Geophysical Journal International

Geophys. J. Int. (2025) **242,** 1–21
Advance Access publication 2025 July 02
Research Paper

https://doi.org/10.1093/gji/ggaf249

Bayesian seismic inversion with implicit neural representations

Juan Romero ,[1] Wolfgang Heidrich[1] and Matteo Ravasi[2,*]

[1]*Physical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955, Kingdom of Saudi Arabia.*
E-mail: juan.romeromurcia@kaust.edu.sa
[2]*Shearwater Geoservices, AI and Monitoring, Gatwick RH6 0PA, United Kingdom*

Accepted 2025 June 25. Received 2025 June 16; in original form 2025 March 10

SUMMARY

Seismic inversion translates seismic data into subsurface elastic property models, enabling geophysicists to better understand underground rocks and fluids. Due to the inherently ill-posed nature of this inverse problem, accurately capturing the uncertainty associated with the solution is essential for reliable interpretations. Traditional Bayesian inversion methods, such as Markov Chain Monte Carlo and Laplace approximations, have been employed for this purpose but face significant limitations in terms of scalability and computational efficiency for large-scale problems. Combined with deep learning, Variational Inference (VI) has emerged as a promising alternative, striking a balance between computational efficiency and flexibility (i.e. the ability to approximate complex posterior distributions). However, selecting an appropriate proposal distribution remains a key challenge, as it directly influences the quality of the estimated posterior distribution. In this study, we extend IntraSeismic, an implicit neural representation (INR)-based framework for seismic inversion applications, to Bayesian inversion using VI with different parametrizations of the proposal distribution. We introduce two methods: B-IntraSeismic (BIS), which uses a mean-field Gaussian proposal, and B-IntraSeismic with Conditional Normalizing Flows (BIS-Flow), which utilizes a mean-field unparametrized proposal distribution to better capture deviations from Gaussianity in the posterior distribution. These methods are evaluated on a synthetic data set (Marmousi) and two field data (Volve and Sleipner). Our results indicate that both BIS and BIS-Flow can accurately capture structural details and produce high-resolution mean models and standard deviation maps. BIS-Flow is also shown to be able to model complex posterior distributions, offering a more comprehensive characterization of uncertainty while maintaining computational feasibility.

Key words: Inverse theory; Planetary seismology; Acoustic properties.

1 INTRODUCTION

Seismic inversion is the foremost technique to extract quantitative information of the subsurface, such as P-wave and S-wave velocities, from seismic data (Tarantola 1984). Having access to a quantitative estimate of the physical properties of the subsurface enables a number of downstream applications in various fields, including civil engineering (Steeples & Miller 1988; Oz & Miller 2015), natural resource exploration (Telford *et al.* 1990; Yilmaz 2001), subsurface hazard analysis and more. Conducted in the seismic image domain (i.e. post-migration) to estimate acoustic impedance models, post-stack seismic inversion is routinely employed for detailed subsurface analysis, particularly in reservoir characterization studies (Houck 2002; Bosch *et al.* 2010). However, seismic inverse problems are inherently challenging due to three main factors: the limited frequency bandwidth acquired in a seismic experiment, various types of noise in the data and the simplifying assumptions in modelling operators that may not fully capture the underlying physics. These challenges make the problem ill-posed, meaning there is no an exact solution to it.

The non-uniqueness inherent in seismic inverse problems leads to uncertainties in the recovered subsurface properties. Uncertainty quantification (UQ) applied to seismic inversion enables not only the estimation of the most likely subsurface model but also the quantification of the range of plausible solutions that fit the observed data (Tarantola & Valette 1981; Ulrych *et al.* 2001). This enhanced understanding of subsurface variability and non-uniqueness is crucial for numerous applications. For example, in reservoir characterization, UQ helps identify regions of high uncertainty that may require additional data acquisition or alternative interpretation strategies (Strutz & Curtis 2024). For resource assessment, uncertainty estimates enable more reliable volumetric calculations and risk analysis, leading to better-informed investment decisions (Talarico

---

[*] Work for this study was conducted while affiliated with KAUST.

et al. 2020). Furthermore, UQ facilitates the integration of seismic inversion results with other geophysical and geological data by providing probability distributions rather than single deterministic solutions.

The probabilistic approach to seismic inversion treats subsurface model parameters as random variables characterized by probability distributions, where the inference process aims to estimate the posterior distribution of these unknown parameters conditioned on the seismic data. Exact computation of the posterior distribution is feasible for small-scale problems as long as the likelihood and prior distributions belong to conjugate families but becomes computationally prohibitive in more general cases. In seismic inversion, simplifying assumptions are often made, such as assuming lateral independence between seismic traces (Buland & Omre 2003), to facilitate covariance computation. Given the inherently large-scale nature of seismic inversion, geophysicists commonly rely on approximate inference algorithms that balance computational efficiency and accuracy. Traditional approaches such as Markov Chain Monte Carlo (MCMC) methods and Laplace approximations have been extensively used in seismic inversion applications to estimate the reliability of predictive models (Mosegaard & Tarantola 1995; Sen & Stoffa 1996; Sambridge & Mosegaard 2002). While effective for small examples, these methods still face significant scalability and computational efficiency challenges, especially for large and high-dimensional data sets.

Variational Inference (VI), on the other hand, reformulates the inference problem as an optimization task, where the true posterior distribution is approximated by the closest member of a family of tractable distributions, whose parameters are obtained by minimizing the Kullback–Leibler divergence with the true posterior (Jordan *et al.* 1998; Blei *et al.* 2017). This approach combines the advantages of optimization techniques with probabilistic modelling, making it particularly suitable for large-scale problems where traditional sampling methods become computationally intractable. Moreover, recent advances in stochastic optimization and deep learning have further increased the practicality of VI for large-scale inverse problems (Zhang *et al.* 2018), like those commonly found in geophysical applications. These advancements enable efficient and accurate computation of gradients, facilitating the optimization of improved variational objectives that can enhance the stability and accuracy of the entire optimization process. One prominent example of the use of VI in deep learning is the variational autoencoder (VAE) (Kingma & Welling 2022), a model designed to learn complex, high-dimensional probability distributions through a combination of probabilistic inference and convolutional neural networks. This success has inspired various applications in inverse problems, where neural networks can learn to capture complex posterior distributions while maintaining computational efficiency. More specifically, these developments have opened new possibilities for UQ in high-dimensional problems like seismic inversion, where traditional methods struggle to balance computational efficiency with accurate uncertainty estimation (Zhang *et al.* 2021; Li *et al.* 2024).

Implicit neural representations (INR) are a novel family of neural networks that learn a direct mapping from the spatial coordinates to one or more output values in a domain of interest (Sitzmann *et al.* 2020; Tancik *et al.* 2020). A notable example is Neural Radiance Fields (NeRF) (Mildenhall *et al.* 2020), which has revolutionized image reconstruction by directly representing a scene using its spatial coordinates as input. Inspired by the capabilities of neural networks to act as pre-conditioners in image reconstruction problems, INR has also been leveraged for seismic inversion tasks (Romero *et al.* 2024; Sun *et al.* 2023). Recent research has shown that INR offer a powerful strategy for parametrizing models (or acting as deep pre-conditioners) for seismic inverse problems, leading to state-of-the-art inversion results with faster convergence compared to conventional methods. One such method, IntraSeismic (Romero *et al.* 2024), offers additional benefits such as fast data access, compression capabilities and the ability to seamlessly conduct UQ via Monte Carlo Dropout (MCD).

In this paper, we present a Bayesian extension of the IntraSeismic framework, termed Bayesian-IntraSeismic (B-IntraSeismic), which extends IntraSeismic's capabilities by providing an estimate of the uncertainty of the inverted model parameters. We develop and evaluate two distinct approaches: (1) a mean-field Gaussian method (BIS), in which IntraSeismic's nonlinear mapping module is modified to predict both the mean and standard deviation of the subsurface model parameters at each location, offering a scalable and computationally efficient solution; and (2) a second approach that uses conditional normalizing flows (BIS-Flow) to estimate nonparametric posterior distributions for each input point, where a feature vector from IntraSeismic's nonlinear mapping module serves as the context for the flow. This paper is organized as follows: we begin with an overview of exact and approximate inference algorithms for post-stack seismic inversion; this is followed by a review of VI and its integration into the B-IntraSeismic framework, with a practical focus on its parametrization through INRs. Next, we describe the implementation of the two approaches—BIS and BIS-Flow—for post-stack seismic inversion. In the results section, we apply BIS and BIS-Flow to synthetic (Marmousi) and field (Volve and Sleipner) data sets, comparing their results against analytical and other approximate solutions to demonstrate their effectiveness in capturing complex posterior distributions and characterizing uncertainty. Finally, we discuss the advantages, limitations and potential future directions for these methods in the context of seismic inversion.

2 THEORY AND METHODS

2.1 Deterministic post-stack inversion

Post-migration angle-dependent seismic data $d_\theta(t)$ can be mathematically represented via the so-called convolutional model (Goupillaud 1961). This entails convolving a source function or wavelet $w(t)$ with the earth P-wave reflectivity series $r_{pp}(t)$:

$$d_\theta(t) = w(t) * r_{pp}(t) \tag{1}$$

The Aki-Richards approximation (Aki & Richards 2002), which builds on the foundational Zoeppritz equation (Zoeppritz 1919), expresses r_{pp} as follows:

$$r_{pp}(t,\theta) = \sum_{i=1}^{3} c_i(t,\theta) \frac{\mathrm{d}}{\mathrm{d}t} \log\left(m_i(t)\right), \tag{2}$$

where the mixing coefficients c_i are given by

$$\begin{aligned} c_1(t,\theta) &= \frac{1}{2}\left(1+\tan^2(\theta)\right) \\ c_2(t,\theta) &= -4\left(\frac{\bar{V}_s(t)}{\bar{V}_p(t)}\right)^2 \sin^2(\theta) \\ c_3(t,\theta) &= \frac{1}{2} - 2\left(\frac{\bar{V}_s(t)}{\bar{V}_p(t)}\right)^2 \sin^2(\theta) \end{aligned} \quad , \tag{3}$$

where θ represents the incidence angle of the seismic ray path, $V_p(t)$ is the P-wave velocity, $V_s(t)$ the S-wave velocity and $\rho(t)$ is the density of the medium of interest. When only zero offset data is considered, that is, $\theta = 0$, eq. (2) becomes:

$$r_{PP} = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\log(m(t)), \tag{4}$$

where r_{PP} is referred in this case as post-stack reflectivity and $m(t)$ is the *P*-wave or acoustic impedance. Post-stack seismic data is hence modelled as:

$$d(t) = \frac{1}{2}w(t) * \frac{\mathrm{d}}{\mathrm{d}t}\log(m(t)) \tag{5}$$

In compact matrix–vector notation:

$$\mathbf{d} = \mathbf{WDm}, \tag{6}$$

where $\mathbf{W} \in \mathbb{R}^{N_y N_x N_t \times N_y N_x N_t}$ is a block-Toeplitz convolution matrix, which encapsulates the seismic wavelet, while $\mathbf{D} \in \mathbb{R}^{N_y N_x N_t \times N_y N_x N_t}$ is a first-order derivative operator. From here onwards, we use $\mathbf{G} = \mathbf{WD}$ for the post-stack modelling operator to simplify our notation.

Post-stack seismic inversion is the process aimed at transforming seismic images or 3-D seismic volumes $\mathbf{d} \in \mathbb{R}^{N_y N_x N_t}$ into quantitative estimates of the subsurface's acoustic impedance $\mathbf{m} \in \mathbb{R}^{N_y N_x N_t}$ (Oldenburg *et al.* 1983; Russell & Hampson 2005); the associated optimization problem can be expressed in the following generic form:

$$\mathbf{m}^* = \arg\min_{\mathbf{m}} \frac{1}{2}\|\mathbf{Gm} - \mathbf{d}\|_2^2 + \alpha R(\mathbf{m}). \tag{7}$$

The objective of eq. (7) is to minimize the discrepancy between the modelled and observed seismic data, adjusted by a regularization term $R(\mathbf{m})$.

2.2 Bayesian post-stack seismic inversion

Bayesian post-stack seismic inversion is formulated using a probabilistic framework to estimate subsurface properties while explicitly accounting for uncertainties in both the data and prior geological knowledge. This approach is grounded in Bayes' theorem, which states that the posterior probability distribution $p(\mathbf{m}|\mathbf{d})$ of model parameters $\mathbf{m}$ given observed seismic data $\mathbf{d}$ is proportional to the product of the likelihood function $p(\mathbf{d}|\mathbf{m})$ and the prior probability distribution $p(\mathbf{m})$, normalized by the evidence $p(\mathbf{d})$:

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m}) \cdot p(\mathbf{m})}{p(\mathbf{d})} \tag{8}$$

The posterior distribution obtained through this Bayesian formulation provides not only the most probable impedance model but also a complete characterization of the uncertainty space, including parameter correlations and multiple modes in the solution space. The deterministic post-stack seismic inversion eq. (7) can be interpreted as finding the maximum *a posteriori* (MAP) estimate in the Bayesian framework. In this interpretation, the data misfit term corresponds to the negative log-likelihood function (under the assumption of Gaussian noise in the data). The regularization term represents the negative logarithm of the prior distribution (which can be represented by any distribution from the exponential family), where the regularization parameters act as the inverse variance of the prior distribution (when the prior is Gaussian). This Bayesian interpretation provides a probabilistic foundation for the deterministic optimization problem and reveals how different choices of regularization terms correspond to different prior beliefs about the subsurface properties. In the following, we provide a detailed analysis of the key elements in the Bayesian seismic inversion formulation, including the likelihood functions, prior distributions and their implications for UQ.

2.2.1 *The prior model*

The prior model, also known as the prior distribution, is a fundamental concept in Bayesian statistics as it represents our initial beliefs or knowledge about unknown parameters before observing any data. In the context of seismic inversion, the prior typically embodies theoretical considerations about subsurface elastic properties (Curtis & Lomax 2001; Scales & Tenorio 2001). These considerations may include assumptions about the proximity to a background or initial velocity model, expectations of smoothness or blockiness in the solution and time-lag dependencies (Buland & Omre 2003), among others. For instance, geophysicists might incorporate prior knowledge that velocities generally increase with depth or that certain geological formations are likely to exhibit specific ranges of elastic properties (Grana *et al.* 2017). Within the definition of the prior, one can also incorporate the degree of certainty or uncertainty in these initial assumptions. In some cases, this might be based on well-log data, regional geological knowledge or results from previous inversions in similar areas.

In this work, we consider three types of priors, which will be combined in different ways for the two methods discussed later. These priors serve to further constrain the solution by incorporating different aspects of our prior knowledge about the subsurface properties.

