Bayesian closed-skew Gaussian inversion of seismic AVO data into elastic material properties

Omid Karimi[1,2], Henning Omre[2], Mohsen Mohammadzadeh[1]

[1]Department of statistics, Tarbiat Modares University, Tehran, Iran.

E-mail: Karimio@modares.ac.ir; mohsen_ m@modares.ac.ir

[2]Department of Mathematical sciences, Norwegian University of Science & Technology,

Trondheim, Norway, www.math.ntnu.no.

E-mail: omre@math.ntnu.no

February 15. 2009

Abstract

Bayesian closed-skew Gaussian inversion is defined as a generalization of traditional Bayesian Gaussian inversion, which is frequently used in seismic AVO inversion. The new model captures skewness in the variables of interest, hence the posterior model for log-transformed elastic properties given seismic AVO data will be a non-Gaussian skew pdf. The model is analytically tractable, and this opens for use on high-dimensional 3D inversion problems. Assessment of the posterior models of high dimensions require numerical approximations, however. The Bayesian closed-skew Gaussian inversion approach is demonstrated on real elastic properties from a well from the Sleipner field. Comparison with results from traditional Bayesian Gaussian inversion shows that the mean-square-error of (v_p, v_s)-predictions are reduced by factor two - somewhat less for ρ-predictions. No significant conclusions concerning the coverage of the 0.8-prediction intervals can be made.

1 Introduction

Inversion of seismic amplitude versus offset (AVO) data into elastic material properties is an important component in petroleum reservoir characterization. The seismic data must be prestack migrated and offset to angle converted prior to the inversion. Based on the inverted elastic properties, reservoir characteristics like lithology-fluid and porosity-permeability can be assessed. Since seismic inversion is an crucial intermediate step in the reservoir evaluation process, the quality of the inversion has large impact on the evaluation.

The inversion is frequently cast in a Bayesian setting, see Mosegaard and Tarantola (1995), Gouveia and Scales (1998) and Buland and Omre (2003). The assessment of the posterior model is crucial for the practical use of the approaches, since iterative simulation and optimization approaches are too computer demanding to be used in large scale 3D studies. The two former references presents models that can only be assessed by iterative algorithms. The Bayesian model presented in the latter reference is subject to analytical evaluation and hence extremely computer efficient, see also Buland *et al.* (2003).

The approach in Buland and Omre (2003) relies heavily on Gaussian assumptions, hence it can be termed Bayesian Gaussian inversion. The prior model of log-transformed elastic material properties must be Gaussian. The log-transform is needed to obtain a linear relation between seismic prestack data and elastic properties. The likelihood model is based on a convolved linearized Zoeppritz relation with an additive Gaussian error term. Hence the likelihood model is Gauss-linear. Under these assumption, the posterior model is Gaussian with associated model parameters analytically assessable from the model parameters of the prior and the likelihood model. Consequently, the posterior model can be determined even for high-dimensional 3D problems, although numerical approximations may be needed to assess the posterior model.

In the current study a Bayesian closed skew (CS) Gaussian inversion approach is introduced. The Gaussian assumptions in Buland and Omre (2003) are replaced by CS-Gaussian assumptions which can represent skewness in the variables involved. The CS-Gaussian model was introduced in González-Farías *et al.* (2004) and our work draws on their results. Under these CS-Gaussian assumptions, the posterior model is also CS-Gaussian with analytically tractable model parameters. More reliable seismic inversion results are expected since Bayesian CS-Gaussian inversion appears as a generalization of Bayesian Gaussian inversion, with the latter being a special case of the former. The CS-Gaussian approach can be used on high-dimensional 3D problems due to its analytical tractability, but numerical approximations are needed to assess the posterior model.

In the paper, vectors are denoted by bold, lower-case letter, for example $\boldsymbol{x}$, while matrices are upper-case letters, for example A. The term $p(\boldsymbol{x})$ is used as a generic term for probability density function (pdf) of the argument in appropriate dimension. The term $p(\boldsymbol{x}|\boldsymbol{o})$ represents the conditional pdf of $\boldsymbol{x}$ given $\boldsymbol{o}$.

Inversion Methodology

We use a Bayesian inversion terminology in this study, i.e

$$p(\boldsymbol{x}|\boldsymbol{o}) = const_o \times p(\boldsymbol{o}|\boldsymbol{x})p(\boldsymbol{x})$$

with $\boldsymbol{x}$ the variable of interest and $\boldsymbol{o}$ the associated observations. The prior model $p(\boldsymbol{x})$ captures general experience about $\boldsymbol{x}$, the likelihood model $p(\boldsymbol{o}|\boldsymbol{x})$ defines the relation between $\boldsymbol{x}$ and the observations $\boldsymbol{o}$, and $const_o$ is a normalizing constant dependent on the value of $\boldsymbol{o}$ only. The ultimate solution is the posterior model $p(\boldsymbol{x}|\boldsymbol{o})$ which may be simulated from or from which optimal predictions with associated prediction intervals can be computed for different loss functions. In Buland and Omre (2003) seismic AVO inversion is defined in a Bayesian setting using a Gaussian model for $p(\boldsymbol{x})$ and a Gauss-linear form on $p(\boldsymbol{o}|\boldsymbol{x})$. Consequently $p(\boldsymbol{x}|\boldsymbol{o})$ will be Gaussian with model parameters analytically obtainable from the model parameters of $p(\boldsymbol{x})$ and $p(\boldsymbol{o}|\boldsymbol{x})$. We denote this Bayesian Gaussian inversion. Note that $p(\boldsymbol{x}|\boldsymbol{o})$ is constrained to be Gaussian and hence symmetrical and light-tailed as the Gaussian pdf. In many problems, these constraints on the model are severe.

The current paper introduces the Bayesian CS-Gaussian inversion approach. This inversion model captures asymmetries or skewness in the distribution of the variable of interest $\boldsymbol{x}$ and in the errors related to the observation $\boldsymbol{o}$. Firstly, general characteristics of the multivariate CS-Gaussian pdf will be discussed, thereafter Bayesian CS-Gaussian inversion is defined.

The multivariate CS-Gaussian model was first introduced in González-Farías *et al.* (2004) and this reference provides a more formal exposure of the model. A CS-Gaussian $(n \times 1)$-vectorial random variable $\boldsymbol{x}$ is easiest defined through the associated $(n \times 1)$-vectorial $\boldsymbol{t}$ and $(q \times 1)$-vectorial $\boldsymbol{v}$ random variables. Let the joint pdf of $(\boldsymbol{t}, \boldsymbol{v})$ be Gaussian as:

$$\begin{bmatrix} \boldsymbol{t} \\ \boldsymbol{v} \end{bmatrix} \sim N_{n+q}(\begin{pmatrix} \boldsymbol{\mu}_t \\ \boldsymbol{\mu}_v \end{pmatrix}, \begin{pmatrix} \Sigma_t & \Gamma_{tv} \\ \Gamma_{vt} & \Sigma_v \end{pmatrix})$$

with expectations $\boldsymbol{\mu}_t$ and $\boldsymbol{\mu}_v$ and covariance matrices Σ_t, Σ_v and Γ_{vt} of proper dimensions. The variable of interest $\boldsymbol{x}$, is then defined by:

$$\boldsymbol{x} = [\boldsymbol{t}|\boldsymbol{v} \geq \boldsymbol{0}]$$

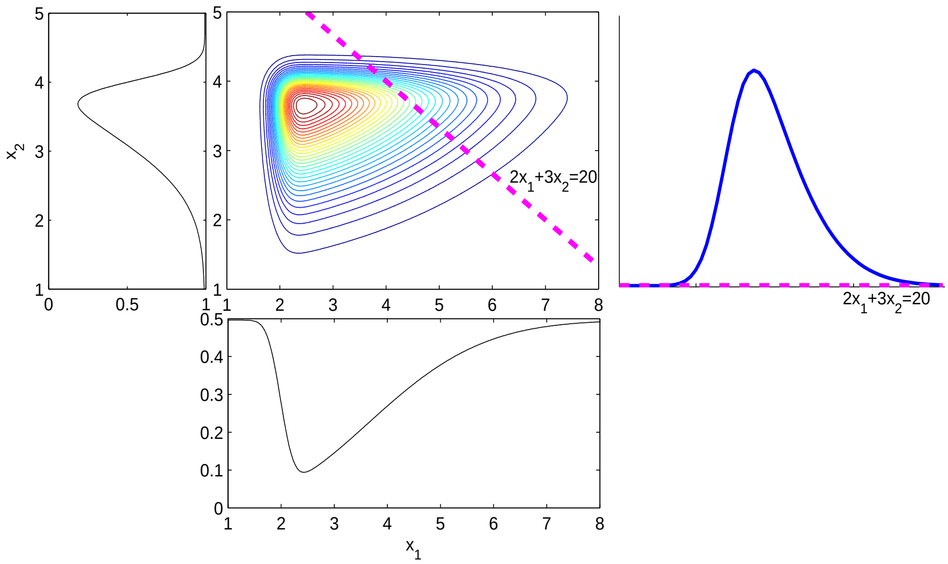

Figure 1: Bivariate CS-Gaussian example: bivariate pdf, marginal pdfs and conditional pdf.

which obviously has a skew pdf if $\Gamma_{vt} \neq \mathbf{0}$. The pdf of $\boldsymbol{x}$ is:

$$\begin{aligned} p(\boldsymbol{x}) = p(\boldsymbol{t}|\boldsymbol{v} \geq \mathbf{0}) &= \frac{p(\boldsymbol{v} \geq \mathbf{0}|\boldsymbol{t})p(\boldsymbol{t})}{p(\boldsymbol{v} \geq \mathbf{0})} \\ &= [1 - \Phi_q(\mathbf{0}; \boldsymbol{\mu}_v, \Sigma_v)]^{-1}[1 - \Phi_q(\mathbf{0}; \boldsymbol{\mu}_{v|t}, \Sigma_{v|t})]\phi_n(\boldsymbol{t}; \boldsymbol{\mu}_t, \Sigma_t) \end{aligned}$$

with

$$\boldsymbol{\mu}_{v|t} = \boldsymbol{\mu}_v + \Gamma_{vt}\Sigma_t^{-1}(\boldsymbol{t} - \boldsymbol{\mu}_t)$$
$$\Sigma_{v|t} = \Sigma_v - \Gamma_{vt}\Sigma_t^{-1}\Gamma_{tv}$$

and $\phi.(\cdot; \boldsymbol{\mu}, \Sigma)$ and $\Phi.(\cdot; \boldsymbol{\mu}, \Sigma)$ being the pdf and cumulative distribution function (cdf) of the Gaussian distribution, respectively. The usual parametrization of the CS-Gaussian distribution is:

$$\begin{aligned} \boldsymbol{x} \sim p(\boldsymbol{x}) &= CSN_{n,q}(\boldsymbol{\mu}, \Sigma, \Gamma, \boldsymbol{\nu}, \Delta) \\ &= [\Phi_q(\mathbf{0}; \boldsymbol{\nu}, \Delta + \Gamma\Sigma\Gamma')]^{-1}\Phi_q(\Gamma(\boldsymbol{x} - \boldsymbol{\mu}); \boldsymbol{\nu}, \Delta)\phi_n(\boldsymbol{x}; \boldsymbol{\mu}, \Sigma) \end{aligned}$$

where $\boldsymbol{\mu} = \boldsymbol{\mu}_t$, $\Sigma = \Sigma_t$, $\Gamma = \Gamma_{vt}\Sigma_t^{-1}$, $\boldsymbol{\nu} = -\boldsymbol{\mu}_v$, $\Delta = \Sigma_v - \Gamma_{vt}\Sigma_t^{-1}\Gamma_{tv}$. It is easy to demonstrate that this is identical to the expression above by using symmetries and standardization properties of Gaussian distribution. The dimension indicator q and the related $(q \times n)$-matrix Γ is closely related to the skewness of the pdf. The integer q can be interpreted as the degree of skewness freedom in the pdf. Note that for $q = 0$ or $\Gamma = \mathbf{0}$ the CS-Gaussian pdf coincide with the Gaussian pdf with model parameters $(\boldsymbol{\mu}, \Sigma)$. The CS-Gaussian random variable is obviously closely related to the Gaussian random variables by construction, and it has inherited several favorable characteristics to be presented below.

To demonstrate the characteristics, consider a particular CS-Gaussian distributed variable $\boldsymbol{x}$ of dimension two, i.e. $n = 2$, and skewness dimension two, i.e. $q = 2$. Furthermore, let:

$$\boldsymbol{\mu} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \Sigma = \begin{pmatrix} 5 & 0.7 \\ 0.7 & 1 \end{pmatrix}, \Gamma = \begin{pmatrix} 5 & 0 \\ 0 & -5 \end{pmatrix}, \boldsymbol{\nu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Delta = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

the associated two dimensional CS-Gaussian pdf is displayed in Figure 1. Other favorable characteristics of the CS-Gaussian model that are relevant for inversion problems are:

- linear combinations of components are also CS-Gaussian random variables. In particular, the marginal pdfs are CS-Gaussian, which is displayed for the example in Figure 1. Note that each component has individual skewness properties controlled by the model parameter Γ.
- CS-Gaussian variables conditional on linear combinations of the components are also CS-Gaussian variables. In the example the CS-Gaussian pdf of $[\boldsymbol{x}|\boldsymbol{a}'\boldsymbol{x} = \boldsymbol{c}]$ with $\boldsymbol{a}' = (2, 3)$ and $\boldsymbol{c} = 20$ is displayed in Figure 1.
- joint distribution of independent blocks of CS-Gaussian random variables are CS-Gaussian distributed.

Based on the characteristics of the multivariate CS-Gaussian random variables the Bayesian CS-Gaussian inversion can be defined:

Proposition: **Bayesian Closed-Skew Gaussian Inversion**

Consider Bayesian inversion of the form

$$p(\boldsymbol{x}|\boldsymbol{o}) = const_o \times p(\boldsymbol{o}|\boldsymbol{x})p(\boldsymbol{x})$$

where $\boldsymbol{x} \in \mathbb{R}^{n_x}$ is the variable of interest; $\boldsymbol{o} \in \mathbb{R}^{n_o}$ is the observation and $const_o$ is the normalizing constant.

Let the prior model $p(\boldsymbol{x})$ be

$$\begin{aligned} \boldsymbol{x} \sim p(\boldsymbol{x}) &= CSN_{n_x,q_x}(\boldsymbol{\mu}_x, \Sigma_x, \Gamma_x, \boldsymbol{\nu}_x, \Delta_x) \\ &= [\Phi_{q_x}(\mathbf{0}; \boldsymbol{\nu}_x, \Delta_x + \Gamma_x\Sigma_x\Gamma_x')]^{-1}\Phi_{q_x}(\Gamma_x(\boldsymbol{x} - \boldsymbol{\mu}_x); \boldsymbol{\nu}_x, \Delta_x)\phi_{n_x}(\boldsymbol{x}; \boldsymbol{\mu}_x, \Sigma_x). \end{aligned}$$

Further let the likelihood model be

$$[\boldsymbol{o}|\boldsymbol{x}] = H\boldsymbol{x} + \boldsymbol{e},$$

with the error term $\boldsymbol{e} \in \mathbb{R}^{n_o}$ independent of $\boldsymbol{x}$, be:

$$\boldsymbol{e} \sim p(\boldsymbol{e}) = CSN_{n_o,q_o}(\mathbf{0}, \Sigma_e, \Gamma_e, \mathbf{0}, I)$$

which entails

$$[\boldsymbol{o}|\boldsymbol{x}] \sim CSN_{n_o,q_o}(H\boldsymbol{x}, \Sigma_e, \Gamma_e, \mathbf{0}, I)$$

with $\mathbf{0}$ the null-vector and I the diagonal identity matrix of proper dimensions.
Then the posterior model $p(\boldsymbol{x}|\boldsymbol{o})$ is

$$[\boldsymbol{x}|\boldsymbol{o}] \sim p(\boldsymbol{x}|\boldsymbol{o}) = CSN_{n_x,q_x+q_o}(\boldsymbol{\mu}_{x|o}, \Sigma_{x|o}, \Gamma_{x|o}, \boldsymbol{\nu}_{x|o}, \Delta_{x|o})$$

where

$$\begin{aligned}
\boldsymbol{\mu}_{x|o} &= \boldsymbol{\mu}_x + \Sigma_x H'[H\Sigma_x H' + \Sigma_e]^{-1}(\boldsymbol{o} - H\boldsymbol{\mu}_x), \\
\Sigma_{x|o} &= \Sigma_x - \Sigma_x H'[H\Sigma_x H' + \Sigma_e)]^{-1} H\Sigma_x, \\
\Gamma_{x|o} &= \left[\begin{bmatrix} \Gamma_x\Sigma_x \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \Gamma_x\Sigma_x H' \\ \Gamma_e\Sigma_e \end{bmatrix} [H\Sigma_x H' + \Sigma_e]^{-1} H\Sigma_x \right] \Sigma_{x|o}^{-1}, \\
\boldsymbol{\nu}_{x|o} &= \begin{bmatrix} -\boldsymbol{\nu}_x \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \Gamma_x\Sigma_x H' \\ \Gamma_e\Sigma_e \end{bmatrix} [H\Sigma_x H' + \Sigma_e]^{-1}(\boldsymbol{o} - H\boldsymbol{\mu}_x), \\
\Delta_{x|o} &= \begin{bmatrix} \Delta_x + \Gamma_x\Sigma_x\Gamma_x' & \mathbf{0} \\ \mathbf{0} & I + \Gamma_e\Sigma_e\Gamma_e' \end{bmatrix} - \begin{bmatrix} \Gamma_x\Sigma_x H' \\ \Gamma_e\Sigma_e \end{bmatrix} [H\Sigma_x H' + \Sigma_e]^{-1} \begin{bmatrix} \Gamma_x\Sigma_x H' \\ \Gamma_e\Sigma_e \end{bmatrix}' \\
&\quad - \Gamma_{x|o}\Sigma_{x|o}\Gamma_{x|o}',
\end{aligned}$$

and the marginal posterior model for $x_i = \boldsymbol{b}_{(i)}'\boldsymbol{x}$, with $\boldsymbol{b}_{(i)}$ a $(n_x \times 1)$-vector with entry one at location i and zero elsewhere, is $p(x_i|\boldsymbol{o})$:

$$[x_i|\boldsymbol{o}] \sim p(x_i|\boldsymbol{o}) = CSN_{1,q_x+q_o}(\mu_{x_i|o}, \sigma^2_{x_i|o}, \boldsymbol{\gamma}_{x_i|o}, \boldsymbol{\nu}_{x_i|o}, \Delta_{x_i|o}),$$

where

$$\begin{aligned}
&\mu_{x_i|o} = \boldsymbol{b}_{(i)}'\boldsymbol{\mu}_{x|o}, \quad \sigma^2_{x_i|o} = \boldsymbol{b}_{(i)}'\Sigma_{x|o}\boldsymbol{b}_{(i)}', \quad \boldsymbol{\gamma}_{x_i|o} = \Gamma_{x|o}\Sigma_{x|o}\boldsymbol{b}_{(i)}'(\boldsymbol{b}_{(i)}'\Sigma_{m|o}\boldsymbol{b}_{(i)})^{-1}, \\
&\boldsymbol{\nu}_{x_i|o} = \boldsymbol{\nu}_{x|o}, \quad \Delta_{x_i|o} = \Delta_{x|o} + \Gamma_{x|o}\Sigma_{x|o}\Gamma_{x|o}' - \Gamma_{x|o}\Sigma_{x|o}\boldsymbol{b}_{(i)}'(\boldsymbol{b}_{(i)}'\Sigma_{m|o}\boldsymbol{b}_{(i)})^{-1}\boldsymbol{b}_{(i)}\Sigma_{x|o}\Gamma_{x|o}'.
\end{aligned}$$

The essence of the Bayesian CS-Gaussian inversion is that with a CS-Gaussian prior model and a CS-Gauss-linear likelihood model, the resulting posterior model will be CS-Gaussian. The model parameters of the posterior model are analytically obtainable from the model parameters of the prior and likelihood models. The demonstration of the proposition can be found in Appendix A.

Note that traditional Bayesian Gaussian inversion is a special case of the proposition above. This can be seen by setting $\Gamma_x = \mathbf{0}$ in the prior model and $\Gamma_e = \mathbf{0}$ in the likelihood model. Then the prior is Gaussian and the likelihood Gauss-linear. From the proposition $\Gamma_{v|o} = \mathbf{0}$ and hence the posterior model is Gaussian with model parameters $(\boldsymbol{\mu}_{x|o}, \Sigma_{x|o})$. This coincides with Bayesian Gaussian inversion. Consequently, Bayesian CS-Gaussian inversion is a generalization of traditional Bayesian Gaussian inversion, not merely an alternative approach.

Seismic Inversion model

Focus of the study is on the elastic material properties (v_p, v_s, ρ), representing P-wave velocity, S-wave velocity and density, along a vertical profile through a reservoir zone. The profile is discretized into a grid, $t = (1, \cdots, T)$, and the grid corresponds to the seismic sampling design. Let the elastic properties be represented by $\boldsymbol{m} = (\boldsymbol{m}_t = (\ln v_{pt}, \ln v_{st}, \ln_{\rho t}); t = 1, \cdots, T)$. The log-transform is used in order to make the relation between $\boldsymbol{m}$ and the seismic AVO data $\boldsymbol{d}$ linear, see Buland and Omre (2003). The prior information about the elastic properties is represented in the prior model $p(\boldsymbol{m})$. The seismic AVO data is collected at angles $(\theta_1, \cdots, \theta_{n_\theta})$ and is represented by $\boldsymbol{d} = (\boldsymbol{d}_t = (d_1^{\theta_1}, \cdots, d_t^{\theta_{n_\theta}}); t = 1, \cdots, T)$. A convolved linearized Zoeppritz model $[\boldsymbol{d}|\boldsymbol{m}] = WAD\boldsymbol{m} + \boldsymbol{e}$, see Buland and Omre (2003), is used as likelihood model. Here, W is the wavelet matrix, A is the linearized Zoeppritz matrix, D is a differential matrix and $\boldsymbol{e}$ is a random error term. The associated likelihood model is $p(\boldsymbol{d}|\boldsymbol{m})$. The ultimate solution is the posterior model $p(\boldsymbol{m}|\boldsymbol{d})$ representing the elastic properties given the available seismic AVO data. The posterior model can be simulated from to provide realization of the elastic properties. Moreover, optimal prediction with associated prediction intervals for different loss functions can be provided.

In traditional Bayesian Gaussian inversion, see Buland and Omre (2003), the posterior model is Gaussian which is a unimodal, symmetric pdf. For the Gaussian pdf the expectation is a natural predictor since it coincide with the value of both the median and the mode. In the current study involving Bayesian CS-Gaussian inversion, the posterior pdf may be skewed and hence the expectation, median and mode can take different values. We have chosen to use the median value which minimizes the L_1-norm loss as the optimal predictor. Hence $\widehat{[\boldsymbol{m}|\boldsymbol{d}]} = Med\{\boldsymbol{m}|\boldsymbol{d}\} = Q_{0.5}\{\boldsymbol{m}|\boldsymbol{d}\}$ where the operators act componentwise on the vectorial argument. There are several reasons for this choice: (i) the median is logically related to the natural definition of the $(1-\alpha)$-prediction interval $[Q_{\alpha/2}\{\boldsymbol{m}|\boldsymbol{d}\}, Q_{1-\alpha/2}\{\boldsymbol{m}|\boldsymbol{d}\}]$; (ii) the median is a linear operator with respect to any monotone function $h(\cdot)$, i.e. $Q_{0.5}\{h(\boldsymbol{m}|\boldsymbol{d})\} = h(Q_{0.5}\{\boldsymbol{m}|\boldsymbol{d}\})$ hence a natural pre-

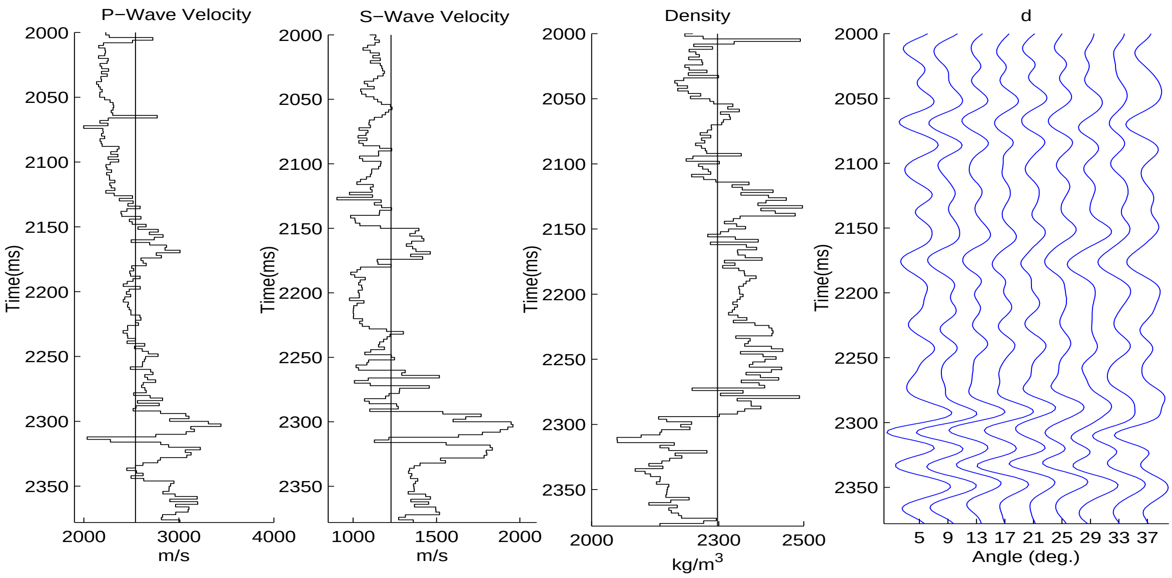

Figure 2: Elastic material property observations from well and corresponding synthetic AVO data.

dictor for the elastic properties themselves are $\exp\{\widehat{[\boldsymbol{m}|\boldsymbol{d}]}\}$; and (iii) the median is robust with respect to deviations in the tail behavior of $p(\boldsymbol{m}|\boldsymbol{d})$. The mean square error, $MSE\big\{\exp\{[\boldsymbol{m}|\boldsymbol{d}]\}, \exp\{\widehat{[\boldsymbol{m}|\boldsymbol{d}]}\}\big\}$, with the square error operator acting componentwise before averaging on each of the elastic properties, will be used to quantily the deviation between the truth and the predictions.

Inversion Case Study

Bayesian Gaussian inversion is used in Buland and Omre (2003) on data from the Sleipner Øst Field in the North Sea. The prior Gaussian model is inferred from elastic properties in one well and a Gaussian probability plot [Buland and Omre, 2003, Fig. 11] is presented to justify this assumption. The empirical curves in this plot deviate considerably from linear, however. The authors recognize this deviation, but states that assuming a Gaussian prior model for $p(\boldsymbol{m})$ and hence a log-Gaussian model for the elastic properties, is necessary to make the inversion analytically tractable. The Bayesian CS-Gaussian inversion approach presented in the previous section makes it possible to capture skewness in the prior $p(\boldsymbol{m})$ and still obtain analytical tractability. Re-evaluating the data in Buland and Omre (2003) would have made an excellent example, but unfortunately the data used in Buland and Omre (2003) is not easily available . Another set of data from the same reservoir, the Sleipner Øst Field, is available however. The well observations of the elastic material properties inferred from relevant well logs are displayed in Figure 2. The depth window

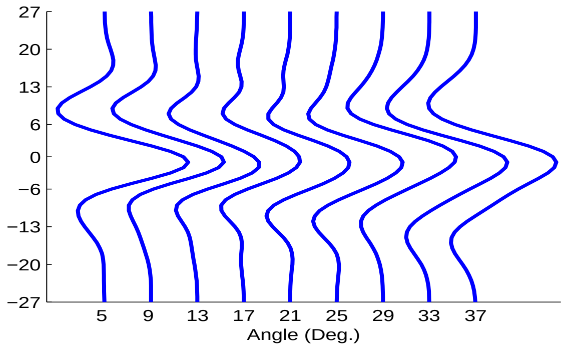

Figure 3: Angle dependent wavelet functions.

used is [2000, 2378] msec with two msec sampling interval making $T = 190$. The associated seismic AVO data is collected at nine angles, i.e. $n_\theta = 9$. The data-owner has provided angle-dependent wavelet functions of width ± 27 units, see Figure 3. These functions define the wavelet matrix W in the likelihood model. The seismic observation error is defined as $\boldsymbol{e} = W\boldsymbol{e}_1 + \boldsymbol{e}_2$ hence wavelet colored, where $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are independent Gaussian with mean zero and variance 5×10^{-5}. This entails a signal-to-noise ratio **. This define the likelihood covariance Σ_e.

Focus of the study is on the improvements of the inversion by generalizing the prior model $p(\boldsymbol{m})$ to account for skewness. Consequences of possible deviations from linearity in the likelihood model is not in focus. This subject is discussed in Buland and Omre (2003). Hence we have chosen to generate synthetic seismic AVO data based on the real elastic properties and the convolved linearized Zoeppritz likelihood model defined above. The synthetic seismic AVO data, $\boldsymbol{d}$, is displayed in Figure 2.

As a reference for comparison a prior Gaussian model is fitted. The trivariate Gaussian marginal pdf at each depth is assumed to be identical and the model parameters are estimated from the compiled observations of elastic properties, see Figure 4. This assessment is in the empirical Bayes tradition as discussed in Buland and Omre (2003). Traditional estimators for the expectation and the covariance matrix provide:

$$\boldsymbol{\mu}^0_{m_i} = \begin{bmatrix} 7.8350 \\ 7.0985 \\ 7.7385 \end{bmatrix}, \quad \Sigma^0_{m_i} = \begin{bmatrix} 0.0125 & 0.0130 & -0.0005 \\ 0.0130 & 0.0258 & -0.0031 \\ -0.0005 & -0.0031 & 0.0016 \end{bmatrix}.$$

The spatial correlation function is assumed to be identical for the three elastic properties and it is estimated from the compiled observations, see Figure 4. The fitted correlation function is:

$$c(h) = \exp\{-\frac{|h|}{13}\},$$

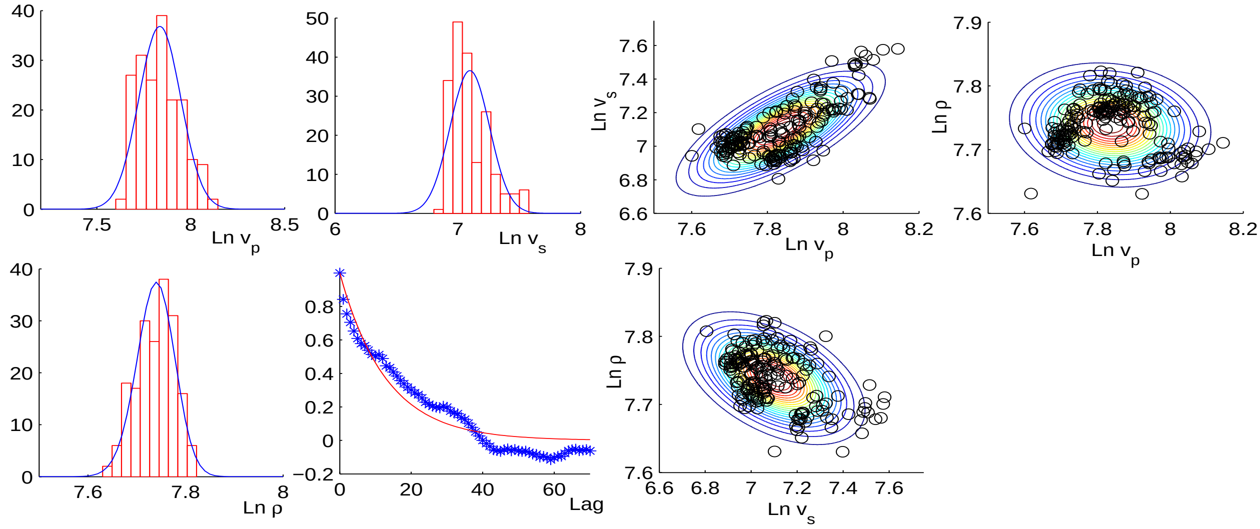

Figure 4: Trivariate prior Gaussian model(solid), spatial correlation function and empirical estimates from compiled elastic property observations.

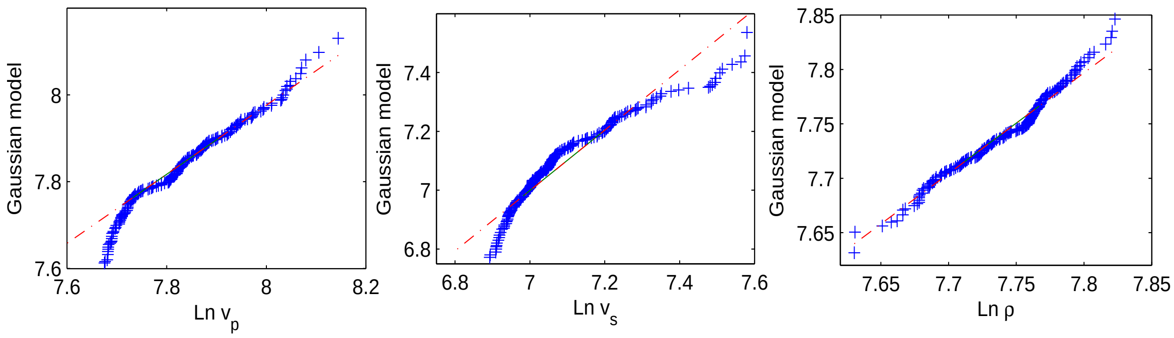

Figure 5: Gaussian probability plot of the compiled elastic property observations.

where h represents the interdistance.

The likelihood model $p(\boldsymbol{d}|\boldsymbol{m})$ defined above is Gauss-linear with given model parameter values. Hence the Bayesian Gaussian inversion, see Buland and Omre (2003), can be used to assess the posterior model. The posterior pdf is Gaussian with model parameters analytically obtainable from the model parameters specified above.

Based on the Gaussian prior model defined above Gaussian probability plots of the compiled elastic property observations can be computed, see Figure 5. The empirical curves deviate somewhat from linearity, although less than for the data set in [Buland and Omre, 2003; Fig. 11]. The $\ln v_s$-variable appears to have a distribution which is skewed and this justified the introduction of the CS-Gaussian model.

The prior CS-Gaussian model $p(\boldsymbol{m})$ is decomposed into one common trivariate CS-

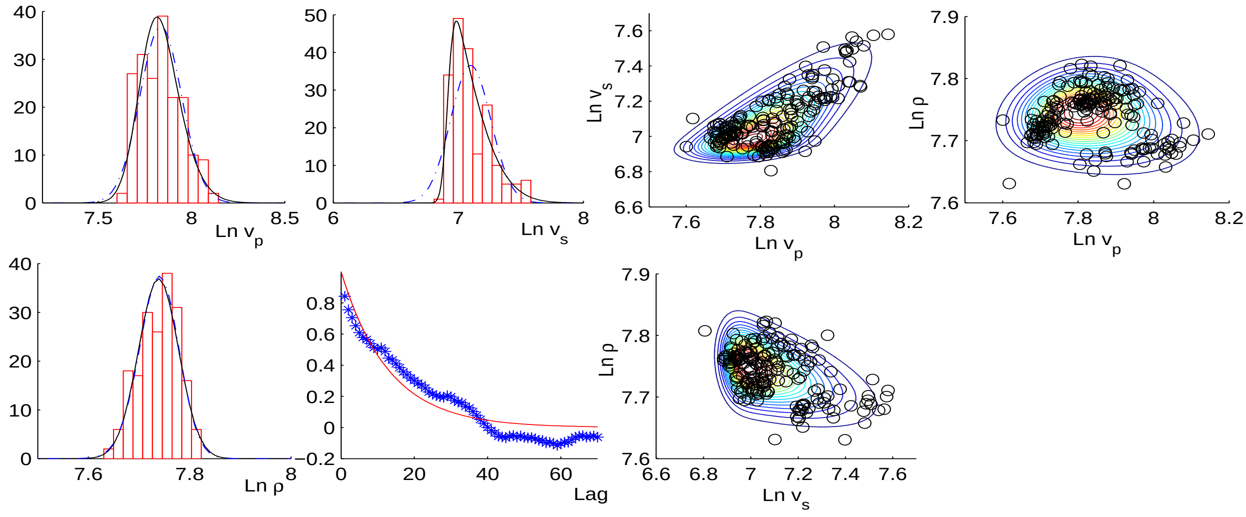

Figure 6: Trivariate prior CS-Gaussian model (solid) and Gaussian model (hatched), spatial correlation function and empirical estimates from compiled elastic property observations.

Gaussian pdf for the marginals $p(\boldsymbol{m}_t)$ and one common spatial correlation function for the variables, see Appendix B. Hence we assume that the elastic properties themselves are log-CS-Gaussian. The trivariate CS-Gaussian pdf for $p(\boldsymbol{m}_t)$ is inferred from the compiled elastic property observations in the well in an empirical Bayes tradition. A maximum likelihood approach is used, see Appendix B, and the model parameter estimates are:

$$\hat{\boldsymbol{\mu}}_{m.} = \begin{bmatrix} 7.763 \\ 6.920 \\ 7.758 \end{bmatrix}, \; \hat{\Sigma}_{m.} = \begin{bmatrix} 0.0620 & 0.1002 & -0.0099 \\ 0.1002 & 0.1811 & -0.0200 \\ -0.0099 & -0.0200 & 0.0035 \end{bmatrix},$$

$$\hat{\Gamma}_{m.} = \begin{bmatrix} -4.57 & 0 & 0 \\ 0 & 24.99 & 0 \\ 0 & 0 & 5 \end{bmatrix}, \;\; \boldsymbol{\nu}_{m.} = \mathbf{0}, \;\; \Delta_{m.} = I.$$

The estimated trivariate CS-Gaussian pdf is displayed in Figure 6 together with the compiled observations. The univariate and bivariate marginal pdfs expose clear skewness whenever the $\ln v_s$-variable is involved. The spatial dependence is modeled by the same correlation function as for the Gaussian prior, see Figure 6. Based on this decomposition and the estimated model parameters, the CS-Gaussian prior model can be defined:

$$p(\boldsymbol{m}) = CSN_{3T,3T}(\boldsymbol{\mu}_m, \Sigma_m, \Gamma_m, \boldsymbol{\nu}_m, \Delta_m)$$

with the model parameters $\boldsymbol{\mu}_m$, Σ_m, Γ_m, $\boldsymbol{\nu}_m$ and Δ_m specified in Appendix B. Note that the skewness indicator $q = 3T$ needs to be identical to the dimensionality of the inversion $n = 3T$ in order to provide sufficient degree of skewness freedom to represent separate skewness for each of the three elastic properties at each depth in the posterior model.

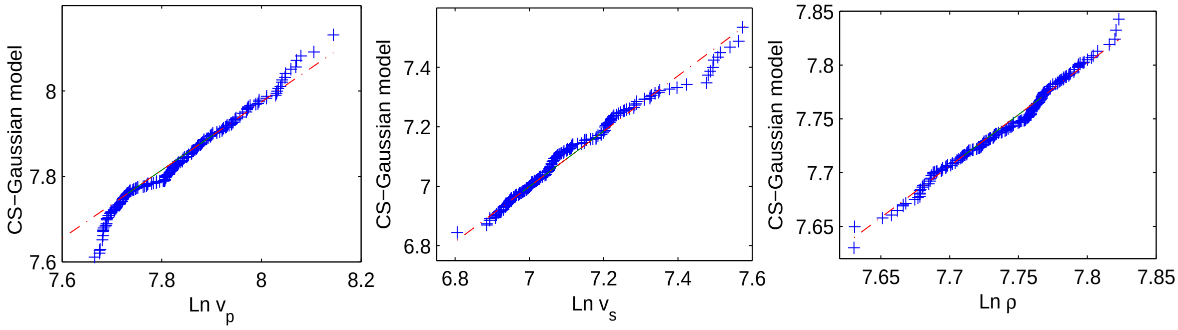

Figure 7: CS-Gaussian probability plot of the compiled elastic property data set.

Based on the CS-Gaussian prior model defined above CS-Gaussian probability plots of the compiled elastic property observations can be computed, see Figure 7. The empirical curves deviate from linearity, but less than for the plots based on the Gaussian model in Figure 5. The skewness in the distribution of the $\ln v_s$-variable appears to be captured by the CS-Gaussian model. The distribution of the $\ln v_p$-variable appears to have lighter tails than the Gaussian pdf, however, but this feature cannot be captured by CS-Gaussian model.

The likelihood model $p(\boldsymbol{d}|\boldsymbol{m})$ is based on the convolved linearized Zoeppritz model with a centered, colored Gaussian error term, identical to the model described above. This likelihood model constitutes a particular degenerate CS-Gauss-linear likelihood model:

$$\begin{aligned} p(\boldsymbol{d}|\boldsymbol{m}) &= N_{3T}(\boldsymbol{\mu}_{d|m}, \Sigma_{d|m}), \\ &= CSN_{3T,0}(\boldsymbol{\mu}_{d|m}, \Sigma_{d|m}, \cdot, \cdot, \cdot), \end{aligned}$$

where

$$\boldsymbol{\mu}_{d|m} = WAD\boldsymbol{m},$$

and $\Sigma_{d|m} = \Sigma_e$ and the three latter parameters unspecified since $q = 0$.

The prior model is CS-Gaussian and the likelihood model is CS-Gauss-linear, hence it follows from Bayesian CS-Gaussian inversion that the posterior model is a CS-Gaussian pdf:

$$p(\boldsymbol{m}|\boldsymbol{d}) = CSN_{3T,3T}(\boldsymbol{\mu}_{m|d}, \Sigma_{m|d}, \Gamma_{m|d}, \boldsymbol{\nu}_{m|d}, \Delta_{m|d}),$$

with model parameters $\boldsymbol{\mu}_{m|d}$, $\Sigma_{m|d}$, $\Gamma_{m|d}$, $\boldsymbol{\nu}_{m|d}$ and $\Delta_{m|d}$ analytically obtained from the model parameters of the prior and likelihood models according to the Proposition for Bayesian CS-Gaussian inversion. Moreover, all posterior marginal models are also CS-Gaussian pdfs

$$p(\boldsymbol{m}_t|\boldsymbol{d}) = CSN_{3,3T}(\boldsymbol{\mu}_{m_t|d}, \Sigma_{m_t|d}, \Gamma_{m_t|d}, \boldsymbol{\nu}_{m_t|d}, \Delta_{m_t|d}),$$

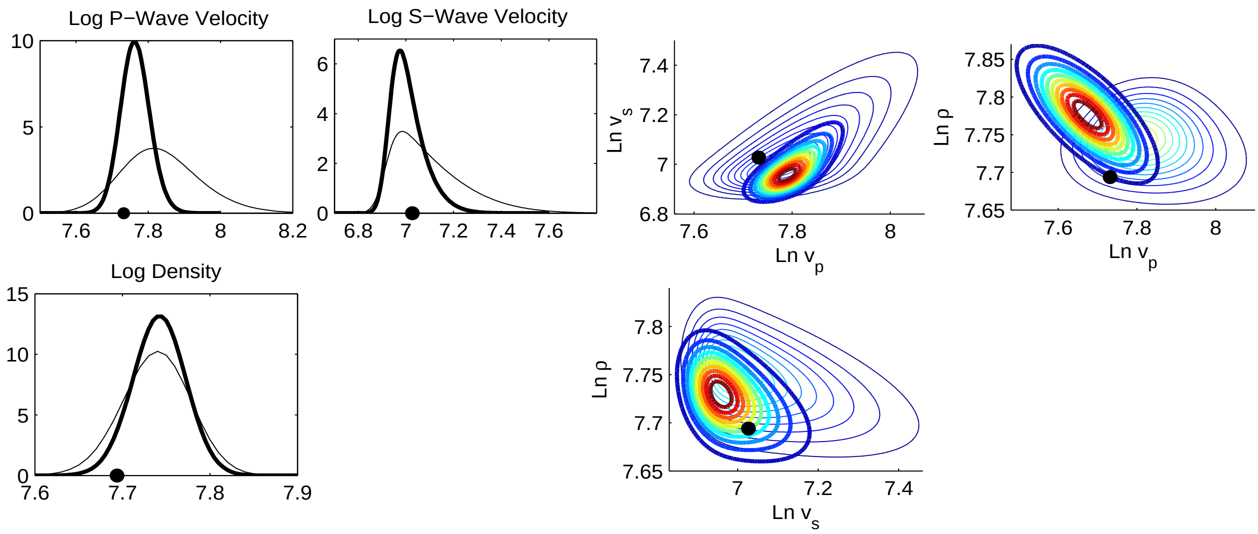

Figure 8: Trivariate posterior CS-Gaussian model at $t = 2318$ (thick solid) and corresponding trivariate prior model (thin solid). Correct value is marked.

with model parameters analytically obtainable according to the Proposition. Note that the solutions involves integrals of $3T$-dimensional Gaussian pdfs which needs to be approximated numerically. See Appendix C for more details on the approach used in this study.

Inversion results with discussion

The posterior model $p(\boldsymbol{m}|\boldsymbol{d})$ is the ultimate solution to the inversion problem. In Figure 8, the marginal posterior model $p(\boldsymbol{m}_t|\boldsymbol{d})$ of depth $t = 2318$ is displayed together with the marginal prior model $p(\boldsymbol{m}_t)$ and the correct value of $\boldsymbol{m}_t$. The center of the posterior model is closer to the correct value than the prior center and the uncertainty is significantly reduced by conditioning on the seismic AVO data $\boldsymbol{d}$. Note also that the posterior model capture skewness in the pdf. In the Gaussian approach this trivariate posterior model will be Gaussian and hence symmetrical. Similar displays could of course have been presented for all depths t. We have chosen to use the median or $0.5-$quantile value, of the posterior model as predictor. Hence $(\widehat{v_p, v_s, \rho}) = \exp\{Q_{0.5}\{\boldsymbol{m}|\boldsymbol{d}\}\}$ with the operators acting componentwise. The associated 0.8-prediction intervals are $[\exp\{Q_{0.1}\{\boldsymbol{m}|\boldsymbol{d}\}\}, \exp\{Q_{0.9}\{\boldsymbol{m}|\boldsymbol{d}\}\}]$.

In Figure 9, prediction of the elastic material properties $(\widehat{v_p, v_s, \rho})$ along the profile with associated 0.8-prediction intervals are displayed for the Gaussian and the CS-Gaussian model.

The predictions based on the CS-Gaussian prior model appear more reliable than

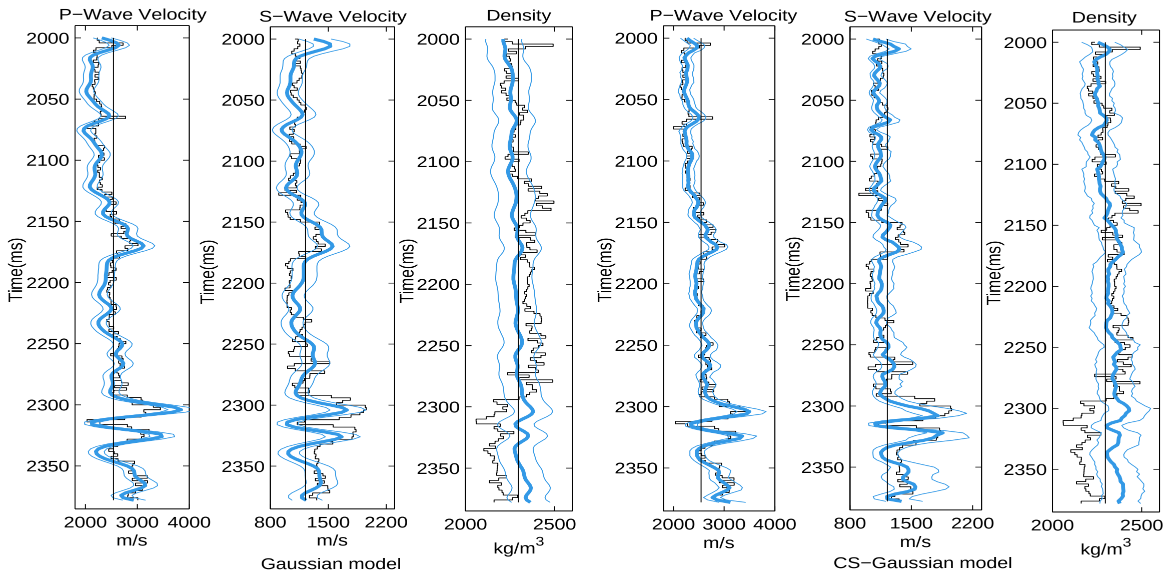

Figure 9: Inversion results from Bayesian Gaussian inversion (left) and Bayesian CS-Gaussian inversion (right). Predictions (thick solid) and 0.8-prediction intervals (thin solid) with elastic property observations. Empirical averages of observations are marked.

predictions from the Gaussian prior. The v_s-predictions in the high-variability depth interval $[2275, 2350]$ examplifies this. The CS-Gaussian predictions reproduces high values better and are not tempted to under-estimate low values as much as the Gaussian-based predictions. The inclusion of skewness in the posterior provides freedom to reproduce patterns like this. The empirical coverages of the prediction intervals for the two models are difficult to compare visually.

In Table 1, the fractions of correct values in the profile that lies below the corresponding estimated $Q_{0.1}$, $Q_{0.5}$ and $Q_{0.9}$ values are listed for each variable, for each of the two models. The corresponding mean-square-error (MSE) of the predictions, are also listed. The precision of the predictions based on the CS-Gaussian prior model is significantly better than for predictions from the Gaussian prior. This can be seen from the $Q_{0.5}$-fraction being closer to 0.5 and the MSE-values. The latter are reduced by a factor two for v_p and v_s predictions. The Gaussian model provide best estimates of the lower 0.8-prediction interval limit, while the CS-Gaussian model provide very reliable estimates of the upper limit.

Model	Variable	Quantile 0.1	0.5	0.9	Deviation MSE
Gaussian	v_p	0.093	0.743	0.814	2.7705×10^4
	v_s	0.207	0.443	0.850	2.9740×10^4
	ρ	0.107	0.800	0.829	9.5302×10^3
CS-Gaussian	v_p	0.071	0.507	0.929	1.1826×10^4
	v_s	0.321	0.486	0.907	1.7613×10^4
	ρ	0.121	0.671	0.900	8.5244×10^3

Table 1: Inversion results from Bayesian Gaussian inversion (upper) and Bayesian CS-Gaussian inversion (lower). Fraction of observations below predicted quantiles and mean-square-errors of predictions.

Conclusion

The Bayesian linearized AVO inversion as introduced in Buland and Omre (2003) relies heavily on Gaussian assumptions and may be termed Bayesian Gaussian inversion. The log-transformed elastic properties must be Gaussian and the error term in the convolved linearized Zoeppritz relation must be additive Gaussian. The advantage of making these assumptions is that the posterior model is Gaussian with model parameters analytically assessable from the prior and likelihood models. The analytical tractability make it possible to solve the seismic inversion problems on large 3D grids with spatial couplings, although numerical approximations may be required to assess the posterior model.

In the current paper we have generalized the model in Buland and Omre (2003) into what we have termed Bayesian CS-Gaussian inversion. The log-transformed elastic properties may be represented by a skew pdf and so do the error terms in the convolved linearized Zoeppritz relation. The inversion is cast in a CS-Gaussian framework, see González-Farías *et al.* (2004). The posterior model is within the CS-Gaussian family of the models with model parameters analytically assessable from the prior and likelihood models. Hence the posterior model can capture skewness in the log-transformed elastic properties. The model in Buland and Omre (2003) appears as a special case of the current model. Since the model is analytically tractable seismic inversion on large 3D grids with spatial couplings is possible, but also here numerical approximations are required to assess the posterior model. Several of these numerical challenges must be further studied.

In the current paper the Bayesian CS-Gaussian inversion is demonstrated on real elastic material properties from one well with corresponding synthetically generated seismic AVO data. A convolved linearized Zoeppritz forward model is used. The results are compared to inversion results based on the Bayesian Gaussian inversion presented in Buland and Omre (2003). The mean-square-error in the (v_p, v_s)-predictions based on the CS-Gaussian approach is reduced by a factor of two relative to the Gaussian predictions. The improvement is somewhat less for the $\rho-$predictions. Conclusions concerning empirical coverage of the estimated 0.8-prediction intervals cannot be significantly made.

Acknowledgement

The research work is supported by the Uncertainty in Reservoir Evaluation (URE) activity of NTNU.

Appendix A. Demonstration of proposition.

Proof

Consider

$$\boldsymbol{x} \sim p(\boldsymbol{x}) = CSN_{n_x,q_x}(\boldsymbol{\mu}_x, \Sigma_x, \Gamma_x, \boldsymbol{\nu}_x, \Delta_x)$$

and

$$\boldsymbol{o} = H\boldsymbol{x} + \boldsymbol{e}$$

with $\boldsymbol{e}$ independent of $\boldsymbol{x}$ and

$$\boldsymbol{e} \sim p(\boldsymbol{e}) = CSN_{n_o,q_o}(\mathbf{0}, \Sigma_e, \Gamma_e, \mathbf{0}, I),$$

where $\mathbf{0}$ is the null-vector and I is the diagonal unit-matrix.

Represent

$$\begin{aligned} \boldsymbol{x} &= [\boldsymbol{t}|\boldsymbol{u} \geq \mathbf{0}], \\ \boldsymbol{e} &= [\boldsymbol{s}|\boldsymbol{v} \geq \mathbf{0}], \end{aligned}$$

then

$$\begin{bmatrix} \boldsymbol{t} \\ \boldsymbol{r} = H\boldsymbol{t} + \boldsymbol{s} \\ \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix} \sim Gauss(\begin{bmatrix} \boldsymbol{\mu}_x \\ H\boldsymbol{\mu}_x \\ -\boldsymbol{\nu}_x \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \Sigma_x & \Sigma_x H' & \Sigma_x \Gamma_x & \mathbf{0} \\ H\Sigma_x & H\Sigma_x H' + \Sigma_e & H\Sigma_x \Gamma_x' & \Sigma_e \Gamma_e' \\ \Gamma_x \Sigma_x & \Gamma_x \Sigma_x H' & \Delta_x + \Gamma_x \Sigma_x \Gamma_x' & \mathbf{0} \\ \mathbf{0} & \Gamma_e \Sigma_e & \mathbf{0} & I + \Gamma_e \Sigma_e \Gamma_e' \end{bmatrix}),$$

$$\begin{bmatrix} \boldsymbol{t}|\boldsymbol{r} \\ \boldsymbol{u}|\boldsymbol{r} \\ \boldsymbol{v}|\boldsymbol{r} \end{bmatrix} \sim Gauss(\begin{bmatrix} \boldsymbol{\mu}_{t|r} \\ \boldsymbol{\mu}_{u|r} \\ \boldsymbol{\mu}_{v|r} \end{bmatrix}, \begin{bmatrix} \Sigma_{t|r} & \Gamma_{tu|r} & \Gamma_{tv|r} \\ \Gamma_{ut|r} & \Sigma_{u|r} & \Gamma_{uv|r} \\ \Gamma_{vt|r} & \Gamma_{vu|r} & \Sigma_{v|r} \end{bmatrix}),$$

where

$$\begin{aligned}
\boldsymbol{\mu}_{t|r} &= \boldsymbol{\mu}_x + \Sigma_x H'\big[H\Sigma_x H' + \Sigma_e\big]^{-1}(\boldsymbol{r} - H\boldsymbol{\mu}_x), \\
\boldsymbol{\mu}_{u|r} &= -\boldsymbol{\nu}_x + \Gamma_x\Sigma_x H'\big[H\Sigma_x H' + \Sigma_e\big]^{-1}(\boldsymbol{r} - H\boldsymbol{\mu}_x), \\
\boldsymbol{\mu}_{v|r} &= \mathbf{0} + \Gamma_e\Sigma_e\big[H\Sigma_x H' + \Sigma_e\big]^{-1}(\boldsymbol{r} - H\boldsymbol{\mu}_x), \\
\Sigma_{t|r} &= \Sigma_x - \Sigma_x H'\big[H\Sigma_x H' + \Sigma_e\big]^{-1}H\Sigma_x, \\
\Sigma_{u|r} &= [\Delta_x + \Gamma_x\Sigma_x\Gamma_x'] - \Gamma_x\Sigma_x H'\big[H\Sigma_x H' + \Sigma_e\big]^{-1}H\Sigma_x\Gamma_x, \\
\Sigma_{v|r} &= [I + \Gamma_e\Sigma_e\Gamma_e'] - \Gamma_e\Sigma_e\big[H\Sigma_x H' + \Sigma_e\big]^{-1}\Sigma_e\Gamma_e', \\
\Gamma_{ut|r} &= \Gamma_x\Sigma_x - \Gamma_x\Sigma_x H'\big[H\Sigma_x H' + \Sigma_e\big]^{-1}H\Sigma_x, \\
\Gamma_{vt|r} &= \mathbf{0} - \Gamma_e\Sigma_e\big[H\Sigma_x H' + \Sigma_e\big]^{-1}H\Sigma_x, \\
\Gamma_{uv|r} &= \mathbf{0} - \Gamma_x\Sigma_x H'\big[H\Sigma_x H' + \Sigma_e\big]^{-1}\Sigma_e\Gamma_e'.
\end{aligned}$$

Let $\boldsymbol{w} = \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix}$, then

$$\begin{bmatrix} \boldsymbol{t}|\boldsymbol{r} \\ \boldsymbol{w}|\boldsymbol{r} \end{bmatrix} \sim Gauss(\begin{bmatrix} \boldsymbol{\mu}_{t|r} \\ \boldsymbol{\mu}_{w|r} \end{bmatrix}, \begin{bmatrix} \Sigma_{t|r} & \Gamma_{tw|r} \\ \Gamma_{wt|r} & \Sigma_{w|r} \end{bmatrix}),$$

where

$$\boldsymbol{\mu}_{w|r} = \begin{bmatrix} \boldsymbol{\mu}_{u|r} \\ \boldsymbol{\mu}_{v|r} \end{bmatrix}, \quad \Sigma_{w|r} = \begin{bmatrix} \Sigma_{u|r} & \Gamma_{uv|r} \\ \Gamma_{vu|r} & \Sigma_{v|r} \end{bmatrix}, \quad \Gamma_{wt|r} = \begin{bmatrix} \Gamma_{ut|r} \\ \Gamma_{vt|r} \end{bmatrix}.$$

Then, from the definition of the CS-Gaussian model one has:

$$[\boldsymbol{x}|\boldsymbol{o}] = [\boldsymbol{t}|\boldsymbol{r}, \boldsymbol{w} \geq \mathbf{0}] \sim CSN_{n_x, q_x+q_o}(\boldsymbol{\mu}_{x|o}, \Sigma_{x|o}, \Gamma_{x|o}, \boldsymbol{\nu}_{x|o}, \Delta_{x|o})$$

where

$$\begin{aligned}
\boldsymbol{\mu}_{x|o} &= \boldsymbol{\mu}_x + \Sigma_x H'[H\Sigma_x H' + \Sigma_e]^{-1}(\boldsymbol{o} - H\boldsymbol{\mu}_x), \\
\Sigma_{x|o} &= \Sigma_x - \Sigma_x H'[H\Sigma_x H' + \Sigma_e)]^{-1}H\Sigma_x, \\
\Gamma_{x|o} &= \left[\begin{bmatrix} \Gamma_x\Sigma_x \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \Gamma_x\Sigma_x H' \\ \Gamma_e\Sigma_e \end{bmatrix}[H\Sigma_x H' + \Sigma_e]^{-1}H\Sigma_x\right]\Sigma_{x|o}^{-1}, \\
\boldsymbol{\nu}_{x|o} &= \begin{bmatrix} -\boldsymbol{\nu}_x \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \Gamma_x\Sigma_x H' \\ \Gamma_e\Sigma_e \end{bmatrix}[H\Sigma_x H' + \Sigma_e]^{-1}(\boldsymbol{o} - H\boldsymbol{\mu}_x), \\
\Delta_{x|o} &= \begin{bmatrix} \Delta_x + \Gamma_x\Sigma_x\Gamma_x' & \mathbf{0} \\ \mathbf{0} & I + \Gamma_e\Sigma_e\Gamma_e' \end{bmatrix} - \begin{bmatrix} \Gamma_x\Sigma_x H' \\ \Gamma_e\Sigma_e \end{bmatrix}[H\Sigma_x H' + \Sigma_e]^{-1}\begin{bmatrix} \Gamma_x\Sigma_x H' \\ \Gamma_e\Sigma_e \end{bmatrix}'
\end{aligned}$$

$$- \Gamma_{x|o}\Sigma_{x|o}\Gamma'_{x|o},$$

Consider $x_i = \boldsymbol{b}'_{(i)}\boldsymbol{x}$, with $\boldsymbol{b}_{(i)}$ a $(n_x \times 1)$-vector with entries zeros except for entry no i which is one, then according to González-Farías *et al.* (2004) the marginal pdf is:

$$[x_i|\boldsymbol{o}] \sim p(x_i|\boldsymbol{o}) = CSN_{1,q_x+q_o}(\mu_{x_i|o}, \sigma^2_{x_i|o}, \boldsymbol{\gamma}_{x_i|o}, \boldsymbol{\nu}_{x_i|o}, \Delta_{x_i|o}),$$

where

$$\mu_{x_i|o} = \boldsymbol{b}'_{(i)}\boldsymbol{\mu}_{x|o}, \quad \sigma^2_{x_i|o} = \boldsymbol{b}'_{(i)}\Sigma_{x|o}\boldsymbol{b}_{(i)}, \quad \boldsymbol{\gamma}_{x_i|o} = \Gamma_{x|o}\Sigma_{x|o}\boldsymbol{b}_{(i)}(\boldsymbol{b}'_{(i)}\Sigma_{x|o}\boldsymbol{b}_{(i)})^{-1},$$
$$\boldsymbol{\nu}_{x_i|o} = \boldsymbol{\nu}_{x|o}, \quad \Delta_{x_i|o} = \Delta_{x|o} + \Gamma_{x|o}\Sigma_{x|o}\Gamma'_{x|o} - \Gamma_{x|o}\Sigma_{x|o}\boldsymbol{b}_{(i)}(\boldsymbol{b}'_{(i)}\Sigma_{x|o}\boldsymbol{b}_{(i)})^{-1}\boldsymbol{b}'_{(i)}\Sigma_{x|o}\Gamma'_{x|o},$$

Appendix B. Construction of prior model

The prior model is

$$p(\boldsymbol{m}) = CSN_{3T,3T}(\boldsymbol{\mu}_m, \Sigma_m, \Gamma_m, \boldsymbol{\nu}_m, \Delta_m).$$

By assuming stationarity, all trivariate marginal pdfs at each depth are identical:

$$p(\boldsymbol{m}_i) = p(\boldsymbol{m}_j) = CSN_{3,3}(\boldsymbol{\mu}_{m.}, \Sigma_{m.}, \Gamma_{m.}, \boldsymbol{\nu}_{m.}, \Delta_{m.}), \quad all \quad i,j = 1, \cdots, T.$$

Moreover, the spatial correlation function is assumed to be identical for all log-transformed elastic properties

$$c(i,j) = c(|i-j|), \quad all \quad i,j = 1, \cdots, T.$$

From this decomposition the prior model can be constructed, see Karimi and Mohammadzadeh (2009):

$$\underset{3T\times 1}{\boldsymbol{\mu}_m} = \begin{bmatrix} \boldsymbol{\mu}_{m.} \\ \vdots \\ \boldsymbol{\mu}_{m.} \end{bmatrix}, \quad \underset{3T\times 3T}{\Sigma_m} = \begin{pmatrix} \Sigma_{m.} & c(1)\Sigma_{m.} & \cdots & c(T-1)\Sigma_{m.} \\ c(1)\Sigma_{m.} & \Sigma_{m.} & \vdots & \vdots \\ \vdots & & \ddots & \vdots \\ c(T-1)\Sigma_{m.} & \cdots & \cdots & \Sigma_{m.} \end{pmatrix},$$

$$\underset{3T\times 3T}{\Gamma_m} = \begin{pmatrix} \Gamma_{m.} & 0 & \cdots & 0 \\ 0 & \Gamma_{m.} & \vdots & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \Gamma_{m.} \end{pmatrix}, \quad \underset{3T\times 1}{\boldsymbol{\nu}_m} = \begin{bmatrix} \boldsymbol{\nu}_{m.} \\ \vdots \\ \boldsymbol{\nu}_{m.} \end{bmatrix},$$

$$\underset{3T\times 3T}{\Delta_m} = \begin{pmatrix} \Delta_{m.} & 0 & \cdots & 0 \\ 0 & \Delta_{m.} & \cdots & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & \Delta_{m.} \end{pmatrix}.$$

The prior model parameters $(\boldsymbol{\mu}_{m.}, \Sigma_{m.}, \Gamma_{m.}, \boldsymbol{\nu}_{m.}, \Delta_{m.})$ are inferred from the observed elastic material properties in the well in an empirical Bayes tradition. The log-transformed elastic properties observed are denoted $\boldsymbol{m}^o = (\boldsymbol{m}_i^o,\ i = 1, \cdots, T)$. The pseudo-likelihood function, considering $\boldsymbol{m}_i^o$; $i = 1, \cdots, T$ as independent, is:

$$\begin{aligned} pL(\boldsymbol{\mu}_{m.}, \Sigma_{m.}, \Gamma_{m.}, \boldsymbol{\nu}_{m.}, \Delta_{m.}|\boldsymbol{m}^o) &= \left[\Phi_3(0; \boldsymbol{\nu}_{m.}, \Delta_{m.} + \Gamma_{m.}\Sigma_{m.}\Gamma'_{m.})\right]^{-T} \\ &\times \prod_{i=1}^{T} \Phi_3(\Gamma_{m.}(\boldsymbol{m}_i^o - \boldsymbol{\mu}_{m.}); \boldsymbol{\nu}_{m.}, \Delta_{m.}) \\ &\times \prod_{i=1}^{T} \Phi_3(\boldsymbol{m}_i^o; \boldsymbol{\mu}_{m.}, \Sigma_{m.}) \end{aligned}$$

The associated maximum-pseudo-likelihood estimator, MpLE, is:

$$(\hat{\boldsymbol{\mu}}_{m.}, \hat{\Sigma}_{m.}, \hat{\Gamma}_{m.}, \hat{\boldsymbol{\nu}}_{m.}, \hat{\Delta}_{m.}) = \underset{\mu_{m.}, \Sigma_{m.}, \Gamma_{m.}, \nu_{m.}, \Delta_{m.}}{\text{argmax}} \left\{pL(\boldsymbol{\mu}_{m.}, \Sigma_{m.}, \Gamma_{m.}, \boldsymbol{\nu}_{m.}, \Delta_{m.}|\boldsymbol{m}^o)\right\}.$$

It is recognized, however, that this MpL-estimate is not well defined since some model parameters are unidentifiable, see Gonzalez-Farias *et al.* (2004). We avoid these problems by using a constrained MpL-estimator with

$$\Gamma_{m.} - \text{diagonal}(3 \times 3) - \text{matrix}, \quad \boldsymbol{\nu}_{m.} = \mathbf{0}, \quad \Delta_{m.} = I,$$

where $\mathbf{0}$ is the null-vector and I is the diagonal unit-matrix of proper dimensions.
The remaining model parameters are then identifiable, and by numerical optimization we obtain:

$$\hat{\boldsymbol{\mu}}_{m.} = \begin{bmatrix} 7.763 \\ 6.920 \\ 7.758 \end{bmatrix}, \hat{\Sigma}_{m.} = \begin{pmatrix} 0.0620 & 0.1002 & -0.0099 \\ 0.1002 & 0.1811 & -0.0200 \\ -0.0099 & -0.0200 & 0.0035 \end{pmatrix},$$

$$\hat{\Gamma}_{m.} = \begin{pmatrix} -4.57 & 0 & 0 \\ 0 & 24.99 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

The prior spatial correlation function $c(|i-j|)$ is also inferred from the well observations. Traditional statistical estimators are used, and a parametric model is fitted to ensure positive definiteness. We obtain:

$$\hat{c}(h) = \exp\{-\frac{|h|}{13}\}.$$

Appendix C. Numerical approximation of posterior model

The marginal posterior model is of the form:

$$\begin{aligned} p(x_i|\boldsymbol{o}) &= CSN_{1,q}(\mu_., \sigma_.^2, \boldsymbol{\gamma}_., \boldsymbol{\nu}_., \Delta_.) \\ &= \left[\Phi_q(\mathbf{0}; \boldsymbol{\nu}_., \Delta_. + \boldsymbol{\gamma}_.\sigma_.^2\boldsymbol{\gamma}_.')\right]^{-1} \\ &\quad \times\Phi_q(\boldsymbol{\gamma}_.(x_i - \mu_.); \boldsymbol{\nu}_., \Delta_.) \\ &\quad \times\phi_1(x_i; \mu_., \sigma_.^2) \end{aligned}$$

where $\Phi_.(\cdot;\cdot)$ and $\phi_.(\cdot;\cdot)$ are Gaussian cdfs and pdfs respectively.
Since q is a large integer, the term complicated to compute is:

$$\Phi_q(\boldsymbol{\gamma}_.(x_i - \mu_.); \boldsymbol{\nu}_., \Delta_.)$$

because it is defined by an integral of a q-dimensional Gaussian pdf. This integral is not analytically tractable and reliable approximation for the general case is hard to obtain.

Note, however, that for elements in the q-dimensional vector which have only diagonal non-zero entries in $\Delta_.$, ie are independent of other elements, the integral factorizes. If, in addition the corresponding elements in $\boldsymbol{\gamma}_.$ are zero, ie regression coefficient zero, then the factor is a constant. Hence the complexity of the expression is related to the dependence structure of $[x|\boldsymbol{o}]$.
Consider the seismic inversion problem under study. Let $\boldsymbol{m}_i = (m_i^1, m_i^2, m_i^3)$ represent the log-transformed elastic properties of depth i. Focus is on the marginal posterior pdf, $p(m_i^j|\boldsymbol{d})$ and the complicated term is:

$$\Phi_{3T}(\boldsymbol{\gamma}_{m_i^j|d}(m_i^j - \mu_{m_i^j|d}); \boldsymbol{\nu}_{m_i^j|d}, \Delta_{m_i^j|d}).$$

Note initially that both the prior Δ_m and Γ_m are diagonal $(3T \times 3T)$-matrices. All coupling in the prior model is through intervariable and spatial dependence in Σ_m. Moreover, coupling is introduced through the conditioning on $\boldsymbol{d}$ through the correlation matrix WAD. Both Σ_m and WAD introduces neighborhood dependence, and not long-distance dependence, however. We have conducted a limited empirical study using different localized approximations, and decided to use the following approximation:

$$p^*(m_i^j|\boldsymbol{d}) = CSN_{1,7}(\mu_{m_i^j|d}, \sigma^2_{m_i^j|d}, \gamma^*_{m_i^j|d}, \nu^*_{m_i^j|d}, \Delta^*_{m_i^j|d}),$$

where the *-marked model parameters is a 7-dimensional subset of the corresponding $3T$-dimensional model parameters, representing the entries $(m_i^1, m_i^2, m_i^3, m_{i-2}^j, m_{i-1}^j, m_{i+1}^j, m_{i+2}^j)$, see Figure 10.

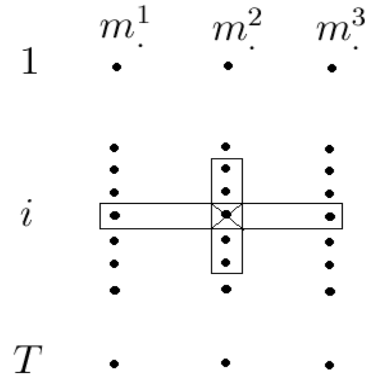

Figure 10: Definition of neighborhood for m_i^2 in approximation for marginal posterior model.

This approximate marginal $p^*(m_i^j|\boldsymbol{d})$ is assessed by using the basic definition of CS-Gaussian variables. Simulation based inference is made by rejection sampling from the appropriate 8-dimensional Gaussian pdf.

We consider the approximations to be reliable, but expect that improvements can be made through further work on the subject.

References

Buland, A., O. Kolbjornsen, and H. Omre, 2003, Rapid spatially coupled AVO inversion in the Fourier domain: Geophysics, **68**, 824-836.

Buland, A. and H. Omre, 2003, Bayesian linearized AVO inversion: Geophysics **68**, 189-198.

González-Farías, J., A. G., Dominguez-Molina, and A. K. Gupta, 2004, The closed skew normal distribution. *in* M. G. Genton, ed. Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality. Boca Raton, FL: Chapman and Hall/CRC, pp. 25-42.

Gouveia, W. and J. A. Scales, 1998, Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis, Journal of Geophysical Research, **103**, 2759–2779.

Karimi, O. and M. Mohammadzadeh, 2008, Bayesian spatial prediction for closed skew Gaussian random field. submitted for publication.

Mosegaard, K. and A. Tarantola, 1995, Monte Carlo sampling of solutions to inverse problems: Journal of Geophysical Research, **100**, 12431-1247.