A Bayesian model for lithology/fluid class prediction using a Markov mesh prior fitted from a training image

Håkon Tjelmeland, Xin Luo and Torstein Fjeldstad

Department of Mathematical Sciences, Norwegian University of Science and Technology

Abstract

We consider a Bayesian model for inversion of observed amplitude variation with offset (AVO) data into lithology/fluid classes, and study in particular how the choice of prior distribution for the lithology/fluid classes influences the inversion results. Two distinct prior distributions are considered, a simple manually specified Markov random field prior with a first order neighborhood and a Markov mesh model with a much larger neighborhood estimated from a training image. They are chosen to model both horizontal connectivity and vertical thickness distribution of the lithology/fluid classes, and are compared on an offshore clastic oil reservoir in the North Sea. We combine both priors with the same linearised Gaussian likelihood function based on a convolved linearised Zoeppritz relation and estimate properties of the resulting two posterior distributions by simulating from these distributions with the Metropolis–Hastings algorithm.

The influence of the prior on the marginal posterior probabilities for the lithology/fluid classes is clearly observable, but modest. The importance of the prior on the connectivity properties in the posterior realisations, however, is much stronger.

1

The larger neighborhood of the Markov mesh prior enables it to identify and model connectivity and curvature much better than what can be done by the first order neighborhood Markov random field prior. As a result, we conclude that the posterior realisations based on the Markov mesh prior appear with much higher lateral connectivity, which is geologically plausible.

Key words: Computing aspects, Inverse problem, Inversion, Mathematical formulation, Seismics.

Introduction

From seismic data one can predict elastic properties and lithology/fluid classes (LFC) in a reservoir. This is an inverse problem and for a given seismic data set many solutions exist. Different methods have been used for inverting seismic data to elastic properties and lithology/fluid classes, both deterministic approaches such as optimisation-based methods (Aster et al., 2011; Sen and Stoffa, 2013) and probabilistic approaches such as Bayesian inversion (Tarantola, 2005). Using the Bayesian framework, a linearised relation between the data and the elastic properties is commonly used and a Gaussian likelihood function is adopted, see for example Buland and Omre (2003), Gunning and Glinsky (2007) and the discussion in Grana et al. (2017). When inverting to elastic properties the prior is also often assumed to be Gaussian, in which case the posterior becomes Gaussian with analytically available mean and covariance, see again Buland and Omre (2003). When inverting to lithology/fluid classes, other priors have to be used. In particular Rimstad and Omre (2010) define a hierarchical prior, where a Markov random field (Kindermann and Snell, 1980; Hurn et al., 2003) is used to model the lithology/fluid classes and conditional on these the elastic properties are assumed to be Gaussian with mean and covariance

functions depending on the lithology/fluid classes. Grana and Della Rossa (2010) consider a Gaussian mixture prior for the elastic attributes to include multi-modality and skewness in the prior model where the effect of the lithology/fluid classes is summed out. With a non-Gaussian prior the posterior is no longer analytically available and Markov chain Monte Carlo (Gilks et al., 1996; Robert and Casella, 1999; Gamerman and Lopes, 2006) must typically be used to estimate properties of the resulting posterior distribution. It is also challenging to specify lateral connectivity and spatial dependency laterally for non-Gaussian priors, and often the inverse problem is solved trace-by-trace before a smoother is applied afterwards (Connolly and Hughes, 2016).

To specify a prior that reflects available prior information in a spatial problem like inversion of seismic data can be difficult. The properties of a Gaussian field is analytically well understood, so by adopting a Gaussian prior as in Buland and Omre (2003) the prior specification process is simplified. Eidsvik et al. (2004) consider a one dimensional problem and assume a Markov chain prior to predict geological attributes from well log data, and Fjeldstad and Omre (2017) use a similar model to predict lithology/fluid classes and elastic attributes from seismic data. The properties of Markov chains are also analytically available, which again simplifies the specification of a reasonable prior. For most non-Gaussian spatial prior models the situation is less favorable. In Rimstad et al. (2012) a discrete Markov random field prior is used for the lithology/fluid classes. The properties of discrete Markov random fields are analytically not available, which makes it difficult to verify the properties of the chosen prior. To cope with the problem of specifying non-Gaussian spatial prior models it has in geostatistics become common practice to estimate the prior model from a so-called training image. A training image is from an outcrop or a constructed scene assumed to have the same spatial structure as the phenomenon under study. The idea is to estimate a prior model from one or more training images, see the discussion in Mariethoz

and Caers (2014). Various multiple-point statistics models (Guardiano and Srivastava, 1993; Strebelle, 2002; Journel and Zhang, 2006; Zhang et al., 2012) have been defined to implement this idea. These models are algorithmically defined. The nodes in a lattice are visited in a random order and when a node is visited, the value in that node is simulated conditional on values in previously visited nodes, where the conditional distribution used is estimated from the training image. There are two serious complications associated with the use of multiple-point statistics models. First, the number of conditional distributions that has to be estimated from the training image is enormous, and the information content in a typical training image is not sufficient to estimate this number of parameters. Emery and Lantuéjoul (2014) are discussing this issue mathematically. Second, the models are only algorithmically defined and no simple to evaluate expressions are available for the estimated model. The implication of this is that if we want to use the estimated model as a prior and generate realisations conditional on some observed data, it is in general not clear how to do this. Since we have no analytical formula for the prior, we neither have an expression for the posterior. This issue is also discussed in Toftaker and Tjelmeland (2013). As alternatives to the multiple-point statistics models Arnesen and Tjelmeland (2017) and Luo and Tjelmeland (2017) introduce procedures for fitting Markov random fields and Markov mesh models, respectively, to a given training image. For these model classes explicit expressions for the distributions are available, so to simulate from a corresponding conditional distribution Markov chain Monte Carlo procedures can for example be employed.

The purpose of this article is to demonstrate how inversion of seismic data into lithology/fluid classes can be accomplished in a Bayesian framework by estimating a prior model for the lithology/fluid classes from a given training image, and combine this with a linearised and Gaussian likelihood function. We fit a Markov mesh prior model to a training

4

image as discussed in Luo and Tjelmeland (2017) and use Markov chain Monte Carlo to simulate from the resulting posterior distribution as discussed in Rimstad et al. (2012). A Markov mesh prior is used for lithology/fluid prediction also in Stien and Kolbjørnsen (2011), but they specify manually the neighborhood and interaction structures and fit only the parameter values to the training image. In our fitting procedure we fit both the neighborhood and interaction structures and the parameter values to the given training image. To focus on the methodological aspects we consider a situation with only two lithology/fluid classes, oil sand and shale. In particular we compare the results from our procedure with what we get by instead using a simpler manually chosen Markov random field prior.

The article has the following layout: First we present the data set and the associated training image, and we analyse and introduce our Bayesian model formulation. A Gaussian likelihood function is defined, and its properties are discussed. We introduce the Markov mesh and Markov random field priors, and in specific focus on how we fit the Markov mesh prior to the given training image. Next, a sampling algorithm for the posterior distribution is discussed. The two priors are tested on a real 2D section case study in the North Sea. Finally, we discuss the posterior properties of the two priors and provide some closing remarks.

Methodology

The objective is to demonstrate and compare two different prior models in a Bayesian framework to predict lithology/fluid classes in the subsurface. In this section, we introduce the data set and formulate the inverse problem in a Bayesian setting, define a likelihood function and the two priors, and discuss posterior simulation.
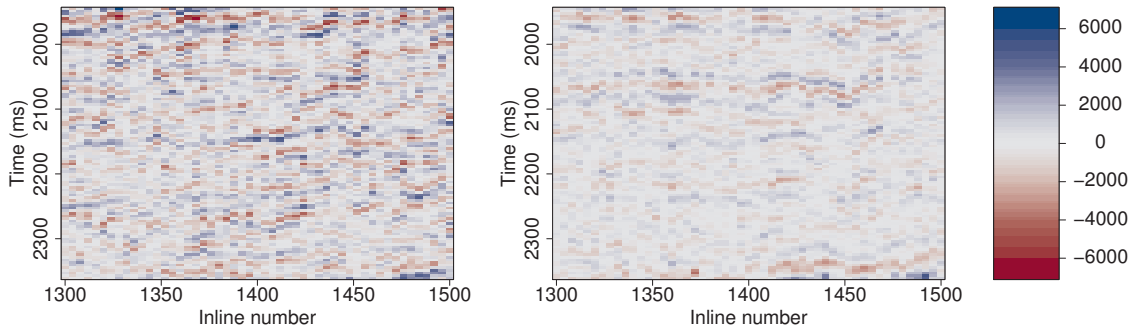
Figure 1: Near (left) and far (right) offset seismic data used for lithology/fluid prediction.