(i) *Proximity to a low-frequency background model* – This prior constrains the solution to remain close to a known low-frequency background model $\mathbf{m}_0$, ensuring stability and incorporating prior geological knowledge into the inversion:

$$\mathbf{m} \sim N(\mathbf{m}_0, \mathbf{\Sigma}_1) \rightarrow p(\mathbf{m}) \propto \exp\left(-\frac{1}{2}(\mathbf{m} - \mathbf{m}_0)^T \mathbf{\Sigma}_1^{-1}(\mathbf{m} - \mathbf{m}_0)\right), \tag{9}$$

where $\mathbf{\Sigma}$ represents the covariance matrix. In this work, we assume an isotropic Gaussian distribution, implying equal variance in all directions. This simplifies the covariance matrix to $\mathbf{\Sigma} = \frac{1}{\lambda_1}\mathbf{I}$, where λ_1 controls the strength of the prior constraint.

(ii) *Smoothness prior (generalized Tikhonov regularization)* – This prior promotes smoothness in the solution by penalizing rapid spatial variations:

$$\mathbf{Dm} \sim N(\mathbf{0}, \mathbf{\Sigma}_2) \rightarrow p(\mathbf{Dm}) \propto \exp\left(-\frac{1}{2}(\mathbf{Dm})^T \mathbf{\Sigma}_2^{-1}\mathbf{Dm}\right), \tag{10}$$

where $\mathbf{D}$ represents the Laplacian operator, enforcing smoothness by penalizing second-order variations in the model. The covariance matrix is assumed to be isotropic, given by $\mathbf{\Sigma}_2 = \frac{1}{\lambda_2}\mathbf{I}$, where λ_2 controls the degree of regularization. This formulation discourages abrupt changes in impedance while still allowing for necessary structural variations.

(iii) *Blockiness prior (total variation regularization)* – This prior promotes piecewise-constant solutions, preserving sharp boundaries in the model, which is particularly useful for detecting geological discontinuities:

$$\nabla\mathbf{m} \sim \mathrm{Laplace}(\mathbf{0}, \frac{1}{\lambda_3}\mathbf{I}) \rightarrow p(\nabla\mathbf{m}) \propto \exp\left(-\lambda_3||\nabla\mathbf{m}||_1\right) \tag{11}$$

where ∇ is the gradient operator, enforcing blocky structures by penalizing large variations in $\mathbf{m}$. The prior follows a multivariate Laplace distribution, which naturally promotes sparsity in the gradient domain. This formulation ensures that abrupt impedance contrasts, such as fault zones and layer boundaries, are preserved while reducing unnecessary smoothness in homogeneous regions.

2.2.2 The Likelihood model

The likelihood function provides a direct link between the observed data and the model parameters. It quantifies the probability of observing the data given a particular set of parameter values in the model. Unlike the prior, which encapsulates our initial beliefs about the parameters, the likelihood encodes the information contained in the data. It measures how well different parameter values explain the observed data, with higher likelihood values indicating a better fit. In practice, the likelihood often incorporates assumptions about the data-generating process and measurement errors. For the seismic inversion case, we usually assume Gaussian distributed noise in the seismic data such that:

$$\mathbf{d} \sim N(\mathbf{Gm}, \sigma_\epsilon^2 \mathbf{I}) \rightarrow p(\mathbf{d}) \propto \exp\left(-\frac{1}{2\sigma_\epsilon^2}||\mathbf{d} - \mathbf{Gm}||_2^2\right), \tag{12}$$

where $\mathbf{G}$ is the seismic modelling operator that computes the seismic data from the logarithm of the acoustic impedance model $\mathbf{m}$ to be compared with the observed data $\mathbf{d}$. Here, we choose to have a fixed noise level with a standard deviation of σ_ϵ. However, the VI framework allows one to choose any data covariance of choice, provided that its inverse is tractable.

2.3 The posterior model

2.3.1 Exact inference

The posterior distribution combines our prior knowledge with observed data to provide an updated belief about the parameters of interest. A closed-form solution for the posterior distribution exists when the prior and likelihood are from conjugate families, for example, Gaussian distributions. In other cases, finding an analytical expression for the posterior becomes impossible as the marginal likelihood often involves intractable integrals. In the specific case of post-stack seismic inversion, when we incorporate the Gaussian priors in eqs (9) and (10), the negative log-posterior becomes:

$$\begin{aligned} J &= -\log p(\mathbf{d}|\mathbf{m}) - \log p_1(\mathbf{m}) - \log p_2(\mathbf{m}) \\ &= \frac{1}{2}||\mathbf{d} - \mathbf{Gm}||_2^2 + \frac{\lambda_1}{2}||\mathbf{m} - \mathbf{m}_0||_2^2 + \frac{\lambda_2}{2}||\mathbf{Dm}||_2^2 \end{aligned} \tag{13}$$

Minimizing this objective function is equivalent to finding the least-squares solution of the following system of equations:

$$\begin{bmatrix} \mathbf{G} \\ \sqrt{\lambda_1}\mathbf{I} \\ \sqrt{\lambda_2}\mathbf{D} \end{bmatrix} \mathbf{m} = \begin{bmatrix} \mathbf{d} \\ \sqrt{\lambda_1}\mathbf{m}_0 \\ \mathbf{0} \end{bmatrix} \tag{14}$$

Because in the case of Gaussian priors and Gaussian likelihood, the posterior is also Gaussian, this also corresponds to the mean of the posterior distribution, which can be written as:

$$\mathbf{m}_{\text{post}} = (\mathbf{G}^T\mathbf{G} + \lambda_1\mathbf{I} + \lambda_2\mathbf{D}^T\mathbf{D})^{-1}(\mathbf{G}^T\mathbf{d} + \lambda_1\mathbf{m_0}) \tag{15}$$

Moreover, an analytical expression exists also for the posterior covariance matrix:

$$Cov(\mathbf{m}_{\text{post}}) = (\mathbf{G}^T\mathbf{G} + \lambda_1\mathbf{I} + \lambda_2\mathbf{D}^T\mathbf{D})^{-1}, \tag{16}$$

which can be computed efficiently for small problems using standard matrix inversion techniques.

2.3.2 Approximate inference: RTO

The 'randomize-then-optimize' (RTO) method is a novel technique for sampling from posterior distributions in Bayesian inference, particularly effective when the data errors and priors are modelled as Gaussian distributions (Bardsley *et al.* 2014). RTO converts a deterministic optimization problem into a stochastic one, facilitating the efficient sampling of high probability models within the posterior distribution. This transformation requires two key adjustments (Blatter *et al.* 2022). First, the data is perturbed to incorporate uncertainty directly into the inversion process, ensuring that data variability is reflected in the solutions. Secondly, the regularization target is also adjusted to a randomly generated model that satisfies the regularization constraints. This randomly generated model does not necessarily minimize the data misfit but aligns with the regularization criteria, thereby capturing the uncertainty inherent in the prior model covariance during the inversion. For the post-stack seismic inversion problem described in eq. (13), the RTO method introduces a modified objective function with perturbed terms:

$$\frac{1}{2}\left\|\mathbf{Gm} - \tilde{\mathbf{d}}\right\|_2^2 + \frac{\lambda_1}{2}\left\|\mathbf{m} - \tilde{\mathbf{m}}_0\right\|_2^2 + \frac{\lambda_2}{2}\left\|\mathbf{D}(\mathbf{m} - \tilde{\mathbf{m}})\right\|_2^2, \tag{17}$$

where

$$\begin{aligned} &\tilde{\mathbf{d}} \sim N(\mathbf{d}, \mathbf{C}_d), \quad \tilde{\mathbf{m}}_0 \sim N\left(\mathbf{m}_0, \frac{1}{\lambda_1}\mathbf{I}\right), \\ &\times\tilde{\mathbf{m}} \sim N\left(\mathbf{0}, \frac{1}{\lambda_2}(\mathbf{D}^T\mathbf{D})^{-1}\right). \end{aligned} \tag{18}$$

In this formulation, $\tilde{\mathbf{d}}$ is drawn from a Gaussian distribution centred at the observed data $\mathbf{d}$ with a data covariance matrix $\mathbf{C}_d$ (in our case $\mathbf{C}_d = \mathbf{I}$), introducing noise into the data. $\tilde{\mathbf{m}}_0$ is sampled from a Gaussian distribution centred at the background model $\mathbf{m}_0$, with covariance $\frac{1}{\lambda_1}\mathbf{I}$, allowing for uncertainty around the prior model. Finally, $\tilde{\mathbf{m}}$ is a zero-mean Gaussian distribution with covariance $\frac{1}{\lambda_2}(\mathbf{D}^T\mathbf{D})^{-1}$, reflecting the uncertainty in the smoothing regularization constraint. One convenient way to draw a realization of $\tilde{\mathbf{m}}$ is to solve the linear system

$$\mathbf{D}\tilde{\mathbf{m}} = \boldsymbol{\eta}, \qquad \boldsymbol{\eta} \sim N(\mathbf{0}, \frac{1}{\sqrt{\lambda_2}}\mathbf{I}). \tag{19}$$

Substituting this definition of $\tilde{\mathbf{m}}$ into eq. (17) yields an equivalent form of the RTO objective function:

$$\frac{1}{2}\left\|\mathbf{Gm} - \tilde{\mathbf{d}}\right\|_2^2 + \frac{\lambda_1}{2}\left\|\mathbf{m} - \tilde{\mathbf{m}}_0\right\|_2^2 + \frac{\lambda_2}{2}\left\|\mathbf{Dm} - \boldsymbol{\eta}\right\|_2^2 \tag{20}$$

Solving this objective for many random draws of $(\tilde{\mathbf{d}}, \tilde{\mathbf{m}}_0, \boldsymbol{\eta})$ produces an ensemble of RTO solutions that sample the posterior. The main limitation of RTO is that the repeated optimization for each randomized instance may be time-consuming as the dimensionality of the problem increases. This makes RTO less practical for extremely large-scale models or when computational resources are limited.

2.3.3 Approximate inference: VI

VI provides a principled framework for approximating intractable probability distributions, particularly in our case, the posterior distribution of the latent variables representing subsurface properties in a domain of interest. Let $\mathbf{d}$ denote the observed seismic data and $\mathbf{m}$ represent the latent subsurface properties. Our goal is to approximate the intractable posterior distribution $p(\mathbf{m}|\mathbf{d})$ with a tractable proposal distribution $q_\Theta(\mathbf{m})$, where Θ represents the optimized free parameters that characterize the proposal distribution. VI achieves this by minimizing the so-called Kullback–Leibler (KL) divergence, a measure of the dissimilarity between two distributions, here the proposal or variational distribution and the true posterior distribution:

$$\begin{aligned}\mathrm{KL}(q_\Theta(\mathbf{m})\|p(\mathbf{m}|\mathbf{d})) &= \mathbb{E}_{\mathbf{m}\sim q_\Theta}\left[\log\frac{q_\Theta(\mathbf{m})}{p(\mathbf{m}|\mathbf{d})}\right]\\ &= \mathbb{E}_{\mathbf{m}\sim q_\Theta}[\log q_\Theta(\mathbf{m})] - \mathbb{E}_{\mathbf{m}\sim q_\Theta}[\log p(\mathbf{m}|\mathbf{d})]\\ &= \mathbb{E}_{\mathbf{m}\sim q_\Theta}[\log q_\Theta(\mathbf{m})] - \mathbb{E}_{\mathbf{m}\sim q_\Theta}\left[\log\frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})}\right]\\ &= \mathbb{E}_{\mathbf{m}\sim q_\Theta}[\log q_\Theta(\mathbf{m})] - \mathbb{E}_{\mathbf{m}\sim q_\Theta}[\log p(\mathbf{d}|\mathbf{m})]\\ &\quad -\mathbb{E}_{\mathbf{m}\sim q_\Theta}[\log p(\mathbf{m})] + \mathbb{E}_{\mathbf{m}\sim q_\Theta}[\log p(\mathbf{d})],\end{aligned} \tag{21}$$

where $\mathbb{E}_{\mathbf{x}\sim q_\Theta}$ denotes the expectation with respect to the variational distribution q_Θ, $p(\mathbf{d}|\mathbf{m})$ corresponds to the likelihood distribution of the data given the model and $p(\mathbf{m})$ is the prior distribution. Since computing the KL divergence as in eq. (21) is often intractable, VI uses a lower bound on the log-evidence, known as the evidence lower bound (ELBO):

$$\mathbb{E}_{\mathbf{m}\sim q_\Theta}[\log p(\mathbf{d}|\mathbf{m})] - \mathrm{KL}(q_\Theta(\mathbf{m})\|p(\mathbf{m})). \tag{22}$$

As maximizing the ELBO is equivalent to minimizing the KL divergence, VI ultimately entails minimizing the negative of the ELBO:

$$\underset{\Theta}{\operatorname{argmin}}\ \mathrm{KL}(q_\Theta(\mathbf{m})\|p(\mathbf{m})) - \mathbb{E}_{\mathbf{m}\sim q_\Theta}[\log p(\mathbf{d}|\mathbf{m})]. \tag{23}$$

By optimizing the ELBO with respect to the variational parameters Θ, we effectively find the variational distribution that best approximates the true posterior while maintaining tractability.

Key factors to the success of VI are: (i) the proposal and prior are chosen such that their KL divergence can be computed efficiently (or, ideally, a closed-form solution expression exists), (ii) the expectation in the second term can be approximated by the sample mean, obtained by sampling multiple models from the proposal distributions and evaluating this term as:

$$\mathbb{E}_{\mathbf{m}\sim q_\Theta}[\log p(\mathbf{d}|\mathbf{m})] \approx -\frac{1}{n_r}\sum_{i=1}^{n_r}\frac{\|\mathbf{d}-\mathbf{G}\mathbf{m}_i\|^2}{2\sigma_\epsilon^2}, \tag{24}$$

where n_r represents the number of samples drawn from the proposal and used to estimate the expectation of the likelihood distribution. Note that when VI is used in the context of machine learning (e.g. training of neural networks), this term can be computed from a batch of training samples: in other words, one leverages stochastic optimization techniques, ensuring computational efficiency even for large-scale problems. Unlike MCMC methods that require numerous iterations to converge and often struggle with high-dimensional spaces (Curtis & Lomax 2001), VI's optimization-based approach enables efficient parameter updates through mini-batch processing and automatic differentiation. Furthermore, VI's deterministic nature ensures reproducible results across multiple runs, while its optimization objective (ELBO) serves as a reliable metric for model comparison and convergence monitoring.

The success of VI critically depends on the choice of the variational family—the class of distributions used to approximate the true posterior. This choice represents a fundamental trade-off between approximation accuracy and computational tractability. The mean-field approximation is a variational family were the proposal distribution $q_\Theta(\mathbf{m})$ is factorized as:

$$q_\Theta(\mathbf{m}) = \prod_{i=1}^{N_y\times N_x\times N_t} q_\Theta(m_i) \tag{25}$$

This approach estimates the diagonal of the covariance matrix, providing high computational efficiency and a general assessment of uncertainty for each inverted parameter. While appealing from a computational perspective, it assumes parameter independence, which can lead to underestimated uncertainties and fails to capture key features of the true posterior, such as multimodality or parameter correlations. To address these limitations, we present two approaches, building on the IntraSeismic framework (Romero *et al.* 2024):

(i) B-IntraSeismic (**BIS**): we first explore the use of a standard mean-field approximation with a Gaussian proposal distribution, where the parameters (mean and standard deviation) are learned through the nonlinear mapping module of IntraSeismic;

(ii) B-IntraSeismic with CNF (**BIS-Flow**): we enhance the mean-field approximation by implementing a 1-D CNF, allowing for the learning of a non-parametric proposal distribution that captures complexities such as asymmetry and multimodality in the inverted parameters.

