ . Then in Section II-G-2, the structure of random field Q is studied marginally from T . In particular, it is rigorously shown that Q is not a MRF (i.e., not with a local neighborhood system), which *a posteriori* justifies the introduction of transition process T .

1) *Extension to Infinite Lattices:* In this subsection, any random field X built on lattice Λ will be noted X_Λ in order to distinguish between random fields with identical structure though defined on different lattices. First, we introduce the definition of *mutual compatibility*, which generalizes the one given in [17].

Definition 1: Let Λ be a finite set, and A any subset of Λ . Let X_Λ and X_A be two random fields, respectively built on lattices Λ and A . X_Λ and X_A are *mutually compatible*, if and only if the restriction of X_Λ to A is identical to X_A , i.e., under a compact form where the measured field is implicit:

$$P_\Lambda(A) \equiv P_A(A). \quad (17)$$

Then, valid random fields on infinite lattices are defined using the following theorem.

Theorem 2: Let $X_{\Lambda_1}, \dots, X_{\Lambda_n}, \dots$ denote random fields defined on a growing series of finite lattices $\Lambda_1 \subset \dots \subset \Lambda_n \subset \dots$. Assume that the series $\Lambda_1, \dots, \Lambda_n, \dots$ converge to an infinite lattice Λ_∞ . If any pair $(X_{\Lambda_n}, X_{\Lambda_{n+1}})$ is mutually compatible, then $X_\infty \triangleq \lim_{n \rightarrow \infty} X_{\Lambda_n}$ defines a valid random field on infinite lattice Λ_∞ . In addition, the restriction of X_∞ to any Λ_n is X_{Λ_n} .

Proof: The assumptions of Theorem 2 allow direct application of the Kolmogorov extension theorem to countable sets [23]. The same argument has already been used by Pickard [16].

Mutual compatibility is not only a point of theoretical interest. In many applications, and particularly in image restoration, the observations represent a small part of a larger structure. Their size is only defined by the imaging process which is not intrinsic to the structure itself. Therefore, the statistical properties of the prior model should not depend upon the size of the observed data, and in particular, the mutual compatibility property should be fulfilled for lattices which represent the possibly observed data sets.

Unfortunately, mutual compatibility is not fulfilled by most MRF's. In general, when restricted to a subset, they lose their Markovian property because all sites located on the subset boundaries get connected in the resulting graph [13]. Pickard random fields (see Section II-B) are a notable exception to this rule, and they have been shown by Goutsias [17] to be the only MRF's built on rectangular lattices which fulfill the mutual compatibility property.

Focusing on MBRF's again, it is easy to check that random fields $\{T, Q\}_\Lambda$ built from rectangular lattices Λ° are not mutually compatible because of edge effects at horizontal boundaries of Λ . However, some other finite subsets of Λ_∞ allow the construction of MBRF's which are mutually compatible with any larger MBRF, so as direct application of Theorem 2 provides the extension to $\{T, Q\}_\infty$. Such subsets will be referred to as compatible subsets (CSS's). In the sequel, they are shown to be fully characterized through an equivalence theorem between compatibility and convexity, the latter notion being defined from the neighborhood system of MBRF's. But before stating the equivalence theorem, we give some preliminary definitions.

Definition 2: For any variable X on Λ_∞ , sets of *left neighbors* $N^l(X)$ and *right neighbors* $N^r(X)$ are defined as follows:

$$\begin{aligned} N^l(Q_{ij}) &\triangleq \{T_{i+1j-1}^l, T_{ij-1}^-, T_{i-1j-1}^\backslash\}, \\ N^l(T_{ij}^l) &\triangleq N^l(T_{ij}^-) \triangleq N^l(T_{ij}^\backslash) \triangleq \{Q_{ij}\}, \\ N^r(Q_{ij}) &\triangleq \{T_{ij}^l, T_{ij}^-, T_{ij}^\backslash\}, \\ N^r(T_{i+1j-1}^l) &\triangleq N^r(T_{ij-1}^-) \triangleq N^r(T_{i-1j-1}^\backslash) \triangleq \{Q_{ij}\}. \end{aligned} \quad (18)$$

Note the obvious mutual implication for any pair of variables X, Y on Λ_∞ :

$$X \in N^l(Y) \Leftrightarrow Y \in N^r(X).$$

Definition 3: An *admissible path* $L(X, Y)$ of length N is a finite series of variables (X_1, \dots, X_N) on Λ_∞ such as $X_1 = X, X_N = Y$, and $X_{n+1} \in N^r(X_n)$ for any $n = 1, \dots, N-1$.

Simple examples based upon specific choices of X and Y show that neither the existence nor the uniqueness of $L(X, Y)$ are guaranteed. Nonetheless, the length N of the path characterizes the relative position of X and Y , in the sense that all admissible paths between X and Y have the same length.

Definition 4: A finite subset A is convex if and only if, for any pair of variables X, Y on A , any admissible path $L(X, Y) = (X_1, \dots, X_N)$ fully belongs to A , i.e., $X_n \in A$ for $n = 1, \dots, N$.

Theorem 3: A finite subset A is a CSS, if and only if it is convex.

Proof: See Appendix C.

In fact, we conjecture that Theorem 3 and Appendix C are generalized to other unilateral Markov fields. The case of PRF's is of special interest; partial characterization of CSS's can be found in [17] under the name of *primary sublattices*. Here, we conjecture the complete characterization for PRF's as follows.

Conjecture: A finite subset A is a CSS for PRF's, if and only if it is both connex and convex; connexity refers to the first-order neighborhood system of PRF's; convexity is given by Definition 4, admissible paths being horizontal paths and vertical paths. Of course, any finite rectangular set is both connex and convex, as well as the primary sublattices defined in [17].

Extension of MBRF's to the infinite lattice Λ_∞ is now straightforward since Theorem 2 can be applied to any growing series of convex lattices $\Lambda_1, \dots, \Lambda_n, \dots$ provided that the series converge to Λ_∞ . For instance, all triangular subsets A

$$A = \{(i, j): 0 \leq j_A - j \leq J_A \text{ and } |i_A - i| \leq j_A - j\}$$

are convex and the family extends to Λ_∞ for any i_A when j_A and J_A take infinite values, which proves the existence of $\{T, Q\}_\infty$. Restriction of $\{T, Q\}_\infty$ to CSS's remains in the class of MBRF's, but this does not hold for rectangular lattices.

2) *Random Fields Q and T:* In this subsection, we first study the structure of random field Q marginally from T . The question is to determine whether the probability distribution of Q presents a factored expression similar to the one derived for $P(t, q)$, in which case introduction of the transition random field T would have been superfluous. Here, we show that Q can be viewed as a vector Markov chain, but is not a MRF with a local neighborhood system. Then the properties of T marginally from Q are investigated, and the transition field T is shown to be a MRF with a local neighborhood system and a unilateral structure.

a) *Random field Q:* First, we show that the first-order vector Markov chain structure of $\{T, Q\}$ also holds for Q alone. Expression (2) can be rewritten as

$$P(t, q) = P(q_1) \prod_{j=1}^{J-1} P(q_{j+1}, t_j | q_j)$$

and summing each conditional term on the right side of (28) with respect to t_j immediately yields

$$P(q) = P(q_1) \prod_{j=1}^{J-1} P(q_{j+1} | q_j), \quad (19)$$

which proves the result.

We now turn to the characterization of the neighborhood system of random field Q . For any site (i, j) of Λ° , define $Q^{ij} = Q \setminus \{Q_{ij}\}$ and $q^{ij} = q \setminus \{q_{ij}\}$ from any possible realization q of Q . The point is to determine which variables the functional form $P(q_{ij} | q^{ij})$ actually depends upon. Using the vector Markov chain structure (19), one can immediately establish that

$$P(q_{ij} | q^{ij}) = \left[1 + \frac{P(q_{j+1} | \bar{q}_j) P(\bar{q}_j | q_{j-1})}{P(q_{j+1} | q_j) P(q_j | q_{j-1})} \right]^{-1}, \quad (20)$$

where \bar{q}_j differs from q_j only at location (i, j) by $\bar{q}_{ij} = 1 - q_{ij}$. Equation (20) shows that $P(q_{ij} | q^{ij})$ is a function of the transition probabilities $P(q_{j+1} | q_j)$ only and, therefore, the neighborhood system of Q is local only if these transition probabilities can be factored into locally dependent terms. In general, it is not possible to derive such a factorization and an example in which the set of neighbors of site (i, j) can take an arbitrarily large size is presented in Appendix D. Thus, Q is not a MRF with a local neighborhood system and introduction of transition field T does provide considerable simplification, as conjectured in Section II-C.

b) *Random field T:* The UMRF property is straightforward to establish for T . First define σ as

$$\begin{aligned} \sigma(t_1^l, t_1^-, t_1^\backslash, t_2^l, t_2^-, t_2^\backslash) \\ \triangleq \sum_q \frac{\tau(q, t_1^l, t_1^-, t_1^\backslash) \tau(q, t_2^l, t_2^-, t_2^\backslash)}{\tau(q)}. \end{aligned}$$

Then, since each factor on the right side of (12) is a function of only one location variable q_{ij} , summation of $P(t, q)$

over q immediately yields the expression of the probability distribution of T in the following factored form:

$$P(\mathbf{t}) = \prod_{j=1}^J \prod_{i=1}^I \frac{\sigma(t_{ij}^+, t_{ij}^-, t_{ij}^{\setminus}, t_{i+1j-1}^-, t_{ij-1}^-, t_{i-1j-1}^{\setminus})}{\tau(t_{ij}^+) \tau(t_{ij}^-) \tau(t_{ij}^{\setminus})} \quad (21)$$

The MRF characteristic property (1), the neighborhood system and the unilateral structure of random field T can be immediately deduced from (21). Therefore, T retains every basic characteristics of the MBRF $\{T, Q\}$.

These results underline the hierarchical structure of compound random field $\{T, Q\}$: while the distribution of Q is naturally defined jointly with (or conditionally to) T , T can easily be handled apart from Q . The fruitfulness of hierarchical constructions has already been stressed by Geman and Geman (see [2, pt. IV]). The MBRF-based method for multichannel deconvolution presented in the next section further illustrates the interest of the approach.

III. APPLICATION TO 2-D SEISMIC DECONVOLUTION

In this section, application to Bayesian multichannel seismic deconvolution is considered. The stratified structure of the unknown data is modeled *a priori* as a MBRF. This representation is completed so as to incorporate amplitude information, and the resulting model is used as priors in a *maximum a posteriori* (MAP) deconvolution procedure. A suboptimal recursive detection-estimation algorithm is derived, and simulation results are presented in order to illustrate the performances and the practicality of the approach.

A. Seismic Exploration

In the field of seismic exploration, as well as in other areas of applied physics such as nondestructive evaluation or biomedical ultrasonic imaging, one of the major issues is to estimate the shape and nature of homogeneous layers in a stratified structure from acoustic measurements performed at different points of its surface (see Fig. 6).

Here, the original three-dimensional problem is approximated by a 2-D one, and it is assumed without loss of generality that the general orientation of the layers is horizontal. Therefore, the 2-D field of interest represents a vertical section of the unknown propagation medium.

In the application domains considered here, the boundary where the measurements are performed is parallel to the general orientation of the layers, and under reasonable assumptions, the observations can be considered as the noise-corrupted convolution product of two signals: the wavelet and the reflectivity. The wavelet is a 1-D vertical signal that represents the incident waveshape that scatters through the propagation medium. The reflectivity characterizes the propagation medium; it is a 2-D field defined as the vertical logarithm derivative of the acoustic impedance. Since the wavelet is a 1-D vertical signal, its convolution product with the 2-D reflectivity can be decomposed into independent 1-D convolution products, so that each column z_j of the observed data is a function of the corresponding vertical reflectivity

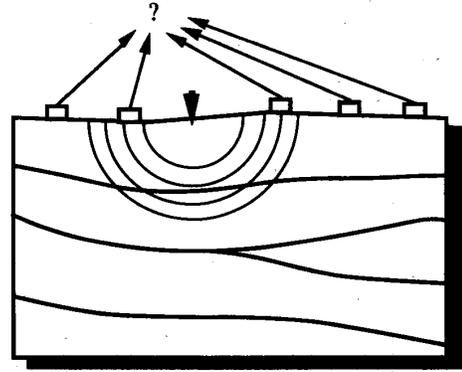


Fig. 6. Seismic exploration: The problem is to characterize a layered structure from acoustic measurements performed at its surface.

sequence \mathbf{r}_j as

$$z_j = \mathbf{h} * \mathbf{r}_j + \mathbf{n}_j,$$

where \mathbf{h} and \mathbf{n}_j denote the known time-invariant wavelet and the white Gaussian observation noise respectively. In addition, under the assumption that the layers are homogeneous, the incident wave only reflects at the boundaries between layers and 1-D vertical reflectivity sequences \mathbf{r}_j appear as sparse spike trains.

B. Main Characteristics and Limitations of the B-G Approach

A common approach to the estimation of 2-D reflectivity sections is based upon the previous remark and only involves 1-D processing: the columns of the reflectivity are estimated independently from one another, via 1-D deconvolution of the corresponding columns of the observed data. A Bayesian approach to such *single-channel* deconvolution has been successfully pioneered by Kormylo and Mendel [7], [8], followed by others [9]–[11], [24]. It is based on Bernoulli–Gaussian (B-G) prior modeling of the reflectivity pulse trains: 1-D Bernoulli sequences Q_j locate the spikes (i.e., reflectors, which correspond to layer boundary sites), while the amplitudes R_j of the spikes are assumed Gaussian. The B-G model of each reflectivity column $\mathbf{X}_j = (Q_j, R_j)$ can be expressed as:

Q_j and R_j both contain independent random variables, respectively Q_{ij} and R_{ij} ($1 \leq i \leq I$)

$$Q_{ij}: \text{Bernoulli random variable } \begin{cases} P(Q_{ij} = 1) = \lambda \\ P(Q_{ij} = 0) = 1 - \lambda \end{cases}$$

$P(R_{ij} | Q_{ij} = q)$: Zero-mean Gaussian random variable with variance $q\sigma^2$.

The parameters (λ, σ^2) which characterize the B-G sequence, as well as the variance of the observation noise, are assumed to be known. Restoration of \mathbf{x}_j can be performed through maximization of the *a posteriori* likelihood $P(\mathbf{x}_j | z_j)$ [7],

[11]. However, it has been empirically shown [8], [11] that the following detection-estimation procedure is preferable:

(1) Detection of the reflectors:

$$\hat{q}_j = \operatorname{argmax}_{\mathbf{q}_j} P(\mathbf{q}_j | \mathbf{z}_j), \quad (22a)$$

(2) Amplitude estimation:

$$\hat{\mathbf{r}}_j = \operatorname{argmax}_{\mathbf{r}_j} P(\mathbf{r}_j | \mathbf{z}_j, \hat{\mathbf{q}}_j). \quad (22b)$$

$P(\mathbf{q}_j | \mathbf{z}_j)$ can be evaluated easily, but its maximization (step 1, (22a)) is a combinatorial exploration problem which can hardly be solved exactly in real-world conditions. However, several efficient algorithms have been proposed to perform the detection step in a suboptimal or near-optimal manner [7]–[11]. Step 2 (22b) is simpler as $P(\mathbf{r}_j | \mathbf{q}_j)$ is Gaussian, and as all phenomena are linear; thus, \mathbf{r}_j is obtained in closed-form and its computation can easily be carried out using standard—linear—estimation algorithms such as Kalman filters [8], [11]. However, the B–G approach is suboptimal in a 2-D setting, since the lateral continuity of the layered reflectivity section is not accounted for.

C. Markov–B–G Model

To make up for the inherent deficiency of single-channel approaches such as B–G deconvolution, we propose a *multichannel* approach based on MBRF prior modeling of the geometric characteristics of the unknown reflectivity section. In order to define the *a priori* model completely, the distribution of the amplitude of the spikes (which are located on the boundary sites) must be specified. This should be done in accordance with the specific characteristics of the problem at hand. Here, it is desirable that 1) for physical reasons, the amplitudes of spikes located on the same boundary be strongly correlated; 2) the 2-D model be a true extension of the B–G representation which showed its efficiency in a 1-D setting; 3) the complete model $\{T, Q, R\}$ still be a UMRF so as to fulfill constraint 4) defined in Section II. Therefore, the whole 2-D prior model is hierarchically defined as follows:

$$P(\mathbf{r} | \mathbf{t}, \mathbf{q}) = P(\mathbf{r}_1 | \mathbf{q}_1) \prod_{j=2}^J P(\mathbf{r}_j | \mathbf{q}_j, \mathbf{t}_{j-1}, \mathbf{q}_{j-1}, \mathbf{r}_{j-1}), \quad (23)$$

where every conditional probability on the right side factors into a product of I locally dependent terms

$$P(\mathbf{r}_j | \mathbf{q}_j, \mathbf{t}_{j-1}, \mathbf{r}_{j-1}) = \prod_{i=1}^I P(r_{ij} | q_{ij}, t_{i+1j-1}^+, t_{ij-1}^-, t_{i-1j-1}^-, r_{i+1j-1}, r_{ij-1}, r_{i-1j-1}) \quad (24)$$

with implicit modifications for $j = 1, i = 1$, and $i = I$ in accordance with the free boundary assumption. Now the conditional probabilities appearing on the right side of (24) are defined as follows.