Data set and Bayesian model formulation

In this article we consider a seismic section from the Alvheim field in the North Sea, which is a clastic oil reservoir. The Alvheim field is characterised by a complex sand lobe geometry and is buried approximately 2 km below the sea floor. In the analysis we use one near and one far offset seismic data represented in a 105×51 lattice $G = \{(i,j)|i = 1,\ldots,105; j = 1,\ldots,51\}$. The stacked sections were generated from pre-stack time migrated (PSTM) common depth gathers (CDP), see Rimstad et al. (2012) for further processing details. The seismic data are shown in Figure 1. The horizontal and vertical resolutions are about 100 m and 4 ms, respectively. We let $d_{ij}, (i,j) \in G$ denote a vector of size two containing the observed near and far offset seismic data in node $(i,j) \in G$, and let $\boldsymbol{d}$ be a vector where all $d_{ij}, (i,j) \in G$ are stacked on top of each other. We model two lithology/fluid classes, oil sand and shale. For each node $(i,j) \in G$ we let $\kappa_{ij} \in \{0,1\}$ denote the lithology fluid class in node (i,j), where $\kappa_{ij} = 0$ and $\kappa_{ij} = 1$ represent shale and oil sand, respectively. We let $\boldsymbol{\kappa}$ be a vector of all $\kappa_{ij}, (i,j) \in G$ stacked on top of each other. To estimate a Markov mesh prior distribution for $\boldsymbol{\kappa}$ we use a training image from Lang and Grana (2017). The
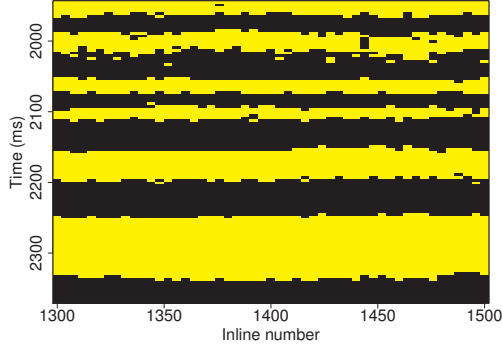
Figure 2: Training image that we use to estimate a Markov mesh prior distribution for the spatial distribution of lithology/fluid classes, $\boldsymbol{\kappa}$. Black and yellow represent shale and oil sand, respectively.

training image is from a reservoir with similar characteristics and is shown in Figure 2.

To model the relation between $\boldsymbol{\kappa}$ and $\boldsymbol{d}$ we first introduce a vector $\boldsymbol{m} = \{\boldsymbol{m_{ij}}, (i,j) \in G\}$ of elastic properties, where $\boldsymbol{m_{ij}}$ is a vector of length two. We let the first element in $\boldsymbol{m_{ij}}$ be the product of the density ρ and the pressure-wave velocity v_p in node (i,j) and let the second element be the v_p/v_s ratio in the same node, where v_s is the shear-wave velocity.

For the three variables $\boldsymbol{\kappa}$, $\boldsymbol{m}$ and $\boldsymbol{d}$ we adopt a Bayesian model. We let $p(\boldsymbol{\kappa})$ denote a prior distribution for $\boldsymbol{\kappa}$ and let $p(\boldsymbol{m}|\boldsymbol{\kappa})$ denote the conditional distribution for the elastic parameters $\boldsymbol{m}$ given the lithology/fluid classes $\boldsymbol{\kappa}$. Finally, we assume the seismic data $\boldsymbol{d}$ to be conditionally independent of $\boldsymbol{\kappa}$ when the elastic properties $\boldsymbol{m}$ are given. We let $p(\boldsymbol{d}|\boldsymbol{m})$ denote the conditional distribution for the seismic data $\boldsymbol{d}$ given elastic properties $\boldsymbol{m}$. The $p(\boldsymbol{d}|\boldsymbol{m})$ represents a probabilistic formulation of the forward model. Bayes theorem then gives

$$p(\boldsymbol{\kappa}|\boldsymbol{d}) \propto p(\boldsymbol{\kappa})p(\boldsymbol{d}|\boldsymbol{\kappa}), \tag{1}$$

where

$$p(\boldsymbol{d}|\boldsymbol{\kappa}) = \int p(\boldsymbol{m}, \boldsymbol{d}|\boldsymbol{\kappa})\mathrm{d}\boldsymbol{m} = \int p(\boldsymbol{m}|\boldsymbol{\kappa})p(\boldsymbol{d}|\boldsymbol{m})\mathrm{d}\boldsymbol{m}. \tag{2}$$

In the following we first outline the details of $p(\boldsymbol{m}|\boldsymbol{\kappa})$ and $p(\boldsymbol{d}|\boldsymbol{m})$, which is used to specify the likelihood $p(\boldsymbol{d}|\boldsymbol{\kappa})$, and thereafter specify the prior $p(\boldsymbol{\kappa})$ before we describe the Markov chain Monte Carlo procedure we use to simulate from $p(\boldsymbol{\kappa}|\boldsymbol{d})$.

Likelihood model

Following Grana et al. (2017) and Fjeldstad and Omre (2017) we adopt a linearised and Gaussian likelihood for the forward model for $\boldsymbol{d}$ given $\boldsymbol{\kappa}$. More specifically, we assume each of $p(\boldsymbol{m}|\boldsymbol{\kappa})$ and $p(\boldsymbol{d}|\boldsymbol{m})$ to be Gaussian, the conditional mean of $\boldsymbol{d}$ given $\boldsymbol{m}$ to be a linear function of $\boldsymbol{m}$, and the conditional covariance matrix of $\boldsymbol{d}$ given $\boldsymbol{m}$ not to be a function of $\boldsymbol{m}$. In the following we outline the distributions $p(\boldsymbol{m}|\boldsymbol{\kappa})$ and $p(\boldsymbol{d}|\boldsymbol{m})$ in more detail, starting with $p(\boldsymbol{m}|\boldsymbol{\kappa})$.

We assume $\boldsymbol{m}|\boldsymbol{\kappa}$ to be Gaussian and

$$\mathrm{E}[\boldsymbol{m_{ij}}|\boldsymbol{\kappa}] = \mu_{\kappa_{ij}} \quad \text{and} \quad \mathrm{Cov}[\boldsymbol{m_{ij}}|\boldsymbol{\kappa}] = \Sigma_{\kappa_{ij}}, \tag{3}$$

where μ_0 and Σ_0 are the conditional mean and covariance for $\boldsymbol{m_{ij}}$ if node (i,j) contains shale ($\kappa_{ij} = 0$), and μ_1 and Σ_1 are corresponding quantities when node (i,j) contains oil sand ($\kappa_{ij} = 1$). Moreover, we assume a separable correlation function $\rho((i,j),(k,l))$ for $\boldsymbol{m}|\boldsymbol{\kappa}$, but we do not allow the correlations to depend on the lithology/fluid classes $\boldsymbol{\kappa}$.

For the forward model $p(\boldsymbol{d}|\boldsymbol{m})$ we use a convolved linearised approximation of the Zoeppritz equation (Buland and Omre, 2003) based on Aki-Richards formulation that is valid for weak vertical contrasts (Aki and Richards, 1980). The vector $\boldsymbol{d}$ is then formed from $\boldsymbol{m}$ in several steps. First, all vertical first-order contrasts or differences $\boldsymbol{m_{ij}} - \boldsymbol{m_{i-1,j}}$ are formed by pre-multiplying $\boldsymbol{m}$ with a matrix D. Thereafter, reflection coefficients are
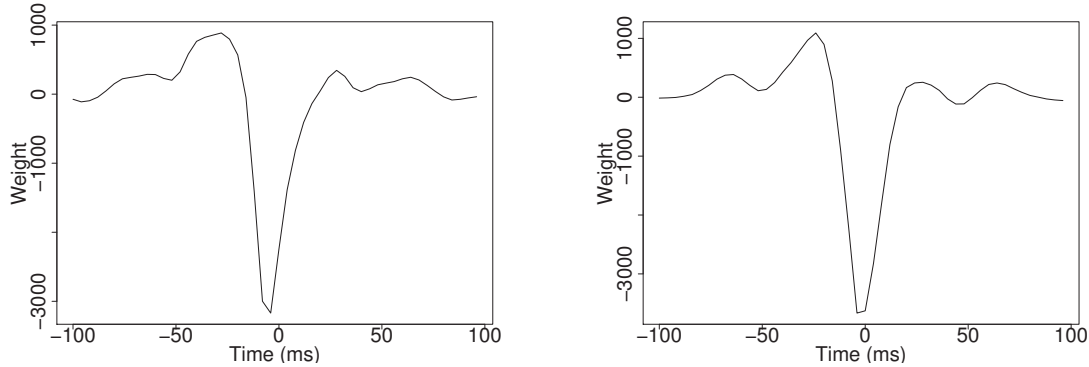
Figure 3: Wavelets used for the near (left) and far (right) offset seismic data. The x- and y- axes show vertical distance and wavelet values, respectively.

formed by pre-multiplying $D\boldsymbol{m}$ with a block diagonal matrix A, where all the blocks are identical 2×2 matrices containing coefficients in the Aki-Richards formulation. The mean value of the seismic data are then formed via a convolution of each column of $AD\boldsymbol{m}$ with wavelets. Different wavelets are used for the near and far offset seismic data as shown in Figure 3. These wavelets are estimated from data in a well in the same reservoir as the seismic data is coming from. This well is, however, located some distance away from the seismic section we are studying. The effect of the convolutions can be written as pre-multiplying $AD\boldsymbol{m}$ with a matrix W. Finally, a zero mean Gaussian error term $\boldsymbol{\varepsilon}$ with a fixed covariance matrix Σ_ε is added to $WAD\boldsymbol{m}$. Thus, $\boldsymbol{d}|\boldsymbol{m}$ is Gaussian with

$$\mathrm{E}[\boldsymbol{d}|\boldsymbol{m}] = WAD\boldsymbol{m} \quad \text{and} \quad \mathrm{Cov}[\boldsymbol{d}|\boldsymbol{m}] = \Sigma_\varepsilon. \tag{4}$$

With the Gaussian distributions specified above for $p(\boldsymbol{m}|\boldsymbol{\kappa})$ and $p(\boldsymbol{d}|\boldsymbol{m})$ it follows from standard properties of the Gaussian distribution that also $p(\boldsymbol{d}|\boldsymbol{\kappa})$ becomes Gaussian. Moreover, expressions are available for the mean $\mathrm{E}[\boldsymbol{d}|\boldsymbol{\kappa}]$ and the covariance $\mathrm{Cov}[\boldsymbol{d}|\boldsymbol{\kappa}]$ as function of $\boldsymbol{\kappa}$, μ_0, μ_1, Σ_0, Σ_1, $\rho(\cdot,\cdot)$, W, A, D and Σ_ε.

9

Prior models

The main purpose of this article is to demonstrate how a Markov mesh model fitted to a training image can be used as prior in a Bayesian model for lithology/fluid prediction. However, we also want to study how the inversion results change when using such a prior relative to what we get by using a simpler manually specified prior. To fit a Markov mesh model to a training image involves extra working and computing time, so there is no reason to do so unless it results in a significant change in the inversion results. In the following we first specify the class of Markov mesh models and briefly discuss the procedure we use to fit the model to the training image in Figure 2. Thereafter we describe a simpler manually specified prior we use for comparison, the profile Markov random field introduced in Ulvmoen and Omre (2010).

Markov mesh prior

An introduction to the class of Markov mesh models can be found in Abend et al. (1965) and the more general class of partially ordered Markov models is defined in Cressie and Davidson (1998). In the following description we limit the attention to binary fields and introduce the necessary notions to define homogeneous Markov mesh models defined on a rectangular lattice.

Let $G = \{(i,j)|i = 1, \ldots, n_1; j = 1, \ldots, n_2\}$ be a rectangular lattice, to each node $(i,j) \in G$ of which we associate a binary variable $\kappa_{ij} \in \{0, 1\}$. We let $\boldsymbol{\kappa} = (\kappa_{ij} : (i,j) \in G)$ denote the collection of all these binary variables and use $\boldsymbol{\kappa}_\lambda = (\kappa_{ij} : (i,j) \in \lambda)$ to denote the collection of binary variables in a set $\lambda \subseteq G$ of nodes. The Markov mesh model is based on numbering the nodes in G from 1 to nm in the lexicographical order. Without

loss of generality, the distribution of $\boldsymbol{\kappa}$ can then be expressed as

$$p(\boldsymbol{\kappa}) = \prod_{(i,j)\in G} p(\kappa_{ij}|\boldsymbol{\kappa}_{\rho_{ij}}), \tag{5}$$

where ρ_{ij} is the set of all nodes coming before node (i, j), i.e.

$$\rho_{ij} = \{(k,l) \in G : nk + l < ni + j\}. \tag{6}$$

The set ρ_{ij} is called the predecessor set of node (i, j). The central assumption in Markov mesh models is that $p(\kappa_{ij}|\boldsymbol{\kappa}_{\rho_{ij}})$ has a Markov property in that

$$p(\kappa_{ij}|\boldsymbol{\kappa}_{\rho_{ij}}) = p(\kappa_{ij}|\boldsymbol{\kappa}_{\nu_{ij}}), \tag{7}$$

where $\nu_{ij} \subseteq \rho_{ij}$ is called the sequential neighborhood of node (i, j). Following Luo and Tjelmeland (2017), we assume that all the sequential neighborhoods are generated via a translation of a template sequential neighborhood τ. The set τ can best be thought of as the sequential neighborhood of node $(0, 0)$ in a infinite lattice. More precisely, τ should contain a finite number of elements and

$$\tau \subset \{(i,j) : i \in \mathbb{Z}^-, j \in \mathbb{Z}\} \cup \{(0,j) : j \in \mathbb{Z}^-\}, \tag{8}$$

where $\mathbb{Z} = \{0, \pm1, \pm2, \ldots\}$ and $\mathbb{Z}^- = \{-1, -2, \ldots\}$ are the sets of all integers and all negative integers, respectively. Given the set τ we assume the sequential neighborhood ν_{ij} to be generated by translating each element in τ a distance (i, j) and, if necessary, dropping elements falling outside the lattice G. Mathematically, ν_{ij} is then given as

$$\nu_{ij} = (\tau \oplus (i, j)) \cap G, \tag{9}$$

where the translation operator $\oplus$ is defined as

$$\tau \oplus (i, j) = \{(k + i, l + j) : (k, l) \in \tau\}. \tag{10}$$

11

Constructing ν_{ij} in this way, the sequential neighborhoods for all nodes sufficiently far away from the lattice borders will have the same form.

Still following Luo and Tjelmeland (2017), we model $p(\kappa_{ij}|\boldsymbol{\kappa}_{\nu_{ij}})$ by assuming the logit transformation of $p(\kappa_{ij} = 1|\boldsymbol{\kappa}_{\nu_{ij}})$ to be given by

$$\text{logit}\left[p(\kappa_{ij} = 1|\boldsymbol{\kappa}_{\nu_{ij}})\right] = \ln\left(\frac{p(\kappa_{ij}|\boldsymbol{\kappa}_{\nu_{ij}})}{1 - p(\kappa_{ij} = 1|\boldsymbol{\kappa}_{\nu_{ij}})}\right) = \theta(\xi(\boldsymbol{\kappa}, \tau, (i, j))), \tag{11}$$

where $\xi(\boldsymbol{\kappa}, \tau, (i, j)) \subseteq \tau$ is the set of elements $(k, l) \in \tau$ associated to a node with oil sand in the sequential neighborhood for node (i, j), i.e.

$$\xi(\boldsymbol{\kappa}, \tau, (i, j)) = \{(k, l) \in \tau : (i + k, j + l) \in G \text{ and } \kappa_{i+k, j+l} = 1\}, \tag{12}$$

and $\theta(\cdot)$ is a pseudo-Boolean function (Hammer and Holzman, 1992; Grabisch et al., 2000) to be specified. One should note that, as we assume the same function $\theta(\cdot)$ for all nodes $(i, j) \in G$ we get a homogeneous model. Moreover, one should note that the definition of $\xi(\boldsymbol{\kappa}, \tau, (i, j))$ implies that for nodes (i, j) close to the boundary of the lattice, so that $(\tau \oplus (i, j)) \setminus G \neq \emptyset$, the conditional distribution $p(\kappa_{ij}|\boldsymbol{\kappa}_{\nu_{ij}})$ becomes as if one had an infinite lattice where all variables associated to nodes outside of G were zero.

The last step in specifying the Markov mesh model is to choose the function $\theta(\cdot)$. This is a real valued function, where the argument is a subset of τ specifying for which sequential neighbors the associated binary variable is equal to one. Without loss of generality, the $\theta(\cdot)$ can be uniquely expressed in terms of a collection of interaction parameters $\{\beta(\lambda) : \lambda \subseteq \tau\}$ by

$$\theta(\lambda) = \sum_{\lambda^\star \subseteq \lambda} \beta(\lambda^\star) \quad \text{for } \lambda \subseteq \tau. \tag{13}$$

The number of interaction parameters is $2^{|\tau|}$, where $|\tau|$ is the number of elements in τ. Unless $|\tau|$ is very small the number of parameters necessary to specify $\theta(\cdot)$ is thereby very large. We still follow Luo and Tjelmeland (2017) and limit the number of model parameters

Figure 4: The sequential neighborhood (left) and the corresponding Markov random field neighborhood (right) for the fitted prior $p(\boldsymbol{\kappa})$. The nodes marked with a dot are (sequential) neighbors of the node marked with a cross.

by restricting many of the interaction parameters to be zero. More specifically, for some $\Lambda \subseteq \Omega(\tau)$, where $\Omega(\tau)$ is the set of all subsets of τ, we assume $\beta(\lambda) = 0$ for all $\lambda \notin \Lambda$. Thus, we specify the Markov mesh model by choosing the sets τ and Λ and the interaction values $\{\beta(\lambda) : \lambda \in \Lambda\}$.

To fit the Markov mesh model specified above to the training image in Figure 2 we adopt the Bayesian procedure introduced in Luo and Tjelmeland (2017), including the hyper-parameter values used in that article. A prior is specified for τ, Λ and $\{\beta(\lambda) : \lambda \in \Lambda\}$ and assuming the training image to be a sample from the specified Markov mesh model, a Metropolis–Hastings algorithm is used to generate samples from the resulting posterior distribution for τ, Λ and $\{\beta(\lambda) : \lambda \in \Lambda\}$. When we conditioned on the training image in Figure 2, the convergence of the Metropolis–Hastings algorithm was so slow that we were unable to obtain convergence within a reasonable computation time. As a pragmatic approach to obtain a reasonable prior $p(\boldsymbol{\kappa})$ we simply run the Metropolis–Hastings algorithm in Luo and Tjelmeland (2017) for a large number of iterations and used the last values for τ, Δ and $\{\beta(\lambda) : \lambda \in \Lambda\}$ in this run to define the prior $p(\boldsymbol{\kappa})$. The resulting $p(\boldsymbol{\kappa})$ prior has $|\tau| = 9$ sequential neighbors and $|\Delta| = 31$ interaction parameters that are allowed to differ from zero. The sequential neighborhood τ is illustrated in Figure 4, while the complete specification of τ, Δ and $\{\beta(\lambda) : \lambda \in \Lambda\}$ is given in an Appendix.
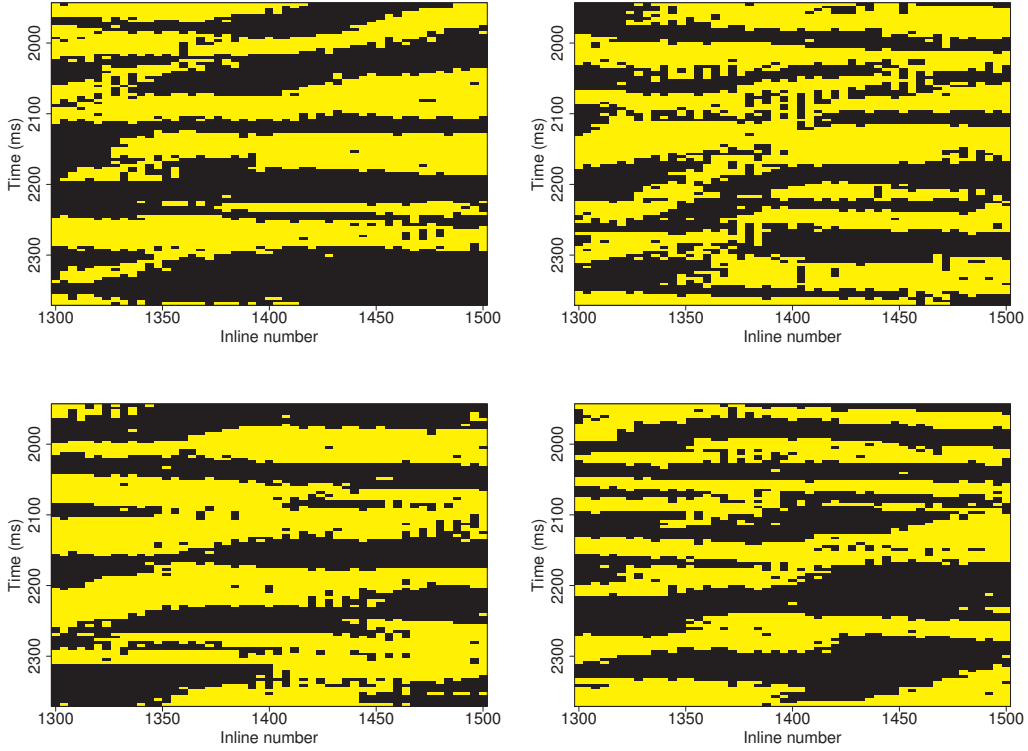
Figure 5: Four independent realisations from the Markov mesh prior fitted to the training image shown in Figure 2. Black and yellow represent shale and oil sand, respectively.

The best way to understand the properties of the prior is perhaps to look at realisations sampled from $p(\boldsymbol{\kappa})$, four of which are shown in Figure 5. We see that the fitted prior is reproducing large continuous areas of shale and oil sand as seen in the training image, but the boundaries between shale and oil sand is less horizontal in the realisations from the prior than in the training image.

Profile Markov random field prior

The profile Markov random field prior was first defined and used for seismic inversion in Ulvmoen and Omre (2010), see also Rimstad and Omre (2010). Even though the prior class

14

is defined for categorical variables, in our description of the model we limit the attention to the binary variable case.

Let again $G = \{(i,j)|i = 1,\ldots,n_1; j = 1,\ldots,n_2\}$ be a rectangular lattice, where in each node we associate a binary variable $\kappa_{ij} \in \{0,1\}$. We let $C_j = \{(i,j) : i = 1,\ldots,n_1\}$ be the set of nodes in profile or column j of the lattice G and let $\boldsymbol{\kappa}_{C_j} = (\kappa_{ij} : (i,j) \in C_j)$ denote the collection of the binary variables associated to this column. The collection of all the binary variables except the ones in column j we denote by $\boldsymbol{\kappa}_{-C_j}$. The profile Markov random field prior is then specified by first adopting the Markov property

$$p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j}) = p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{C_{j-1}},\boldsymbol{\kappa}_{C_{j+1}}), \tag{14}$$

i.e. given the values in columns $j-1$ and $j+1$, the values in column j are independent of the values in the remaining columns. Secondly, the profile Markov random field prior assumes $p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{C_{j-1}},\boldsymbol{\kappa}_{C_{j+1}})$ to be a Markov chain down along the column,

$$p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{C_{j-1}},\boldsymbol{\kappa}_{C_{j+1}}) = p(\kappa_{(1,j)}|\kappa_{(1,j-1)},\kappa_{(1,j+1)}) \times \prod_{i=2}^{n} p(\kappa_{(i,j)}|\kappa_{(i-1,j)},\kappa_{(i,j-1)},\kappa_{(i,j+1)}), \tag{15}$$

where the conditional distribution $p(\kappa_{(i,j)}|\kappa_{(i-1,j)},\kappa_{(i,j-1)},\kappa_{(i,j+1)})$ is the same for all values of i and j. The values we have used for the transition probabilities are defined in Table 1. Rimstad and Omre (2010) describe the structure used to specify these values. The basic idea is that these values should represent high probability for lateral continuity of oil sand and shale. The initial distribution $p(\kappa_{(1,j)}|\kappa_{(1,j-1)},\kappa_{(1,j+1)})$ are set equal to $p(\kappa_{(i,j)}|\kappa_{(i-1,j)} = 0,\kappa_{(i,j-1)},\kappa_{(i,j+1)})$, i.e. conditioning on shale being present above the lattice. Correspondingly, $p(\kappa_{(i,j)}|\kappa_{(i-1,j)},\kappa_{(i,j-1)},\kappa_{(i,j+1)})$ for the left and rightmost columns $j = 1$ and $j = n$ are defined by conditioning on shale being present outside the lattice.

Table 1: Values used for $p(\kappa_{(i,j)}|\kappa_{(i-1,j)}, \kappa_{(i,j-1)}, \kappa_{(i,j+1)})$ in the specification of the profile Markov random field prior.

$\kappa_{i,j-1} = 0, \kappa_{i,j+1} = 0$

	$\kappa_{ij} = 0$	$\kappa_{ij} = 1$
$\kappa_{i-1,j} = 0$	0.9877	0.0123
$\kappa_{i-1,j} = 1$	0.8339	0.1661

$\kappa_{i,j-1} = 0, \kappa_{i,j+1} = 1$

	$\kappa_{ij} = 0$	$\kappa_{ij} = 1$
$\kappa_{i-1,j} = 0$	0.6539	0.3461
$\kappa_{i-1,j} = 1$	0.1056	0.8944

$\kappa_{i,j-1} = 1, \kappa_{i,j+1} = 0$

	$\kappa_{ij} = 0$	$\kappa_{ij} = 1$
$\kappa_{i-1,j} = 0$	0.6539	0.3461
$\kappa_{i-1,j} = 1$	0.1056	0.8944

$\kappa_{i,j-1} = 1, \kappa_{i,j+1} = 1$

	$\kappa_{ij} = 0$	$\kappa_{ij} = 1$
$\kappa_{i-1,j} = 0$	0.0425	0.9575
$\kappa_{i-1,j} = 1$	0.0028	0.9972

Posterior model and simulation algorithm

For each of the Markov mesh and profile Markov random field priors we obtain a posterior distribution for the lithology/fluid classes $\boldsymbol{\kappa}$, given in (1). To explore and estimate properties of the two posterior distributions we adopt the Metropolis–Hastings algorithm (Gilks et al., 1996; Robert and Casella, 1999; Gamerman and Lopes, 2006). Since the wavelets in the likelihood model induce strong dependencies between different κ_{ij}'s in the same column, a simple single-site updating scheme would give a Markov chain with a long burn-in and slow mixing. We therefore adopt the proposal scheme previously used in Rimstad and Omre (2010) and propose in each iteration new values for all lithology/fluid classes in one column. Using notation from the discussion of the profile Markov random field prior, the joint full conditional for the lithology/fluid classes in column j is

$$p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j}, \boldsymbol{d}) \propto p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j})p(\boldsymbol{d}|\boldsymbol{\kappa}). \tag{16}$$

To sample from this distribution is, however, computationally very expensive due to the long range dependencies in $\boldsymbol{\kappa}_{C_j}$ induced by the wavelets in the likelihood model. Still following Rimstad and Omre (2010) we therefore adopt the approximation scheme specified in Larsen et al. (2006) to construct an approximation $p_\nu^\star(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j}, \boldsymbol{d})$ to $p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j}, \boldsymbol{d})$, where ν is an algorithmic tuning parameter, and generate potential new values for $\boldsymbol{\kappa}_{C_j}$ by sampling from $p_\nu^\star(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j}, \boldsymbol{d})$. The $p_\nu^\star(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j}, \boldsymbol{d})$ is a higher-order Markov chain and thereby by construction easy to sample from. In general the approximation quality grows with ν, but so does the computation time required for simulating one realisation from $p_\nu^\star(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j}, \boldsymbol{d})$. Based on preliminary runs of the Metropolis–Hastings algorithm we find a value for ν which gives reasonable acceptance rates for the Metropolis–Hastings algorithm.

To run the Metropolis–Hastings scheme discussed above we first need to have expres-

sions for $p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j})$ for each of the two priors. For the profile Markov random field prior this is by construction given by (14) and the values in Table 1. To obtain $p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j})$ for the Markov mesh prior we first need to reformulate the Markov mesh model as a Markov random field. The resulting Markov random field has a neighborhood system where the set of neighbors to node $(i,j) \in G$ is

$$\partial_{ij} = \nu_{ij} \cup \left(\bigcup_{(k,l) \in G : (i,j) \in \nu_{kl}} (\nu_{kl} \cup \{(k,l)\}) \right). \tag{17}$$

For nodes sufficiently far away from the lattice borders the ∂_{ij} becomes as shown in Figure 4(b). When the prior $p(x)$ is formulated as a Markov random field it is straightforward to find the corresponding $p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j})$ by first ignoring potential functions for cliques which do not include any node in C_j and thereafter plugging in values for x_{-C_j} in the remaining potential functions. In particular, for the Markov mesh sequential neighborhood used in this study, $p(\boldsymbol{\kappa}_{C_j}|\boldsymbol{\kappa}_{-C_j})$ becomes a third-order Markov chain.

The second factor in (16) is a high dimensional multivariate Gaussian density. To be able to evaluate this efficiently it is essential that we have chosen the correlation structure of $\boldsymbol{d}|\boldsymbol{\kappa}$ to be separable. For each of the two priors we run the Metropolis–Hastings algorithm scheme for the resulting posterior distribution for a large number of iterations. We use standard output analysis to identify and discard a burn-in period. In the next section we use the $\boldsymbol{\kappa}$ realisations after the burn-in period to estimate and compare properties of the two posterior distributions.

North Sea Case study

Recall that the objective is to assess the posterior of the lithology/fluid classes $\kappa_{ij} \in G$ given seismic AVO data $\boldsymbol{d}$ in a clastic oil reservoir in the North Sea. That is, we want to

assess the posterior $p(\boldsymbol{\kappa}|\boldsymbol{d})$ given in (1), for the two prior models discussed earlier. Note that the two posteriors will not be identical since the priors are different.

To study and compare the properties of the two posterior distributions we can first look at the posterior realisations shown in Figure 6. The left and right columns show four realisations from each of the two posteriors. Realisations from the posterior when using the Markov mesh prior are shown in the left column, whereas the realisations in the right column is based on a model with the profile Markov random field prior. The eight realisations are quite similar, but when studying them in more detail one can observe that with the Markov mesh prior there seems to be more skewed and curved structures than when using the profile Markov random field prior. Since the Markov mesh prior has much larger neighborhoods than the profile Markov random field prior, this is not really surprising. With larger neighborhoods, and corresponding larger cliques, it becomes possible for the model to identify skewed and curved structures.

The upper row in Figure 7 shows the result of estimating in each node the marginal posterior probability of oil sand. In each node the probability is estimated as the fraction of the realisations where the node has oil sand. Again the left and right images are results when using the Markov mesh and profile Markov random field priors, respectively. The two probability maps are similar, but somewhat more continuity of skewed and curved high probability areas can be observed when using the Markov mesh prior. In the middle row of Figure 7 the probabilities in the upper row is rounded to the nearest integer to get an estimate of the most probable lithology/fluid class in each node. Again we can observe somewhat more continuity of skewed and curved oil sand areas when using the Markov mesh prior. The histograms in the lower row of Figure 7 are simply probability histograms of the estimated marginal posterior probabilities shown in the upper row of the same figure. We can observe that when using the profile Markov random field prior, somewhat more
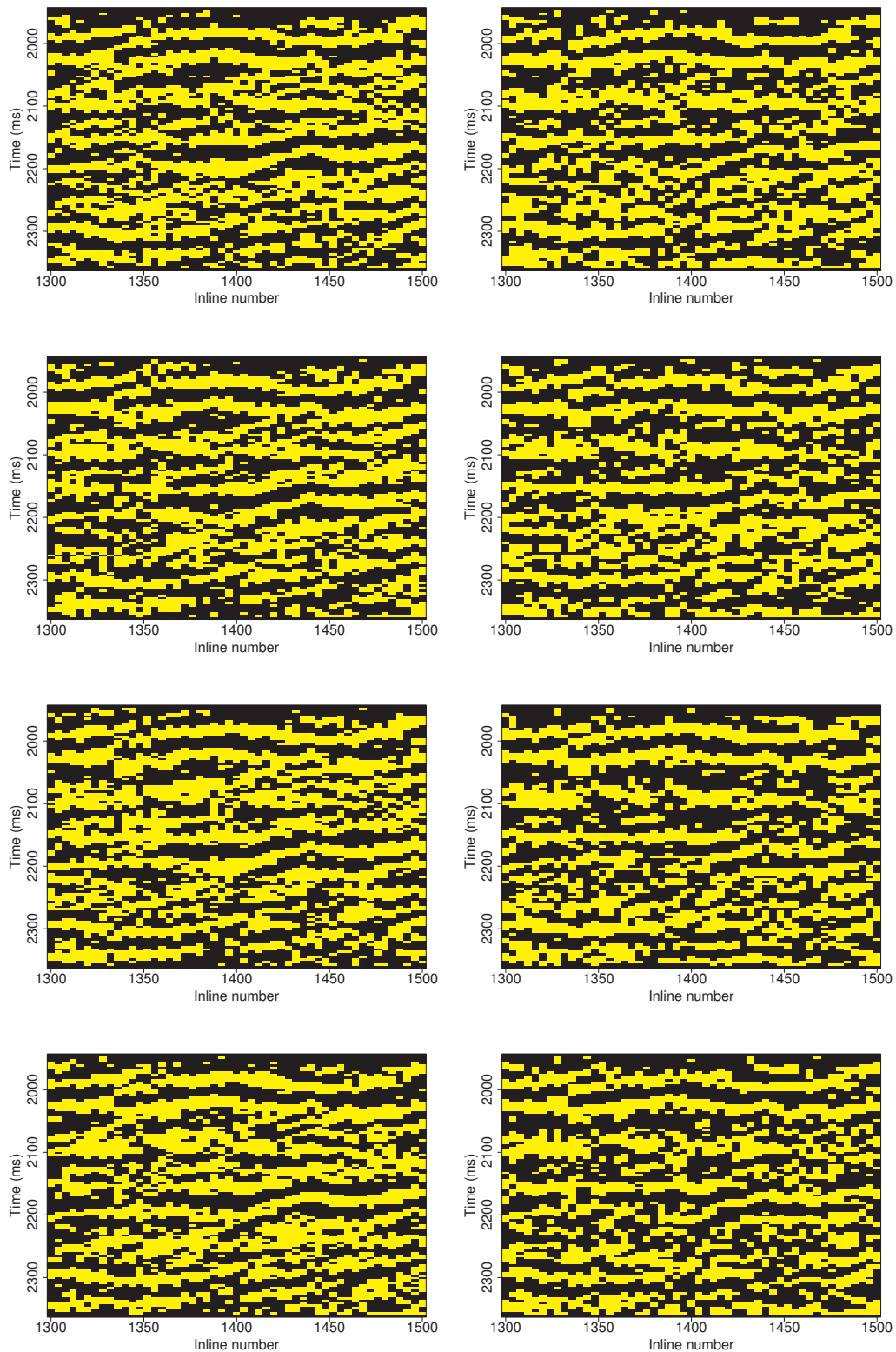
Figure 6: The left column shows four independent realisations from the posterior distribution when using the Markov mesh prior. The right column shows correspondingly four independent realisations from the posterior distribution when using the profile Markov random field prior. Black and yellow represent shale and oil sand, respectively.
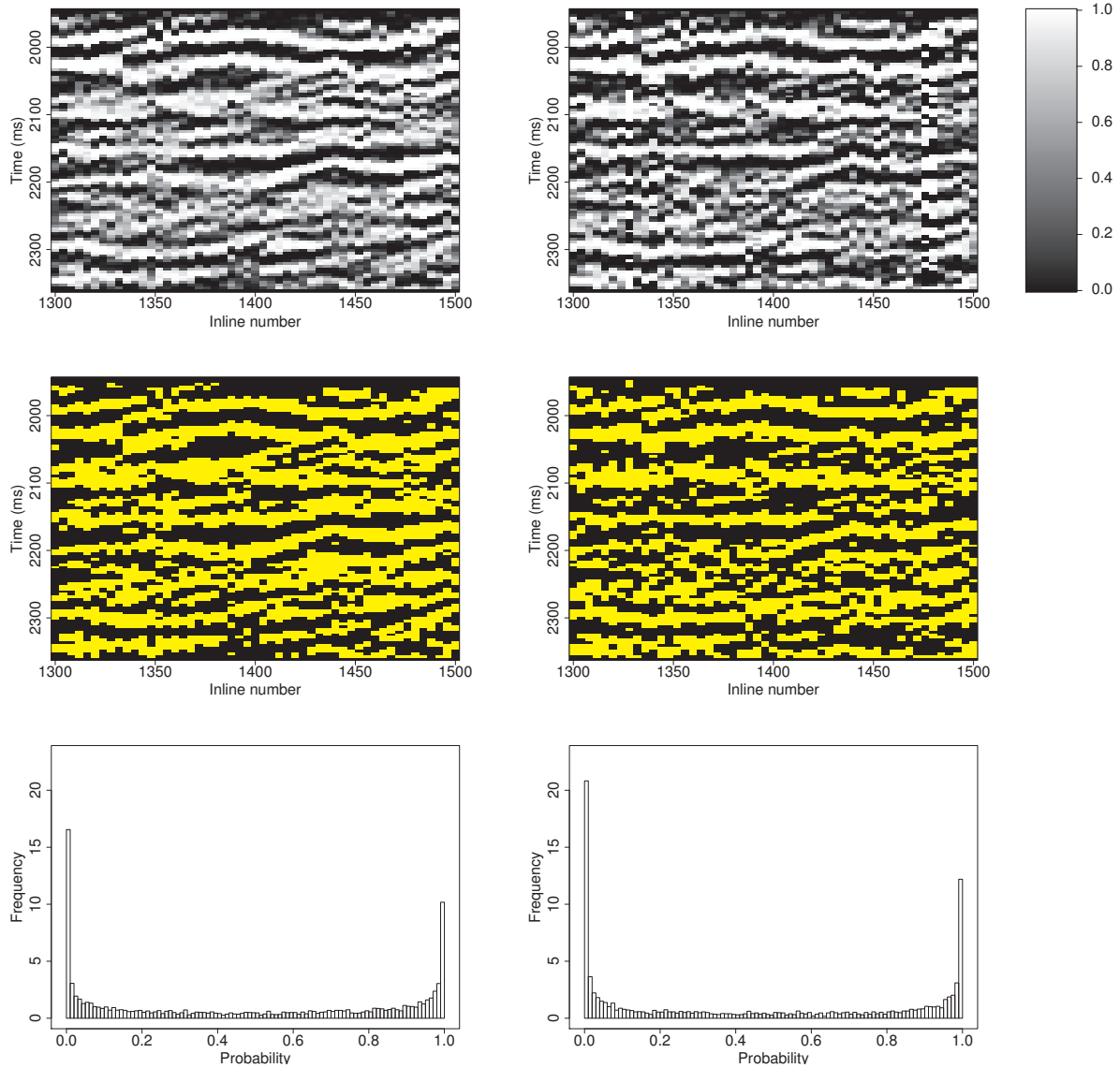
Figure 7: Upper row: Estimated posterior marginal probabilities for oil sand when using the Markov mesh (left) and the profile Markov random field (middle) priors. The colour scale is shown in the rightmost plot. Middle row: Estimated marginal posterior mode for each node. Black and yellow represent shale and oil sand, respectively. Lower row: Probability histograms of estimated posterior marginal probabilities when using the Markov mesh (left) and the profile Markov random field (right) priors.
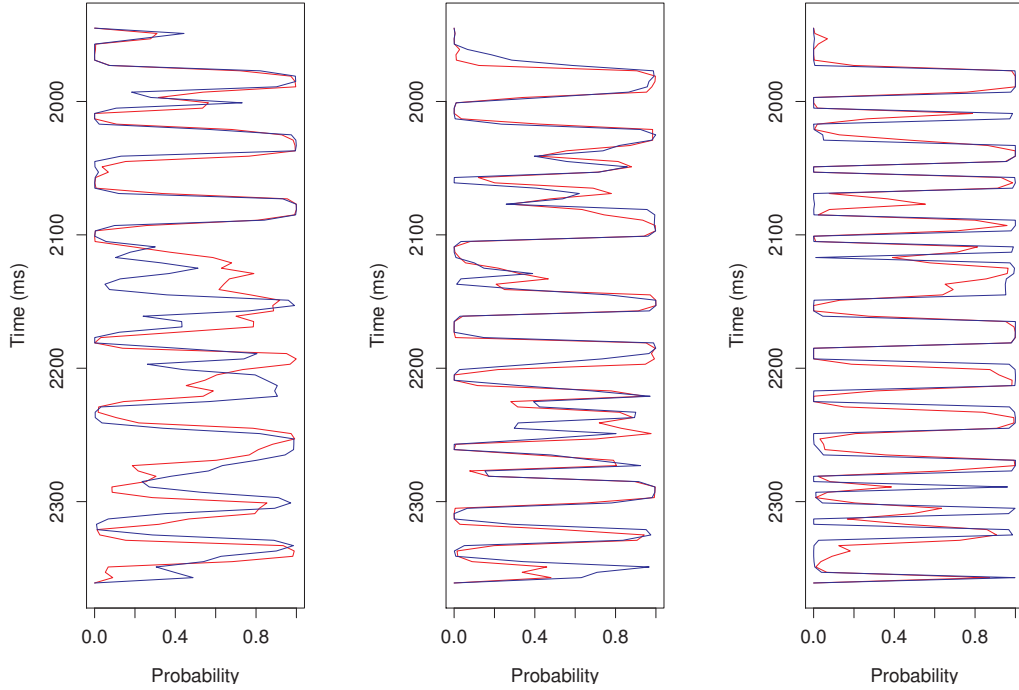
Figure 8: Estimated marginal probabilities for oil sand in three traces. Results when using the Markov mesh and profile Markov random field priors are shown by red and blue lines, respectively. The left, middle and right plots show the marginal probabilities in traces $j = 15$, $j = 30$ and $j = 45$, respectively.

marginal posterior probabilities are close to zero and one than when using the Markov mesh prior.

To study the marginal probabilities a little more we have chosen three traces, or columns, $j = 15$, 30 and 45, and in Figure 8 plotted the marginal probabilities. The estimated posterior marginal probabilities when using the Markov mesh and profile Markov random field priors are plotted in red and blue, respectively. More than in Figure 7 we can here see how close the two posterior probabilities are for most of the nodes. In a few of the

nodes, however, the difference is quite clear.

The continuity of oil sand is very important for fluid flow in a petroleum reservoir. We can get some understanding of how the prior influences this continuity by studying Figure 7, but to study the continuity in more detail we need to summarise how this continuity is in each posterior realisation. To do this, we have manually picked four nodes with very high posterior probability for oil sand both when using the Markov mesh and the profile Markov random field priors. These four nodes are marked with a red bullet in Figure 9, one row for each of the four chosen nodes. For each of these four nodes and for each posterior realisation we identified all other nodes with oil sand which through other oil sand nodes had contact with the chosen node. Thereby we could estimate the posterior probability that any node was in contact with the chosen node as the fraction of the realisations where this occurred. The resulting estimated probabilities are visualised in Figure 9. The left and right columns are again the results when using the Markov mesh and the profile Markov random field priors, respectively. In the three upper rows we can see a lot more continuity in the posterior realisations when using the Markov mesh prior than when using the profile Markov random field prior. In the lower row the situation is for some reason reversed. To study this type of continuity more generally, not only for the four hand picked nodes used in Figure 9, we finally repeat the exercise of finding all nodes in a realisation with oil sand connected to a particular node, but now the particular node is sampled at random among all nodes with oil sand. For each realisation and each particular node we find the number of oil sand nodes connected to the particular node. In Figure 10 we show the resulting estimated posterior probabilities for the randomly chosen particular node to be connected to more than η other oil sand nodes, as a function of η. The red and blue curves are the results when using the Markov mesh and the profile Markov random field priors, respectively. We see that the curve related to the Markov mesh prior lies consistently
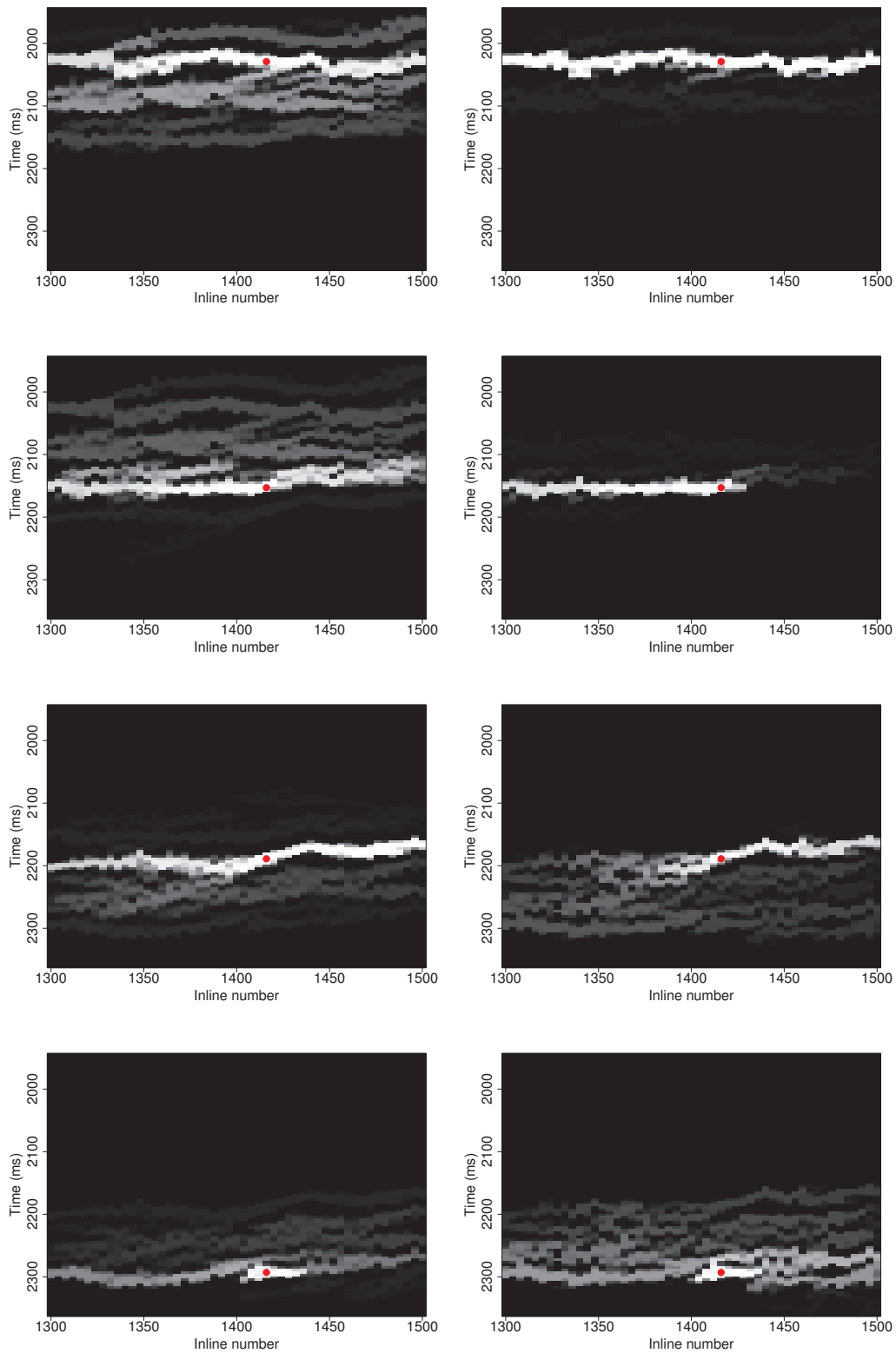
Figure 9: For each node, estimated posterior probabilities that there is contact (via other oil sand nodes) between this node and the node marked with a red filled circle. The plots in the left and right columns are when using the Markov mesh and the profile Markov random field priors, respectively. The colour scale is as defined by the legend in Figure 7.
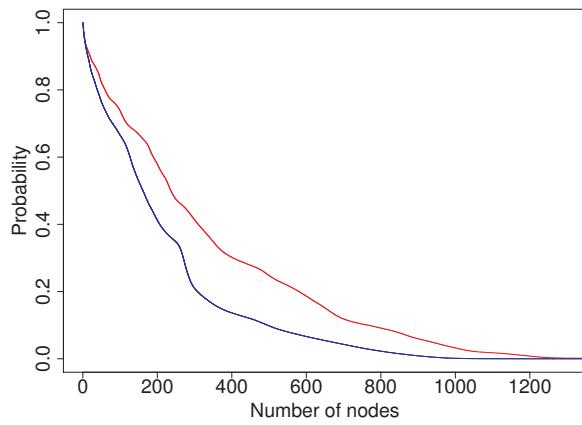
Figure 10: As a function of a number of nodes η, estimated posterior probability for a random oil sand node to be connected (via other oil sand nodes) to at least η other oil sand nodes. The η is along the x-axis and the estimated probability is along the y-axis. Results when using the Markov mesh and profile Markov random field priors are shown in red and blue, respectively.

clearly above the curve related to the profile Markov random field prior, showing that the Markov mesh prior produces more posterior continuity of oil sand than the profile Markov random field prior.

Discussion

In this article we have, for a particular seismic data set in the North Sea and two particular prior models for the lateral connectivity of the lithology/fluid classes, studied how the prior influences the posterior properties. When focusing on the posterior marginal probabilities we found, for most nodes, that the prior had little influence. When focusing on posterior continuity of oil sand, however, we found that the prior had a quite strong influence on the results. Not surprisingly, the prior with the largest neighborhoods produced the largest posterior continuity. When evaluating whether to use a simple prior with a small neighborhood or to use a more complicated prior with a larger neighborhood one should therefore first decide what posterior properties that are of interest. If the focus is only on the posterior marginal probabilities, a simple prior is perhaps sufficient. If the focus is on fluid flow, however, spatial continuity is crucial and it may be beneficial with a more complicated prior which is better able to capture spatial continuity.

When deciding what prior to use one should also take into account the computational resources necessary to simulate from the posterior distribution. A more complicated prior typically gives a posterior which requires more computation time to explore. With our implementations, sampling from the posterior when using the Markov mesh prior required approximately 20 times more computation time compared to when using the profile Markov random field prior. However, our implementation of the sampling when using the Markov mesh prior was partly in Matlab and partly in C++ and a lot of the computation time here was just overhead in the communication between Matlab and C++. Our implementation

of the sampling algorithm when using the profile Markov random field prior was entirely in Matlab, so we did not have the same overhead in this case. If we had implemented also the sampling algorithm when using the Markov mesh prior entirely in Matlab we expect this algorithm would have required a factor between 3 and 5 more computation time than that for the profile Markov random field prior. That the sampling when using the Markov mesh prior requires more computation time than when using the profile Markov random field prior should come as no surprise, since the Markov mesh prior has a much larger neighborhood than the profile Markov random field prior.

The study presented in this article is quite limited. The model includes only two lithology/fluid classes and we have considered only one seismic section. It is of interest also to study the effect of using a prior with a larger neighborhood when the model represent more than two lithology/fluid classes. The profile Markov random field prior is already defined with more than two lithology/fluid classes and the Markov mesh construction used here can easily be extended to such a situation. To condition on a seismic cube is also of interest. Then the lithology/fluid classes need to be represented on a three dimensional lattice and the prior models need to be defined for such a situation. Again the profile Markov random field prior is already formulated in such a situation. The Markov mesh formulation used here can also be extended to a three dimensional lattice, but the effectiveness of such a formulation need to be further studied.

Acknowledgments

References

Abend, K., Harley, T., and Kanal, L. (1965). "Classification of binary random patterns." *IEEE Transactions on Information Theory*, 11, 538–544.

Aki, K. and Richards, P. G. (1980). *Quantitative Seismology: Theory and Methods*. New York: W. H. Freeman and Co.

Arnesen, P. and Tjelmeland, H. (2017). "Prior specification of neighbourhood and interaction structure in binary Markov random fields." *Statistics and Computing*, 27, 737–756.

Aster, R., Borchers, B., and Thurber, C. H. (2011). *Parameter estimation and inverse problems*. Amsterdam: Elsevier.

Buland, A. and Omre, H. (2003). "Bayesian linearized AVO invserion." *Geophysics*, 68, 185–198.

Connolly, P. and Hughes, M. (2016). "Stochastic inversion by matching to large numbers of pseudo-wells." *Geophysics*, 81, 2, M7–M22.

Cressie, N. and Davidson, J. (1998). "Image analysis with partially ordered Markov models." *Computational Statistics and Data Analysis*, 29, 1–26.

Eidsvik, J., Mukerji, T., and Switzer, P. (2004). "Estimation of geological attributes from a well log: An application of hidden Markov chains." *Mathematical Geology*, 36, 379–397.

Emery, X. and Lantuéjoul, C. (2014). "Can a training image be a substitute for a random field model?" *Mathematical Geosciences*, 46, 133–147.

Fjeldstad, T. and Omre, H. (2017). "Bayesian inversion of convolved hidden Markov models with applications in reservoir prediction." Tech. rep., ArXiv e-prints 1710.06613v1, Available from http://arxiv.org/abs/1710.06613v1.

Gamerman, D. and Lopes, H. F. (2006). *Markov chain Monte Carlo: Stochastic simulation for Bayesian inference*. 2nd ed. London: Chapman & Hall/CRC.

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996). *Markov chain Monte Carlo in practice*. London: Chapman & Hall.

Grabisch, M., Marichal, J. L., and Roubens, M. (2000). "Equivalent representations of set functions." *Mathematics of Operations Research*, 25, 157–178.

Grana, D. and Della Rossa, E. (2010). "Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion." *Geophysics*, 75, 3, O21–O37.

Grana, D., Fjelstad, T., and Omre, H. (2017). "Bayesian Gaussian mixture linear inversion for geophysical inverse problems." *Mathematical Geosciences*, 49, 493–515.

Guardiano, F. and Srivastava, R. (1993). "Multivariate geostatistics: beyond bivariate moments." In *Geostatistics Tróia '92*, ed. A. Soares, 133–144. Kluwer, Dordrecht. Proceedings to the 1992 International Geostatistics Congress, Tróia, Portugal.

Gunning, J. and Glinsky, M. E. (2007). "Detection of reservoir quality using Baysian seismic inversion." *Geophysics*, 72, R37–R39.

Hammer, P. L. and Holzman, R. (1992). "Approximations of pseudo-Boolean functions; applications to game theory." *Methods and Models of Operation Research*, 36, 3–21.

Hurn, M., Husby, O., and Rue, H. (2003). "A tutorial on image analysis." In *Spatial statistics and computational methods*, ed. J. Møller, vol. 173 of *Lecture Notes in Statistics*, 87–139. Springer.

Journel, J. and Zhang, T. (2006). "The necessity of a multiple-point prior model." *Mathematical Geology*, 38, 591–610.

Kindermann, R. and Snell, J. L. (1980). *Markov random fields and their applications*. Providence, R.I.: American Mathematical Society.

Lang, X. and Grana, D. (2017). "Geostatistical inversion of prestack seismic data for the joint estimation of facies and impedances using stochastic sampling from Gaussian mixture posterior distribution." *Geophysics*, 82, M55–M65.

Larsen, A. L., Ulvmoen, M., Omre, H., and Buland, A. (2006). "Bayesian lithology/fluid prediction and simulation on the basis of a Markov-chain prior model." *Geophysics*, 71, R69–R78.

Luo, X. and Tjelmeland, H. (2017). "Prior specification for binary Markov mesh models." Tech. rep., ArXiv e-prints 1707.08339v1, Available from http://arxiv.org/abs/1707.08339v1.

Mariethoz, G. and Caers, J. (2014). *Multiple-point geostatistics: Stochastic modeling with training images*. John Wiley & Sons.

Rimstad, K., Avseth, P., and Omre, H. (2012). "Hierarchical Bayesian lithology/fluid prediction: A North Sea case study." *Geophysics*, 77, B69–B856.

Rimstad, K. and Omre, H. (2010). "Impact of rock-physics depth trends and Markov random fields on hierarchical Bayesian lithology/fluid prediction." *Geophysics*, 75, R93–R108.

Robert, C. P. and Casella, G. (1999). *Monte Carlo statistical methods*. Berlin: Springer.

Sen, M. K. and Stoffa, P. L. (2013). *Global optimization methods in geophysical inversion*. Cambridge: Cambridge University Press.

Stien, M. and Kolbjørnsen, O. (2011). "Facies Modeling Using a Markov Mesh Model Specification." *Mathematical Geosciences*, 43, 6, 611.

Strebelle, S. (2002). "Conditional simulation of complex geolgical structures using multiple-point statistics." *Mathematical Geology*, 34, 1–21.

Tarantola, A. (2005). *Inverse problem theory*. Philidelphia: SIAM.

Toftaker, H. and Tjelmeland, H. (2013). "Construction of binary multi-grid Markov random field prior models from training images." *Mathematical Geosciences*, 45, 383–409.

Ulvmoen, M. and Omre, H. (2010). "Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations: Part 1-Methodology." *Geophysics*, 75, R21–R35.

Zhang, T., Pedersen, S. I., Knudby, C., and McCormick, D. (2012). "Memory-efficient categorical multi-point statistics algorithms based on compact search trees." *Mathematical Geosciences*, 44, 863–879.

Fitted Markov mesh prior, $p(\boldsymbol{\kappa})$

The Markov mesh model fitted to the training image in Figure 2 has template sequential neighborhood $\tau = \{(-1,0), (0,-1), (-1,2), (0,-2), (-3,-1), (0,-3), (-1,4), (0,-4), (-2,-4)\}$ and Δ and $\{\beta(\lambda) : \lambda \in \Lambda\}$ are as specified in Table 2.

Table 2: The elements λ in the set Λ and the associated interaction parameters $\beta(\lambda)$.

$\lambda \in \Lambda$	$\beta(\lambda)$
$\emptyset$	-4.33884
$\{(-1,0)\}$	3.27479
$\{(0,-1)\}$	2.96595
$\{(-1,0),(0,-1)\}$	-0.460735
$\{(-1,2)\}$	1.49237
$\{(-1,2),(0,-1)\}$	-1.10759
$\{(0,-2)\}$	1.99035
$\{(-3,-1)\}$	-1.43573
$\{(0,-3)\}$	3.06786
$\{(-1,0),(0,-3)\}$	-3.44258
$\{(0,-3),(0,-1)\}$	-2.03335
$\{(-1,0),(0,-3),(0,-1)\}$	1.95605
$\{(0,-3),(0,-2)\}$	-1.02729
$\{(-1,4)\}$	2.90431
$\{(-1,0),(-1,4)\}$	-3.42674
$\{(-1,4),(0,-1)\}$	-0.404195
$\{(-1,2),(-1,4)\}$	0.268767
$\{(-1,4),(0,-3)\}$	-2.73426
$\{(-1,0),(-1,4),(0,-3)\}$	2.96929
$\{(-1,4),(0,-3),(0,-1)\}$	1.95346
$\{(0,-4)\}$	2.1858
$\{(-1,0),(0,-4)\}$	-0.355664
$\{(0,-4),(0,-2)\}$	-1.61185
$\{(0,-4),(0,-3)\}$	-1.23267
$\{(-1,0),(0,-4),(0,-3)\}$	0.606075
$\{(0,-4),(0,-3),(0,-2)\}$	2.03717
$\{(-1,4),(0,-4)\}$	-4.01512
$\{(-1,0),(-1,4),(0,-4)\}$	3.80173
$\{(-1,4),(0,-4),(0,-3)\}$	2.6053
$\{(-1,0),(-1,4),(0,-4),(0,-3)\}$	-1.64379
$\{(-2,-4)\}$	-0.717159