A detailed methodology and implementation of each variant are presented in the following section.

2.4 BIS

Inspired by the implementation in NeRF by Shen *et al.* (2021), we first parametrize our Gaussian proposal distribution with parameters Θ, which are the weights of an INR (Fig. 1). This INR takes coordinates of the subsurface model $\mathbf{x}_i = [x_i, y_i, z_i]$ as input and utilizes a nonlinear mapping module to predict the mean μ_Θ and standard deviation σ_Θ of a Gaussian distribution, which aims to approximate the model's posterior distribution:

$$m_i \sim N(\mu_\Theta(\mathbf{x}_i), \sigma^2_\Theta(\mathbf{x}_i)) = q_\Theta(m_i), \tag{26}$$

where $\mathbf{x}$ has been added to both the mean and standard deviation to remark the fact that different values are predicted from different input coordinates:

$$[\mu_\Theta(\mathbf{x}_i), \sigma^2_\Theta(\mathbf{x}_i)] = F_\Theta(\mathbf{x}_i), \tag{27}$$

where F_Θ represents our network. For the prior distributions, we first implement two Gaussian priors as detailed in eqs (9) and (10), which ensure proximity to the background model and smoothness, respectively. As we assume these priors to be independent from each other, the overall prior equals to the product of each individual prior:

$$p(\mathbf{m}) = p_1(\mathbf{m})p_2(\mathbf{m}), \tag{28}$$

where $p_1(\mathbf{m})$ and $p_2(\mathbf{m})$ are the two selected priors. This combination allows for direct comparison with the analytical solution, as it uses the same priors. Subsequently, we employ the blockiness prior in combination with the proximity prior to showcase the flexibility of the proposed methods. In both cases, when the prior term is inserted in eq. (23), p_1 is used alongside the proposal to compute the KL divergence (which in the case of a Gaussian prior and a Gaussian proposal has a closed-form solution). On the other hand, p_2 is taken out of the KL divergence and computed in a similar manner to the likelihood term. For the Gaussian case:

$$\mathbb{E}_{\mathbf{m}\sim q_\Theta}[\log p_2(\mathbf{m})] \approx -\frac{1}{n_r}\sum_{i=1}^{n_r}\frac{\lambda_2}{2}\|\mathbf{D}\mathbf{m}_i\|^2 \tag{29}$$

and for the Laplace term:

$$\mathbb{E}_{\mathbf{m}\sim q_\Theta}[\log p_2(\mathbf{m})] \approx -\frac{1}{n_r}\sum_{i=1}^{n_r}\lambda_3\|\nabla\mathbf{m}_i\|_1. \tag{30}$$

To conclude, the BIS workflow consists of the following steps:

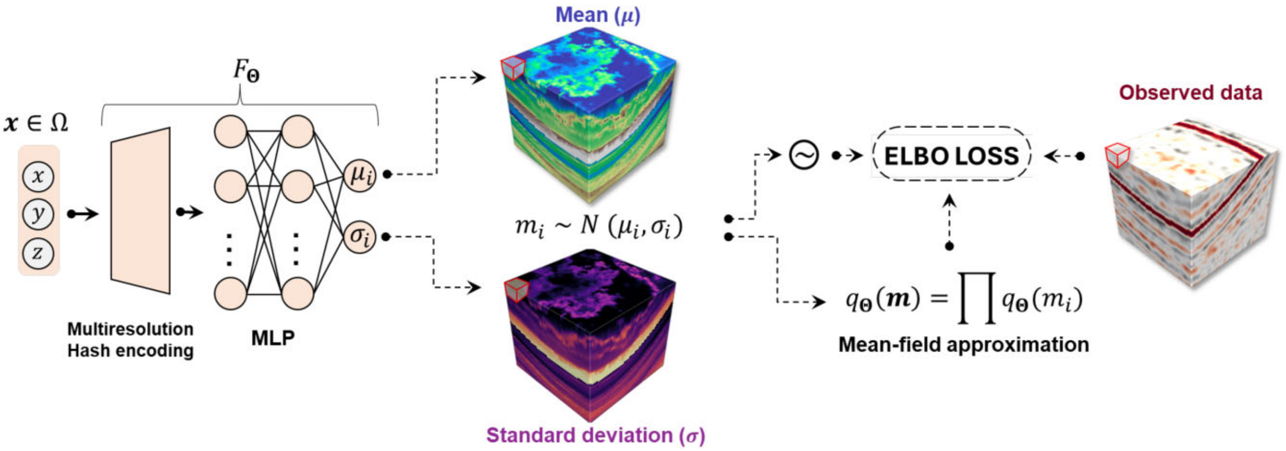

Figure 1. Schematic representation of the BIS framework. Input coordinates are processed through a nonlinear mapping module to produce the mean and standard deviation of a point-wise Gaussian distribution. Multiple acoustic-impedance models are sampled from this Gaussian proposal; those samples are then used to compute the likelihood term and one of the priors in the ELBO loss.

(i) **Parameter prediction:** The nonlinear mapping module processes input coordinates to predict the mean and standard deviation of the posterior distribution at each coordinate point (mean-field approximation). This module employs a trainable multigrid-hash encoding scheme followed by a two-layer Multi-Layer Perceptron (MLP) with 64 neurons and ReLU activation functions.

(ii) **Posterior sampling:** Using the predicted mean and standard deviation for all the input points $\mathbf{x}_i$, we draw N samples from the proposal distribution, where each sample represents a possible realization of the subsurface model parameters.

(iii) **Forward modelling:** Each of the n_r sampled models is fed to the seismic modelling operator, generating n_r predicted seismic data sets corresponding to each model realization.

(iv) **Loss computation:** The predicted seismic data are incorporated into the Gaussian-likelihood term of the loss function, and the predicted Gaussian distribution is used to compute the KL divergence with one of the priors. Additionally, the n_r realizations are used to compute the sample log probability for the other prior distribution.

(v) **Optimization:** Through backpropagation, the parameters of the nonlinear mapping module are optimized to maximize the ELBO.

2.5 BIS-Flow

Normalizing flows (NF) is a class of generative models that transform a simple base distribution through a series of learnable, invertible mappings, enabling the representation of multimodal, skewed and otherwise complex distributions (Kobyzev *et al.* 2021; Papamakarios *et al.* 2021). The key idea behind NFs is to apply a sequence of bijective transformations to a simple base distribution, making it possible to compute both the probability density and generate samples efficiently. The flow can be defined as:

$$\mathbf{x} = f(\mathbf{z}) = f_n(f_{n-1}(\dots f_1(\mathbf{z}))), \tag{31}$$

where

$$\mathbf{z} \sim N(0, \mathbf{I}) = p_z(\mathbf{z}) \tag{32}$$

and $\mathbf{x}$ is a sample from an unknown complex distribution $p_x(\mathbf{x})$. Each transformation f_i must be both invertible, so the original data can be recovered from the transformed data, and differentiable, to be able to compute the probability density function of the output variable. The transformation from a simple distribution $p_z(\mathbf{z})$ to a more complex distribution $p_x(\mathbf{x})$ is computed using the change of variables formula:

$$p_x(\mathbf{x}) = p_z(\mathbf{z}) \left| \det \frac{\partial f^{-1}(\mathbf{x})}{\partial \mathbf{x}} \right| = p_z(\mathbf{z}) \left| \det \frac{\partial f(\mathbf{z})}{\partial \mathbf{z}} \right|^{-1}, \tag{33}$$

where $\mathbf{z} = f^{-1}(\mathbf{x})$ is the latent variable, and the term $\left|\det \frac{\partial f^{-1}(\mathbf{x})}{\partial \mathbf{x}}\right|$ is the determinant of the Jacobian of the transformation, which accounts for how the transformation changes volume. This flexibility proves especially valuable in scenarios with intricate posterior landscapes, where standard VI might miss critical features. The iterative nature of flow-based transformations allows for the progressive refinement of the approximation, potentially capturing finer details of the true posterior while maintaining computational tractability. By proper mixing of variables by the transform used, the NF automatically discover relevant correlation structures in the posterior, eliminating the need for manual specification of variational families. Several recent studies have demonstrated the use of normalizing flows (NF) for Bayesian inference in seismic applications. Siahkoohi *et al.* (2020) propose a two-step Bayesian framework in which a conditional normalizing flow (CNF) is first trained on paired data to approximate the posterior distribution and subsequently reused as a prior or warm start for inference under new observation distributions. Zhao *et al.* (2022) apply NF to fully nonlinear seismic tomography, showing that it can produce accurate posterior approximations with lower computational cost compared to Monte Carlo methods. In a follow-up work, Siahkoohi *et al.* (2023) employ conditional NFs for amortized variational inference in seismic imaging, introducing a latent space correction where the standard Gaussian prior is replaced by a learnable Gaussian distribution with optimized mean and diagonal covariance to better match the physics-based posterior. Finally, Rizzuti & Vasconcelos (2024) present a multiscale variational inference framework for seismic uncertainty quantification, based on a wavelet-structured NF architecture that generates posterior samples hierarchically, from coarse to fine scales.

In this work, we leverage NF to create a more expressive proposal distribution whilst still remaining within the mean-field approximation regime (Fig. 2). For each input coordinate, this approach involves passing a single scalar drawn from a Gaussian distribution into a CNF, where the condition is composed of a feature vector generated by the nonlinear mapping module of IntraSeismic. The

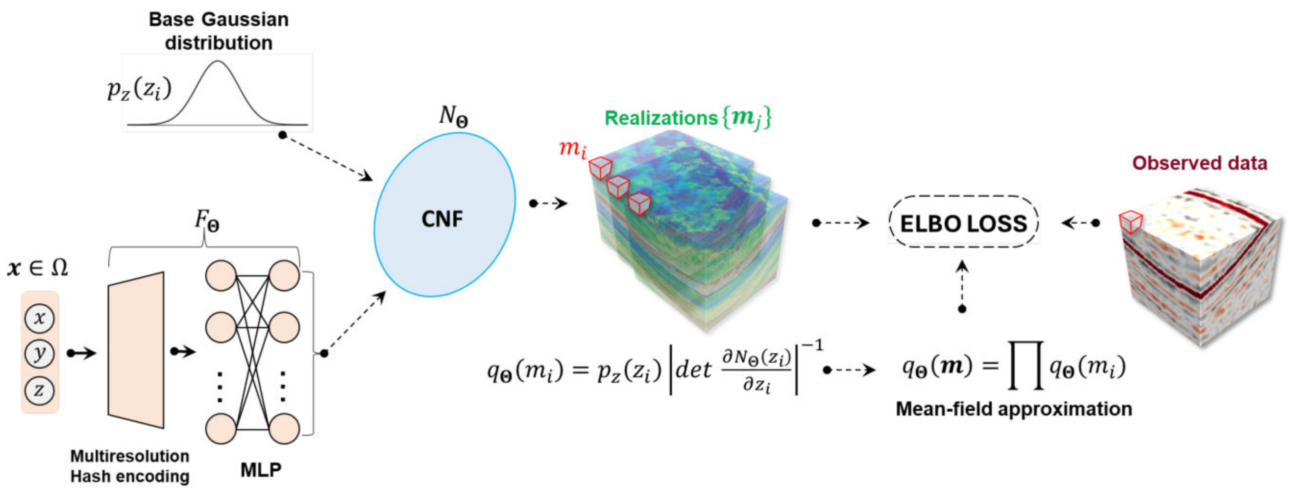

Figure 2. Schematic representation of the BIS-Flow framework. Input coordinates are processed through a nonlinear mapping module to produce a context vector, which is fed into the CNF along with samples from the base distribution. The output of the CNF provides both samples and log probabilities from a non-parametric proposal distribution, which are used to compute the terms of the ELBO loss.

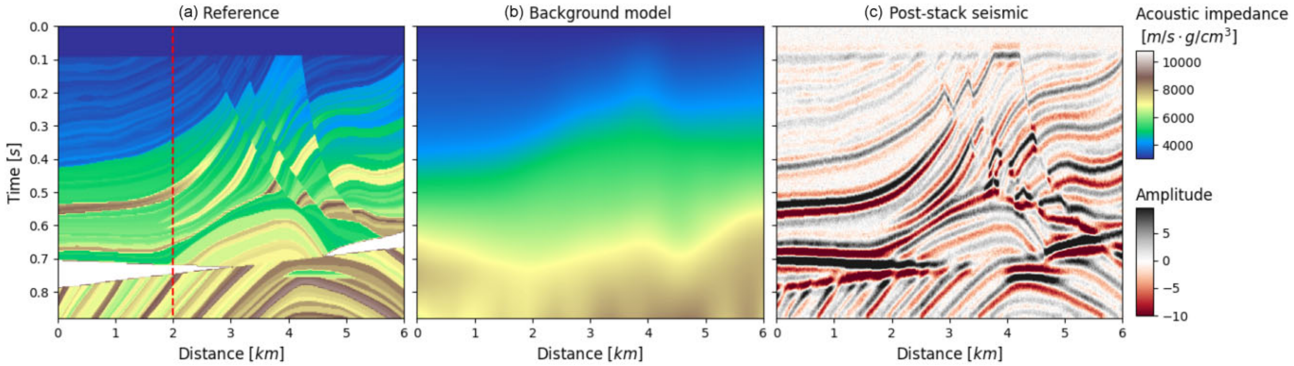

Figure 3. Marmousi acoustic impedance model used as the reference solution, background model used as the mean for the proximity prior and post-stack seismic data with added Gaussian noise. The vertical dashed line in the reference model corresponds to the location of the profile shown in Fig. 6.

input is then transformed by the flow into a sample of the non-parametric proposal distribution (Fig. 2). The log probability of the proposal is then computed for each sample using the change of variables formula in eq. (33). In this study, we employ Gaussianization flows (Meng *et al.* 2020), a powerful method for distribution modelling that maps non-Gaussian data to a Gaussian distribution. The fundamental concept involves applying monotonic, invertible transformations to the input data, producing outputs that resemble samples from a standard normal distribution. The transformation is defined as:

$$f(x) = \Phi^{-1}\left(\frac{1}{K}\sum_{i=1}^{K}\Phi(\exp(a_i)x + b_i)\right), \tag{34}$$

where Φ is the cumulative distribution function (CDF) of the standard normal distribution $N(0, 1)$, Φ^{-1} is the inverse CDF mapping probabilities to corresponding standard normal values, a_i (scale) and b_i (shift) are transformation parameters, and K is the number of components in the transformation. In our implementation, $\mathbf{a}$ and $\mathbf{b}$ are parametrized by a MLP N_Θ as:

$$[\mathbf{a}, \mathbf{b}] = N_\Theta(\mathbf{c}), \tag{35}$$

where Θ represents the learnable parameters and $\mathbf{c}$ is the context vector. Here, the context vector is a 64-dimensional feature output from the IntraSeismic MLP module. For the prior distributions, in this case, we use the proximity and blockiness promoting prior (eqs 9 and 11). The BIS-Flow workflow consists of the following steps:

(i) **Context vector:** The nonlinear mapping module processes the input coordinates, mapping each point to a 64-dimensional feature vector that serves as the context vector for the CNF, following the approach in Shen *et al.* (2022).

(ii) **Base distribution sampling:** The base distribution of the NF is sampled n_r times for each input point.

(iii) **Flow transformation:** The n_r sampled points, along with the context vector, are passed through the CNF Gaussianization flow, producing outputs that represent values from a complex, non-parametric distribution.

(iv) **Forward modelling:** Each of the n_r sampled model parameter realizations is fed to the seismic modelling operator, generating n_r predicted seismic data sets corresponding to each realization.

(v) **Loss computation:** The predicted seismic data are used in the Gaussian-likelihood term of the loss function, while the n_r realizations contribute to the computation of the proposal and prior terms. Note that here we do not rely on a closed-form solution for the KL divergence, instead the negative log-probability of the proposal is computed using eq. (33) and those from the prior terms are computed using eqs (29) and (30).

(vi) **Optimization:** The parameters of the nonlinear mapping module and the CNF are optimized using backpropagation to maximize the ELBO.

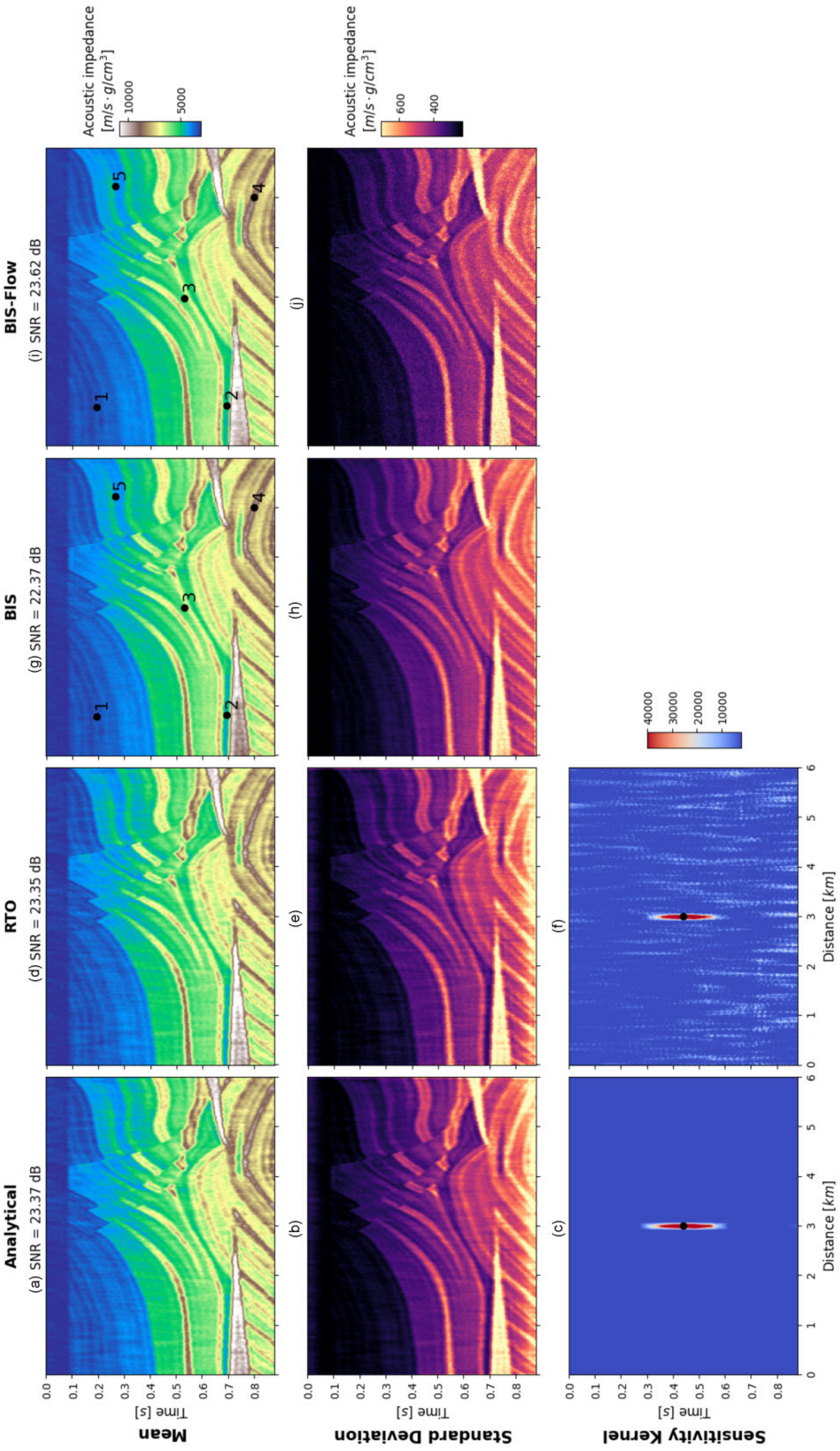

Figure 4. Comparison of the analytical, RTO, BIS and BIS-Flow for the Bayesian inversion of the post-stack modelled Marmousi data.

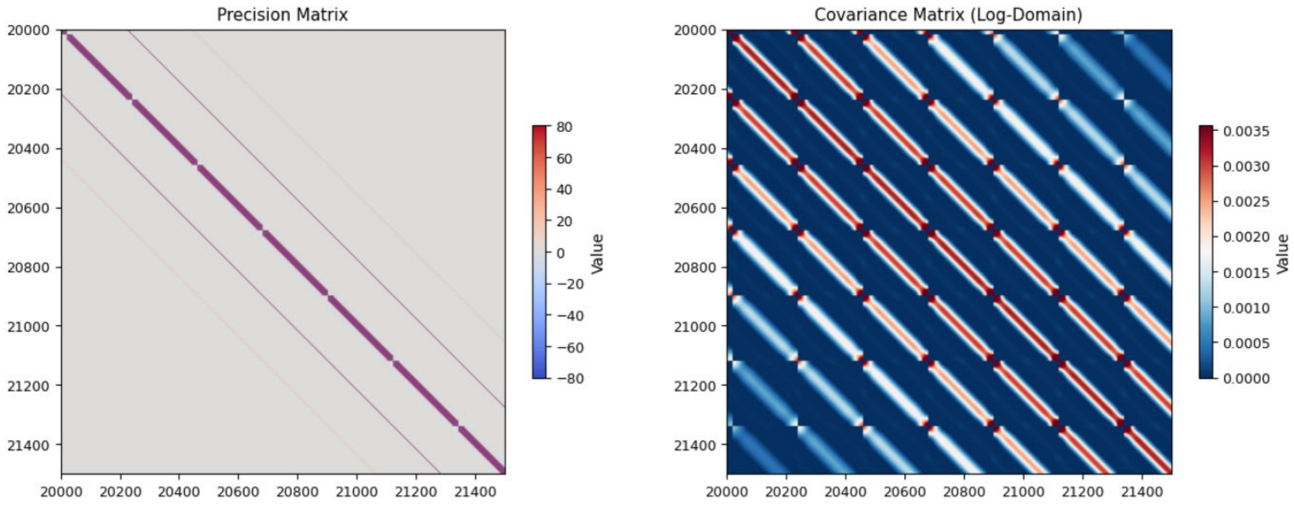

Figure 5. (a) Precision matrix (inverse of the covariance) and (b) covariance matrix in the log domain for the post-stack seismic inversion problem with Gaussian priors.

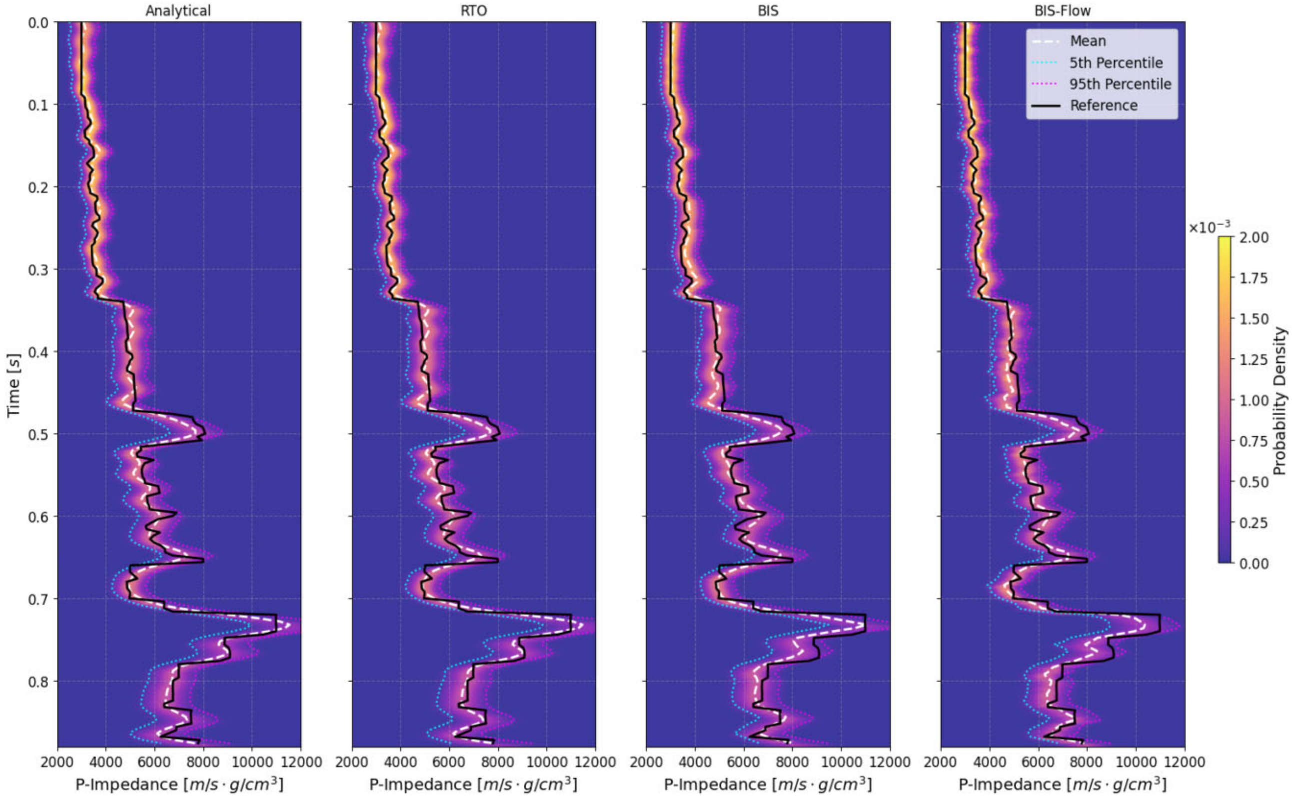

Figure 6. 1-D comparison of the analytical, RTO, BIS and BIS-Flow methods for the Bayesian inversion of the post-stack modeled Marmousi data. The extracted profile corresponds to the dashed vertical line in Fig. 3.

This workflow enables the model to approximate more flexible posterior distributions, capturing complexities such as asymmetry and multimodality.

2.6 Implementation details

For the experiments presented in the next section, we use $n_r = 100$ samples during training to evaluate the negative log-likelihood and the negative log-prior distributions. For all the experiments of BIS and BIS-Flow (synthetic and field data), we use Adam optimizer with a learning rate of 10^{-2}. The entire framework is implemented in PyTorch 2.0.1. All our experiments are conducted on a single machine equipped with an AMD EPYC 7713 64 Core Processor and a single NVIDIA TESLA A100 GPU. The post-stack seismic modelling operator is implemented using the PyLops library (Ravasi & Vasconcelos 2020) and integrated into the PyTorch framework to enable seamless automatic differentiation across multiple independent computational frameworks.

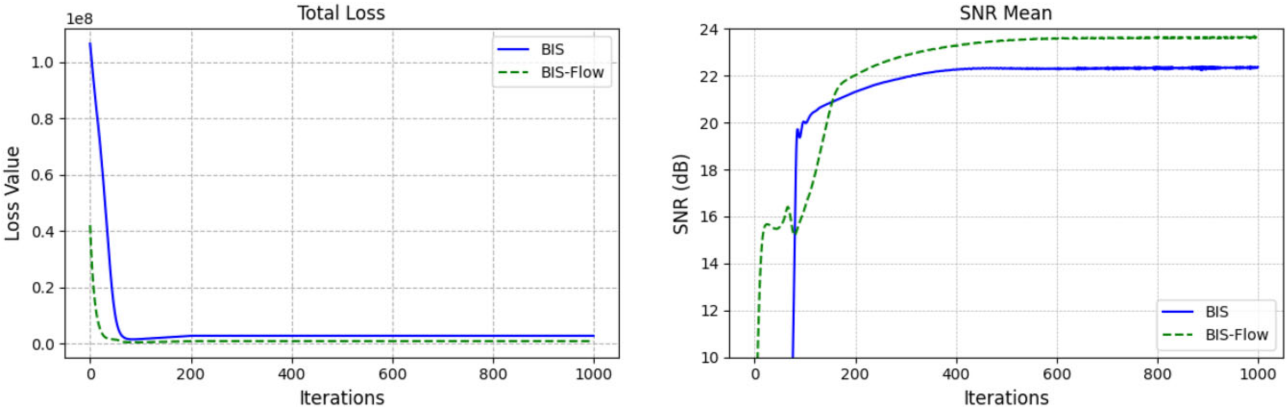

Figure 7. Loss and SNR of the mean through iterations of BIS and BIS-Flow for the Bayesian inversion Marmousi data.

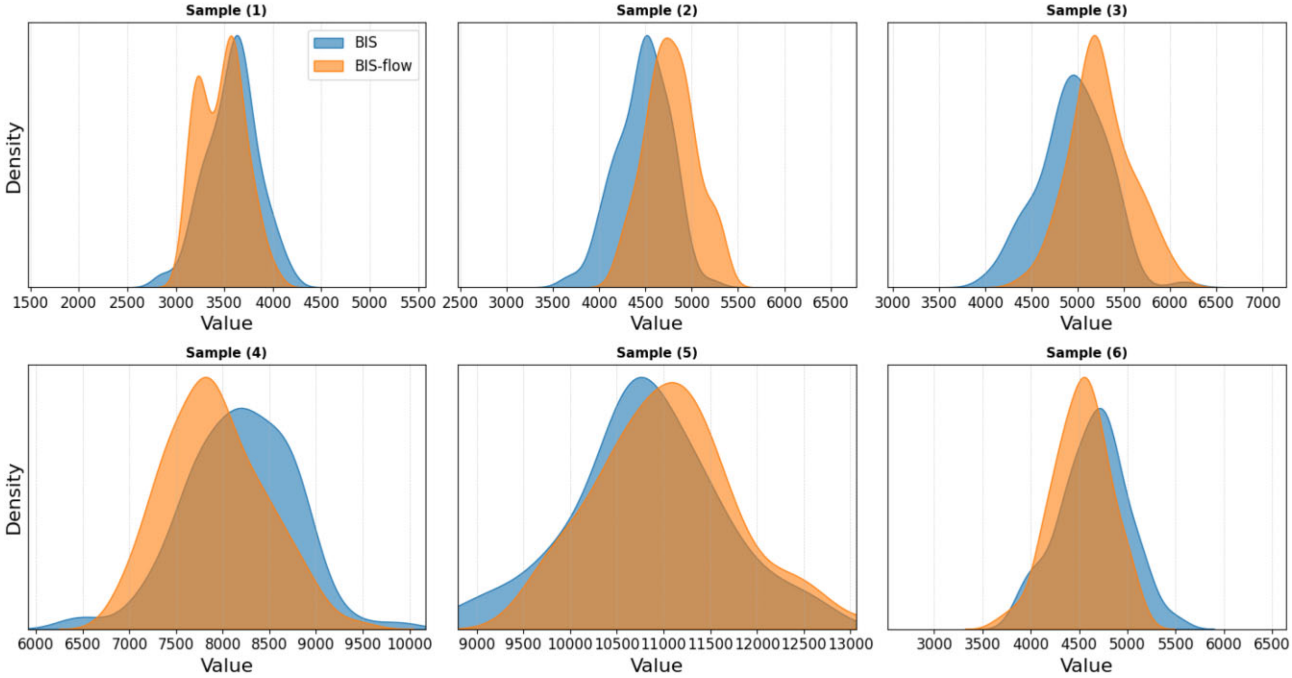

Figure 8. Kernel density estimates of the proposal distributions predicted by BIS and BIS-Flow at the six locations highlighted in Figs 4(g) and 4(i).

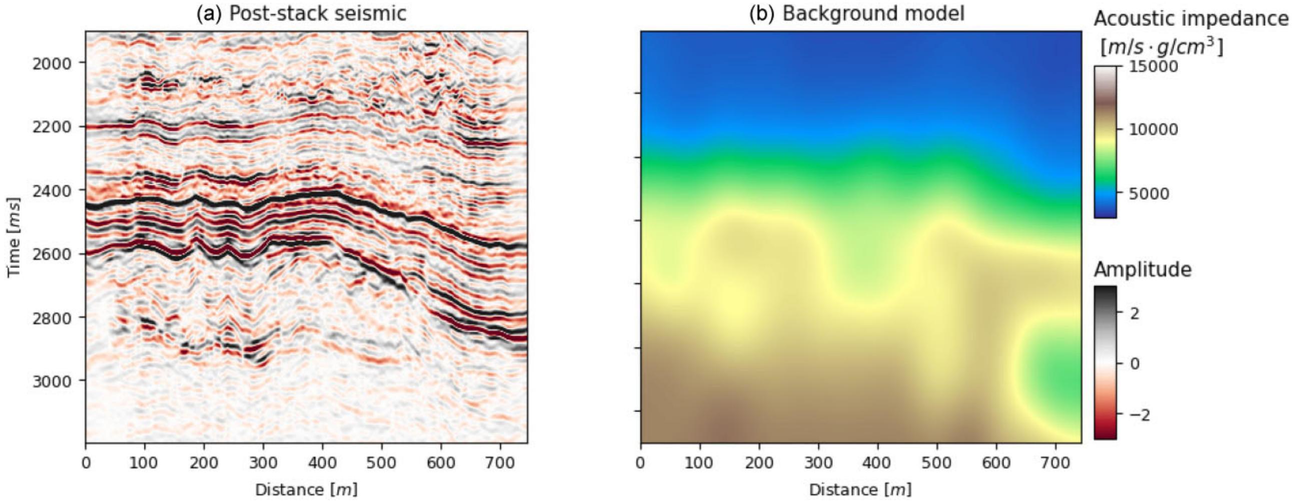

Figure 9. Volve post-stack seismic data and background model used in this study.

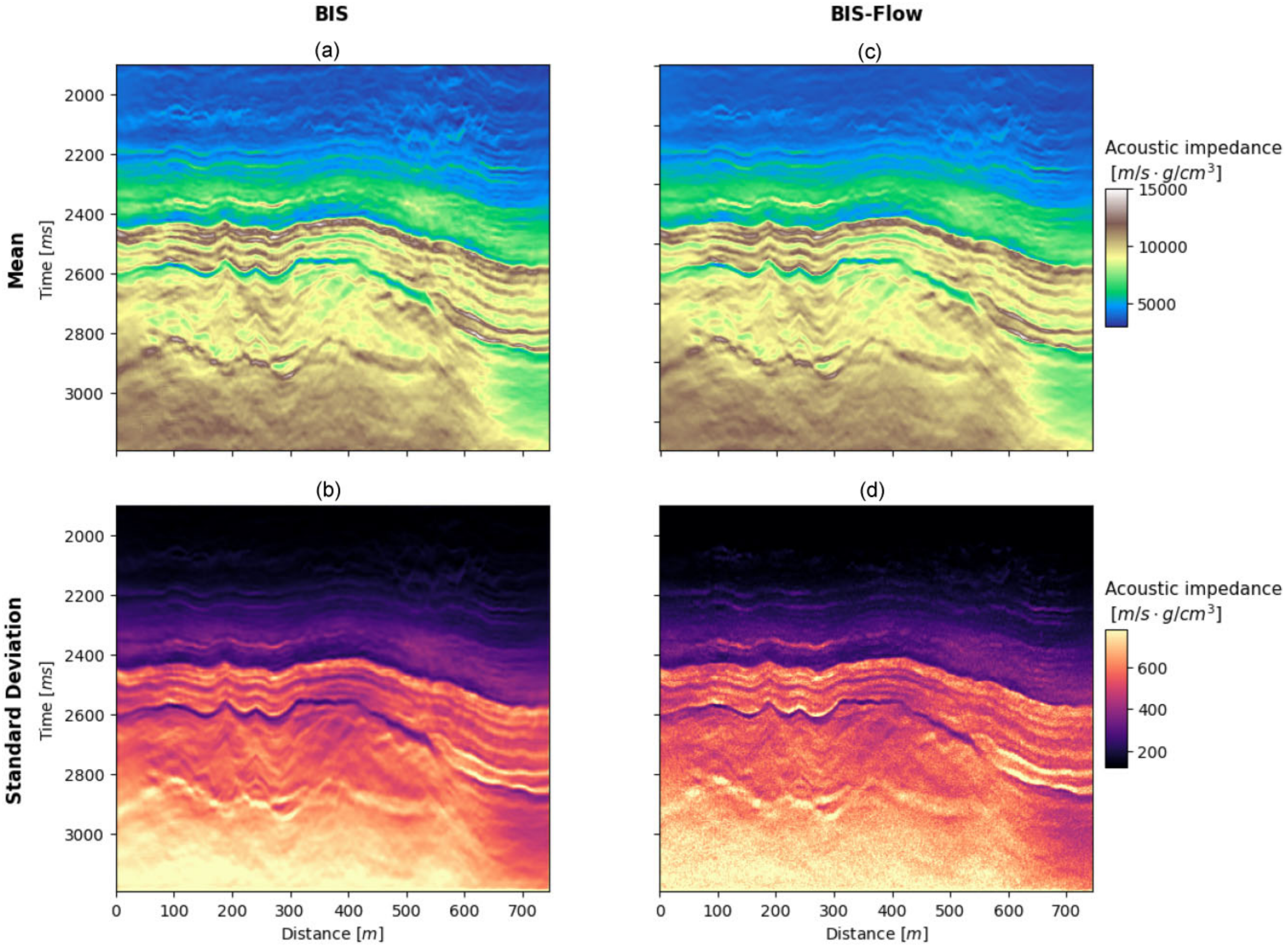

Figure 10. Bayesian inversion results for the Volve data set. (a)–(b) display the posterior mean and standard deviation estimated with BIS, whereas (c)–(d) show the corresponding statistics obtained with BIS-Flow.

3 RESULTS

3.1 Synthetic example: Marmousi model

The Marmousi model (Brougois *et al.* 1990) is a synthetic 2-D subsurface velocity model that mimics complex geological structures characterized by sharp velocity contrasts (Fig. 3a). We utilize a smooth version of the Marmousi model as the initial background model for the inversion process (Fig. 3b). To generate synthetic post-stack seismic data, we apply the linear post-stack modelling operator to the logarithm of the acoustic impedance model (computed by scaling the provided velocity model by a constant density value) using a 15 Hz peak-frequency Ricker wavelet. The data and wavelet are then scaled and the former is contaminated with Gaussian noise having a standard deviation of 1.0 (Fig. 3c).

3.1.1 *Analytical solution*

First, we calculate the analytical mean and covariance matrix of the acoustic impedance model from the post-stack data, following eqs (15) and (16). This calculation utilizes two Gaussian priors: one that enforces proximity to a background model and another that ensures smoothness in the solution (see Section 2.2.1). Figs 4(a) and (b) display the analytical mean and standard deviation (computed as the square root of the covariance matrix diagonal). The recovered mean effectively reconstructs the structural features and impedance values of the ground truth model, albeit with a lower signal-to-noise ratio (SNR) compared to the deterministic solution using advanced regularization terms (Romero *et al.* 2024). The standard deviation exhibits strong structural conformance, with maximum variance corresponding to regions of high impedance contrast.

Fig. 5 presents the precision and covariance matrices for the least-squares post-stack seismic inversion problem. The precision matrix (i.e. the inverse of the covariance matrix) provides insights into the conditional dependencies between variables, given the influence of the modelling operator and the priors. In this case, the precision matrix has dominant diagonal elements that correspond to the modelling operator and off-diagonal bands representing the smoothing prior (eq. 10). The covariance matrix in Fig. 5(b) presents a blocky structure, where each block displays the correlation between two traces. Parameters within the same block (or trace) are more strongly correlated, with diagonal dominance highlighting that uncertainty (variance) is highest for individual parameters and decreases with distance from the diagonal. This suggests that nearby parameters influence each other to some extent.

The sensitivity kernel in Fig. 4(c) (one row of the covariance matrix) provides an alternative visualization of parameter uncertainty and spatial correlation of the inverted parameters. For a selected point in the model (black dot), the kernel reveals high sensitivity in the surrounding region, indicating that parameter changes in this area significantly impact the predicted data at the selected point, while changes in the blue areas have minimal effect. The vertical

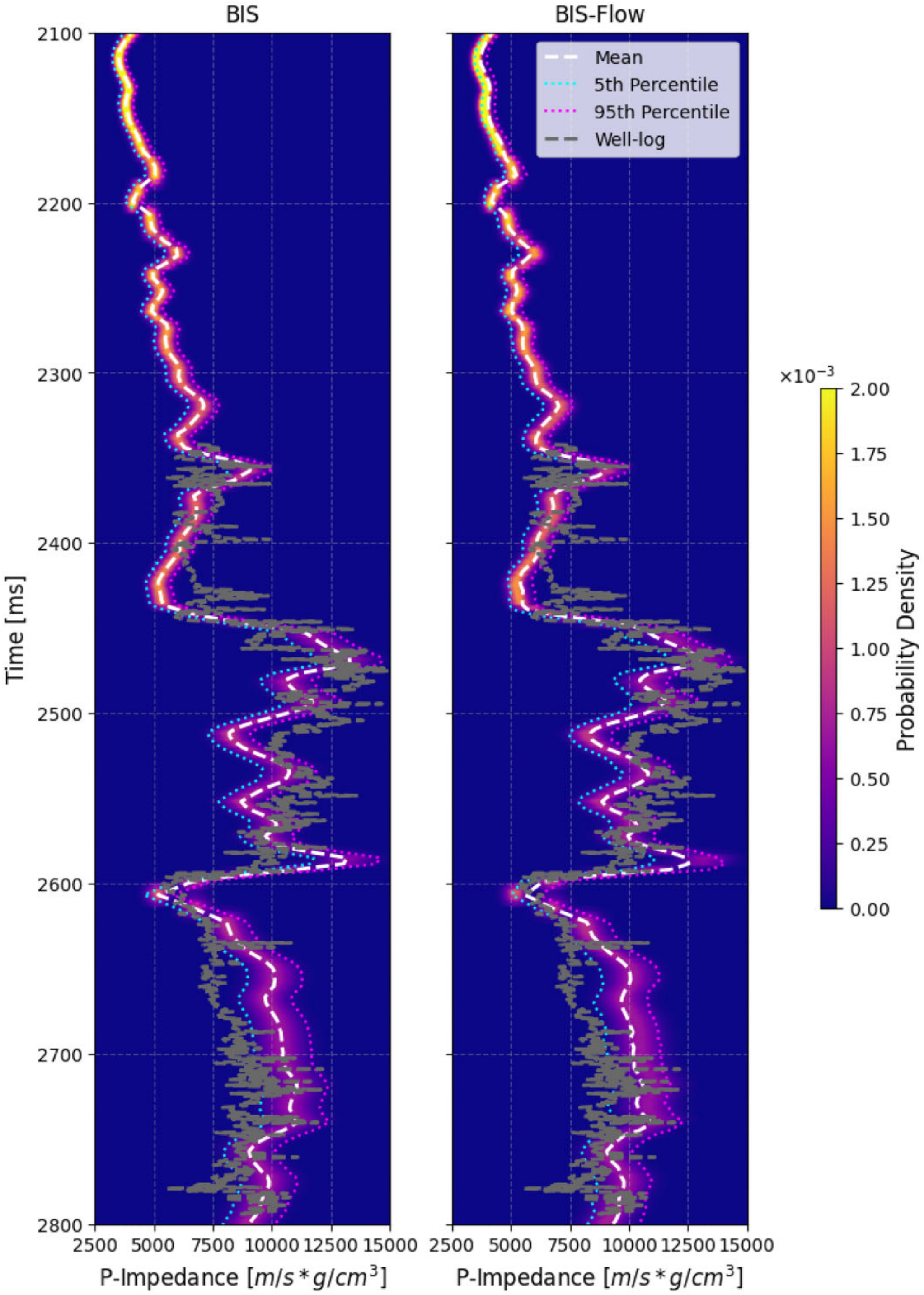

Figure 11. Profile view along a well trajectory for the Bayesian inversion results for the Volve data set using BIS and BIS-Flow, compared to acoustic impedance well log.

elongation is primarily due to the modelling operator, characterized by a Toeplitz convolutional matrix with a filter size of 60 samples. The overall shape of the sensitivity kernel reflects the Laplacian regularization, which promotes smoother models.

A more detailed comparison of the results is presented in Fig. 6, which illustrates the marginal posterior distributions of the acoustic impedance model along a single vertical profile (indicated by a red dashed line in Fig. 3) for the various inference methods shown in Fig. 4.

3.1.2 *RTO solutions ensemble*

Figs 4(d)–(f) display the mean, standard deviation and sensitivity kernel of the RTO solutions obtained for the post-stack seismic inversion of the Marmousi synthetic data. In this analysis, 1000 perturbations of the data and priors were performed, and each of the 1000 inverse problems is solved using a conjugate gradient solver with 200 iterations. The posterior mean and standard deviation closely match the analytical solution, successfully reproducing

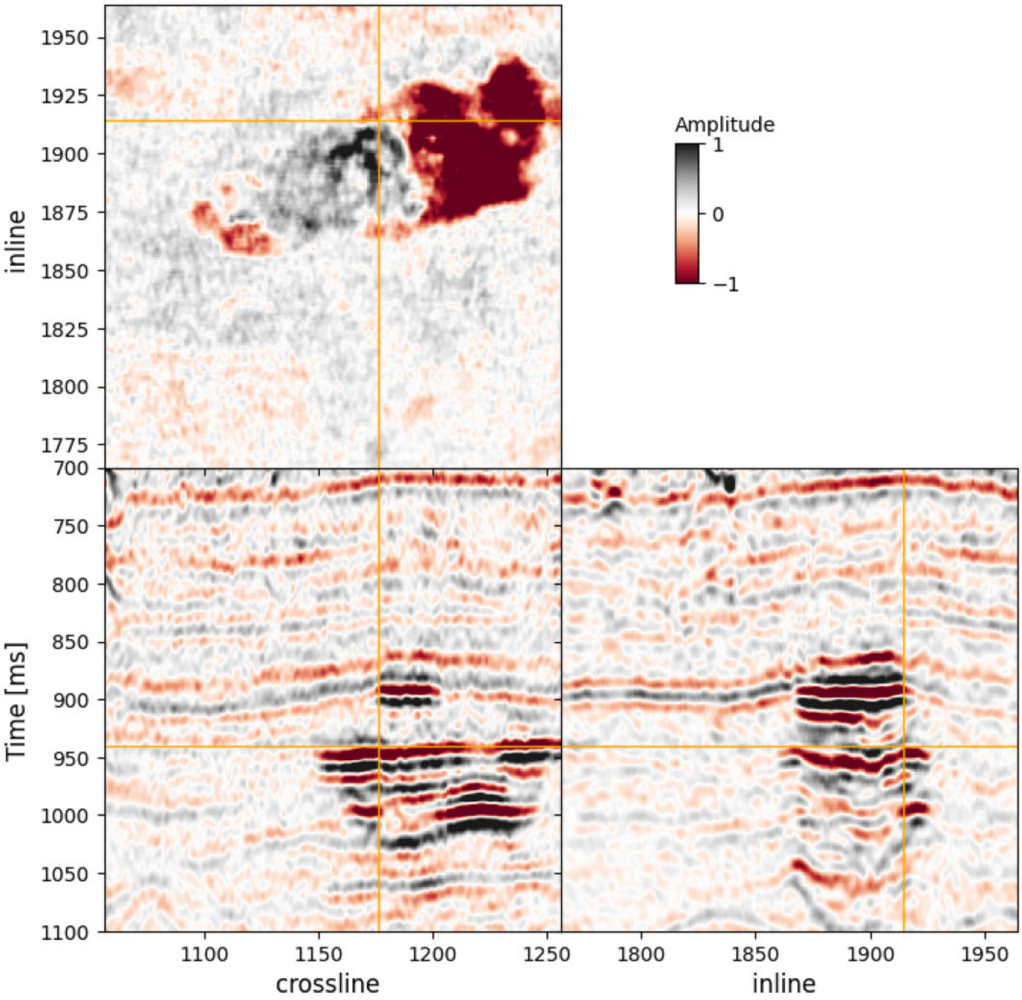

Figure 12. Selected subvolume from the full stack 2001 Sleipner data set.

the key features of the Marmousi model. The RTO sensitivity kernel (Fig. 4f) likewise demonstrates the strong agreement between the RTO-estimated covariance and the analytical covariance, with the former only slightly noisier owing to the finite number of realizations and the limited iterations of the linear solver.

3.1.3 Gaussian posterior and Gaussian priors

Next, we utilize BIS to estimate the parameters of our Gaussian proposal distribution. For this estimation, we have also used a combination of two Gaussian priors as applied in the analytical case (eqs 9 and 10). The predicted mean and standard deviation are illustrated in Figs 4(g)–(h) and are shown to closely resemble those from the analytical posterior distribution. However, since each point is parametrized independently in BIS, the resulting covariance matrix is diagonal, and therefore, it does not make sense to display the resulting sensitivity kernel.

3.1.4 Flexible posterior and non-Gaussian priors

We finally employ BIS-Flow to estimate a non-parametric version of the proposal distribution for the Marmousi model; in this case, we use a blockiness-promoting prior (as shown in 11). This prior encourages higher resolution realizations of the posterior distribution while maintaining the same computational cost as when using purely Gaussian priors. Figs 4(i)–(j) present the predicted mean and standard deviation for the inversion of the Marmousi post-stack seismic data. Fig. 7 illustrates the loss and SNR evolution over iterations for both BIS and BIS-Flow corresponding to the results in Figs 4(i)–(j). The loss in both cases exhibits a steep decline during the first 100 iterations, with SNR converging at different points: BIS reaches a plateau after 400 iterations, whereas BIS-Flow converges more quickly, matching the SNR of BIS at 200 iterations and showing a marginal increase until 800 iterations.

Fig. 8 compares the proposal distributions predicted by BIS and BIS-Flow at the locations marked in Figs 4(g) and 4(i). Whereas BIS uses a Gaussian proposal distributed and therefore explicitly produces a Gaussian approximation of the posterior distribution, BIS-Flow is allowed to produce more flexible distributions—for example, distributions with slight skewness and sharper modes, a direct consequence of the blockiness-promoting prior. This additional flexibility enables BIS-Flow to approximate the true posterior more faithfully, free from the restrictive Gaussian assumption. By accommodating local asymmetries and pronounced peaks, BIS-Flow provides a more realistic description of the uncertainty and variability in the inverted parameters, which is essential for modelling the complex behaviour of subsurface properties in seismic inversion.

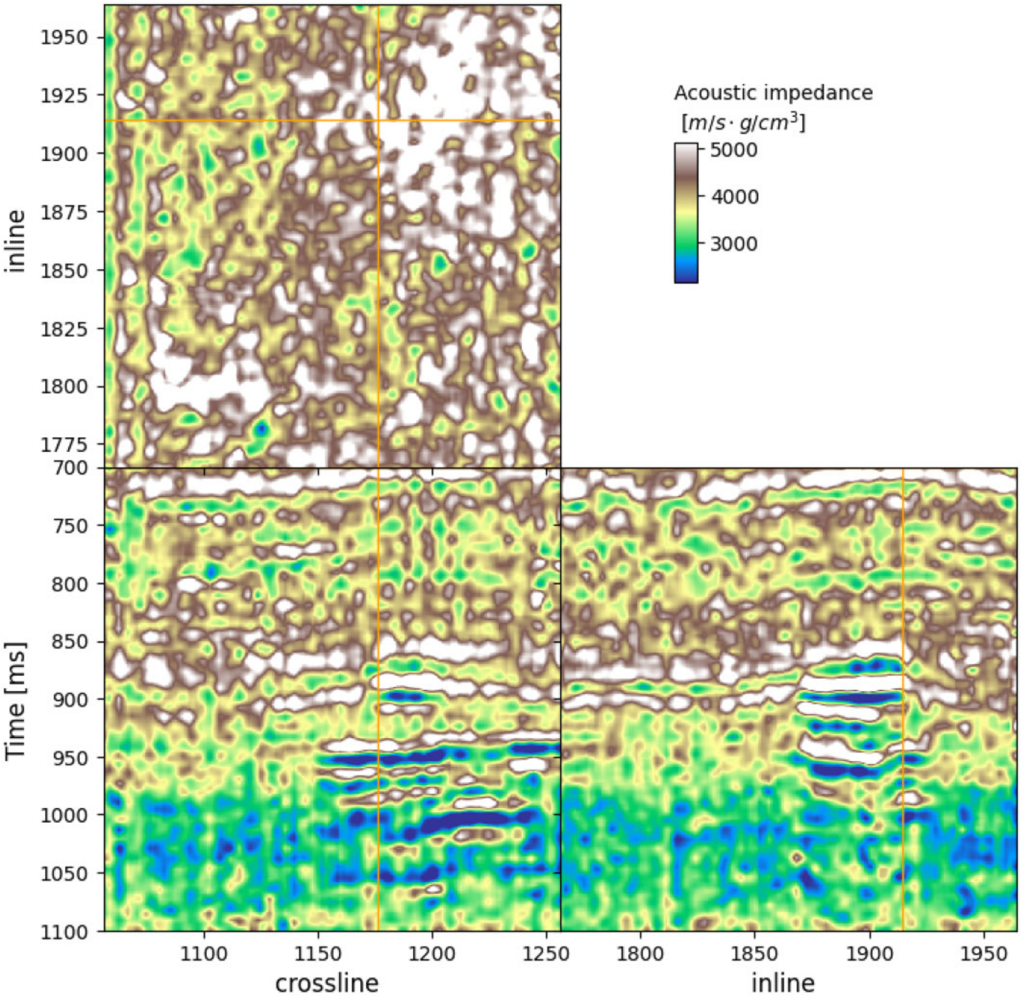

Figure 13. Mean acoustic impedance model obtained by BIS from the full stack 2001 Sleipner data set.

3.2 Field data example 1: Volve data set

Next, we evaluate the performance of BIS and BIS-Flow on a field data set. We use a 2-D composite section extracted along a well of the Volve data set (refer to Ravasi & Birnie 2021 for details about the Volve data set processing). Fig. 9 shows the post-stack seismic data along the well fence and the background acoustic impedance model constructed from well logs, which serves as the mean for the proximity prior. Fig. 10 presents the predicted mean and standard deviation for both BIS and BIS-Flow. Although the two models are similar, we note that BIS-Flow produces a mean model with better horizontal continuity and improved vertical resolution compared to the mean model predicted by BIS. The standard deviation maps for BIS and BIS-Flow exhibit similar value ranges; however, the BIS-Flow standard deviation map displays higher resolution, with more detailed layering and defined zones of uncertainty. These standard deviation values are consistent with state-of-the-art Bayesian inversion results on the Volve data set obtained using multiscale wavelet flows (Rizzuti & Vasconcelos 2024).

Finally, Fig. 11 presents the predicted mean and standard deviation along the well impedance log (NO/15-9 19 BT2). Both BIS and BIS-Flow face challenges in accurately predicting the mean, with the well impedance log falling outside the 95 per cent confidence interval in certain regions. These discrepancies likely arise from two main factors: the presence of significant non-Gaussian noise in the Volve seismic data, which is not fully captured by BIS and BIS-Flow, and misalignment in depth between the well log and seismic data.

3.3 Field data example 2: 3-D Sleipner data set

We further evaluate the performance of BIS and BIS-Flow on a second field data set—that is, a subvolume of dimensions of 200 × 200 × 200 extracted from the 2001 full-stack seismic data of the Sleipner field (Fig. 12). This data is composed of predominantly horizontal geological structures intersected by strong reflection responses due to CO_2 injection within the main reservoir. Details regarding the creation of the background model, statistical wavelet extraction and other pre-processing steps can be found in Romero *et al.* (2023b). In this case, we use 50 samples to compute the log likelihood and the sample log probability for the priors. In the case of the BIS-Flow, allocating all of the samples at the same time on a GPU requires substantial memory; this can be managed by processing in batches of points in 3-D domain, utilizing multiple GPUs or reducing the model complexity. In this case, we follow the first approach, dividing the domain into two subvolumes of 200 × 200 × 100 voxels and processing them sequentially.

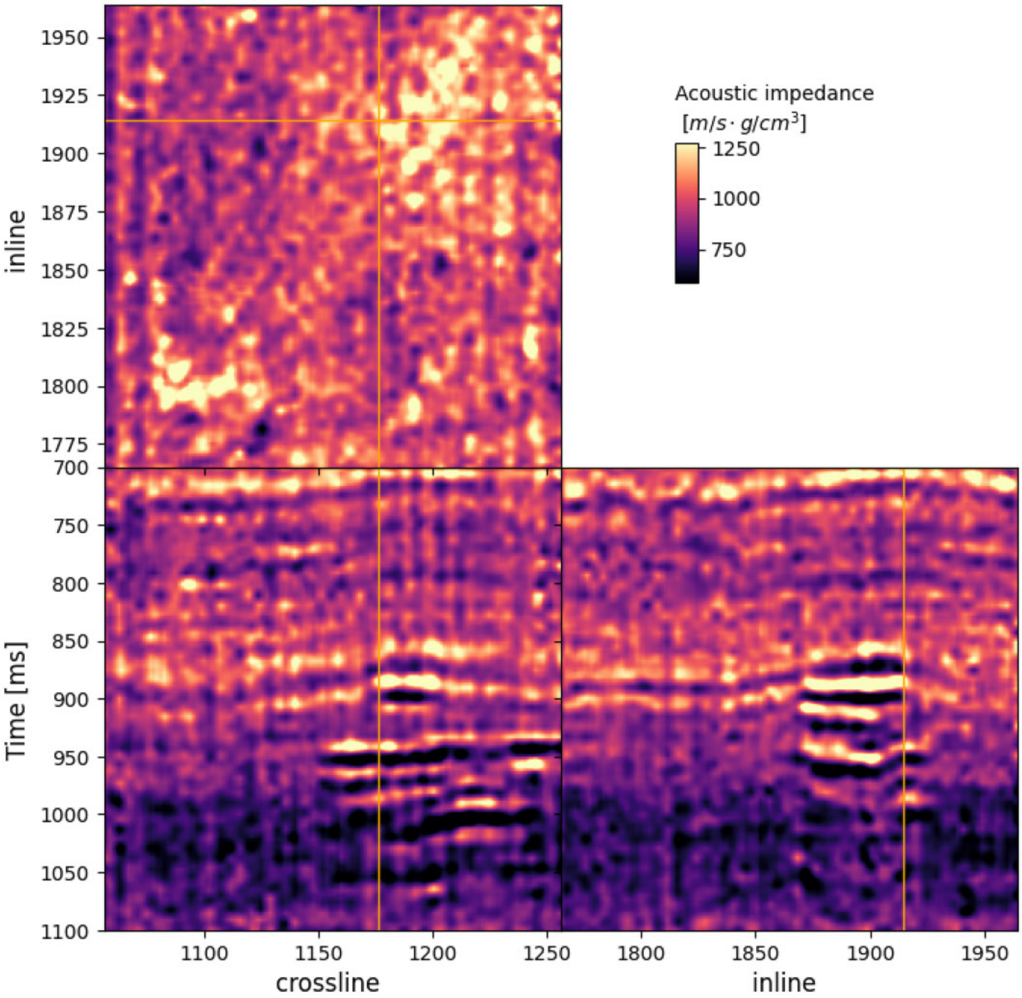

Figure 14. Standard deviation of the BIS solution from the full stack 2001 Sleipner data set.

The predicted mean and standard deviation of the acoustic impedance model obtained using BIS for the Sleipner data are shown in Figs 13 and 14. Both the mean and standard deviation appear noisy, lacking the clear horizontal layering observed in the deterministic inversion results reported by Romero *et al.* (2024). This suggests that BIS struggles to produce a reliable, high-resolution acoustic impedance model for this data set. The primary reason for this limitation is attributed to the presence of anomalous amplitude values in the upper layers of the data, likely associated with gas chimneys. Since BIS employs a smoothing prior, these anomalies may be preventing the inversion from converging effectively across the entire volume, leading to reduced model accuracy and resolution.

On the other hand, the predicted mean of BIS-Flow (Fig. 15) exhibits a robust representation of the model parameters, producing a result visually similar to the model originally obtained with the deterministic IntraSeismic inversion method (Romero *et al.* 2024). This suggests that BIS-Flow effectively captures the key features of the subsurface structure. The BIS-Flow standard deviation (Fig. 16) highlights that zones with the highest uncertainty primarily correspond to layers with high impedance values, such as the thick layer at the top and some thin layers within the CO_2 reservoir (potentially due to wavelet distortions). Comparing the predicted acoustic impedance mean and standard deviation for Sleipner with the impedance well log (Fig. 17) shows that the well log generally falls within the 5–95 per cent confidence interval. This suggests that BIS-Flow provides robust and reliable uncertainty estimates also in this 3-D field data example. Fig. 18 presents the BIS-Flow posterior distributions at the six positions marked in Fig. 17. In this case, some distributions are nearly Gaussian, whereas others exhibit a much sharper peak, highlighting the influence of the blockiness-promoting prior.

4 DISCUSSION

This work presents the adaptation of IntraSeismic, an INR-based framework for deterministic seismic inversion, to the realm of Bayesian inversion using VI with different parametrizations for the proposal distribution. More specifically, we explored a proposal based on the mean-field approximation with point-wise Gaussian distributions (BIS) and a mean-field approximation with a non-parametric distributions (BIS-Flow). Both approaches offer computationally efficient alternatives to state-of-the-art methods for assessing the uncertainty of the post-stack seismic inversion problem. BIS-Flow, in particular, demonstrated the ability to capture complexities such as multimodality and asymmetry in the posterior distribution, providing a valuable option for obtaining high-resolution

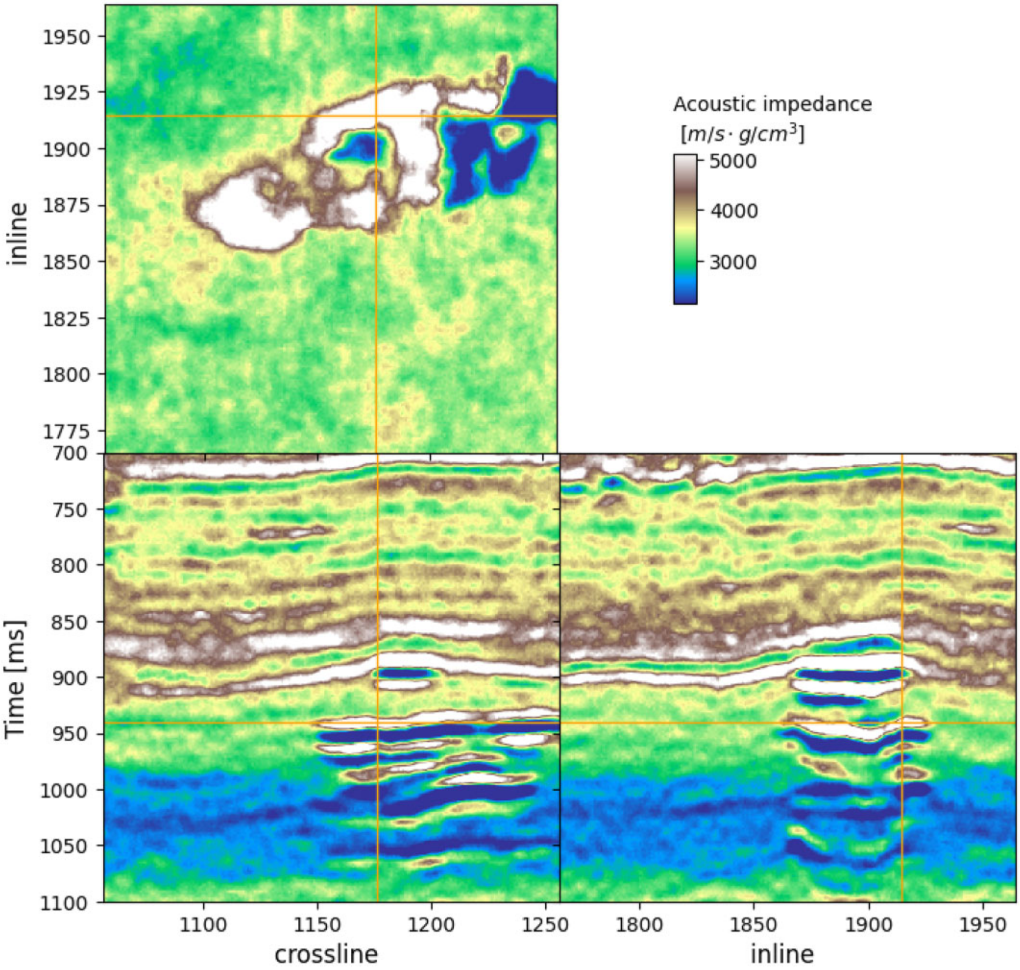

Figure 15. Mean acoustic impedance model obtained by BIS-Flow from the full stack 2001 Sleipner data set.

estimates of the uncertainty of each inverted parameter whilst producing a high-resolution mean model. The primary limitation of both methods is their assumption of parameter independence, which restricts their usefulness for generating realistic posterior samples where interparameter correlations are important.

Fig. 4 compares the mean, standard deviation and a row of the covariance matrix for a single point using three approximate inference algorithms (RTO, BIS, BIS-Flow) against the exact analytical solution for the posterior distribution of the inverted acoustic impedance from post-stack Marmousi seismic data. The different approximate inference methods predicted mean models that visually capture the key features of the Marmousi model, such as fault locations and layer boundaries, resulting in models with SNR close to the analytical one (23.3 dB) and comparable to the deterministic case (24.73 dB) (Romero *et al.* 2024). BIS produced the lowest SNR for the mean, while BIS-Flow achieved the highest SNR due to its blockiness-promoting prior and additional flexibility in the proposal distribution that allowed for higher resolution estimates. The standard deviation maps in Fig. 4 approximate the analytical results reasonably well, showing higher values in regions with greater impedance contrasts. Fig. 6 provides a 1-D view of the mean and the 5th and 95th percentile confidence intervals, illustrating that for all methods, the true Marmousi model mostly lies within this range, with higher impedance contrast layers located near the interval edges. Note that the sensitivity kernel is not shown for the BIS and BIS-Flow methods, as the mean-field approximation assumes parameter independence.

For the Bayesian inversion of field data, we present two examples: the 2-D inversion of the Volve data set and the 3-D inversion of the Sleipner data set. In the Volve data set example, both BIS and BIS-Flow demonstrated strong performance, producing high-resolution estimates of the mean and standard deviation. While there is no ground truth for direct comparison, the standard deviation map aligns well with state-of-the-art results, such as those produced by multiscale wavelet flows (Rizzuti & Vasconcelos 2024). The Sleipner data set highlights a significant advantage of the proposed BIS and BIS-Flow methods: scalability. Using INR that converges more quickly than traditional variational and deep learning-assisted inversion methods (Romero *et al.* 2024), combined with VI, we achieved reliable uncertainty estimation while maintaining low computational costs.

Estimating the covariance matrix for Gaussian priors or the scale matrix for Laplace priors is a crucial step in Bayesian seismic inversion. Although some studies model these parameters as additional latent random variables, this approach typically increases the computational cost significantly due to the added dimensionality of the problem. For the field data used in this work, we chose to set these parameters based on available well-log data, enabling us to incorporate prior information without increasing the complexity of the inversion. For the proximity prior, we selected the standard deviation by calculating the maximum deviation of the background model from the impedance values in the well log. This provides a reasonable bound on variability based on the well data. For the Laplace prior, we computed the gradient of the impedance from

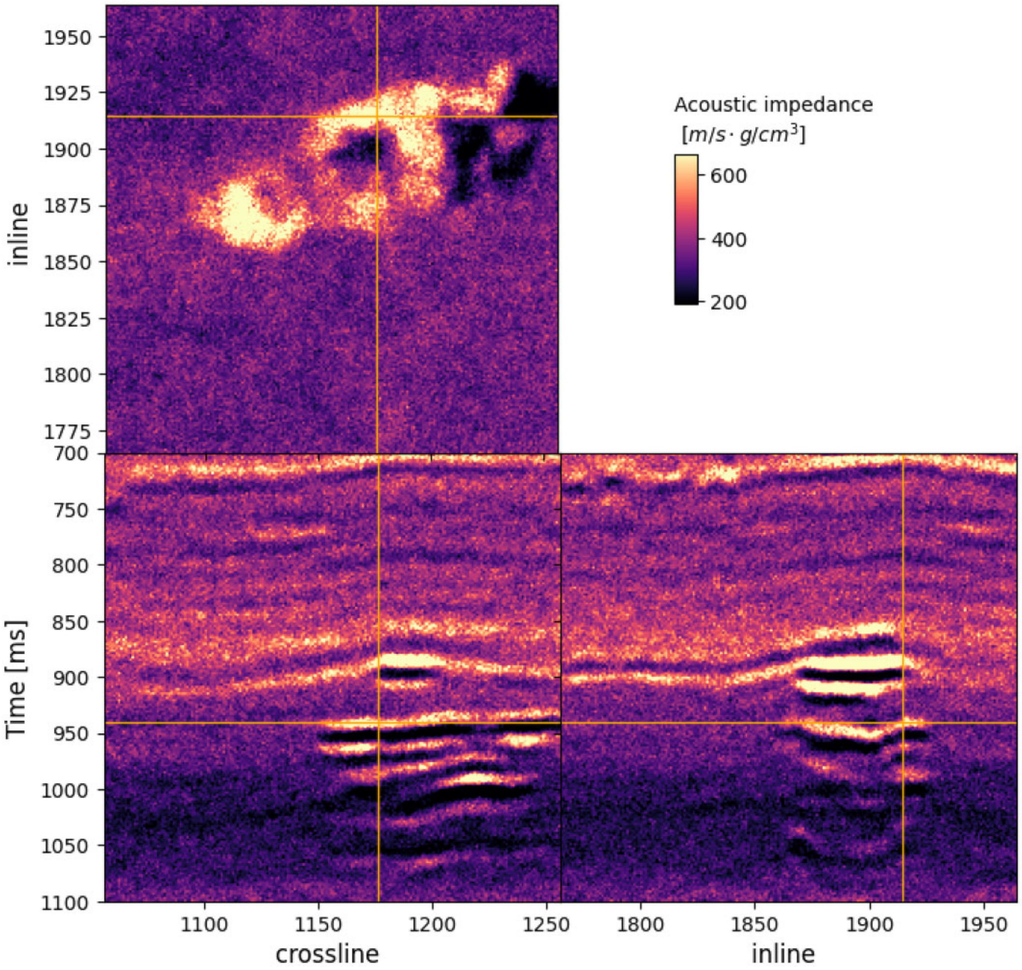

Figure 16. Standard deviation of the BIS-Flow solution from the full stack 2001 Sleipner data set.

the well log and selected an interval for the scale parameter of the Laplace distribution that best reflects this gradient information.

4.1 Computational cost

In this section, we compare the four methods used to perform uncertainty quantification in post-stack seismic inversion from a computational standpoint.

4.1.1 Analytical

To estimate the covariance matrix of the posterior distribution, the precision matrix must be inverted. Moreover, to compute the mean, two additional matrix–vector multiplications must be performed. Since the direct inversion of a dense matrix of size $N_{\text{params}} \times N_{\text{params}}$ has a complexity of $O(N^3_{\text{params}})$, this represents the most computationally expensive step of this approach, and it becomes prohibitively expensive as the size of the problem grows. As an example, for a problem of size 220 by 600 (Marmousi model used in this study), the inversion process took approximately 2.3 hr in the machine specified in 2.6.

4.1.2 RTO

The RTO method requires solving an inverse problem for each perturbation (sample); this is typically carried out using an iterative solver like the Conjugate Gradient (CG) method. The computational cost of solving a system of equations of size $N_{\text{params}} \times N_{\text{params}}$ is approximately equal to $O(N_{\text{iter}} \cdot N^2_{\text{params}})$, where N_{iter} is the number of CG iterations required to reach convergence. Since this process is repeated for N_{samples} number of samples, the overall cost is given by: $O(N_{\text{samples}} \cdot N_{\text{iter}} \cdot N^2_{\text{params}})$.

4.1.3 BIS and BIS-Flow

The computational cost of each iteration of the BIS method is composed of three main parts. First, the forward and backward passes through the MLP network, which is proportional to the size of the model and that of each matrix of the MLP layers. Given a two-layer MLP with 64 neurons and N_{params} parameters, this is equal to $O(2 \cdot (2 \cdot 64 \cdot N^2_{\text{params}}) + 64 \cdot N^2_{\text{params}}))$ where the first 2 is based on the assumption the forward and backward passes have the same cost. Secondly, there is the multiresolution hash-encoding process, where each hash lookup has a cost of $O(1)$, and the bilinear interpolation conducted at each resolution level L has a cost of $O(L \cdot 4 \cdot 2)$, with 4 representing the neighbouring points and 2 representing the encoding dimensionality. For all the points in multiresolution hash encoding has a rough estimate of $O(N_{\text{params}} \cdot (L \cdot 8 + 1))$. Thirdly,

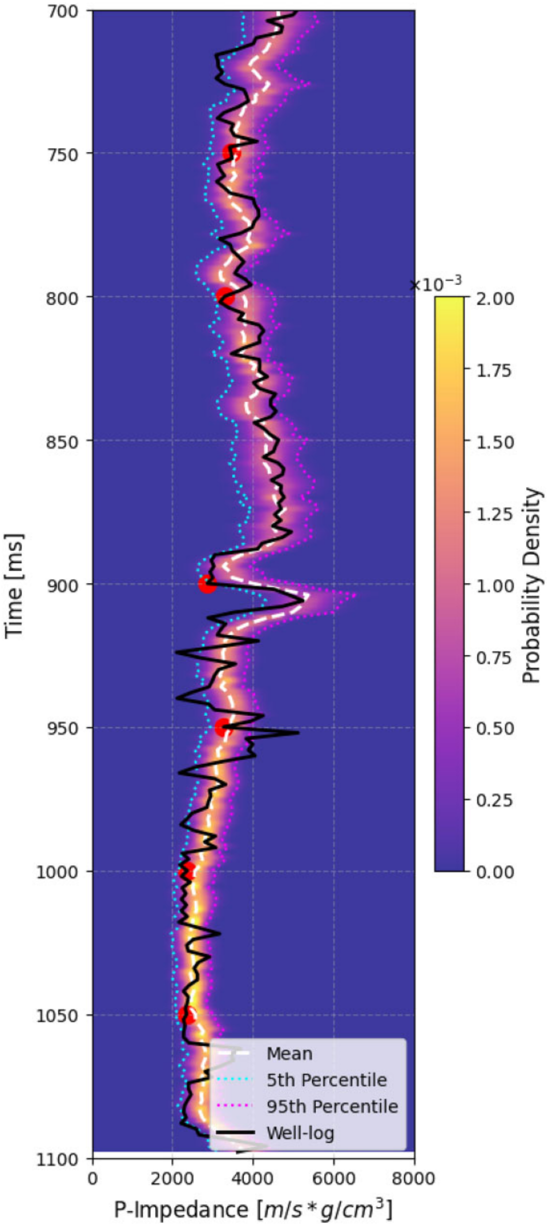

Figure 17. Profile view along a well trajectory for the Bayesian inversion results for the Sleipner data set using BIS-Flow and comparing it to acoustic impedance well log. The red dots show the vertical location of the histograms shown in Fig. 18.

the cost of computing the modelling operator, which is equal to $O(n_\mathrm{r} \cdot N^2_\mathrm{params})$. Note that although $n_\mathrm{r} < N_\mathrm{samples}$, BIS takes a larger number of iterations to converge than each single RTO inversion. Overall, the cost of BIS grows with N^2_params compared to N^3_params for the other two methods, making it more suitable to large-scale applications. The computational cost of BIS-Flow is similar to that of BIS, with the additional overhead of the CNF, which for a 1-D flow is $O(N_\mathrm{params})$. The total computational time for both BIS and BIS-Flow in the Marmousi data inversion was approximately 2 min on the hardware specified in Section 2.6.

The main limitation of the BIS and BIS-Flow methods compared to RTO is their memory footprint; these method require in fact one to approximate the expectations in the loss function via a sample mean over a number of realizations. When the optimization process is carried out on GPUs, this requires that all realizations are be stored in GPU memory at the same time (in the simplest implementation of the algorithm). This constraint becomes particularly challenging for high-dimensional problems, such as the 3-D Sleipner example presented earlier. To address this issue, several strategies can be employed: (1) Batch Processing: performing the inversion in smaller batches rather than processing all input points simultaneously, allowing results to accumulate over multiple iterations and reducing peak memory usage; (2) Leveraging Multiple GPUs: distributing the different realizations across multiple GPUs to increase available memory and improve processing efficiency, and reducing over their gradients before updating the parameters of the MLP network and hash-encoding table and (3) Reducing Model Complexity: simplifying the parametrization of the INR block used in the inversion to decrease the overall memory footprint. These strategies offer viable solutions for managing memory requirements without compromising inversion quality. A summary of the computational efficiency for the methods studied in this work is presented in Table 1.

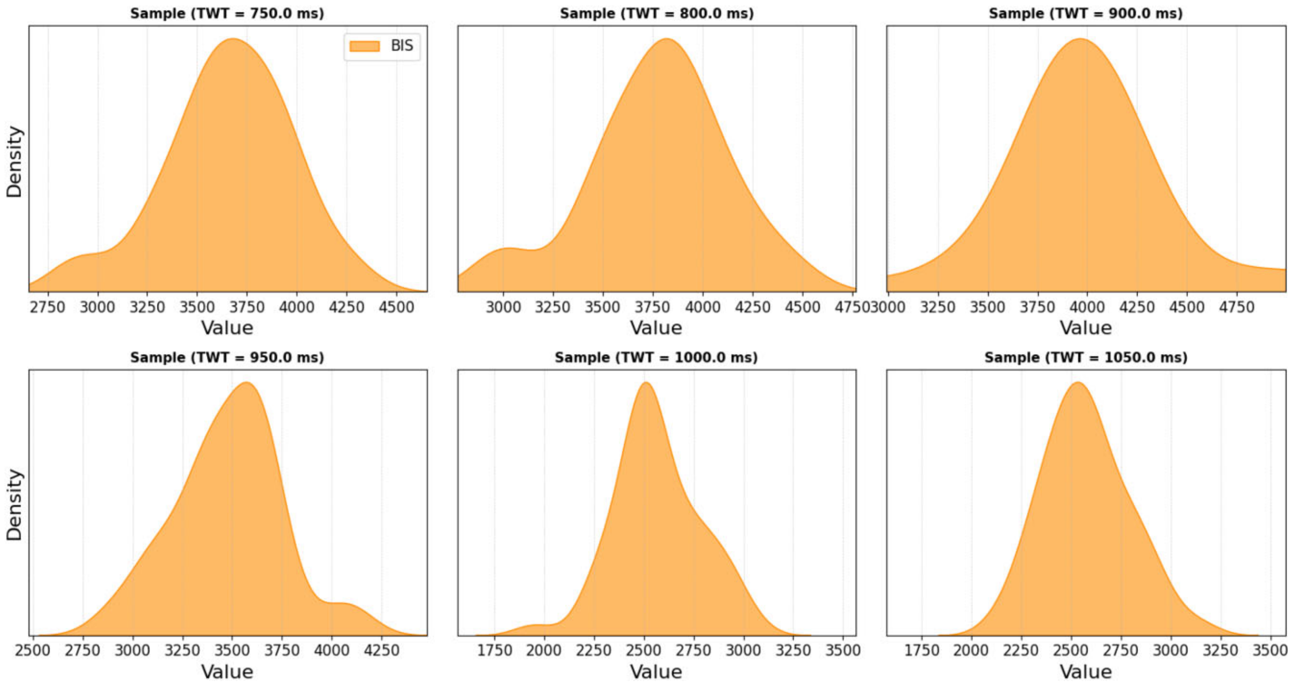

Figure 18. Kernel density estimation of BIS-Flow predicted proposal distribution samples at six locations along the well-trajectory in the Sleipner volume, shown by red dots in Fig. 17.

Table 1. Computational efficiency of the four Bayesian–inference schemes. Runtimes correspond to inversion of the 2-D 220×600 Marmousi synthetic data (Fig. 4), run on an AMD EPYC 7713 (64 cores) and a single NVIDIA A100.

Method	Peak memory	Scalability	Accuracy considerations	Runtime[z] (min)
Analytical	High (dense $n \times n$ matrix)	Poor	Exact mean / covariance; prohibitive for large n	134
RTO	Moderate (storage of realisations in CPU / GPU)	Poor	Asymptotically exact; for each sample, an inverse problem is solved	25.5
BIS	Moderate (realizations and network weights on GPU)	Good	Mean-field Gaussian; no spatial covariance between parameters	2.2
BIS-Flow	Moderate (realizations and network weights on GPU)	Good	Mean-field, non-Gaussian marginals via 1-D normalizing flows; no spatial covariance	2.5

[z] Wall-clock times for: Analytical; RTO–1000 posterior samples; BIS and BIS-Flow–full training run (500 epochs, 100 samples / epoch). All timings include forward-model evaluations but exclude I/O.

4.2 Beyond mean-field approximation

Accurately modelling interparameter correlations across the subsurface while maintaining computational feasibility is essential for comprehensive posterior exploration and remains a significant challenge in uncertainty quantification for seismic inversion. Subsurface parameters are neither independent nor isolated; rather, they exhibit spatial dependencies due to geological structures and stratigraphic continuity. In a fully Gaussian Bayesian setting, the posterior covariance matrix can capture these spatial correlations, providing insight into how changes in one area of the subsurface might influence or be influenced by adjacent regions. When dealing with more complex prior distributions or nonlinear modelling operators and/or model parametrization, however, no analytical expression exists for the mean and covariance of the posterior distribution. It is therefore essential for approximate inference methods to predict posteriors distributions that can be sampled leading to realistic, spatially coherent models that reflect the natural variability and connectivity of subsurface properties. Access to geologically plausible samples from the posterior distribution is crucial for risk assessment and decision-making in applications like resource exploration and CO_2 storage monitoring. Realistic samples that respect spatial correlations provide a more accurate representation of possible subsurface configurations, allowing geoscientists to evaluate a broader range of scenarios and anticipate potential geological risks.

This paper introduces an innovative approach using INR combined with 1-D CNF to model point-wise probability distributions effectively. Future work should focus on extending this framework to capture spatial correlations between input points more comprehensively. Preliminary experiments by the authors indicate that by smartly sharing the input noise samples from CNF base distribution across various spatial location we may be able to model dependencies between selected points while maintaining low computational costs. These findings are preliminary and will be detailed in a future publication.

5 CONCLUSIONS

In this paper, we introduce two innovative methods to perform uncertainty quantification in seismic inversion: B-IntraSeismic (BIS), which employs a mean-field Gaussian proposal for efficient yet simplified uncertainty quantification, and B-IntraSeismic with Conditional Normalizing Flows (BIS-Flow), which utilizes a non-parametric proposal distribution to capture complex posterior characteristics. Our results on synthetic and field data sets demonstrate that both BIS and BIS-Flow can produce high-resolution mean models and accurate standard deviation maps, with BIS-Flow excelling at capturing more complex posterior distributions. The enhanced capability of BIS-Flow to model complex uncertainties while maintaining computational feasibility underscores its potential for more comprehensive subsurface analysis. Additionally, we demonstrate the scalability of both BIS and BIS-Flow to large-scale applications. By using a 3-D field example, we show that these methods can provide reliable 3-D uncertainty estimates at a reasonable computational cost. The implications of this work are significant for post-stack seismic inversion, as they offer a scalable approach for obtaining reliable uncertainty estimates and more informative posterior representations. Future research should build on this foundation by developing methods that explore shared base sampling strategies to capture parameter correlations in larger and higher dimensional settings. These advancements will contribute to a more nuanced understanding of subsurface properties, supporting more accurate decision-making in exploration and risk management.

ACKNOWLEDGMENTS

This research was supported by King Abdullah University of Science and Technology (KAUST) and the DeepWave Consortium. We are deeply grateful to the KAUST Supercomputing Laboratory (KSL) for providing the high-performance computing resources that were essential for this work.

AUTHOR's CONTRIBUTION

J.R. performed the research and analysis and wrote the paper. W.H. and M.R. served as advisors and supervised the work.

DATA AVAILABILITY

All data sets utilized in this study are publicly available. For the synthetic data, please refer to the SEG Wiki page on Open Data at https://wiki.seg.org/wiki/Open_data. Similarly, for the Sleipner data set, information can be found at https://co2datashare.org/data set/sleipner-4d-seismic-data set. We are grateful to Equinor and their partners for providing open access to these valuable data sets.

The code for this paper is open source and available at https://github.com/DeepWave-KAUST/B-IntraSeismic-pub.

REFERENCES

Aki, K. & Richards, P.G., 2002. *Quantitative Seismology,* W. H. Freeman and Company.

Bardsley, J.M., Solonen, A., Haario, H. & Laine, M., 2014. Randomize-then-optimize: a method for sampling from posterior distributions in nonlinear inverse problems, *SIAM J. Sci. Comput.,* **36**(4), A1895–A1910. .

Blatter, D., Morzfeld, M., Key, K. & Constable, S., 2022. Uncertainty quantification for regularized inversion of electromagnetic geophysical data—Part I: motivation and theory, *Geophys. J. Int.,* **231**(2), 1057–1074. .

Blei, D.M., Kucukelbir, A. & McAuliffe, J.D., 2017. Variational inference: a review for statisticians, *J. Am. Stat. Assoc.,* **112**(518), 859–877. .

Bosch, M., Mukerji, T. & Gonzalez, E.F., 2010. Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review, *Geophysics,* **75**(5), 75A165–75A176. .

Brougois, A., Bourget, M., Lailly, P., Poulet, M., Ricarte, P. & Versteeg, R., 1990. Marmousi, model and data, in *EAEG Workshop - Practical Aspects of Seismic Data Inversion,* cp, European Association of Geoscientists and Engineers.

Buland, A. & Omre, H., 2003. Bayesian linearized AVO inversion, *Geophysics,* **68**(1), 185–198. .

Curtis, A. & Lomax, A., 2001. Prior information, sampling distributions, and the curse of dimensionality, *Geophysics,* **66**(2), 372–378. .

Goupillaud, P.L., 1961. An approach to inverse filtering of near-surface layer effects from seismic records, *Geophysics,* **26**(6), 754–760. .

Grana, D., Fjeldstad, T. & Omre, H., 2017. Bayesian Gaussian mixture linear inversion for geophysical inverse problems, *Math. Geosci.,* **49**(4), 493–515. .

Houck, R.T., 2002. Quantifying the uncertainty in an AVO interpretation, *Geophysics,* **67**(1), 117–125. .

Jordan, M.I., Ghahramani, Z., Jaakkola, T.S. & Saul, L.K., 1998. An introduction to variational methods for graphical models, in *Learning in Graphical Models,* pp. 105–161, ed. Jordan, M.I., Springer.

Kingma, D.P. & Welling, M., 2022. Auto-encoding variational Bayes,arXiv:1312.6114 [stat].

Kobyzev, I., Prince, S.J.D. & Brubaker, M.A., 2021. Normalizing flows: an introduction and review of current methods, *IEEE Trans. Pattern Anal. Mach. Intell.,* **43**(11), 3964–3979. ,

Li, P., Grana, D. & Liu, M., 2024. Bayesian neural network and Bayesian physics-informed neural network via variational inference for seismic petrophysical inversion, *Geophysics,* **89**(6), M185–M196. .

Meng, C., Song, Y., Song, J. & Ermon, S., 2020. Gaussianization flows,arXiv:2003.01941.

Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R. & Ng, R., 2020. NeRF: representing scenes as neural radiance fields for view synthesis,arXiv:2003.08934 [cs].

Mosegaard, K. & Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems, *J. geophys. Res.: Solid Earth,* **100**(B7), 12 431–12 447. .

Oldenburg, D., Scheuer, T. & Levy, S., 1983. Recovery of the acoustic impedance from reflection seismograms, *Geophysics,* **48,** 1318–1337.

Oz, Y. & Miller, R.D., 2015. *Engineering Seismology: with Applications to Geotechnical Engineering,* Society of Exploration Geophysicists.

Papamakarios, G., Nalisnick, E., Rezende, D.J., Mohamed, S. & Lakshminarayanan, B., 2021. Normalizing flows for probabilistic modeling and inference,arXiv:1912.02762 [cs, stat].

Ravasi, M. & Birnie, C., 2021. A joint inversion-segmentation approach to assisted seismic interpretation, *Geophys. J. Int.,* **228**(2), 893–912. .

Ravasi, M. & Vasconcelos, I., 2020. Pylops—a linear-operator python library for scalable algebra and optimization, *SoftwareX,* **11,** 100361. https://doi.org/10.1016/j.softx.2019.100361.

Rizzuti, G. & Vasconcelos, I., 2024. Multiscale uncertainty quantification for post-stack seismic inversion with wavelet flows. in *85th EAGE Annual Conference and Exhibition (including the Workshop Programme),* pp. 1–5, European Association of Geoscientists and Engineers, Oslo, Norway.

Romero, J., Luiken, N. & Ravasi, M., 2023b. Seeing through the co2 plume: Joint inversion-segmentation of the sleipner 4d seismic data set, *The Leading Edge,* **42,** 446–516.

Romero, J., Heidrich, W., Luiken, N. & Ravasi, M., 2024. Seismic reservoir characterization with implicit neural representations, *J. geophys. