Hamiltonian Monte Carlo algorithm for target-oriented and interval-oriented AVA inversions

ABSTRACT

A reliable assessment of the posterior uncertainties is a crucial aspect of any Amplitude Versus Angle (AVA) inversion due to the severe ill-conditioning of this inverse problem. To accomplish this task numerical Markov chain Monte Carlo algorithms are usually employed when the forward operator is non-linear. The downside of these algorithms is the considerable number of samples needed to attain stable posterior estimations especially in high-dimensional spaces. To overcome this issue, we assess the suitability of Hamiltonian Monte Carlo (HMC) algorithm for non-linear target-oriented and interval-oriented AVA inversions for the estimation of elastic properties and associated uncertainties from pre-stack seismic data. The target-oriented approach inverts the AVA responses of the target reflection by adopting the non-linear Zoeppritz equations, whereas the interval-oriented method inverts the seismic amplitudes along a time interval using a 1D convolutional forward model still based on the Zoeppritz equations. HMC uses an artificial Hamiltonian system in which a model is viewed as a particle moving along a trajectory in an extended space. In this context, the inclusion of the derivatives information of the misfit function makes it possible long-distance moves with a high probability of acceptance from the current position towards a new independent model. In our application we adopt a simple Gaussian a-priori distribution that allows for an analytical inclusion of geostatistical constraints into the inversion framework and we also propose a strategy that replaces the numerical computation of the Jacobian with a matrix operator analytically derived from a linearization of the Zoeppritz equations. Synthetic and field data inversions demonstrate that the HMC is a very promising approach for Bayesian AVA inversion that guarantees an efficient sampling of the model space and retrieves reliable estimations and accurate uncertainty quantifications with an affordable computational cost.

INTRODUCTION

One of the main goals of seismic inversion is the estimation of the elastic subsurface properties around the investigated area from the acquired seismic data. From a mathematical point of view this process is an ill-conditioned inverse problem (Menke, 2018; Aster et al. 2018) in which many models can fit the observed data equally well. For this reason, the Bayesian approach is usually adopted to quantify the uncertainties affecting the recovered solution. The posterior probability density (PPD) function is the final solution of a Bayesian inversion (Tarantola, 2005) and expresses the conditional probability of model parameters given the observed data. An analytical derivation of the PPD is only possible in case of linear forward operators and Gaussian, Gaussian-mixture, or generalized Gaussian distributions of unknown parameters. Otherwise, numerical methods must be employed to sample the posterior model. To accomplish this task Markov Chain Monte Carlo methods are often used. These methods convert the inverse problem into a sampling problem in which the sampling density is proportional to the posterior, so that the sampled models can be used to approximate the statistical properties of the PPD (Sen and Stoffa, 1996; Sambridge and Mosegaard, 2002). The first stage of the MCMC sampling (usually called the burn-in period) can be viewed as a global optimization that moves from a random starting model to a high-probability region of the model space. The second stage is often called the sampling stage in which the small fluctuations of the misfit value indicate that the MCMC algorithm reaches the stationary regime. Usually the samples accepted during the burn-in period do not accurately represent the target density and for this reason they are disregarded when computing the posterior density.

On the one hand, MCMC methods have been successfully applied to solve many geophysical inverse problems (Sambridge and Mosegaard, 2002; Malinverno, 2002; Bosh et al. 2007; Aleardi et al. 2018) as they can theoretically assess the posterior uncertainties in cases of non-linear forward operators and non-parametric prior distributions. On the other hand, these methods require a considerable computational effort with respect to the analytical inversion approach and for this reason specific

MCMC recipes tailored to the problem at hand are usually needed to efficiently sample the PPD. Suboptimal recipes usually result in low acceptance rates and slow convergence of the chain especially in case of high-dimensional model spaces with posterior density highly localized within specific model dimensions. To partially overcome this issue many approaches have been proposed over the last decades: For examples adopting self-adaptive MCMC algorithms (Tierney and Mira 1999; Haario et al. 1999; Haario et al. 2001; Malinverno 2002), hybridizing MCMC algorithms with global search methods (Turner and Sederberg 2012; Vrugt 2016; Aleardi and Mazzotti 2017), or employing sophisticated updating schemes that ensure an efficient exploration of the model space (Haario et al. 2006; ter Braak and Vrugt 2008). Another crucial issue of standard MCMC algorithms is the correlation between the sampled models. Indeed, it is known that for independent samples the approximation error of MCMC is inversely proportional to the square root of the number of sampled models (MacKay, 2003). When samples are correlated, not only the convergence is usually much slower, but there is risk to derive biased PPD estimations. For this reason, not all the samples collected after the burn-in period are usually used to numerically estimate the posterior, but several iterations of the algorithm are allowed to elapse in between successive samples.

Hamiltonian Monte Carlo (HMC) was designed to circumvent the previously mentioned issues of MCMC algorithms. This method was particularly implemented for problems where the derivative of the PPD with respect to the unknown parameters can be computed quickly (Neal 2011; Betancourt 2017). Introducing the derivative into the sampling framework, helps the algorithm to focus on the most promising model space regions (i.e., regions with more plausible models and then high PPD values), thus wasting fewer samples and computational resources. The HMC method was originally developed in the context of lattice quantum chromodynamics (e.g., Duane et al. 1987). Afterwards, the method was extended to Bayesian neural networks (Neal, 1996) and included in textbooks (MacKay, 2003; Bishop 2006).

HMC has been applied to various domains for more than 20 years (e.g., Neal, 1996) but only over the very last years it has been applied to solve geophysical inverse problems (e.g. Muir, and Tkalcic, 2015; Sen and Biswas 2017; Fichtner and Simuté, 2018; Fichtner et al. 2019). The growing attention to the HMC algorithm within the geophysical community can be easily explained by considering that this approach instead of completely relying on stochastic criteria (such as popular global search methods as genetic algorithms or particle swarm optimization), is built upon a solid theoretical foundation that makes it particularly well suited to tackle high-dimensional problems. More in detail, HMC considers a model as a particle that moves from its current position to a new position along a given trajectory, which is uniquely determined by a set of the particle mass, the kinetic energy, and the potential energy. In particular, the potential energy is interpreted as the misfit function. Note that the kinetic energy and the mass matrix are artificially introduced as auxiliary quantities and allow for the inclusion of the derivative information of the misfit function into the sampling framework. Similarly to standard MCMC algorithms, the ensemble of accepted HMC models after the burn-in period is used to numerically compute the PPD. The main benefits of HMC with respect to standard MCMC algorithms are its rapid convergence toward the stationary regime, so that the burn-in phase is drastically reduced, and its ability to make long jumps in the model space, so that the independence of the sampled models is guaranteed. These aspects reflect on more accurate posterior estimations and in a substantial reduction of the sampling stage. Indeed, for an optimal tuning of the hyperparameters, the acceptance rate in standard MCMC sampling (such as for the Metropolis-Hastings method) lies between 0.2-0.4 (Sambridge and Moseggard, 2002), while in case of optimal tuning the acceptance rate of HMC is around 0.6 (Neal, 2011).

In this paper we apply an HMC algorithm for target- and interval-oriented amplitude versus angle inversions in which the elastic properties of P-wave velocity (*Vp*), S-wave velocity (*Vs*) and density (ρ) are estimated from pre-stack seismic data. The target-oriented approach inverts the AVA responses of the target reflection extracted along a previously interpreted 2D stratigraphic horizon (Adriansyah, and McMechan, 2001; Mazzotti and Zamboni 2003; Zhu and McMechan, 2012; Aleardi

and Mazzotti 2014; Gong and McMechan 2016; Aleardi et al. 2017). This method is less computationally expensive than the interval-oriented inversion, but it requires a previous accurate geological interpretation phase that accurately identifies the target reflection throughout the inverted 3D volume. Another, crucial step is the extraction of the reflection coefficients for the target reflection and many methods have been proposed to reliably accomplish this task (Grion et al. 1998). Instead, the interval-oriented approach inverts the pre-stack seismic response extracted along a given time-interval (Buland and Omre, 2003) assuming a 1D convolutional forward operator. In our case, both the target- and interval-oriented inversions adopt the non-linear, exact Zoeppritz equations as the forward modeling (Zoeppritz, 1919). This set of four equations expresses the complete relationship between the angle dependent reflection coefficients and the elastic contrasts at the reflecting interface. The Zoeppritz equations are non-linear and impede an analytical assessment of the posterior model. For this reason, linear approximations of the full Zoeppritz equations (see Thomas et al. 2016) are routinely used to solve the AVA inversion problem, but a linear forward model might not be sufficiently accurate to describe the relation between seismic data and elastic parameters at far source-receiver offsets (i.e., incidence angles higher than 30°-35°) and significant elastic contrasts at the reflecting interface. In these cases, oversimplified forward operators could result in biased model parameter estimations. For these reasons, over the years many MCMC recipes have been proposed for casting the non-linear AVA inversion into a solid Bayesian framework (see, among many others, Bosch et al. 2007; Bosch et al. 2010; Rimstad and Omre, 2010; Zunino et al. 2015; Aleardi and Salusti, 2019). This work, instead of proposing another MCMC recipe targeted to non-linear Bayesian AVA inversion, is aimed at assessing the suitability of the HMC approach for solving this kind of inverse problem. In particular, the use of an analytical forward operator (i.e., the Zoeppritz equations) makes it possible the computation of the derivative of the PPD with an affordable computational cost. In addition, the AVA inversion is severely ill-conditioned (Avseth et al. 2010) and for this reason it is crucial including as much a-priori information as possible (e.g., expected lateral variability and mutual interdependence of the inverted parameters) to successfully reduce the null-space of solutions

(de Figueiredo et al. 2018; Aleardi et al. 2019). In this work we assume a Gaussian prior model that makes it possible the inclusion of a variogram model into the misfit function.

We start by presenting the HMC algorithms and the AVA inversion approaches we implemented. Afterward, synthetic data inversions are used to illustrate the reliability of the HMC inversion. Then, we focus on the synthetic interval-oriented inversion and detail the differences between a Metropolis-Hastings MCMC and the HMC samplings. We also propose a strategy that can be used to avoid the numerical computation of the Jacobian in case of AVA inversion. Finally, the implemented algorithms are applied to invert field onshore seismic data acquired over a quite complex geological setting where a gas-saturated reservoir is hosted in a shale-sand sequence. To the best of our knowledge this paper discusses the first application of HMC algorithm to target-oriented and interval-oriented AVA inversions

METHODS

The implemented Hamiltonian Monte Carlo algorithm

In this section we briefly review the basic theory of HMC before discussing in more detail our implementation. We refer the interested reader to Neal (2011) and Betancourt (2017) for more theoretical insights into the HMC algorithm.

The HMC is based on the Bayesian inversion framework. In this context the solution of an inverse problem is the posterior probability density (PPD) function that is defined as follows:

$$p(\mathbf{q}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{q})p(\mathbf{q})}{p(\mathbf{d})}, \quad (1)$$

where $\mathbf{d}$ is the N-dimensional observed data vector, and $\mathbf{q}$ is the Q-dimensional model parameter vector. The left-hand side term of equation 1 represents the target PPD that could be numerically estimated from the ensemble of models sampled after the burn-in phase by a Monte Carlo algorithm.

HMC improves the performances of MCMC by proposing models that are distant from the current state but still characterized by a high probability of acceptance. The proposed models are found by simulating the Hamiltonian dynamics, in which a model is viewed as a moving particle with a physical state defined by the momentum vector and position vector: these vectors lie in the so-called phase space. HMC samples an auxiliary distribution defined over the 2Q-dimensional phase space, from which the PPD samples are drawn by neglecting the momentum component. The particle trajectory in the phase space is determined by the kinetic energy (K), the potential energy (U), and the mass matrix (**M**). The potential energy is the negative natural logarithm of the posterior (see equation 1) or in other terms is the misfit function associated to the inverse problem. In this context more plausible models with large PPD values are associated to low potential energies. Generally speaking, the potential energy is given by:

$$U(\mathbf{q}) = -\ln\big(p(\mathbf{q}|\mathbf{d})\big). \quad (2)$$

HMC determines the kinetic energy by introducing an auxiliary variable (momentum variable) **p** defined over a Q-dimensional space. It usually assumes that the auxiliary momentum variable has a multivariate normal distribution with zero mean and a covariance matrix equal to the so-called mass matrix:

$$K(\mathbf{p}) = N(\mathbf{p}; 0, \mathbf{M}), \quad (3)$$

where N represents the Gaussian distribution, and **M** is the $Q \times Q$ mass matrix (note that the number of rows and columns of **M** are equal to the number of model parameters) that must be accurately set to ensure the convergence of the HMC algorithm (Fichtner et al. 2019). The vectors **p** and **q** define the 2Q-dimensional phase space, whereas the Hamiltonian $H(\mathbf{p}, \mathbf{q})$ gives the total energy of the moving particle:

$$H(\mathbf{p}, \mathbf{q}) = U(\mathbf{q}) + K(\mathbf{p}). \quad (4)$$

After defining the kinetic and the potential energies, the Hamiltonian dynamics can be simulated. In this context, the model **q** moves through the 2Q-phase space according to Hamilton's equations:

$$\frac{dq_i}{d\tau} = \frac{\partial K}{dp_i}, \qquad \text{with } i = 1, 2, \dots, Q, \quad (5)$$

$$\frac{dp_i}{d\tau} = -\frac{\partial U}{dq_i}, \quad \text{with } i = 1, 2, \dots, Q, \quad (6)$$

where τ indicates the artificially introduced time variable. Note that the right term of equation 6 contains the partial derivative of the potential energy (i.e., the misfit function) with respect to the considered model **q**. This makes it possible introducing information about the gradient of the misfit function into the Monte Carlo sampling framework.

For each current model **q**, and for each iteration, HMC executes the following steps:

1. Determine the Q momenta p_i by drawing random realizations from the normal distribution $N(\mathbf{p}; 0, \mathbf{M})$;
2. Derive the proposed model $\mathbf{q}(\tau)$ and the new momenta $\mathbf{p}(\tau)$ by solving Hamilton's equations 5 and 6 for a given propagation time τ. In this work we use the leap-frog method as the numerical integration method (Betancourt, 2017);
3. Accept the proposed model with a probability α given by:

$$\alpha = \min\left[1, \frac{\exp\left(-H\big(\mathbf{p}(\tau), \mathbf{q}(\tau)\big)\right)}{\exp\left(-H(\mathbf{p}, \mathbf{q})\right)}\right]. \quad (7)$$

   If accepted, the proposed $\mathbf{q}(\tau)$ point constitutes the starting model for the next trajectory ($\mathbf{q} = \mathbf{q}(\tau)$). Otherwise, the current model **q** is again used as the starting point in the next iteration;
4. Return to step 1 and iterate until a maximum number of HMC iterations is reached.

Note that standard MCMC algorithms explore model space slower than HMC, because in HMC all the model parameters are updated in each iteration, so that long distances in the phase space can be

traversed with a single move. This promotes both a high level of acceptance and the independence of the sampled models. Differently, standard MCMC usually modifies only a subset of the model parameters in each iteration. This strategy ensures a high level of acceptance but at the expense of a high degree of correlation between successively sampled models.

In this work we use equation 3 to define the momentum distribution. Since it is known that standard HMC exhibits poor performance in sampling multimodal target densities (see the discussion section), we employ a Gaussian prior model. In particular, we assume Gaussian-distributed *Vp*, *Vs*, and density in the target-oriented inversion, whereas on the line of Buland and Omre (2003) the interval-oriented inversion assumes log-Gaussian distributed elastic parameters. Let $\mathbf{q}$ represents the vector of elastic parameters in the target-oriented inversion and the vector containing the natural logarithm of *Vp*, *Vs*, and density in the interval-oriented approach. In both cases the prior model can be compactly written as follows:

$$p(\mathbf{q}) = N\left(\mathbf{q};\ \mathbf{q}_{prior}, \mathbf{C}_q\right), \quad (8)$$

where $\mathbf{q}_{prior}$ and $\mathbf{C}_q$ are the a-priori mean vector and the a-priori model covariance matrix that can be derived, for example, from available borehole information. In both the target- and interval-oriented inversions we assume that the prior model is laterally and temporarily invariant. The analytical Gaussian prior allows us to easily include geostatistical constraints into the inversion kernel, but an outstanding benefit of the HMC method is that it can be applied also for non-parametric prior models. In our case the matrix $\mathbf{C}_q$ expresses both the covariances of model parameters and their spatial relationship. In particular, the matrix $\mathbf{C}_q$ is computed as a Kronecker product between a stationary model covariance matrix and the spatial and/or temporal correlation functions expressing the spatial dependency of model parameters (e.g., Buland and Omre, 2003).

In this work the potential energy is defined as:

$$U(\mathbf{q}) = -\ln\left(p(\mathbf{q}|\mathbf{d})\right) = \frac{1}{2}\left[\left(\mathbf{d} - G(\mathbf{q})\right)^T \mathbf{C}_d^{-1}\left(\mathbf{d} - G(\mathbf{q})\right) + \left(\mathbf{q} - \mathbf{q}_{prior}\right)^T \mathbf{C}_q^{-1}\left(\mathbf{q} - \mathbf{q}_{prior}\right)\right], \quad (9)$$

The partial derivative of equation 9 with respect to the model parameters is equal to:

$$\frac{\partial U(\mathbf{q})}{\partial \mathbf{q}} = \mathbf{J}\mathbf{C}_d^{-1}\big(\mathbf{d} - G(\mathbf{q})\big) + \mathbf{C}_q^{-1}\big(\mathbf{q} - \mathbf{q}_{prior}\big), \quad (10)$$

In equations 9 and 10 G is the non-linear forward operator based on the exact Zoeppritz equations, $\mathbf{d}$ is the observed data (P-wave reflection coefficients for the target-oriented approach and partially stacked common-midpoint "CMP" gathers for the interval-oriented method), $\mathbf{C}_d$ is the data covariance matrix that expresses the error on the observed data (under the assumption of a Gaussian distribution) and/or modelling errors (Menke, 2018). $\mathbf{J}$ is the Jacobian matrix that we can compute with a forward finite-difference approach (Aster et al. 2011), in which the partial derivative of the i-th data points with respect to the k-th model parameter can be derived as follows:

$$\frac{\partial d_i}{\partial q_k} = \frac{d_i(q_k + h_k) - d_i(q_k)}{h_k}, \quad (11)$$

where h_k is the increment for the k-th model parameter. If needed a more accurate finite-difference scheme (for example a central-difference approach) can be used at the expense of an additional computational cost. In the following examples, computing the Jacobian with a finite-difference approach is computationally feasible because the Zoeppritz equations can be analytically solved with a minimum computational effort. However, in the following we also propose a different approach to significantly speeds up the computational cost of the AVA-HMC inversion, that replaces the Jacobian with a linear matrix operator computed from the Aki and Richards (1980) approximation of the full Zoeppritz equations.

A proper setting of the mass matrix ($\mathbf{M}$) is of crucial importance in any HMC sampling (Fichtner et al. 2019) because this parameter controls the speed with which the algorithm traverses the phase space. In other words, the mass matrix decorrelates the target distribution (Betancourt, 2017). In practice, a proper setting of this matrix maximizes the exploration of independent models and prevents the exploration of similar model space regions. The optimal setting of the $\mathbf{M}$ matrix is

strongly case dependent, but typically **M** is diagonal and a scalar multiple of the identity matrix. However, this strategy often hampers an efficient sampling of the parameter space. Along the lines of Fichtner et al. (2019) we compute the mass matrix as a local approximation (around the considered model) of the inverse of the posterior covariance matrix:

$$\mathbf{M} = \mathbf{J}^T \mathbf{C}_d^{-1} \mathbf{J} + \mathbf{C}_q^{-1}. \quad (12)$$

Note that the mass matrix brings information on the local curvature of the misfit function. In the leapfrog method, we update the momentum and position variables sequentially. First, we simulate the momentum dynamics (changing momentum) by $\delta/2$ time units, then we simulate the position dynamics (moving in model space) for δ time units, then again completing the momentum simulation for the remaining half-time units, $\delta/2$. This process is repeated for a total of L times after which the algorithm reaches a new state. This leapfrog integration method is described by the following equations:

$$p_i(\tau + \delta/2) = p_i(\tau) - \frac{\delta}{2} \frac{\partial U}{\partial q_i}|_\tau, \quad (13)$$

$$q_i(\tau + \delta) = q_i(\tau) + \delta \frac{\partial K}{\partial p_i}|_{\tau + \frac{\delta}{2}}, \quad (14)$$

$$p_i(\tau + \delta) = p_i\left(\tau + \frac{\delta}{2}\right) - \frac{\delta}{2} \frac{\partial U}{\partial q_i}|_{\tau+\delta}, \quad (15)$$

where τ is the time variable. Although, many other integrator methods exist (Blanes et al. 2014), here we use the leapfrog method for its easy implementation. However, note that the integrator methods used for HMC inversion have to fulfill certain properties such as reversibility and volume preservation (Neal, 2011).

In addition to the mass matrix, also the choice of L and δ plays a crucial role in the efficiency of the sampling and for this reason they need to be set properly to get an optimal acceptance rate (around 0.6). In particular, they determine the trajectory of the sampling in the augmented model space. A too

short trajectory generates proposal very close to the current model, thus slowing down the exploration of the parameter space. Differently, if the trajectory is too long the algorithm resamples points that have been already visited, thereby wasting computing time. To prevent the locking of the sampling in periodic trajectories we follow the strategy discussed in Mackenzie (1989). According to this approach, in each HMC iteration the L parameter is randomly drawn from a previously defined uniform distribution (see the discussion section).

In Appendix A we summarize the different HMC hyperparameters and the statistical characteristics of the prior model used in the following inversion experiments.

Target- and interval-oriented AVA inversions

In the target-oriented inversion the observed data are the PP-wave reflection coefficients R_{pp} pertaining to the target reflection. We consider a single reflecting interface that separates two homogeneous, and isotropic half spaces. The aim is to infer the properties of the layer below the reflecting interface (the target layer) by exploiting its AVA response. For a single partially stacked CMP gather and following a matrix formalism, the forward equation of the target-oriented inversion is given by:

$$\mathbf{d} = G(\mathbf{q}) + \mathbf{n}, \qquad (16)$$

where $\mathbf{d}$ are the observed data, $\mathbf{n}$ is the noise (usually assumed to be Gaussian distributed), $\mathbf{q}$ represents the elastic parameters of the underlying layer, and G is the forward operator. More in detail:

$$\mathbf{d} = [Rpp(\theta_1), \dots, Rpp(\theta_P)]^T, \qquad (17)$$

$$\mathbf{q} = [Vp_2, Vs_2, \rho_2]^T, \qquad (18)$$

where the subscript 2 refers to the layer below the reflecting interface, P is the number of data points, θ is the incidence angle, whereas G is given by the exact Zoeppritz equations. For a target-oriented

AVA inversion, the data vector contains all the AVA responses estimated for each CMP gather, whereas the model vector expresses the *Vp*, *Vs*, and density of the target layer at each CMP location. In this context we can write:

$$\mathbf{d} = [\mathbf{d}^1, \mathbf{d}^2, \dots, \mathbf{d}^M]^T , \qquad (19)$$

$$\mathbf{q} = [\mathbf{q}^1, \mathbf{q}^2, \dots, \mathbf{q}^M]^T , \qquad (20)$$

where the superscripts identify the CMP gather position, and *M* is the total number of CMP gathers considered. Lateral constraints along the horizontal *x* and *y* directions are included into the target-oriented approach to promote the lateral continuity of the elastic properties in the sampled models. In the inversion, the elastic properties of the overlying layer are kept fixed to values that in the field data inversion have been determined by interpolating available well-log data. However, the uncertainties in the definition of the elastic properties of the overlying layer are properly propagated into the estimated PPD. To this end, we follow Aleardi et al. (2019) who included into the data covariance matrix both the random noise affecting the observed AVA response, and the uncertainty on the assumed elastic properties of the overlying layer that are treated as modelling AVA errors (Madsen and Hansen 2018). In other terms we consider the forward operator to be an inexact theory due to the inaccuracy of the assumed elastic attributes of the overlying layer (see Menke 2018). In the field data example, the starting point for the HMC sampling is derived by a geostatistical interpolation of the available well-log data integrated by the geological knowledge about the investigated area.

In the interval-oriented approach the forward modeling convolves the angle-dependent R_{pp} time series determined by the Zoeppritz equations, with the source wavelet **s,** thus deriving the pre-stack CMP gather:

$$\mathbf{d}(\theta, t) = \mathbf{Rpp}(\theta, t) * \mathbf{s}(\theta, t), \qquad (21)$$

where the $*$ symbol indicate the convolution, and *t* is the time. In this case each CMP gather is inverted separately under the assumption of a local 1D subsurface model and by imposing a 1D vertical

constraint in the sampled models. As previously introduced, the model vector is now constituted by the natural logarithm of the elastic properties along the considered time interval of *M* time samples:

$$\mathbf{q} = \ln[Vp^1, Vp^2, \dots, Vp^M, Vs^1, Vs^2, \dots, Vs^M, \rho^1, \rho^2, \dots, \rho^M]^T. \quad (22)$$

In both inversions, a Gaussian variogram model is used to include spatial/temporal constraints into the target- and interval-oriented inversions. These constraints act as regularization terms and contribute to decrease the ill-conditioning of the inverse problem. The lateral spatial constraints in the target-oriented inversion and the vertical temporal constraint in the interval-oriented approach can be generically defined as:

$$\tau = \exp\left(-\frac{h^2}{\alpha^2}\right), \quad (23)$$

where h is the spatial or temporal axis of the autocorrelation function, and α is the parameter that defines the spatial/temporal dependency. The lateral amplitude variability of the seismic data, integrated by the available well-log information and by the geological knowledge of the investigated area, can be used to define the spatial constraints. In the interval-oriented inversion the autocorrelation of available well-log data is usually employed to properly set the temporal constraints. However, in field data inversions due to the very complex spatial relationship of the inverted parameters (i.e., non-stationary variogram model) a quality control of the inversion results and manual adjustments of the so derived autocorrelation functions are often needed to obtain the desired spatial continuity of the results.

In the interval-oriented approach the starting model for the HMC inversion corresponds to the so called low-frequency elastic background model (Buland and Omre, 2003) that is derived by interpolating available well-log data.

Avoiding the numerical computation of the Jacobian

We now describe an approach that significantly reduces the computational effort of the HMC-AVA inversion because avoids the numerical computation of the Jacobian matrix. This strategy can be applied to both the interval- and the target-oriented inversion, but for brevity in the following we only consider the former approach. In this case we take the time-interval extension of the single-interface Aki and Richards formula (Buland and Omre, 2003):

$$Rpp(t,\theta) = \frac{1}{2}(1+\tan^2(\theta)\,)\frac{\partial}{\partial t}\ln Vp(\mathrm{t}) + 4\frac{Vs^2(t)}{Vp^2(t)}\sin^2(\theta)\frac{\partial}{\partial t}\ln Vs(t)$$
$$+\frac{1}{2}\left(1-4\frac{Vs^2(t)}{Vp^2(t)}\sin^2(\theta)\right)\frac{\partial}{\partial t}\ln\rho(t)$$
$$=\alpha_{Vp}(t,\theta)\frac{\partial}{\partial t}\ln Vp(t)+\alpha_{Vs}(t,\theta)\frac{\partial}{\partial t}\ln Vs(t)+\alpha_{\rho}(t,\theta)\frac{\partial}{\partial t}\ln\rho(t). \quad (24)$$

If we consider the convolutional model and adopt the matrix formalism, the linear forward matrix operator linking the model to the data can be analytically computed as:

$$\mathbf{G} = \mathbf{SAD}, \quad (25)$$

where **S** is the wavelet matrix, **A** contains the numerical coefficients $\alpha_{Vp}(t)$, $\alpha_{Vs}(t)$ and $\alpha_{\rho}(t)$ of equation 24, **D** is the first-order numerical derivative operator, and **G** is the linear matrix operator that substitutes the Jacobian. If **G** replaces the Jacobian, the mass matrix can be derived as:

$$\mathbf{M} = \mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{G} + \mathbf{C}_q^{-1}. \quad (26)$$

Note that in the target-oriented approach equation 24 is replaced by the single-interface Aki and Richards equation (Aki and Richards, 1980).

According to equation 24, the knowledge of the *Vp*/*Vs* ratio along the time interval to be inverted is needed to derive the numerical values forming the **G** matrix and for this reason we propose to recompute this matrix operator for each sampled model and the associated *Vp*/*Vs* ratio (i.e., the computational cost for computing **G** is negligible). Note that if we replace the Jacobian with the matrix **G**, we are inherently assuming that the curvature of the misfit function is constant over the entire

model space. In addition, the validity of the linear operator depends on the considered angle range and on the elastic contrasts at the interface and for this reason the suitability of this strategy should be evaluated case-by-case (see the discussion section).

SYNTHTETIC INVERSIONS

Target-oriented inversion

The true model is derived from a geostatistical simulation driven by well-log data investigating a productive gas field located offshore. The model represents a slice of a stratigraphic grid where a meandering sand delta-channel system is hosted in a shale sequence. Due to the large dimension of the simulated 2D model (400 in-line and 400 cross-line sections defined over a regular grid of 25×25 m) and for computational feasibility reasons, we split the entire model in blocks with dimension of 50 cross-lines and 50 in-lines. This result in a total number of unknowns for each block equal to $50 \times 50 \times 3 = 7500$. In this high-dimensional model space the inclusion of the geostatistical constraints is particularly useful to reduce the null-space of solution, that is the ensemble of physically plausible models that equally fit the observed data. The inversion is independently run for each block and the final 2D maps of elastic properties are obtained by merging the inversion results retrieved in each block**.** For each block we run the HMC algorithm for 50000 iterations, whereas the models sampled after 100 iterations are used to numerically compute the posterior model. In this case we assume that the elastic parameters of the overlying medium are perfectly known. The Zoeppritz equations are applied to the true model to simulate the observed P-wave reflection coefficients for each CMP gather and for an angle range between 0° and 45°. Gaussian random noise is added to the simulated data to get a signal-to-noise ratio (SNR) of 5.

In Figure 1a we show the true model that simulates a complex meandering system with isolated and interconnected sand bodies surrounded by shales. The variogram model needed to define the lateral

constraints has been directly determined from the spatial autocorrelation of the true model and it is the same for the three elastic properties we invert for. Figure 1b shows the starting point for the HMC inversion that is a heavily smoothed version of the true model obtained by applying a simple moving average filter to the true parameter values. Figure 1c shows the mean a-posteriori solution estimated by the HMC algorithm, where we can appreciate how the inversion correctly estimates the lateral variability of the three elastic properties, while the lateral constraints efficiently preserve the lateral transition boundaries between the shale and sand formations.

Figure 2 shows some examples of 1D marginal posterior distributions for *Vp*, *Vs* and density pertaining to different CMP gather positions. Note that the peak of the posterior is always very close to the true model. In this case the posterior is very close to a Gaussian distribution because the inversion is weakly non-linear in the angle range we consider. Finally, Figure 3 shows that less than 100 iterations are required to converge to the stationary regime after which the misfit value fluctuates around a stable value.

We implement a parallel Matlab code running on two deca-core intel E5-2630 at 2.2 GHz (128 Gb RAM). The parallelization of the code is particularly useful because each column of the Jacobian matrix can be independently computed, so that the computation of **J** can be distributed easily across different processors. The acceptance rate for this inversion was higher than 0.9 in the pre-burn in phase in which the chain rapidly moves from the starting model toward a high probability region in the model space. The acceptance rate stabilizes around 0.6 in the sampling, post-burn-in stage. The computational time to invert a single cell of 50 in-line and 50 cross-line is 15 minutes, approximately. A more efficient and scalable code is needed to invert the entire model of 400 in-line and 400 cross-line at once. We leave a more detailed discussion about the convergence of the implemented interval-oriented HMC algorithm for the field data inversion.

Interval-oriented inversion

In the following synthetic example, the true model is derived from the borehole information pertaining to a single well investigating a productive gas field located onshore, whereas the a-priori information are derived from five other wells drilled in the same area. To compute the synthetic seismic data, we employ a Ricker wavelet with a peak frequency of 50 Hz, a time sampling of 0.001 s, and an angle range of 5°-45°. The time-interval we consider is composed of 91 time samples, thus resulting in $91 \times 3 = 273$ parameters to be determined. The vertical constraints have been derived by approximating the variogram of the upscaled well-log data used to compute the prior model. Gaussian random noise is added to the synthetic data to simulate a SNR of 5. We run the inversion for 5000 iterations in which the first 100 models are disregarded from the computation of the PPD to properly burn-in.

Figure 4 shows that the inversion recovers the vertical variability of the actual property values and provides mean a-posteriori solutions with a good match with the actual *Vp*, *Vs*, and density values. Note that the vertical correlation included into the misfit function, efficiently constraints the vertical variability of the sampled elastic models, thus providing a final solution with a realistic vertical variability of the three elastic properties. As for the target-oriented inversion, the inclusion of the geostatistical a-priori constraints is particularly useful to reduce the length of the sampling stage, because these vertical constraints restrict the model space region containing physically plausible models that are in accordance with the prior assumptions. As expected, if we move from *Vp*, to *Vs* and to density we observe an increase of uncertainty and an overall decrease of the match between the recovered solution and the true model. The good match between the observed data and the data computed on the estimated mean model illustrates the convergence of the algorithm to the stationary regime. From the evolution of the misfit function (Figure 5), we note that only 80 iterations are required to reach the stationary regime, after which the misfit oscillates around a stable value. The acceptance rate was higher than 0.95 in the burn-in phase and around 0.6 in the post burn-in stage. The total computational time is 10 minutes using a parallel Matlab code running on the same hardware configuration previously described in the target-oriented example. Again, we leave a more detailed

discussion about the convergence of the implemented interval-oriented HMC algorithm and the properties of the sampled models to the following sections.

A comparison of MCMC and HMC

We now discuss in more detail the difference between a Metropolis-Hastings MCMC and the implemented HMC inversion. For brevity we limit our discussion to the interval-oriented inversion, but similar conclusions would have been drawn for the target-oriented approach. The true model is the same used in the interval-oriented inversion discussed previously. For a more quantitative comparison of the HMC and MCMC samplings we evaluate the convergence of the two algorithms toward a stable posterior model by computing the potential scale reduction factor (PSRF) for some model parameters. This computation requires the use of multiple chains from which the so called within-walk and between-walk variances can be evaluated for each unknown. The PSRF value decreases to 1.0 as the number of drawn samples increases. A PSRF value higher than 1.2 (Gelman et al. 2013) usually indicates that the within-walk variance is small compared to that between the walks and that additional MCMC iterations are needed to achieve a stable PPD estimation.

We implement a Metropolis-Hastings MCMC in which at each iteration we randomly select a given temporal location *t* where the current *Vp, Vs* and density model must be perturbed. The perturbation follows a multivariate Gaussian proposal with a zero mean and a covariance properly set to ensure an acceptance rate around 0.2-0.3. To preserve the geostatistical constraints in all the sampled models, a kriging interpolation is used to propagate the perturbation for the *t* time instant to the neighboring time positions (Aleardi and Salusti, 2019). Note that we only perturb a small subset of the whole model parameters vector (the parameters pertaining to the selected time position and the neighboring ones) to maintain a sufficiently high acceptance rate. Perturbing multiple time positions permits a reduction of the correlation of the sampled models but at the expense of a substantial decrease of the acceptance rate. We use four independent MCMC chains starting from models randomly generated according to the a-priori assumptions and running in parallel for 100000 iterations with a burn-in of

6000 iterations. For the HMC we replace the numerical computation of the Jacobian used in the previous synthetic tests with the analytical matrix operator **G** as given in equation 25. Similarly, for the MCMC inversion, we start the sampling from randomly generated models, and adopt four independent chains running in parallel for 5000 iterations with a burn-in of 300. The parameter setting for the HMC is equal to that used in the synthetic inversion discussed in the previous section, the only difference is that in this case we use a longer burn-in period.

Figures 6 and 7 compare the PPDs estimated by the two algorithms. We observe that both MCMC and HMC provide congruent posterior estimates with a mean value in good agreement with the true model, whereas the data computed on the predicted mean posterior models satisfactory match the amplitudes of the observed seismic data. A comparison of the evolution of L2 norm data errors for the MCMC and HMC samplings (Figure 8) shows that 6000 iterations, approximately, are needed by the MCMC to converge toward the stationary regime, while the HMC requires less than 300 iterations. Figure 9 shows examples of evolutions of the PSRF values for some model parameters (extracted at different time positions) for both the MCMC and HMC inversions. The P-wave velocity is the elastic property that exerts the major influence on observed seismic amplitudes, and for this reason we observe that for both MCMC and HMC a lower number of iterations is needed to achieve stable posterior estimates of *Vp* with respect to *Vs* and density. For the MCMC inversion (Figures 9a-c) usually 50000 and 70000 iterations are enough to obtain accurate *Vp*, and *Vs* PPDs, respectively, while for some time positions 100000 iterations are yet not enough to achieve reliable uncertainty quantifications for the density. The evolution of the PSRF for the same model parameters considered in Figures 9a-c, shows that the HMC attains stable posterior estimations for *Vp*, *Vs*, and density in less than 2000 iterations (Figures 9d-f).

Figure 10 illustrates some examples of the autocorrelations of the sampled *Vp*, *Vs* and density models at a given time position along the inverted 1D profile and computed after the burn-in phase. We observe that 1000 iterations, approximately, are needed by the MCMC to sample uncorrelated

models, while in the HMC inversion the autocorrelation value drops to zero in less than 20 iterations. This result indicates that the long-distance moves of the HMC provides highly uncorrelated samples.

This example also illustrates that in the considered case replacing the Jacobian with the analytical matrix operator **G** is a viable and successful strategy that provides accurate model predictions and uncertainty quantifications with a substantial reduction of the computational cost. Indeed, replacing the numerical computational of the Jacobian with the linear matrix operator reduces the total computing time of a single chain HMC inversion from 10 minutes to about 1.5 minutes on the same hardware previously described. For comparison we point out that the computational cost of a gradient-based, Gauss-Newton inversion of the same synthetic data performed with a Matlab code running on the same hardware configuration is 5 s, approximately, whereas a Metropolis-Hasting MCMC sampling runs in about 11 minutes. This value is similar to an HMC run in which the numerical approach to the Jacobian computation is employed, but it is much higher than that of the previous HMC inversion in which the matrix **G** replaces the Jacobian. However, note that the convergence toward the stationary regime is slower when the Jacobian is replaced by the linear operator (compare Figure 8b with Figure 5) because in this case (if we neglect the role of the data covariance matrix) we are assuming that the curvature of the misfit function is constant over the entire model space. This oversimplification decreases the convergence speed of the algorithm and increases the degree of correlation between successively sampled models. In other words, successively sampled models are maximally decoupled when the numerical approximation of the Jacobian is used instead of the analytical matrix operator (see the next sections).

Finally, we point out that in this example a more efficient MCMC sampling could be obtained by implementing more sophisticated sampling strategies. For examples, MCMC recipes that allow for better mixing of the different chains (e.g., differential evolution Markov Chain, or differential evolution adaptive Metropolis; Vrugt 2016). However, an optimal compromise between exploration

and exploitation, or in other terms between the independence of the sampled models and the convergence speed toward the stationary regime, must be found.

FIELD DATA INVERSIONS

The target- and interval-oriented HMC inversions were tested on an onshore seismic data investigating a gas field located within a clastic, shale-sand sequence. The seismic data are characterized by a maximum source-to-receiver offset of 4 km and have been processed following a true-amplitude processing sequence. The data have a peak frequency around 40 Hz and a time sampling of 4 ms. Figure 11 displays an in-line section extracted from the 3D seismic stack volume, where we can observe the high amplitude anomaly marking the transition between the shale (cap-rock) and the target sands (indicated by the black arrow). The seismic data along this in-line section will be considered in the interval-oriented inversion, whereas the top reservoir AVA response will be used by the target-oriented inversion.

Figure 12 represents the root-mean-square (RMS) amplitude map extracted around the time-slice representing the top reflection of the investigated reservoir interval. Low amplitudes correspond to shale while high amplitudes identify the sand bodies. The green rectangle encloses the area considered in the target-oriented inversion, whereas the black arrow indicates the main gas sand body object of this study characterized by significant elastic contrasts with respect to the overlying layer. This map shows some high-frequency noise without spatial coherency that we attribute to residual noise contamination in the data. The lateral constraints included into the target-oriented approach are devoted at attenuating the effect of this residual noise in the final solution.

The a-priori information used to define the vertical and lateral variability, the cross-correlation of the elastic properties, the elastic properties of the overlying layer (for the target-oriented inversion), and the low-frequency elastic background model (for the interval-oriented inversion) are derived from the

available well-log data investigating the target layer. We consider the covariance and the spatial correlation of this prior model to be stationary that is spatially and temporarily invariant over the entire inverted domain. The noise standard deviation has been derived by comparing adjacent AVA responses (in the target-oriented inversion) and by comparing adjacent CMP gathers (in the interval-oriented inversion), therefore assuming that these variations are only related to noise contamination (see Aleardi et al. 2018). In both the target- and interval-oriented inversions we consider an angle range between 15° and 40°. The angle-dependent wavelets input for the inversion stage have been estimated through a least-square truncated SVD inversion in which the reflectivity matrix has been derived from available borehole data (Bianco, 2016). In the following inversions we use the numerically computed Jacobian to derive the mass matrix. Appendix A shows the HMC parameter setting and the principal characteristics of the prior model employed in the following inversion experiments.

Target-oriented inversion

In this case the AVA responses of the top reflections of the investigated reservoir have been extracted from pre-stack Kirchhoff time-migrated CMP gathers. The reliability of the available pre-stack seismic data has been assessed by performing seismic-well-tie and by comparing the observed and the synthetic AVA responses derived from available well-log data.

Figures 13a-b show close-up images of two CMP gathers extracted from the seismic volume and located near the main gas sand body. Note the strong negative amplitude anomalies marking the transition between the overlying shale and the reservoir sand. Figures 13c-d illustrate the AVA responses of the top reservoir reflections extracted from the two seismic gathers shown in Figures 13a-b.

Figure 14 represents the actual autocorrelation functions computed from the lateral variation of the seismic amplitudes along the in-line and cross-line directions, together with the approximated

(analytical) autocorrelation functions computed by assuming a Gaussian variogram model. The inclusion of this theoretical variogram model into the HMC inversion imposes lateral constraints to the elastic models sampled by the HMC algorithm, thus reducing the model null space. The elastic properties of the upper layer have been derived from a geostatistical interpolation of the logged elastic values pertaining to the cap rock shale. The uncertainty resulting from this interpolation has been included into the data covariance matrix as previously discussed. The same interpolation method has been used to derive the starting point for the HMC inversion (Figures 15a-c). In the inversion we use a single HMC chain running for 70000 iterations. All the models sampled after a burn-in period of 100 iterations are considered in the computation of the final results (mean a posteriori model, 1D marginal distribution for the elastic parameters, and standard deviation maps).

Figures 15d-f show the estimated *Vp, Vs* and density posterior mean models around the top of the considered reservoir layer. High *Vp*, *Vs*, and density values pertain to shale, while low elastic property values identify the sand bodies. The lowest *Vp*, *Vs*, and density values pertain to the gas-saturated sand body (see the black arrows in Figures 15d-f). Note that the lateral constraints included into the inversion algorithm efficiently attenuate the scattering visible in the RMS map of Figure 12 and provide final estimations with realistic lateral continuity of the elastic parameters.

Figure 16 represents some examples of marginal PPDs for *Vp, Vs*, and density pertaining to different CMP gather positions. As expected, the uncertainties increase moving from *Vp*, to *Vs* and to density. Note that the mildly non-linear forward operator makes these distributions very close to a Gaussian model.

In Figure 17 we represent the estimated posterior standard deviation map for *Vp*, *Vs* and density numerically computed from the ensemble of sampled models. Note, that the uncertainty for a given model parameter is inversely proportional to the curvature of the misfit function and so it depends on the likelihood function and on the prior information infused into the inversion (see equation 9 and

Menke, 2018). In our case, the forward modeling is non-linear and for this reason such curvature, and then the model uncertainty, is expected to vary over the model space.

Figure 18 shows the normalized autocorrelation coefficients computed for 2000 successively sampled *Vp*, *Vs* and density models. The fast drop of the autocorrelation to 0 demonstrates that the HMC algorithms samples highly uncorrelated models, or in other words that the long-distance HMC moves guarantee the mixing of the Monte Carlo chain.

Figure 19a displays the evolution of the misfit function during the HMC iterations. It emerges that less than 100 iterations are needed to attain the stationary regime. However, note that both the extremely fast convergence rate and the independency of the sampled models are also favored by the weakly non-linearity of the inversion procedure and by the use of the Jacobian in the computation of the mass matrix. The independence of the HMC samples is also illustrated by the strongly variable misfit values with iterations (close-up of Figure 19b). Indeed, we observe that the variation of the misfit value for successively sampled models is comparable to the full range of misfit variation attained after the burn-in period.

To finally illustrate the reliability of the HMC predictions, we run a more standard Bayesian non-linear AVA inversion without lateral constraints (see Aleardi et al. 2017 for more details) on the same AVA responses previously considered. The comparison of the mean a-posteriori solutions provided by the HMC algorithm (Figure 15) and by the more standard AVA inversion approach (Figure 20) shows that the two methods yield congruent results. In particular, the lateral constraints included into the HMC inversion efficiently attenuate the noise propagation from the data to the model space and promote the lateral continuity of the results.

During the HMC inversion the acceptance rate in the post-burn in phase was around 0.55, whereas, the total computational cost was 35 minutes on the same hardware configuration used in the synthetic tests.

Interval-oriented inversion

Now we apply the 1D interval-oriented inversion along the in-line section shown in Figure 11. In this case we employ a single HMC chain running for 5000 iterations with a burn-in period of 100 iterations. The low-frequency elastic model that acts as the starting point for the HMC sampling has been derived from a geostatistical interpolation of well log data (Figures 21a-c). Figures 21d-f show the estimated posterior mean values for *Vp, Vs,* and density. Again, significant decreases of *Vp, Vs* and density mark the transition between the encasing shale and the target gas-sand interval (black arrows in Figure 11). We observe that the lateral continuity of the results decreases from *Vp*, to *Vs* and is particularly low for density. This can be explained by considering the relative influence played by each elastic parameter in the observed reflected amplitudes (see for example Aleardi 2015), and that each CMP gather is inverted separately with no lateral constraints included into the inversion kernel. In particular, it is well known that the *Rpp* coefficients are mainly influenced by the *Vp* parameter, while the density cannot be reliably estimated with realistic noise levels (Buland and Omre, 2003) and without ultra-far source-receiver offsets. For these reasons, we are less confident on the estimated density model.

Figure 22 represents the 1D marginal distributions estimated for the CMP located at the horizontal coordinate of 1 km (see Figure 21). Again, we note that the uncertainties increase from *Vp*, to *Vs* and are particularly high for density. However, we also observe that the estimated posterior means are very different from the starting model thus demonstrating that the observed data brings valuable information about the subsurface properties. This is also confirmed by the good match between the observed CMP gather and the data computed on the estimated mean model and by the substantial difference between the predicted data and the data computed on the starting model.

The estimated standard deviations along the considered in-line section are represented in Figure 23. Similarly to the target-oriented inversion, the standard deviation value is inversely proportional to the curvature of the misfit function and increases moving from sand to shale intervals. The lateral

scattering affecting the estimated standard deviation maps is probably related to the different signal-to-noise ratios of adjacent CMPs. Indeed, we remind that the data covariance matrices are independently estimated for each inverted seismic gather.

The evolution of the misfit value for a single CMP gather inversion shows a very fast convergence toward the stationary regime that is reached in 60 iterations (Figure 24a). Figure 24b shows that the interval-oriented HMC algorithm is able to perform long jumps in the model space, thus producing very low correlation coefficients between successively sampled models. For brevity, we limit to plot the autocorrelation of the *Vs* models sampled at a given time position, but similar conclusions would have been drawn for the other elastic properties and for different time samples. The use of the numerical approximation of the Jacobian promotes the independence of the sampled models and speeds up the convergence of the algorithm toward the stationary regime, at the expense of an increased computational effort with respect to the case in which the Jacobian is replaced by the matrix **G**. For example, we can compare the independency of the sampled models for this field data inversion and for the synthetic inversion discussed previously in which the Jacobian was replaced by the analytical matrix **G** (see Figures 7-10). Both inversions consider a comparable number of unknowns, but we observe that when the Jacobian is employed, the autocorrelation value drops to zero in less than five iterations. Otherwise, when the matrix **G** replaces the Jacobian, 15-20 iterations are needed to observe an autocorrelation value close to zero.

Finally, we compare the mean *Vp* model provided by the HMC inversion and the logged *Vp* values measured along three wells investigating the target reservoir (Figure 25). The good match between the sonic log and the *Vp* field provided by the HMC algorithm illustrates the reliability and the applicability of the implemented HMC algorithm. The acceptance rate during the post burn-in stage oscillates around the 50-60%, thus demonstrating an optimal setting of the HMC user-defined parameters. The total computational cost for inverting the entire 2D section was 13 hours, approximately, on the same hardware previously described.

DISCUSSION

The implemented HMC algorithm is aimed at efficiently sampling highly dimensional posterior distributions in non-linear AVA inversions. This ability rests on the exploitation of the derivative information of the misfit function within the sampling framework. This additional information is not considered by standard Monte Carlo methods, such the well-known Metropolis-Hastings algorithm. The main computational requirements for HMC with respect to MCMC comes from the need to compute derivatives. In cases of linear problems this is not an issue because the derivative is explicitly represented by the forward matrix operator. In this context, the outstanding benefit of HMC over the standard analytical inversion is the possibility to accurately sample the PPD even in cases of non-Gaussian prior models.

Jacobian matrix

In cases of non-linear problems, a parallel implementation could be useful to speed up the numerical computation of the Jacobian matrix or of the gradient of the potential energy. In our AVA inversion examples the analytical forward modeling based on the Zoeppritz equations allows for a fast computation of the Jacobian both in the target- and interval-oriented inversions. However, we also proposed a possible strategy to significantly reduce the computational cost of the HMC inversion, in which the numerical approximation of the Jacobian is replaced by a matrix operator **G** analytically derived from a linearized approximation of the full Zoeppritz equations. In the synthetic interval-oriented inversion this approach made it possible to reduce the computational cost of a single HMC inversion from 10 to 1.5 minutes, approximately, and still provided satisfactory model predictions and uncertainty quantifications. In any case this reduction of the computational effort occurs at the expense of a decreased convergence toward the stationary regime and to an overall decrease of the independence of successively sampled models. In any case, using this strategy an AVA-HMC inversion becomes much less computationally expensive than a MCMC run. Additional field data

inversions on the same dataset used in this work confirmed these conclusions. We also remind that the validity of the analytical forward modeling **G** depends on the considered angle range and on the magnitude of the elastic contrasts at the reflecting interface. From the above considerations it emerges that the choice of replacing the Jacobian with the matrix **G** must be considered case-by-case and should constitute a compromise between the sampling efficiency and the computational cost of the HMC inversion.

Geostatistical constraints and alternative model space parameterizations

Due to the high-dimensional parameter spaces sampled in the previous AVA inversions, we found particularly useful the infusion of geostatistical constraints into the prior model. This inclusion mitigates the so-called curse of dimensionality issue, because significantly reduces the hyper-volume of the model space to be explored. In other words, this strategy limits the exploration of the model space to only those regions in which the models satisfy the a-priori assumptions. A parsimonious reparameterization of the model in terms of orthogonal basis functions (e.g., Legendre polynomials, Chebyshev polynomials, Discrete Cosine Transform) could be also useful to mitigate the curse of dimensionality problem. In this context the unknowns will become the numerical coefficients associated to each new basis function. We are now working on this research topic because we believe that this kind of reparameterizations of the model space could be useful to extend the time-interval HMC-AVA inversion to 2D or even 3D models. Indeed, preliminary experiments we carried out show that the current HMC implementation is too inefficient and not scalable to large 2D or even 3D inversions. However, note that this scalability issue (related to the curse of dimensionality issue) also affects the more standard MCMC approach and for this reason many strategies have been proposed such as starting the MCMC sampling from the results of a previous computationally fast analytical inversion (e.g., de Figueiredo et al. 2018).

Parameter setting

A critical aspect of HMC inversion is the choice of the phase-space trajectory length. To this end, in each iteration we draw the L parameter, which controls the number of time integration steps and the trajectory length, from a uniform distribution. However, this parameter can be adaptively set by adopting the so called no U-turns sampling method (Gelman et al. 2013; Hoffman and Gelman et al. 2014) that stops the integration process when the trajectory returns towards its starting position. In cases of linear problems with an optimal choice for the mass matrix, the optimal trajectory length is π (Fictner et al. 2019). Obviously, the setting of L is more complicated in case of non-linear problems. In the examples discussed here we found that 5–10 time steps are usually optimal. Fewer time steps leaded to very high acceptance rates but at the expense of a slow convergence of the algorithm and a limited exploration of the parameter space. Otherwise, more time steps resulted in low acceptance rates and in a decreased accuracy of the numerical integration. A critical element of HMC is the choice of a suitable mass matrix. We found that setting this matrix equal to the inverse of the locally approximated posterior covariance guarantees stable posterior estimations using only few thousands of samples. Similarly, deriving this matrix from a linearization of the Zoeppritz equations demonstrated to be a viable strategy. However, as previously mentioned, the applicability of this alternative approach is strongly case dependent because it assumes a constant curvature of the misfit function over the entire model space. This could be an oversimplified assumption in case of wide-angle ranges and significant elastic contrasts at the reflecting interface. In practical terms this translates into slower convergence and not-null correlation values between successively sampled models. We experimented that other choices of the mass matrix, for examples a scalar multiple of the identity matrix, result in very slow convergence and unreliable posterior assessments because with this approach we are totally neglecting both the correlation and the spatial continuity of the inverted parameters.

We point out that similarly to other sampling methods, the efficiency of the HMC algorithm is heavily context-dependent and in particular it depends on the computational effort for the Jacobian computation and on the parameter tuning (e.g. choice of the mass matrix and of the trajectory length in the phase-space). There are no general rules to optimally set the HMC parameters and for this reason we opted for a trial and error procedure aimed at finding a good compromise between the converge of the algorithm toward the stationary regime and its sampling efficiency. In particular, two main principles guided us in setting the HMC user-defined parameters: obtain an acceptance rate around 0.6 in the post burn-in phase and promote as much as possible the independency of successively sampled models.

Future developments

We are now extending the implemented HMC method to other parametric and even non-parametric prior models: for example the Cauchy or Huber distributions on model parameters could be used to preserve sharp geological boundaries, while a kernel density estimation of the prior model would permit the inclusion of more complex prior distributions within the inversion framework. We are also working on extending the approach to a seismic-petrophysical inversion for a direct estimation of the petrophysical rock properties of interest from seismic data (i.e., porosity, fluid saturation). In this case a properly calibrated rock-physics model should be included into the forward modeling operator to directly link the seismic response to the contrasts in the petrophysical rock-properties. In view of applications of the implemented method to reservoir characterization studies, we are extending the HMC to sample from a multimodal target distribution. Indeed, in this case it would be essential properly modeling the facies-dependent behavior of petrophysical properties. However, it is known that HMC faces major problems when the PPD is multimodal. Making the HMC work in cases of multimodal target distributions is an active research field and improved HMC implementations have been developed over the very last years. These implementations reformulate the geometry of the

phase-space (e.g. Riemann manifold HMC, Girolami and Calderhead, 2011; wormhole HMC, Lan et al. 2014), or exploit thermodynamic principles to jump between different modes of the target distribution (see for example the continuously tempered HMC; Graham and Storkey, 2017), or include these two strategies into a single hybrid sampling algorithm (e.g., the geometrically tempered HMC; Nishimura and Dunson, 2016). Preliminary tests we carried out with a Gaussian mixture prior model show that the our actual HMC implementation can reliably sample from a multimodal distribution only if the Gaussian components of the mixture overlap. The sampling become inefficient when the Gaussian components are separated by high energy barriers (low probability regions). A possible strategy to extend the implemented HMC approach to multimodal, non-parametric posterior models could be introducing a normal score transformation into the sampling framework that converts the non-parametric prior into a Gaussian distribution (Aleardi and Salusti 2019).

Another challenge is to extend the HMC inversion for solving a mixed continuous-discrete inverse problem in which the continuous properties are the elastic or petrophysical parameters while the litho-fluid facies constitute the discrete property. In addition, we are extending the HMC approach to other highly non-linear geophysical inverse problems such as the full-waveform inversion of seismic data. In this context a computationally efficient method to compute the Jacobian (i.e. the adjoint state method) is needed to maintain the computational cost affordable.

CONCLUSIONS

We apply Hamiltonian Monte Carlo (HMC) sampling algorithms for target-oriented and interval-oriented elastic AVA inversions. The aim was to efficiently derive accurate posterior uncertainty estimations for a non-linear forward modeling operator based on the exact Zoeppritz equations. The assumption of a simple Gaussian prior model ensured an efficient sampling of the model space and allowed an easy inclusion of spatial constraints into the inversion framework. The synthetic and field data examples demonstrated the applicability, the reliability, and the efficiency of the implemented

algorithm. In particular, the HMC algorithm (if properly set) showed extremely fast convergence rates and retrieved accurate posterior assessments, together with reliable model parameter estimations with a limited number of sampled models when compared to a Metropolis-Hasting MCMC. However, a crucial aspect of HMC inversion is the need of a computationally efficient strategy to compute the derivative of the misfit function. For this reason, to drastically reduces the computational cost of an AVA-HMC inversion we proposed to replace the numerical computational of the Jacobian with a matrix operator that can be analytically derived from a linear approximation of the full Zoeppritz equations. In our tests, this strategy still provided accurate model predictions and posterior estimations but at the expense of an overall decrease of the sampling efficiency of the HMC algorithm (e.g. slower convergence toward the stationary regime and increase of the correlation value between successively sampled models). The applicability of this approach should be evaluated case-by-case because the validity of the linearization of the Zoeppritz equations depends on the considered angle range and on the elastic contrasts at the reflective interfaces.

This work demonstrates that the HMC method is a very promising approach for non-linear AVA inversion that could constitute a valid alternative to the more popular MCMC algorithms. For this reason, the HMC is surely worth of additional investigations, such as the extension of the implemented HMC to multimodal PPDs.

REFERENCES

Adriansyah, and G.A. McMechan, 2001, AVA analysis and interpretation of a carbonate reservoir: northwest Java basin, Indonesia: Geophysics, **66**, no. 3, 744-754.

Aster, R.C., B. Borchers, and C.H. Thurber, 2018, Parameter estimation and inverse problems: Elsevier.

Aleardi, M., and A. Mazzotti, 2014, A feasibility study on the expected seismic AVA signatures of deep fractured geothermal reservoirs in an intrusive basement: Journal of Geophysics and Engineering, **11**, no. 6, 065008.

Aleardi, M., and A. Mazzotti, 2017, 1D elastic full-waveform inversion and uncertainty estimation by means of a hybrid genetic algorithm–Gibbs sampler approach: Geophysical Prospecting, **65**, no. 1, 64-85.

Aleardi, M., F. Ciabarri, and A. Mazzotti, 2017, Probabilistic estimation of reservoir properties by means of wide-angle AVA inversion and a petrophysical reformulation of the Zoeppritz equations: Journal of Applied Geophysics, **147**, 28-41.

Aleardi, M., F. Ciabarri, and T. Gukov, 2018, A two-step inversion approach for seismic-reservoir characterization and a comparison with a single-loop Markov-chain Monte Carlo algorithm: Geophysics, **83**, no. 3, R227-R244.

Aleardi, M., 2015, The importance of the Vp/Vs ratio in determining the error propagation, the stability and the resolution of linear AVA inversion: a theoretical demonstration: Bollettino di Geofisica Teorica ed Applicata, **56**, no. 3, 357-366.

Aleardi, M. A. and Salusti, 2019, Markov chain Monte Carlo algorithms for target-oriented and

interval-oriented amplitude versus angle inversions with non-parametric priors and non-linear forward modelings: Geophysical Prospecting. In press. doi: 10.1111/1365-2478.12876.

Aleardi, M., F. Ciabarri, and T. Gukov, 2019, Reservoir Characterization Through Target-Oriented AVA-Petrophysical Inversions with Spatial Constraints: Pure and Applied Geophysics, **176**, no. 2, 901-924.

Aki, K., and P.G. Richards, 1980, Quantitative seismology: Theory and methods: New York, 801.

Avseth, P., T. Mukerji, and G. Mavko, 2010, Quantitative seismic interpretation: Applying rock physics tools to reduce interpretation risk: Cambridge university press.

Betancourt, M., 2017, A conceptual introduction to Hamiltonian Monte Carlo: preprint, arXiv:1701.02434.

Bianco, E., 2016, Tutorial: Wavelet estimation for well ties: The Leading Edge, **35**, no. 6, 541-543.

Bishop, C.M., 2006, Pattern Recognition and Machine Learning: Springer.

Blanes, S., F. Casas, and J.M. Sanz-Serna, 2014, Numerical integrators for the Hybrid Monte Carlo method: SIAM Journal on Scientific Computing, **36**, no. 4, A1556-A1580.

Bosch, M., L. Cara, J. Rodrigues, A. Navarro, and M. Díaz, 2007, A Monte Carlo approach to the joint estimation of reservoir and elastic parameters from seismic amplitudes: Geophysics, **72**, no. 6, O29-O39.

Bosch, M., T. Mukerji, and E.F. Gonzalez, 2010, Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review: Geophysics, **75**, no. 5, A165-A176.

Buland, A., and H. Omre, 2003, Bayesian linearized AVO inversion: Geophysics, **68**, no. 1, 185-198.

de Figueiredo, L.P., D. Grana, F.L. Bordignon, M. Santos, M. Roisenberg, and B.B. Rodrigues, 2018, Joint Bayesian inversion based on rock-physics prior modeling for the estimation of spatially correlated reservoir properties: Geophysics, **83**, no. 5, M49-M61.

Duane, S., A.D. Kennedy, B.J. Pendleton, and D. Roweth, 1987, Hybrid Monte Carlo: Physics letters B, **195**, no. 2, 216-222.

Fichtner, A., A. Zunino, and L. Gebraad, 2019, Hamiltonian Monte Carlo solution of tomographic inverse problems: Geophysical Journal International, **216**, no. 2, 1344-1363.

Fichtner, A., and S. Simutė, 2018, Hamiltonian Monte Carlo inversion of seismic sources in complex media: Journal of Geophysical Research: Solid Earth, **123**, no. 4, 2984-2999.

Gelman, A., J.B. Carlin, H.S Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin, 2013, Bayesian Data Analysis, 3rd edition: CRC Press.

Girolami, M. and B. Calderhead, 2011, Riemann manifold Langevin and Hamiltonian Monte Carlo methods: Journal of the Royal Statistical Society, Series B, **73**, no. 2,123–214.

Gong, T., and G.A McMechan, 2016, Target-oriented linear least squares and nonlinear, trust-region Newton inversions of plane waves using AVA and PVA data for elastic model parameters: Geophysics, **81**, no. 5, R325-R338.

Graham, M.M., and A.J. Storkey, 2017, Continuously tempered Hamiltonian Monte Carlo: arXiv preprint arXiv:1704.03338.

Grion, S., A. Mazzotti, and U. Spagnolini, 1998, Joint estimation of AVO and kinematic parameters: Geophysical Prospecting, **46**, no. 4, 405-422.

Haario, H., E. Saksman, and J. Tamminen, 1999, Adaptive proposal distribution for random walk Metropolis algorithm: Computational Statistics, **14**, no. 3, 375-396.

Haario, H., E. Saksman, and J. Tamminen, 2001, An adaptive Metropolis algorithm: Bernoulli, **7**, no. 2, 223-242.

Haario, H., M. Laine, A. Mira, and E. Saksman, 2006, DRAM: efficient adaptive MCMC: Statistics and computing, **16**, no. 4, 339-354.

Hoffman, M.D., and A. Gelman, 2014, The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo: Journal of Machine Learning Research, **15**, no. 1, 1593-1623.

Lan, S., J. Streets, and B. Shahbaba, 2014, Wormhole Hamiltonian Monte Carlo. In 28th AAAI Conference on Artificial Intelligence, 1953–1959.

MacKay, D.J., 2003. Information Theory, Inference and Learning Algorithms: Cambridge University Press.

Mackenze, P.B., 1989, An improved hybrid Monte Carlo method: Physics Letters B, **226**, no.3-4, 369-371.

Madsen, R.B., and T.M. Hansen, 2018, Estimation and accounting for the modeling error in probabilistic linearized amplitude variation with offset inversion: Geophysics, **83**, no. 2, N15-N30.

Mazzotti, A., and E. Zamboni, 2003, Petrophysical inversion of AVA data: Geophysical Prospecting, **51**, no. 6, 517-530.

Malinverno, A., 2002, Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem: Geophysical Journal International, **151**, no. 3, 675-688.

Menke, W., 2018, Geophysical data analysis: Discrete inverse theory. Academic press.

Muir, J.B. and H. Tkalcic, 2015, Probabilistic joint inversion of lowermost mantle P-wave velocities and core mantle boundary topography using differential travel times and hierarchical Hamiltonian Monte-Carlo sampling: in AGU 2015 Fall meeting, San Francisco, S14A-03.

Neal, R.M., 2011, MCMC using Hamiltonian dynamics: in Handbook of Markov Chain Monte Carlo, Brooks, S., A. Gelman, G. Jones, and X. Meng: Chapman and Hall, 113–162.

Neal, R.M., 1996, Bayesian Learning for Neural Networks: Springer.

Nishimura, A., D. and Dunson, 2016, Geometrically tempered Hamiltonian Monte Carlo: arXiv preprint arXiv:1604.00872.

Rimstad, K., and H. Omre, 2010, Impact of rock-physics depth trends and Markov random fields on hierarchical Bayesian lithology/fluid prediction: Geophysics, **75**, no. 4, R93-R108.

Sambridge, M., and K. Mosegaard, 2002, Monte Carlo methods in geophysical inverse problems: Reviews of Geophysics, **40**, no. 3, 3-1.

Sen, M.K., and P.L Stoffa, 1996, Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion: Geophysical Prospecting, **44**, no. 2, 313-350.

Sen, M.K. and R. Biswas, 2017, Transdimensional seismic inversion using the reversible jump Hamiltonian Monte Carlo algorithm: Geophysics, **82**, no. 3, R119–R134.

Tarantola, A. 2005, Inverse problem theory and methods for model parameter estimation: SIAM.

ter Braak, C. J., and J.A Vrugt, 2008, Differential evolution Markov chain with snooker updater and fewer chains: Statistics and Computing, **18**, no. 4, 435-446.

Thomas, M., V. Ball, J.P. Blangy, and L. Tenorio, 2016, Rock-physics relationships between inverted elastic reflectivities: The Leading Edge, **35**, no. 5, 438-444.

Tierney, L., and A. Mira, 1999, Some adaptive Monte Carlo methods for Bayesian inference: Statistics in medicine, **18**, no. 17-18, 2507-2515.

Turner, B.M., and P.B. Sederberg, 2012, Approximate Bayesian computation with differential evolution: Journal of Mathematical Psychology, **56**, no. 5, 375-385.

Vrugt, J.A., 2016, Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation: Environmental Modeling & Software, **75**, 273-316.

Zhu, X., and G.A. McMechan, 2012, Elastic inversion of near-and postcritical reflections using phase variation with angle: Geophysics, **77**, no. 4, R149-R159.

Zoeppritz, K., 1919, Erdbebenwellen vii. Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen: Mathematisch-Physikalische Klasse, 57-65.

Zunino, A., K. Mosegaard, K. Lange, Y. Melnikova, T. and Mejer Hansen, 2015, Monte Carlo reservoir analysis combining seismic reflection data and informed priors: Geophysics, **80**, no. 1, R31-R41.

Appendix A: parameter setting

The following tables summarize the main user-defined parameters of the HMC algorithm, and some statistical properties of the prior model used in the synthetic and field data tests.

Prior in the *Vp*, *Vs* and density model space	Gaussian with spatially variable mean value and stationary covariance matrix
Number of unknowns (for each block)	$50 \times 50 \times 3 = 7500$
Number of iterations (for each block)	50000
Burn-in	100
L parameter for HMC	Uniformly distributed within [5, 10]
δ parameter for HMC	0.7
CMP gridding size	25 m along the cross-line and in-line directions
α parameter for the variogram along the in-line direction	60 m
α parameter for the variogram along the cross-line direction	45 m

Table 1: Parameters for the target-oriented synthetic inversion test.

Prior in the *Vp*, *Vs* and density model space	Log-Gaussian with temporarily variable mean value and stationary covariance matrix
Number of unknowns	$91 \times 3 = 273$
Number of iterations	5000
Burn-in	100
L parameter for HMC	Uniformly distributed within [5, 10]
δ parameter for HMC	0.06
Time sampling	0.001 s
α parameter for the vertical variogram	0.003 s

Table 2: Parameters for the interval-oriented synthetic inversion test.

Prior in the *Vp*, *Vs* and density model space	Gaussian with spatially variable mean value and stationary covariance matrix
Number of unknowns	$71 \times 80 \times 3 = 17040$
Number of iterations	70000
Burn-in	100
L parameter for HMC	Uniformly distributed within [5, 10]
δ parameter for HMC	0.5
CMP gridding size	25 m along the cross-line and in-line directions
α parameter for the variogram along the in-line direction	60 m
α parameter for the variogram along the cross-line direction	40 m

Table 3: Parameters for the field target-oriented inversion.

Prior in the *Vp*, *Vs* and density model space	Log-Gaussian with temporarily variable mean value and stationary covariance matrix
Number of unknowns for each CMP gather	$75 \times 3 = 225$
Number of iterations	5000
Burn-in	100
L parameter for HMC	Uniformly distributed within [5, 10]
δ parameter for HMC	0.1
Time sampling	0.004 s
α parameter for the vertical variogram	0.010 s

Table 4: Parameters for the field interval-oriented inversion.

FIGURES

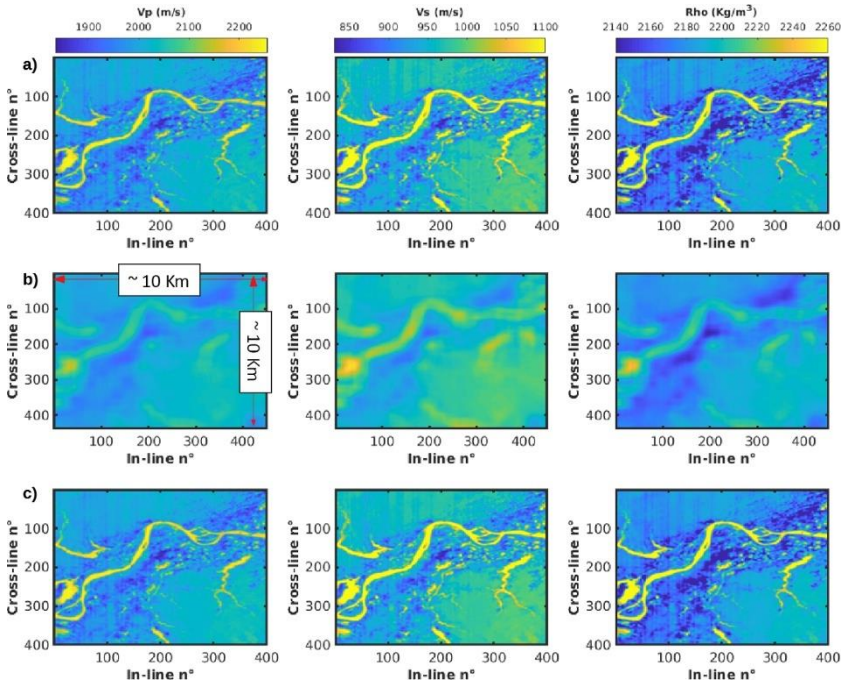

Figure 1: a) The true model. b) A heavily smoothed version of the true model used as starting point for the HMC inversion. c) The estimated posterior mean. In a), b), and c) the Vp, Vs, and density are represented from left to right columns, respectively.

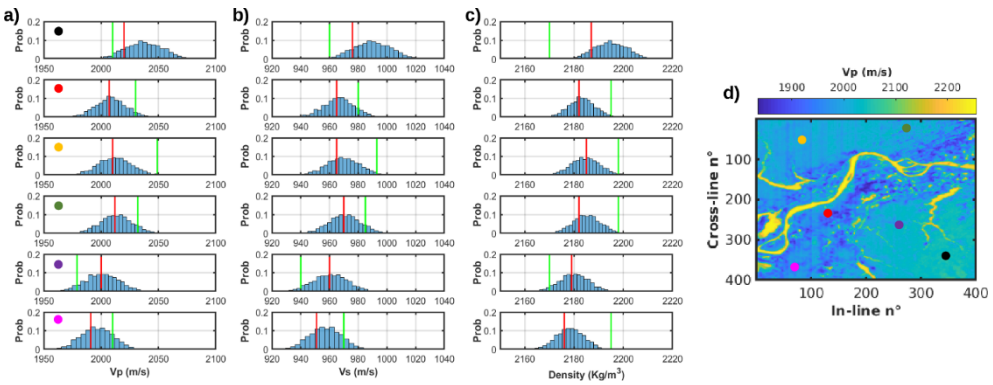

Figure 2: Examples of 1D marginal posterior distributions for Vp (a), Vs (b), and density (c) pertaining to different CMP gather positions. The red and green lines represent the true elastic properties and the starting model, respectively. d) The CMP locations are identified by the colored circles overlapped to the estimated mean Vp model.

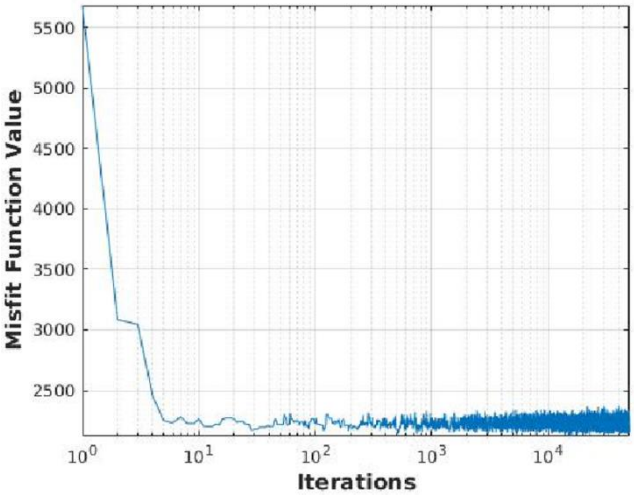

Figure 3: Evolution of the misfit function value for the synthetic target-oriented inversion.

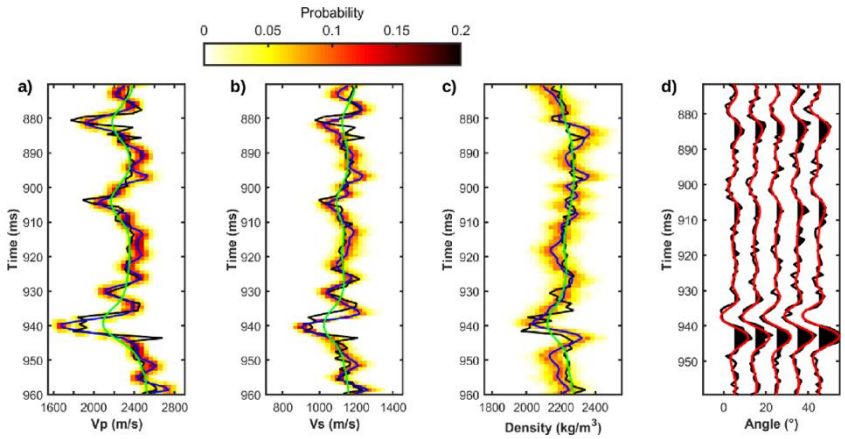

Figure 4: Results for the interval-oriented inversion on synthetic data. In a), b) and c) the black lines represent the true property values, the blue lines are the estimated mean models, the green lines show the starting model, whereas the colormap codes the estimated PPD. d) Comparison of observed (black) and predicted (red) seismic data computed on the a-posteriori mean model.

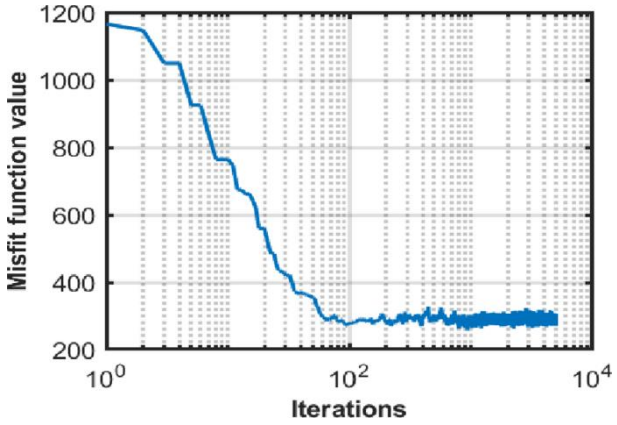

Figure 5: Evolution of the misfit function value for the synthetic interval-oriented inversion.

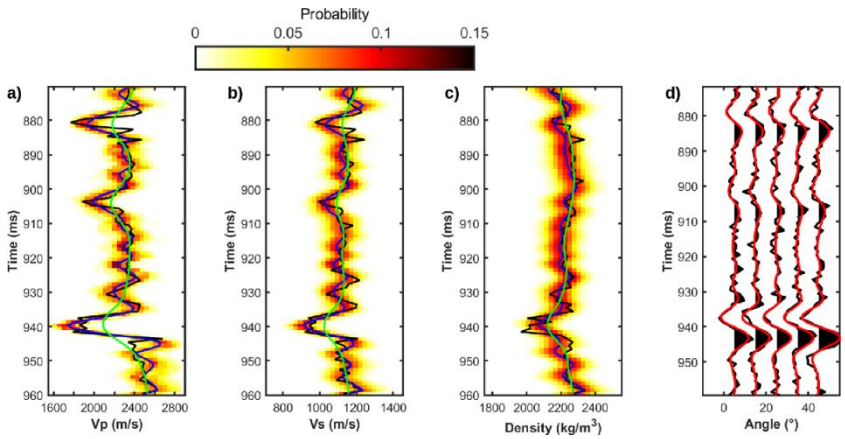

Figure 6: Results for the interval-oriented MCMC inversion on synthetic data. In a), b) and c) the black lines represent the true property values, the blue lines are the estimated mean models, the green lines show the starting model, whereas the colormap codes the estimated PPD. d) Comparison of observed (black) and predicted (red) seismic data computed on the a-posteriori mean model.

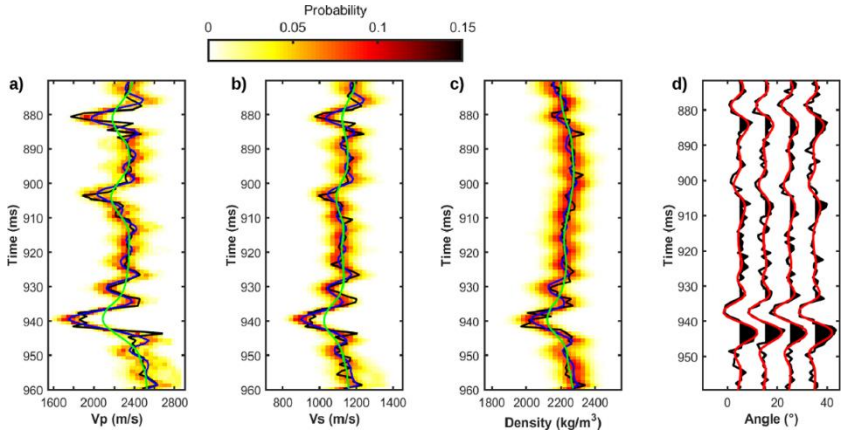

Figure 7: As in Figure 6 but for the HMC inversion in which the Jacobian is replaced by the analytical matrix operator **G**.

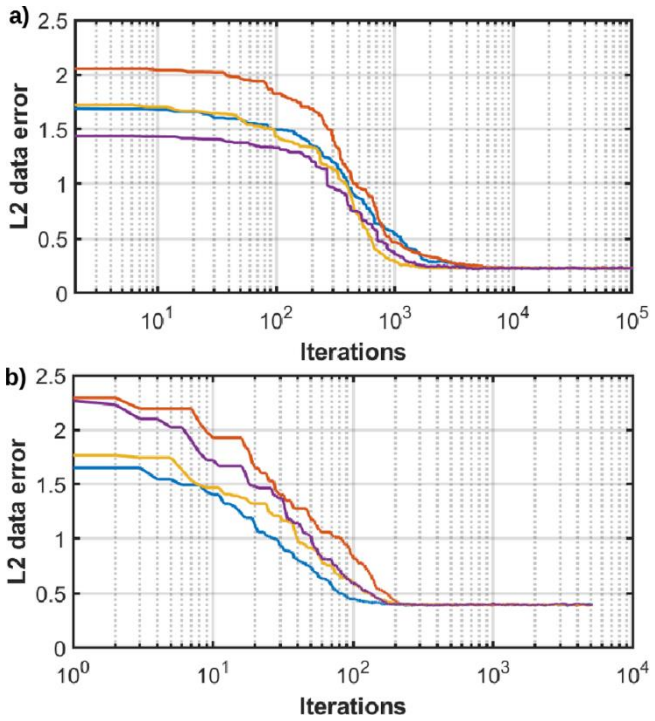

Figure 8: Evolution of the L2 norm data misfit for the four chains. a) MCMC. b) HMC.

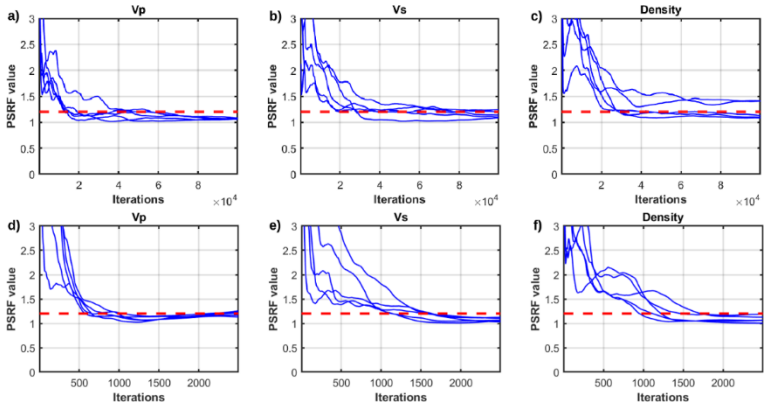

Figure 9: Examples of evolution of the PSRF value computed for five different time positions and for Vp, Vs and density. a), b), and c) pertain to MCMC. d), e) and f) pertain to HMC. The red dotted lines represent the theoretical threshold value of 1.2 that indicates the convergence toward a stable posterior model.

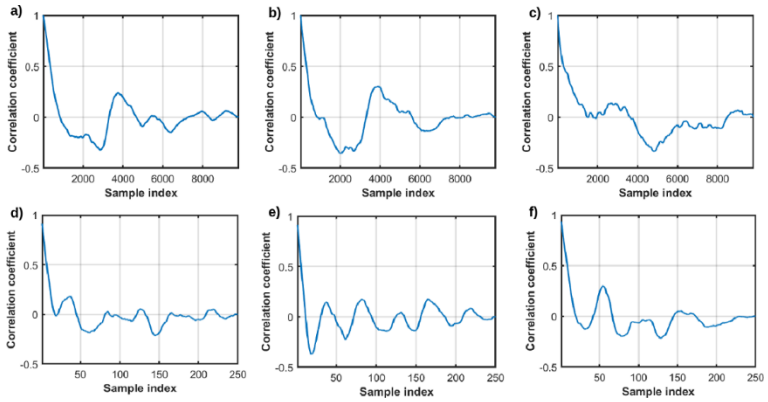

Figure 10: Examples of normalized autocorrelations functions computed from the Vp (a), Vs (b), and density (c) models sampled by the MCMC at a given time position. Examples of normalized autocorrelations functions computed from the Vp (d), Vs (e), and density (f) models sampled by the HMC at a given time position.

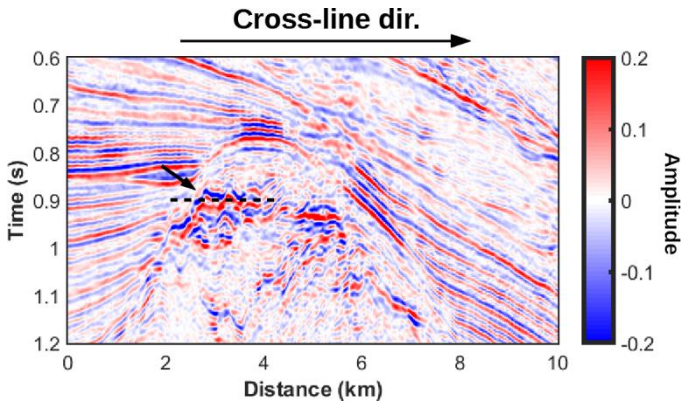

Figure 11: In-line section extracted from the 3D seismic volume. The black arrow shows the top reflection of the investigated reservoir. The dotted line evidences the time slice considered in the target-oriented inversion.

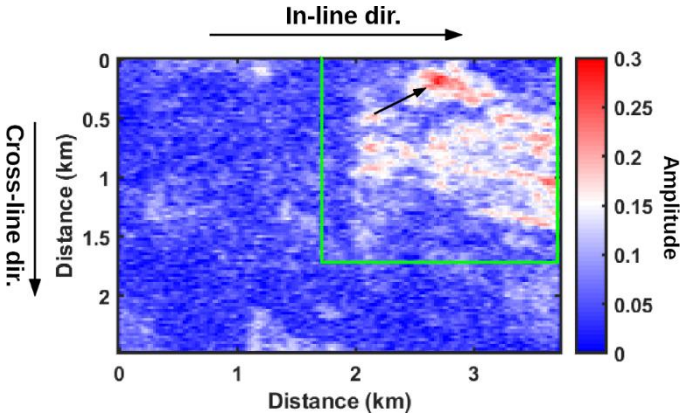

Figure 12: RMS amplitude map at the top of the investigated reservoir. The green rectangle encloses the area considered in the target-oriented inversion. The black arrow points toward the main gas sand body.

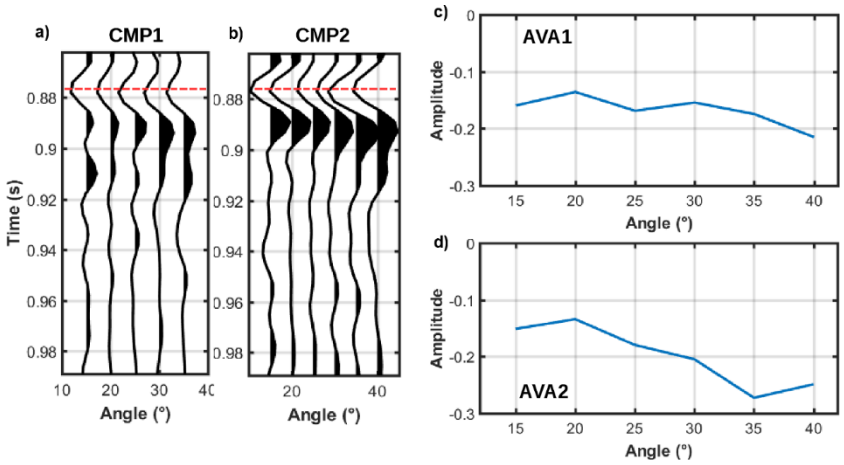

Figure 13: a) and b) show close-ups of two pre-stack seismic gathers. The strong amplitude anomaly at 0.88 s, approximatively (red dotted lines), identifies the top reservoir reflection. c), and d) represent the AVA responses of the target reflection extracted from the CMPs shown in a) and b), respectively.

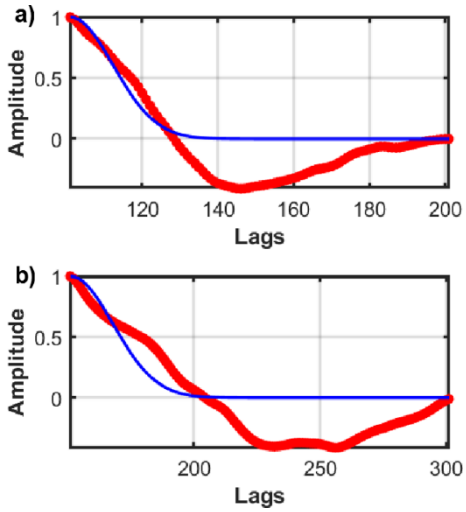

Figure 14: Actual (red lines) and approximated (blue lines) lateral autocorrelation functions for the cross-line (a) and in-line (b) directions.

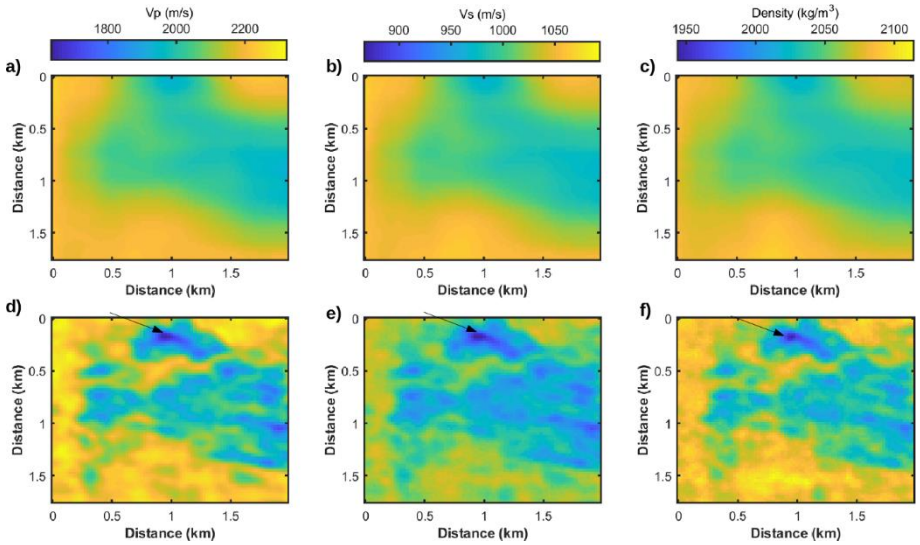

Figure 15: a), b) and c) show the starting Vp, Vs and density models, respectively. d), e), and f) show the a-posteriori mean Vp, Vs and density models, respectively. The black arrows highlight the very low elastic property values associated to the main gas-saturated sand body.

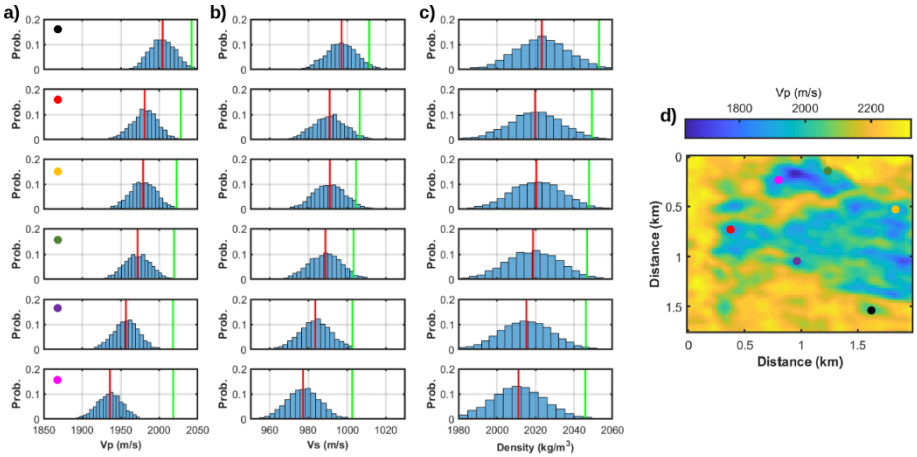

Figure 16: a), b) and c), show some examples of 1D Marginal PPDs for Vp, Vs, and density, respectively, extracted for different CMP gather positions. Red and green lines represent the a-posteriori mean and the starting model values, respectively. d) The CMP locations are identified by the colored circles overlapped to the estimated mean Vp model.

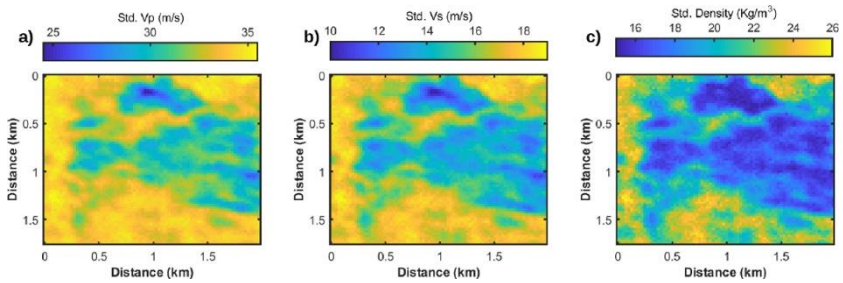

Figure 17: Maps of estimated standard deviations for Vp (a), Vs (b), and density (c) represented for each inverted CMP position.

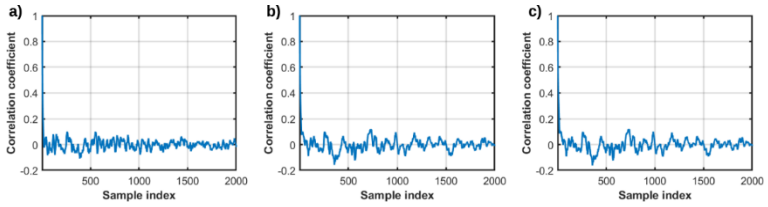

Figure 18: Examples of normalized autocorrelation coefficients for 2000 successively sampled models of Vp (a), Vs (b), and density (c) at a given CMP location.

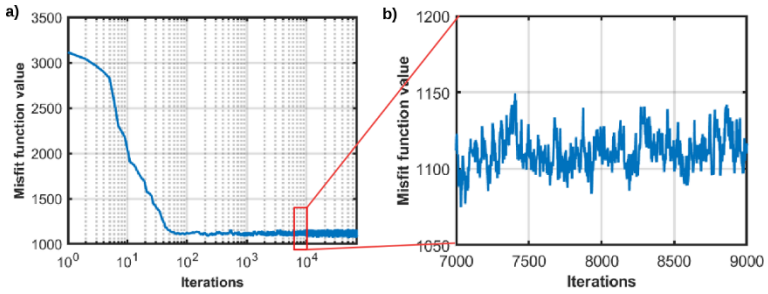

Figure 19: a) Evolution of the misfit function during the HMC sampling. b) Close-up of a) showing the highly variable misfit value versus the iteration number.

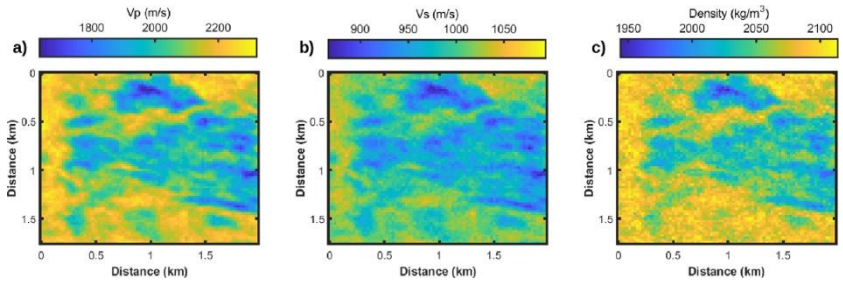

Figure 20: The a-posteriori mean values for Vp (a), Vs (b), and density (c) provided by a standard non-linear Bayesian AVA inversion at the top of the target reservoir.

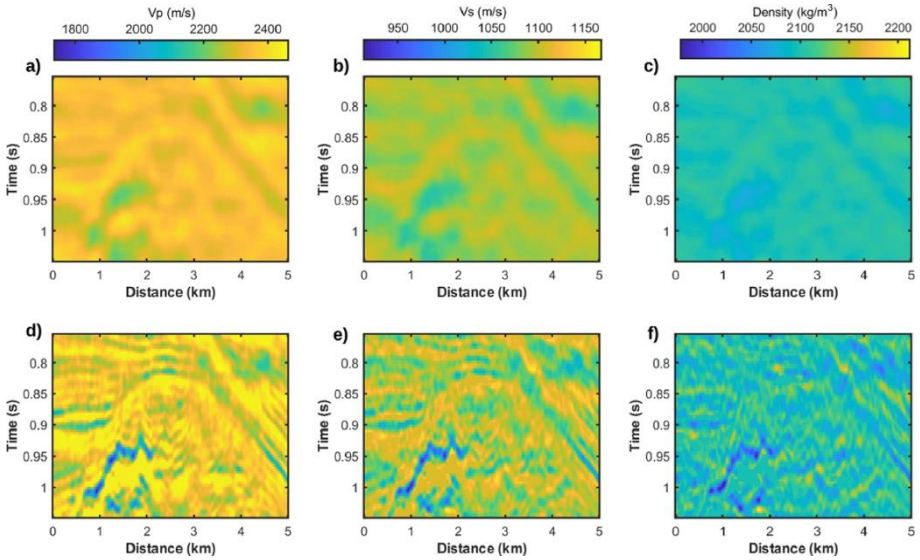

Figure 21: a), b, and c) represent the low frequency Vp, Vs and density models, respectively, used as starting point for the HMC inversion. d), e), and f) show the estimated a-posteriori Vp, Vs, and density mean models, respectively.

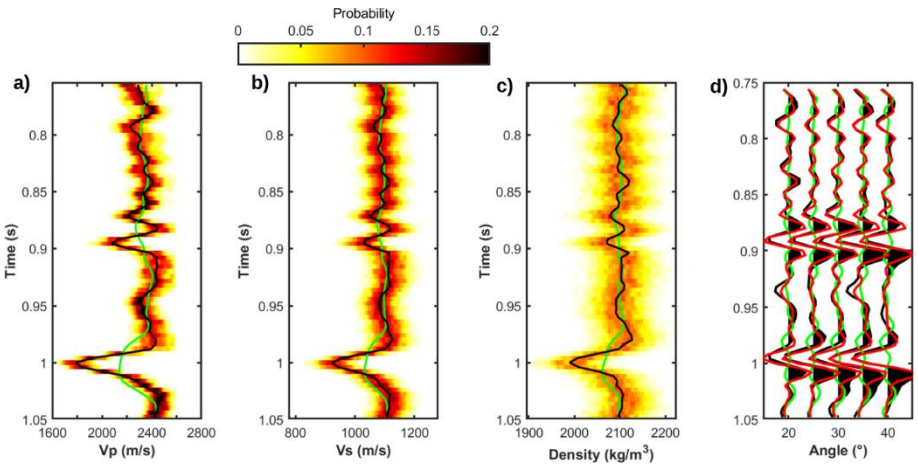

Figure 22: a)-c) Comparison between the estimated marginal distribution for Vp, Vs and density (color scale), the starting models (green line), and the posterior mean (black line). d) Comparison between the observed data (black), the predicted data computed on the posterior mean (red) and the data computed on the starting model (green). These results pertain to the CMP gather located at the horizontal coordinate of 1 km (see Figure 21).

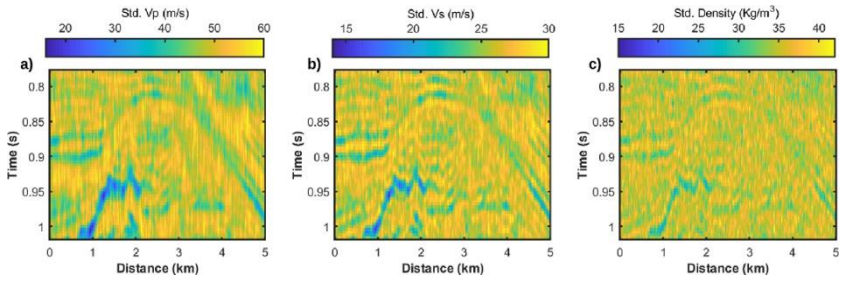

Figure 23: Estimated standard deviations for Vp (a), Vs (b), and density (c).

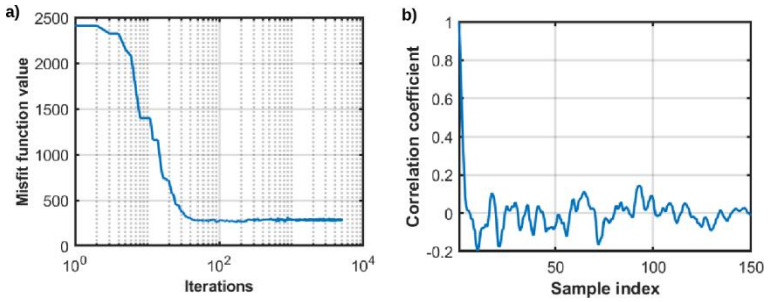

Figure 24: a) Example of evolution of the misfit function during the HMC sampling for a given CMP position. b) Example of autocorrelation for the Vs parameter.

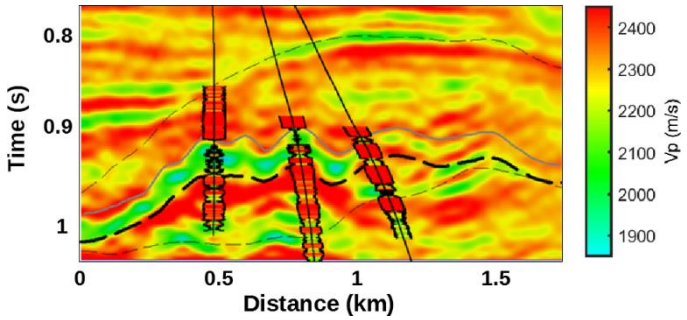

Figure 25: A Comparison of the mean Vp solution provided by the HMC inversion and the logged Vp values recorded in three wells. The dotted and continuous black and gray lines delineate the target zone.