- 1) If $q_{ij} = 0$, then $r_{ij} = 0$.
- 2) If one (and only one) transition variable is set among t_{i+1j-1}^+ , t_{ij-1}^- , and t_{i-1j-1}^- , then the reflector r_{ij} is said to have a unique predecessor $r_{i+di,j-1}$ ($-1 \leq di \leq 1$),

in which case r_{ij} is sampled from a first-order AR process: $r_{ij} = ar_{i+di,j-1} + n_r \cdot n_r$ is Gaussian with zero-mean and variance $(1 - a^2)\sigma^2$ and a is adjustable between 0 and 1 to control the degree of correlation between spikes along the same boundary.

- 3) If r_{ij} has either more than one predecessor or no predecessor at all, it is sampled from the Gaussian distribution with zero-mean and variance σ^2 .

It is very easy to check that such specifications yield B–G distributed reflectivity columns, while also providing suited lateral correlation. Using (23) and (24) in conjunction with (2) and (3), the joint probability $P(\mathbf{t}, \mathbf{q}, \mathbf{r}) = P(\mathbf{r} | \mathbf{t}, \mathbf{q})P(\mathbf{t}, \mathbf{q})$ can be fully factored. Therefore, the resulting compound random field $\mathbf{X} = \{T, Q, R\}$ is still a UMRF. We call it a *Markov–Bernoulli–Gaussian* (M–B–G) field.

D. Bayesian Multichannel Deconvolution using M–B–G Priors

We now use the M–B–G model defined above as priors for MAP estimation of a 2-D reflectivity section. Let $\mathbf{X} = \{T, Q, R\}$ be a M–B–G field. Since \mathbf{X} can still be interpreted as a vector Markov chain, the most efficient way to implement an optimal MAP estimator is to use a Viterbi algorithm [25]. Unfortunately such a procedure is intractable, due to the huge size of the corresponding state space. Instead, we propose a scheme similar to the one used for single-channel deconvolution: $\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_J$ are computed recursively in a two-step suboptimal MAP procedure:

1) Detection step:

$$(\hat{\mathbf{t}}_{j-1}, \hat{\mathbf{q}}_j) = \operatorname{argmax}_{\mathbf{t}_{j-1}, \mathbf{q}_j} P(\mathbf{t}_{j-1}, \mathbf{q}_j | \mathbf{z}_j, \hat{\mathbf{q}}_{j-1}, \hat{\mathbf{r}}_{j-1}), \quad (25a)$$

2) Estimation step:

$$\hat{\mathbf{r}}_j = \operatorname{argmax}_{\mathbf{r}_j} P(\mathbf{r}_j | \mathbf{z}_j, \hat{\mathbf{t}}_{j-1}, \hat{\mathbf{q}}_j, \hat{\mathbf{r}}_{j-1}), \quad (25b)$$

for $j = 2, \dots, J$.

The first step (25a) locates the reflectors in column $\hat{\mathbf{x}}_j$. To do so, estimation of the previous column $\hat{\mathbf{x}}_{j-1} = (\hat{\mathbf{q}}_{j-1}, \hat{\mathbf{r}}_{j-1})$ is taken into account, via the introduction of the transition column $\hat{\mathbf{t}}_{j-1}$. Because of the unilateral structure of the M–B–G model, the probability to be maximized in (25a) is factorable into locally dependent terms. This detection step remains a combinatorial problem which has about the same level of complexity as (22a), and similar deterministic search algorithms can be used to compute a suboptimal solution [27].

The estimation step (25b) is also very similar to the B–G case (22b), since $P(\mathbf{r}_j | \hat{\mathbf{t}}_{j-1}, \mathbf{q}_j, \mathbf{r}_{j-1})$ is Gaussian, and since all phenomena are linear. Therefore, estimation of \mathbf{r}_j can be carried out optimally using standard least-square algorithms or Kalman filters.

Implementation of (25) is very similar to (22), and it requires neither more memory nor processing time; the algorithm itself is given in [27]. Note that the procedure is not self-starting: evaluation of $\hat{\mathbf{x}}_1$ must be carried out using a B–G deconvolution algorithm. Finally, it should be mentioned that the B–G approach of Section III-B appears as a degenerate

case of M-B-G deconvolution, for which lateral correlation is ignored by choosing an underlying MBRF with parameters

$$\begin{cases} \mu^l = \mu^- \mu^h = 0 \\ \epsilon = \lambda \end{cases}$$

E. Simulation Results

Many tests have already been carried out in the field of seismic exploration, but also in nondestructive testing and in medical ultrasonic imaging. The simulation results of Fig. 7 provide a comparison between the B-G method described in Section III-B and the M-B-G approach proposed here.

Fig. 7(a) shows the synthetic reflectivity section, sampled on 65 columns of 51 points. Conventionally, reflectivity pulses are represented as triangles of variable height, and positive areas of signals are shaded in black so as to enhance the figure. The wavelet depicted in Fig. 7(b) is sampled on 17 points. It presents all the characteristics of actual seismic wavelets: nonminimum phase and poor spectral content, which makes the solution of the seismic deconvolution problem quite difficult. In Fig. 7(c), the convolution product has been computed as explained in Section III-A, and white observation noise has been added. The signal-to-noise ratio (SNR), defined as the ratio of the mean power of the noiseless observations to the noise variance, was set to 10 dB.

The results obtained with the B-G deconvolution method are shown in Fig. 7(d). The parameter values were chosen empirically as follows:

$$\begin{cases} \text{SNR} = 10 \text{ dB} \\ \lambda = 0.1 \end{cases}$$

The reflectivity section is roughly restored, but many details are lost: thin layers or weak reflectors are hardly detected; some columns are strongly distorted by artefacts such as false alarms or doubled detection.

Fig. 7(e) depicts the M-B-G estimate, obtained for the following values of parameters, which were chosen empirically:

$$\begin{cases} \text{SNR} = 6 \text{ dB} \\ \mu^l = \mu^h = 0.029, \mu^- = 0.045, \epsilon = 0 \Rightarrow \lambda = 0.1 \\ a = 0.999 \end{cases} \quad (26)$$

Compared to Fig. 7(d), the accuracy of the estimates is greatly improved. Differences between Fig. 7(e) and the original reflectivity section (Fig. 7(a)) are actually negligible. Even small details, such as the double junctions in columns 24 and 55, are restored properly. This result illustrates the superiority of the multichannel M-B-G approach which accounts for lateral correlation, over the usual single-channel B-G approach in which 1-D columns are processed independently from one another. It should be stressed that the computational cost of both methods is about the same.

Finally, M-B-G restorations of the same reflectivity section (Fig. 7(a)) for observation noise levels corresponding to SNR values of 7, 5, and 2 dB are shown in Fig. 8(a)-(c), respectively. The values of the parameters were the same as in (26), except that the SNR is decreased accordingly (respectively, 3, 1, and -2 dB). Compared to Fig. 7(e), the quality of the results

is only slightly altered in Fig. 8(a) and (b), whereas it becomes very poor in Fig. 8(c). In our opinion, this sharp degradation of the performance of the method should be attributed to the implementation of the MAP estimator rather than to the prior model itself. As mentioned in Section III-D, the algorithm proposed for M-B-G deconvolution is suboptimal. More precisely, it performs recursive estimation of reflectivity columns, and decisions are taken without delay nor possibility of reversion. In difficult conditions such as a SNR of 2 dB, the drawback of such a recursive *decision-directed* procedure is that a single error can jeopardize the whole estimation by being repeated and amplified from the erroneous column to the next. In this respect, global (i.e., nonrecursive) optimization algorithms are more robust and yield further improvements, but at the expense of very high computational costs [28]. For more precise indications about the performances of M-B-G deconvolution, the reader is referred to [27] and [28], where comprehensive simulation and real data processing results are presented.

IV. CONCLUSION

In this paper, we addressed the problem of modeling 2-D layered media. Restoration of such structures is a general issue in many application fields such as seismic exploration, nondestructive evaluation or biomedical ultrasonic imaging. In these domains, standard approaches are usually one-dimensional, which means that they cannot take lateral continuity of stratified media into account. The object of this paper was to design an adequate and theoretically sound statistical model of 2-D layered structures and to use it as priors for *maximum a posteriori* estimation, especially in the case of seismic exploration.

To do so, we introduced and studied Markov-Bernoulli random fields (MBRF's). This original subclass of MRF's combines a compound structure including a line process [2] and a unilateral approach which is reminiscent of the work of Pickard on curious lattice processes [16]. A theoretical study was conducted in order to investigate specific properties of MBRF's. The resulting model provides a very general framework to Bayesian restoration of 2-D stratified media.

Then application to Bayesian multichannel seismic deconvolution was considered. We proposed a true extension of Bernoulli-Gaussian single-channel deconvolution [8], based on MBRF prior modeling of the 2-D layered media to be estimated. The practicability and the efficiency of the new Markov-Bernoulli-Gaussian approach were illustrated by a synthetic seismic example.

Future extensions include the derivation of other types of algorithms for Bayesian estimation based on the same MBRF model. Finally, the problem of estimating the parameters of the model should also be addressed.

APPENDIX A

MRF CHARACTERISTIC PROPERTY OF $\{T, Q\}$

Let us first consider the case of lattice Λ° . For any interior site (i, j) of Λ° , define $Q^{ij} = Q \setminus \{Q_{ij}\}$ and $q^{ij} = q \setminus \{q_{ij}\}$

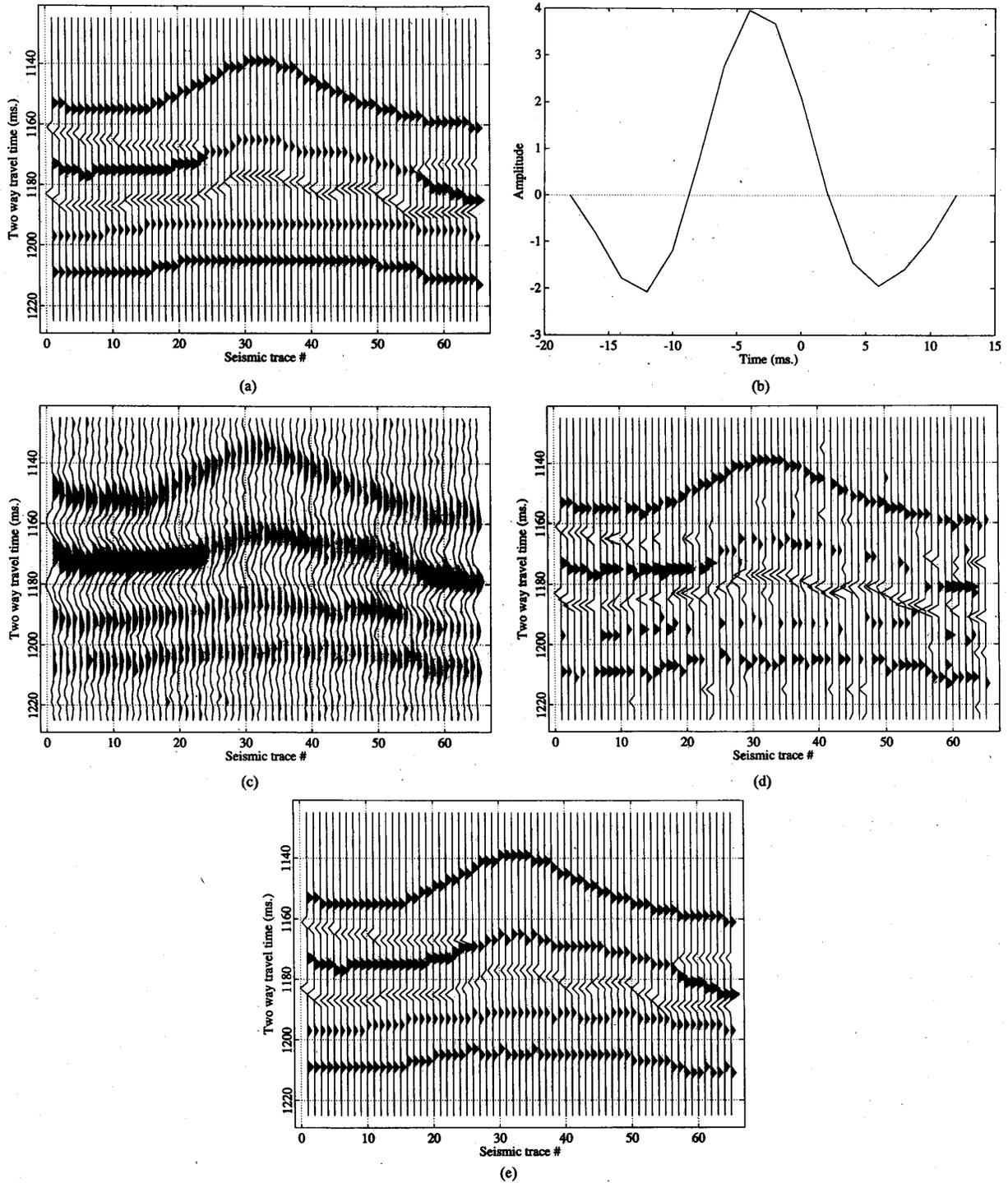


Fig. 7. Comparison of single-channel and multichannel deconvolution. (a) Synthetic reflectivity used for the simulations. (b) Seismic wavelet (i.e., vertical blur function). (c) Observation sequences obtained by convolution between the reflectivity columns (a) and the wavelet (b), and addition of white noise (SNR=10 dB). (d) Single-channel B-G deconvolution using the algorithm presented in Section III-B. (e) Multichannel M-B-G deconvolution based on MBRF modeling (Section III-D).

for any possible realization q of Q . In order to characterize the set of neighbors N_{ij}^c of site (i, j) and the expression

of $P(q_{ij} | N_{ij}^c)$, one must determine which variables the functional $P(q_{ij} | t, q^{ij})$ depends upon. Application of Bayes'

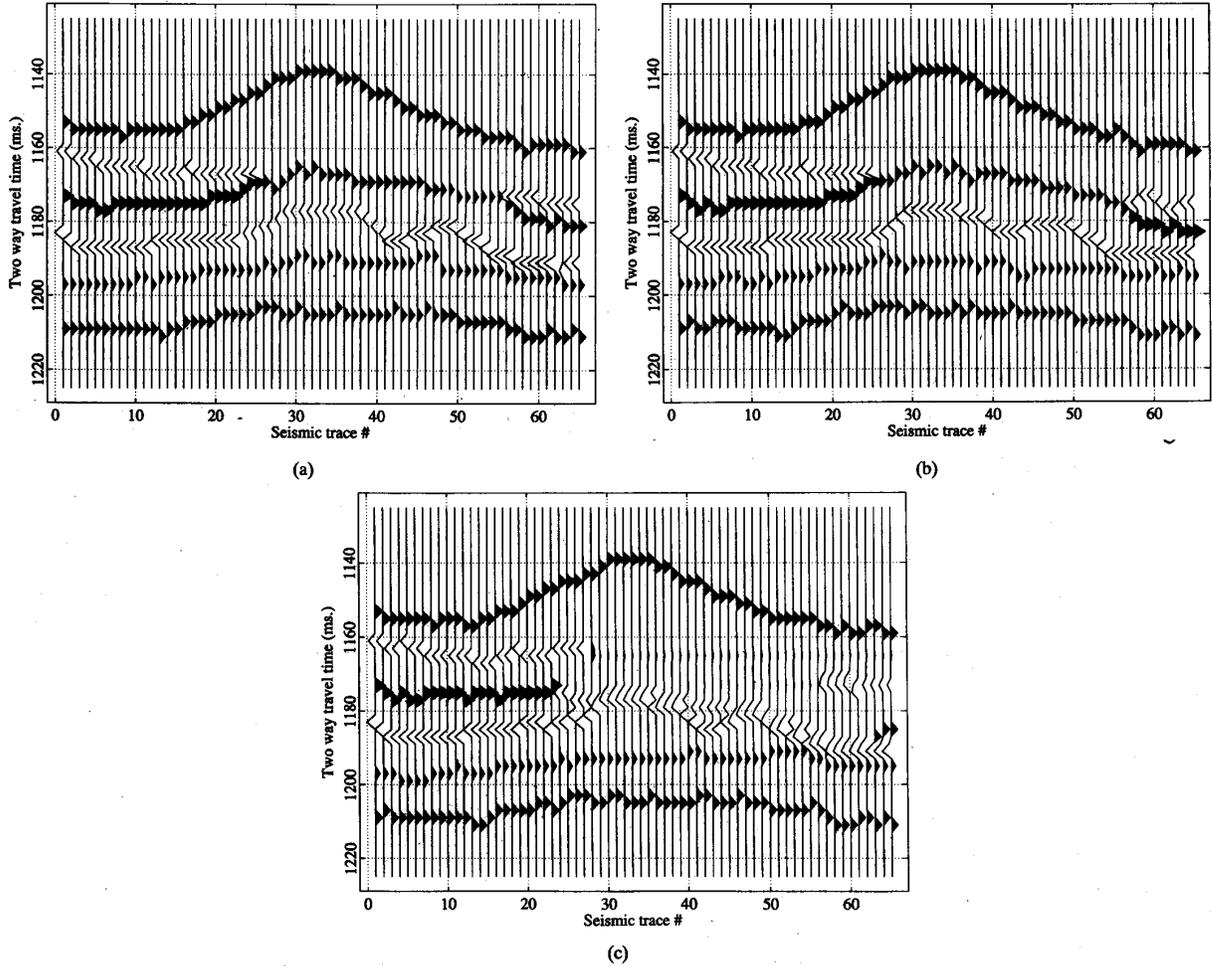


Fig. 8. Effect of the observation noise level. The synthetic data were obtained as in Fig. 7, except that the observation noise level was raised so as to obtain SNRs of 7 dB, 5 dB, and 2 dB. The corresponding multichannel deconvolution results are shown in (a), (b), and (c), respectively. One observes a sharp degradation of the performance for SNR=2 dB, which may be attributed to the suboptimal recursive implementation of the MAP estimator.

rule to $P(t, q)$ yields

$$P(q_{ij} | t, q^{ij}) = \frac{P(t, q)}{P(t, q^{ij})} = \left[1 + \frac{P(t, \bar{q})}{P(t, q)} \right]^{-1}, \quad (A.1)$$

where \bar{q} differs from q only at location (i, j) by $\bar{q}_{ij} = 1 - q_{ij}$. Quantity $P(t, \bar{q})/P(t, q)$ which appears in (A.1) can be developed using (6b). Almost all factors cancel out between numerator and denominator, and the resulting expression only contains terms which explicitly depend upon q_{ij} . We have

$$\frac{P(t, \bar{q})}{P(t, q)} = \frac{\tau(t'_{ij}, t^-_{ij}, t^{\setminus}_{ij} | \bar{q}_{ij}) \tau(\bar{q}_{ij} | t'_{i+1j-1}, t^-_{i-1j-1}, t^{\setminus}_{i-1j-1})}{\tau(t'_{ij}, t^-_{ij}, t^{\setminus}_{ij} | q_{ij}) \tau(q_{ij} | t'_{i+1j-1}, t^-_{i-1j-1}, t^{\setminus}_{i-1j-1})}$$

which immediately yields the form of N°_{ij} given in (7) and the expression of $P(q_{ij} | N^{\circ}_{ij})$ given in (8). The case of interior sites of Λ' , Λ^- , and Λ^{\setminus} can be treated in a similar manner. The case of sites at the boundaries of Λ does not present

further difficulties, except that one is required to consider the fully written form of $P(t, q)$ given by (6a) instead of (6b). By doing so, one shows that the sets of neighbors of such sites are simply those given in (7), once the undefined sites have been removed.

APPENDIX B
PROOF OF THEOREM 1 AND COROLLARY 1

Let us first define the following properties.

- P0) The restriction of P to maximal cliques of UMRF $\{T, Q\}$ is invariant and equal to τ .
- P1) τ fulfills the following property: for any triplet of 0-1 binary variables (t', t^-, t^{\setminus}) , $\tau(t', t^-, t^{\setminus}) = \tau(t')\tau(t^-)\tau(t^{\setminus})$.
- P2(j)) $P(q_j) = \prod_{i=1}^I \tau(q_{ij})$.
- P3(j)) $P(q_j, t_j) = \tau(q_{1j}, t^-_{1j}, t^{\setminus}_{1j}) \tau(q_{Ij}, t^{\setminus}_{Ij}, t^-_{Ij}) \prod_{i=2}^{I-1} \tau(q_{ij}, t^{\setminus}_{ij}, t^-_{ij})$
- P4(j)) $P(t_j) = \tau(t^-_{1j}, t^{\setminus}_{1j}) \tau(t^{\setminus}_{Ij}, t^-_{Ij}) \prod_{i=2}^{I-1} \tau(t^{\setminus}_{ij}, t^-_{ij}, t^{\setminus}_{ij})$.

$$\begin{aligned}
P5(j)) P(t_j) &= \prod_{i=2}^I \tau(t_{ij}^{\prime}) \prod_{i=1}^I \tau(t_{ij}^{\bar{}}) \prod_{i=1}^{I-1} \tau(t_{ij}^{\setminus}). \\
P6(j)) P(q_{i+1}, t_j) &= \tau(q_{1j+1}, t_{2j}^{\prime}, t_{1j}^{\bar{}}) \tau(q_{Ij+1}, t_{Ij}^{\prime}, t_{I-1j}^{\setminus}) \\
&\prod_{i=2}^{I-1} \tau(q_{ij+1}, t_{i+1j}^{\prime}, t_{ij}^{\bar{}}) \tau(t_{i-1j}^{\setminus}).
\end{aligned}$$

In order to establish Theorem 1, i.e., the equivalence between P0 and P1, we prove the following set of implications:

$$\begin{aligned}
& \left. \begin{array}{l} P2(j) \Rightarrow P3(j) \Rightarrow P4(j) \\ P1 \end{array} \right\} \\
& \Rightarrow \left. \begin{array}{l} P5(j) \\ P1 \end{array} \right\} \Rightarrow P6(j) \Rightarrow P2(j+1). \quad (B.1)
\end{aligned}$$

- P2(j) \Rightarrow P3(j): We first use Bayes' rule to decompose $P(q_j, t_j)$ as $P(q_j, t_j) = P(t_j | q_j)P(q_j)$. Then, using (3a), (5a), and P2(j), we obtain

$$\begin{aligned}
P(q_j, t_j) &= \tau(t_{1j}^{\bar{}}, t_{1j}^{\setminus} | q_{1j}) \tau(t_{Ij}^{\prime}, t_{Ij}^{\bar{}} | q_{Ij}) \\
&\cdot \prod_{i=2}^{I-1} \tau(t_{ij}^{\prime}, t_{ij}^{\bar{}}, t_{ij}^{\setminus} | q_{ij}) \prod_{i=1}^I \tau(q_{ij}) \\
&= \tau(q_{1j}, t_{1j}^{\bar{}}, t_{1j}^{\setminus}) \tau(q_{Ij}, t_{Ij}^{\prime}, t_{Ij}^{\bar{}}) \\
&\cdot \prod_{i=2}^{I-1} \tau(q_{ij}, t_{ij}^{\prime}, t_{ij}^{\bar{}}, t_{ij}^{\setminus}).
\end{aligned}$$

- P3(j) \Rightarrow P4(j): Obvious by summation of $P(q_j, t_j)$ over q_j .
- P1 and P4(j) \Rightarrow P5(j): Immediate by substitution of the factored form of τ into the expression of $P(t_j)$.
- P1 and P5(j) \Rightarrow P6(j): First, $P(q_{j+1}, t_j)$ is decomposed using Bayes' rule

$$P(q_{j+1}, t_j) = P(q_{j+1} | t_j)P(t_j). \quad (B.2)$$

Equations (3b) and (5b) allow $P(q_{j+1} | t_j)$ to be expressed as

$$\begin{aligned}
P(q_{j+1} | t_j) &= \tau(q_{1j+1} | t_{2j}^{\prime}, t_{1j}^{\bar{}}) \tau(q_{Ij+1} | t_{Ij}^{\prime}, t_{I-1j}^{\setminus}) \\
&\cdot \prod_{i=2}^{I-1} \tau(q_{ij+1} | t_{i+1j}^{\prime}, t_{ij}^{\bar{}}, t_{i-1j}^{\setminus}) \quad (B.3)
\end{aligned}$$

and the separability property P1 can be used backwards to rewrite P5(j) as

$$P(t_j) = \tau(t_{2j}^{\prime}, t_{1j}^{\bar{}}) \tau(t_{Ij}^{\prime}, t_{I-1j}^{\setminus}) \prod_{i=2}^{I-1} \tau(t_{i+1j}^{\prime}, t_{ij}^{\bar{}}, t_{i-1j}^{\setminus}). \quad (B.4)$$

Substitution of (B.3) and (B.4) into (B.2) yields the result.

- P6(j) \Rightarrow P2(j+1): Immediate by summation of $P(q_{j+1}, t_j)$ over t_j .

Proof of Theorem 1 (Equivalence between P0 and P1):

- Necessary condition P0 \Rightarrow P1: Observe that (4) and (5c) imply that P2(1) holds. Therefore, from (B.1), P4(1) also holds. We now evaluate $P(t_{i+1,1}^{\prime}, t_{i1}^{\bar{}}, t_{i-1,1}^{\setminus})$, $2 \leq i \leq I-1$ in two different ways.

On one hand, for any i such as $2 \leq i \leq I-1$, summation of P4(1) over all possible values of $t_1 \setminus \{t_{i+1,1}^{\prime}, t_{i1}^{\bar{}}, t_{i-1,1}^{\setminus}\}$ yields

$$P(t_{i+1,1}^{\prime}, t_{i1}^{\bar{}}, t_{i-1,1}^{\setminus}) = \tau(t_{i+1,1}^{\prime}) \tau(t_{i1}^{\bar{}}) \tau(t_{i-1,1}^{\setminus}). \quad (B.5)$$

On the other hand, since random variables $\{T_{i+1,1}^{\prime}, T_{i1}^{\bar{}}, T_{i-1,1}^{\setminus}\}$ are placed on a same clique, P0 implies that

$$P(t_{i+1,1}^{\prime}, t_{i1}^{\bar{}}, t_{i-1,1}^{\setminus}) = \tau(t_{i+1,1}^{\prime}, t_{i1}^{\bar{}}, t_{i-1,1}^{\setminus}). \quad (B.6)$$

Since $(t_{i+1,1}^{\prime}, t_{i1}^{\bar{}}, t_{i-1,1}^{\setminus})$ denotes any binary triplet, (B.5) and (B.6) show that P1 is actually a necessary condition for the fulfillment of P0.

- Sufficient condition P1 \Rightarrow P0: Since P2(1) and P1 are fulfilled, then, by induction on (B.1), P2(j)—P6(j) hold for any value of j . P0 is immediately obtained by summation of the proper variables in P3(j) and P6(j). \square

Proof of Corollary 1: If τ is separable, then P2(j) and P5(j) hold for any value of j . Summation of the proper variables in P2(j) and P5(j) yields

$$\begin{aligned}
\forall (i, j) \in \Lambda^{\circ}, P(q_{ij}) &= \tau(q_{ij}), \\
\forall (i, j) \in \Lambda^{\prime}, P(t_{ij}^{\prime}) &= \tau(t_{ij}^{\prime}), \\
\forall (i, j) \in \Lambda^{\bar{}}, P(t_{ij}^{\bar{}}) &= \tau(t_{ij}^{\bar{}}), \\
\forall (i, j) \in \Lambda^{\setminus}, P(t_{ij}^{\setminus}) &= \tau(t_{ij}^{\setminus}).
\end{aligned}$$

Therefore, the distribution of each type of location or transition variable is invariant. Since these variables are binary, they necessarily follow Bernoulli distributions whose characteristic parameters are defined by (10) and (11).

Moreover, P2(j) and P5(j) show that for any j , random variables belonging to Q_j (respectively T_j^{\prime} , $T_j^{\bar{}}$, and T_j^{\setminus}) are independent. Thus, Corollary 1 holds. \square

VII. APPENDIX C

CHARACTERIZATION OF COMPATIBLE SUBSETS

In order to prove Theorem 3, preliminary definitions and properties are stated. Then, it is shown that any MBRF $\{T, Q\}_A$ built on a convex set A is compatible with any MBRF $\{T, Q\}_\Lambda$ on a larger finite subset Λ of Λ_∞ . Finally, for any nonconvex finite subset A of Λ_∞ , another finite set Λ can be found so as to contain A although $\{T, Q\}_\Lambda$ and $\{T, Q\}_A$ are not mutually compatible. To begin with, let us define the notion of *boundary*, which is central in the proof of Theorem 3.

Definition 5: The *left boundary* $B^l(\Lambda)$ (resp. the *right boundary* $B^r(\Lambda)$) of any finite subset Λ of Λ_∞ contains the elements of Λ with no left (resp. right) neighbors in Λ (see Definition 2):

$$\begin{aligned}
X \in B^l(\Lambda) &\Leftrightarrow X \in \Lambda, N^l(X) \cap \Lambda = \emptyset, \\
X \in B^r(\Lambda) &\Leftrightarrow X \in \Lambda, N^r(X) \cap \Lambda = \emptyset. \quad (C.1)
\end{aligned}$$

The *boundary* of Λ is $B(\Lambda) = B^l(\Lambda) \cup B^r(\Lambda)$.

Property 1: For any X on Λ , there exists at least one admissible path $L(X_1, X_N)$ (see Definition 3) which connects X to $B^l(\Lambda)$ and $B^r(\Lambda)$ in Λ in the following sense:

$$\begin{cases} X_1 \in B^l(\Lambda), X_N \in B^r(\Lambda), \\ X \in L(X_1, X_N) \subset \Lambda. \end{cases}$$

A direct consequence is

$$\Lambda \neq \emptyset \Rightarrow B^l(\Lambda) \neq \emptyset \quad \text{and} \quad B^r(\Lambda) \neq \emptyset.$$

Proof: Such an admissible path $L(X_1, X_N)$ is obtained by construction by both sides of X . From (C.1), whether $X \in B^l(\Lambda)$, in which case $X_1 = X$, or else $X = X_n$ (n is unknown at this point) and there exists $X_{n-1} \in N^l(X_n) \cap \Lambda$. Repeated application of the procedure provides a growing series of sites on Λ , all different, so $B^l(\Lambda)$ must be reached to end the procedure because Λ is finite. Symmetrically, $B^r(\Lambda)$ is reached the same way and $L(X_1, X_N)$ is finally obtained by concatenation of the two finite series (X_1, \dots, X_n) and (X_n, \dots, X_N) . \square

Property 2: For any $A \subseteq \Lambda$, we have:

$$A \text{ convex} \Rightarrow B(\Lambda) \not\subseteq A.$$

Proof: Since $A \subseteq \Lambda$, there exists X in Λ and not in A . From Property 1, at least one admissible path $L(X_1, X_N)$ connects X to $B^l(\Lambda)$ and $B^r(\Lambda)$. Now if $B(\Lambda) \subset A$, the particular path $L(X_1, X_N)$ violates the convexity of A since $X_1 \in A$, $X_N \in A$, and $X \notin A$. \square

Property 3: For any $X \in B(\Lambda)$, $\{T, Q\}_\Lambda$, and $\{T, Q\}_{\Lambda \setminus \{X\}}$ are mutually compatible.

Proof: We have to prove (17) with $A = \Lambda \setminus \{X\}$. Since $\{T, Q\}_\Lambda$ is a MBRF on Λ , the expression of $P_\Lambda(\Lambda)$ is given by (14) and $P_A(\Lambda)$ is obtained by

$$P_A(\Lambda) = \sum_{X=0}^1 P_\Lambda(\Lambda). \quad (\text{C.2})$$

To proceed further, it is necessary to investigate separately the four possible cases: $X = Q_{ij}$, $X = T'_{ij}$, $X = T^-_{ij}$, and $X = T^-_{ij}$. Here, only one of these cases is detailed. The result for the other three can be obtained in a similar manner.

Assume that $X = T^-_{ij}$ and $X \in B^l(\Lambda)$. From (13), (18), and (C.1), we have $Q_{ij} \notin C^r_{ij}$, so as separability applies to $\tau(C^r_{ij})$ in (14):

$$\tau(C^r_{ij}) = \tau(t^-_{ij})\tau(C^r_{ij} \setminus \{t^-_{ij}\}).$$

Hence, in (14), T^-_{ij} only appears in $\tau(C^l_{ij+1})$, which is the unique term to be summed in (C.2):

$$\sum_{X=0}^1 \tau(C^l_{ij+1}) = \begin{cases} 1, & \text{if } C^l_{ij+1} = \{T^-_{ij}\}, \\ \tau(C^l_{ij+1} \setminus \{t^-_{ij}\}), & \text{otherwise.} \end{cases}$$

The resulting form of $P_A(\Lambda)$ is nothing but $P_A(A)$ as it would have been directly written according to (14). A similar proof can be derived for the three other cases, and also for $X \in B^r(\Lambda)$. \square

Proof of Theorem 3:

- **Sufficient condition (A convex $\Rightarrow A$ CSS):** Let Λ be any finite subset of Λ_∞ and A any convex subject of Λ ($A \subseteq \Lambda$). From Property 2, there exists at least one element X_1 of $B(\Lambda)$ not in A and from Property 3, $\{T, Q\}_\Lambda$ and $\{T, Q\}_{\Lambda \setminus \{X_1\}}$ are mutually compatible.

Now either $A = \Lambda \setminus \{X_1\}$, hence the result, or there exists at least one element X_2 of $B(\Lambda \setminus \{X_1\})$ not in A . In this case, the procedure is repeated until A is reached. Mutual compatibility between $\{T, Q\}_\Lambda$ and $\{T, Q\}_A$ results from successive descending compatibility.

- **Necessary condition (A nonconvex $\Rightarrow A$ not CSS):** Let A denote any nonconvex finite subset of Λ_∞ (Fig. 9 depicts the case of a rectangular subset). Among all the admissible paths which violate the convexity of A , some of them have a minimal length N . Let $L(X_1, X_N)$ one of the minimal paths. Obviously, $L(X_1, X_N) \cap A$ is reduced to $\{X_1, X_N\}$, otherwise a (shorter) subpath of $L(X_1, X_N)$ could be substituted for $L(X_1, X_N)$. For the same reason, we have also

$$\begin{aligned} N^l(X_n) \cap A &= \emptyset, & \text{for } n &= 3, \dots, N-1, \\ N^r(X_n) \cap A &= \emptyset, & \text{for } n &= 2, \dots, N-2. \end{aligned}$$

Now, consider $\Lambda = A \cup L(X_1, X_N)$; the object of the proof is to show that $P_\Lambda(A) \neq P_A(A)$. Using (14) to express both $P_\Lambda(\Lambda)$ and $P_A(A)$, one gets the relation

$$P_\Lambda(\Lambda) = P_A(A) \frac{P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N-1}, \mathbf{x}_N)}{\tau(\mathbf{x}_1)\tau(\mathbf{x}_N)} \quad (\text{C.3})$$

with $\mathbf{X}_1 = (N^r(X_1) \cup N^l(X_2)) \cap A$, $\mathbf{X}_N = (N^r(X_{N-1}) \cup N^l(X_N)) \cap A$ (note that $L(X_1, X_N) \cap \mathbf{X}_1 = \{X_1\}$ and $L(X_1, X_N) \cap \mathbf{X}_N = \{X_N\}$) and

$$\begin{aligned} P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N-1}, \mathbf{x}_N) \\ = \tau(\mathbf{x}_1)\tau(\mathbf{x}_2 | \mathbf{x}_1)\tau(\mathbf{x}_N | \mathbf{x}_{N-1}) \prod_{n=3}^{N-1} \tau(\mathbf{x}_N | \mathbf{x}_{n-1}). \end{aligned} \quad (\text{C.4})$$

Equation (C.4) shows that $P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N-1}, \mathbf{x}_N)$ is the measure of a nonhomogeneous Markov chain $\mathbf{X} = (X_1, X_2, \dots, X_{N-1}, X_N)$. Actually, (C.4) is of the form of (6), so \mathbf{X} is a particular MBRF. Now $P_\Lambda(A)$ is given by summation of $P_\Lambda(\Lambda)$ over all the joint realization of (X_2, \dots, X_{N-1}) . From (C.3), the summation yields

$$P_\Lambda(A) = P_A(A) \frac{P(\mathbf{x}_1, \mathbf{x}_N)}{\tau(\mathbf{x}_1)\tau(\mathbf{x}_N)}$$

and $P_\Lambda(A) \neq P_A(A)$ turns out to be equivalent to a much simpler condition regarding the nonhomogeneous Markov chain \mathbf{X} only:

$$\exists \mathbf{x}_1, \mathbf{x}_N \text{ such as } P(\mathbf{x}_1, \mathbf{x}_N) \neq \tau(\mathbf{x}_1)\tau(\mathbf{x}_N). \quad (\text{C.5})$$

Whether from (C.4) or from the invariance property of MBRF's, we have $P(\mathbf{x}_1) = \tau(\mathbf{x}_1)$ and $P(\mathbf{x}_N) = \tau(\mathbf{x}_N)$, so (C.5) means that \mathbf{X}_1 and \mathbf{X}_N are mutually dependent.

Let us introduce the transition matrix $\mathbf{P}_{1N} = \{P(X_N = j | X_1 = i)\}_{ij}$ between \mathbf{X}_1 and \mathbf{X}_N (it is understood that i and j , respectively, describe the state spaces of \mathbf{X}_1 and \mathbf{X}_N). Then, classical properties of Markov chains [29] indicate that an equivalent condition of (C.5) and \mathbf{P}_{1N} is

$$\text{rank}(\mathbf{P}_{1N}) > 1. \quad (\text{C.6})$$

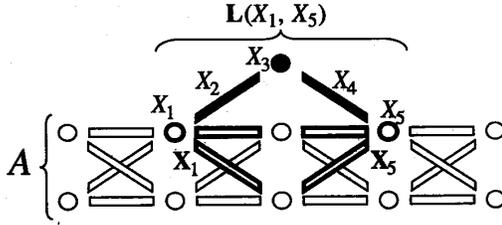


Fig. 9. Example of a nonconvex set A . Pixels in the finite rectangular set A are represented by the white circles and bars. The path $L(X_1, X_5) = (X_1, \dots, X_5)$ is one of the admissible paths of minimum length ($N = 5$) that violate convexity. In such a situation, it is shown that $\{T, Q\}_A$ is not mutually compatible with $\{T, Q\}_\Lambda$, where $\Lambda = A \cup L(X_1, X_5)$.

On the other hand, since X is a Markov chain, we have $P_{1N} = P_2 P_3 \dots P_N$, where

$$\begin{aligned} P_2 &= \{P(X_2 = j | X_1 = i)\}_{ij}, \\ P_n &= \{P(X_n = j | X_{n-1} = i)\}_{ij}, \quad n = 3, \dots, N-1, \\ P_N &= \{P(X_N = j | X_{N-1} = i)\}_{ij} \end{aligned}$$

are the successive transition matrices between X_1, X_2, \dots, X_N . Since X_2, \dots, X_{N-1} are binary variables, the size of P_3, \dots, P_{N-1} is 2×2 . In addition, the maximum number of rows (resp. columns) of P_2 (resp. P_N) is also equal to two. Thus, the rank of all matrices P_2, \dots, P_N is less than or equal to two, and (C.6) is equivalent to

$$\text{rank}(P_n) = 2, \quad \text{for all } n = 2, \dots, N,$$

which in turn is equivalent to

$$\begin{aligned} \exists x_1, x_2 \text{ such as } \tau(x_1, x_2) &\neq \tau(x_1)\tau(x_2), \\ \exists x_{n-1}, x_n \text{ such as } \tau(x_{n-1}, x_n) &\neq \tau(x_{n-1})\tau(x_n) \\ \text{for all } n = 3, \dots, N-1, \\ \exists x_{N-1}, x_N \text{ such as } \tau(x_{N-1}, x_N) &\neq \tau(x_{N-1})\tau(x_N). \end{aligned} \quad (C.7)$$

Finally, (C.7) is fulfilled because from (11) and (13), we have $\tau(q, t) \neq \tau(q)\tau(t)$ for any $t \in \{(t'), (t^-), (t^\backslash), (t', t^-), (t^-, t^\backslash), (t', t^\backslash), (t', t^-, t^\backslash)\}$ at least for the instance of $(q, t', t^-, t^\backslash) = (0, 0, 0, 0)$, provided that we are in the general case: $\mu' \neq 0, \mu^- \neq 0, \mu^\backslash \neq 0$. \square

APPENDIX D

VIOLATION OF THE MRF PROPERTY FOR Q

The object of Appendix D is to build a counter-example which shows that the neighborhood system of random field Q is not local. Let (i, j) be an interior site of Λ° , $2 \leq i \leq I-1, 2 \leq j \leq J-1$. The idea consists of computing $P(q_{ij} | q^{ij})$ for particular realizations of Q , and of showing that the resulting expression depends on sites located at arbitrarily large distances from (i, j) . From (20), it is clear that $P(q_{ij} | q^{ij})$ only depends upon realizations of columns Q_{j-1}, Q_j , and Q_{j+1} . In the sequel, only the realizations of these three columns will be explicitly defined, and it will be assumed that the rest of the configuration is not in a forbidden state, i.e., $P(q) > 0$. For an arbitrary length parameter L such

that $0 \leq L \leq I-i$, we first define particular realizations of columns Q_{j-1}, Q_j and Q_{j+1} . This configuration is made to depend on L such that arbitrarily increasing values of L only modify variables located at arbitrarily increasing distances from site (i, j) . Then, through computation of transition probabilities between columns and use of (20), we show that the corresponding value of $P(q_{ij} | q^{ij})$ varies with L even when L becomes arbitrarily large. Hence, dependence of $P(q_{ij} | q^{ij})$ on q^{ij} is not restricted to a local neighborhood of (i, j) .

Define the following configuration (q_a, q_b) of two adjacent columns:

if L is even:

$$(q_a, q_b) = \begin{bmatrix} 0 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 0 & 1 & i \\ 1 & 0 & \\ 0 & 1 & \vdots \\ \vdots & \vdots & \vdots \\ 0 & 1 & i+L \\ 0 & 0 & \\ \vdots & \vdots & \vdots \\ 0 & 0 & I \end{bmatrix},$$

if L is odd:

$$(q_a, q_b) = \begin{bmatrix} 0 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 0 & 1 & i \\ 1 & 0 & \\ 0 & 1 & \vdots \\ \vdots & \vdots & \vdots \\ 1 & 0 & i+L \\ 0 & 0 & \\ \vdots & \vdots & \vdots \\ 0 & 0 & I \end{bmatrix},$$

and assume that the realization of (Q_{j-1}, Q_j, Q_{j+1}) is (q_a, q_b, q_a) . An essential step for the computation of $P(q_{ij} | q^{ij})$ is the evaluation of $P(q_b | q_a)$. In order to simplify the derivations, the parameters of the MBRF $\{T, Q\}$ are chosen according to

$$\begin{cases} \mu' = \mu^- = \mu^\backslash = \mu, & 0 < \mu < 1, \\ \epsilon = 0, \end{cases} \quad (D.1)$$

which implies that

$$\lambda = 1 - (1 - \mu)^3. \quad (D.2)$$

Evaluation of $p_L \triangleq P(q_b | q_a)$ can be carried out recursively on L as follows: we have

$$p_L = \sum_{t_{ab}} P(q_b | t_{ab}) P(t_{ab} | q_a),$$

where t_{ab} denotes the realizations of the transition column between q_a and q_b . Then, using Bayes' rule yields

$$p_L = \sum_{t_{ab}} P(q_b | t_{ab}) P(q_a | t_{ab}) \frac{P(t_{ab})}{P(q_a)}. \quad (D.3)$$

Since $\epsilon = 0$, $P(q_b | t_{ab})$ and $P(q_a | t_{ab})$ are equal to zero when t_{ab} is not compatible with the configuration (q_a, q_b) , and are equal to one otherwise. Therefore, (D.3) can be written as

$$p_L = \sum_{t_{ab}} \frac{P(t_{ab})}{P(q_a)}$$

where the sum extends over all the configurations of transition variables which are compatible with (q_a, q_b) . Finally, Corollary 1 in conjunction with (D.1) and (D.2) yield the following explicit recursive expression

$$\begin{cases} p_{2l+1} = \frac{\mu(1-\mu)^2}{\lambda} (p_{2l-1} + p_{2l}), \\ p_{2l+2} = \frac{\mu(1-\mu)}{\lambda} (\mu p_{2l-1} + p_{2l}), \end{cases} \quad \text{for any } l \geq 0.$$

It can be easily verified that the initial conditions are given by $p_{-1} = 1$ and $p_0 = 0$. We now proceed with the evaluation of $P(q_{ij} | q^{ij})$. Using (20) and Bayes' rule, we get

$$P(q_{ij} | q^{ij}) = \left[1 + \frac{P(q_a | \bar{q}_b)^2 P(\bar{q}_b) P(q_b)}{P(q_b | q_a)^2 P(q_a)^2} \right]^{-1}$$

and noticing that $P(q_a | \bar{q}_b)$ is equal to p_{L-1} , we can rewrite $P(q_{ij} | q^{ij})$ as

$$P(q_{ij} | q^{ij}) = \left[1 + \frac{\lambda}{1-\lambda} (u_L)^2 \right]^{-1},$$

where (u_L) is defined by

$$(u_L) = \begin{cases} \left(\frac{p_{L-1}}{p_L} \right), & \text{if } L \text{ is even,} \\ \frac{1-\lambda}{\lambda} \left(\frac{p_{L-1}}{p_L} \right), & \text{if } L \text{ is odd,} \end{cases} \quad (L > 0).$$

From (D.2), it is straightforward to show that series (u_L) is characterized by the unique recursive equation

$$u_{L+1} = \frac{1-\mu}{\mu} \frac{1}{1+u_L}$$

with initial condition $u_1 = 0$. General properties of series of the form

$$\nu_{n+1} = \frac{k}{1+\nu_n}, \quad k > 0$$

indicate that if $\nu_1 > -1$, the series converges toward a limit ν_∞ defined by

$$\nu_\infty = \frac{\sqrt{1+4k} - 1}{2}.$$

Furthermore, if $\nu_1 \neq \nu_\infty$, then the series $|\nu_n - \nu_\infty|$ strictly decreases to zero. Therefore, since $u_1 \neq u_\infty$ and since $P(q_{ij} | q^{ij})$ is a continuous and strictly monotonous function of u_L , $P(q_{ij} | q^{ij})$ converges to the limit

$$P_\infty(q_{ij} | q^{ij}) = \left[1 + \frac{\lambda}{1-\lambda} (u_\infty)^2 \right]^{-1},$$

but never reaches it. Consequently, only the length of the columns limits the size of the neighborhoods of random field Q , i.e., the neighborhood system of Q is not local.

REFERENCES

- [1] J. Besag, "On the statistical analysis of dirty pictures," *J. Roy. Statist. Soc. B*, vol. 48, pp. 259-302, 1986.
- [2] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 721-741, Nov. 1984.
- [3] R. Chellappa and R. L. Kashyap, "Digital image restoration using spatial interaction models," *IEEE Trans. Acoust., Speech Signal Processing*, vol. ASSP-30, pp. 461-472, June 1982.
- [4] F. C. Jeng and J. W. Woods, "Compound Gauss-Markov random fields for image estimation," *IEEE Trans. Signal Processing*, vol. 39, pp. 683-697, Mar. 1991.
- [5] C. Y. Chi and J. M. Mendel, "Multichannel maximum-likelihood deconvolution," presented at the *Proc. 55th Ann. Meeting of the Soc. of Exploration Geophysicists*, Washington, DC; also in Internal Rep., Univ. of Southern California, Los Angeles, 1985.
- [6] J. Brac, P. Y. Déquierez F. Hervé, C. Jacques, P. Lailly, V. Richard, and D. Tran Van Nhieu, "Inversion with a priori information: An approach to integrated stratigraphic interpretation," presented at the *Proc. 58th Ann. Meeting Soc. of Exploration Geophysicists*, Anaheim, CA, 1988.
- [7] J. Kormylo and J. M. Mendel, "Maximum-likelihood detection and estimation of Bernoulli-Gaussian processes," *IEEE Trans. Information Theory*, vol. IT-28, pp. 482-488, May 1982.
- [8] J. M. Mendel, *Optimal Seismic Deconvolution*. New York: Academic Press, 1983.
- [9] A. K. Mahalanabis, S. Prasad, and K. P. Mohandas, "Recursive decision-directed estimation of reflection coefficients for seismic data deconvolution," *Automatica*, vol. 18, pp. 721-726, 1982.
- [10] S. D. Kollias and C. C. Halkias, "An instrumental variable approach to minimum-variance seismic deconvolution," *IEEE Trans. Geoscience Remote Sensing*, vol. GE-23, pp. 778-788, Nov. 1985.
- [11] Y. Goussard and G. Demoment, "Recursive deconvolution of Bernoulli-Gaussian processes using a MA representation," *IEEE Trans. Geoscience Remote Sensing*, vol. GE-27, pp. 384-394, July 1989.
- [12] R. Godfrey, F. Muir, and F. Rocca, "Modeling seismic impedance with Markov chains," *Geophysics*, vol. 45, pp. 1351-1372, 1980.
- [13] H. Derin and P. A. Kelly, "Discrete-index Markov-type random processes," *Proc. IEEE*, vol. 77, pp. 1485-1510, Oct. 1989.
- [14] R. Kinderman and J. L. Snell, *Markov Random Fields and Their Applications*. Providence, RI: Amer. Math. Soc., 1980.
- [15] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Roy. Statist. Soc. B*, vol. 36, pp. 192-235, 1974.
- [16] D. K. Pickard, "Unilateral Markov fields," *Adv. Appl. Probab.*, vol. 12, pp. 655-671, 1980.
- [17] J. Goutsias, "Mutually compatible Gibbs random fields," *IEEE Trans. Inform. Theory*, vol. 35, pp. 1233-1249, Nov. 1989.
- [18] D. K. Pickard, "A curious binary lattice process," *J. Appl. Probab.*, vol. 14, pp. 717-731, 1977.
- [19] K. Abend, T. J. Harley, and L. N. Kanal, "Classification of binary random patterns," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 538-544, Oct. 1965.
- [20] L. N. Kanal, "Markov mesh models," in *Image Modeling*. New York: Academic Press, 1980, pp. 239-243.
- [21] H. Derin, H. Elliott, R. Cristi, and D. Geman, "Bayes smoothing algorithms for segmentation of binary images modeled by Markov random fields," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 707-720, Nov. 1984.
- [22] J. Goutsias, "Unilateral approximation of Gibbs random field images," *GVGIP: Graphical Models Image Processing*, vol. 53, pp. 240-257, 1991.
- [23] L. Breiman, *Probability*. Reading, MA: Addison Wesley, 1968.
- [24] Y. Goussard, G. Demoment, and J. Idier, "A new algorithm for iterative deconvolution of sparse spike trains," *Proc. Int. Conf. ASSP*, Albuquerque, NM, 1990, pp. 1547-1550.
- [25] G. D. Forney, "The Viterbi algorithm," *Proc. IEEE*, vol. PROC-61, pp. 268-278, Mar. 1973.
- [26] J. Idier and Y. Goussard, "Multi-channel seismic deconvolution," LSS Internal Rep. #GPI-90/09, 1990. To appear in *IEEE Trans. Geoscience Remote Sensing*.
- [27] J. Idier, "Imagerie des milieux stratifiés: modélisation markovienne, application à la déconvolution sismique," Thèse de Doctorat, Université de Paris-Sud, Centre d'Orsay, 1991.
- [28] W. Feller, *An Introduction to Probability Theory and Its Applications*. New York: Wiley, 1950.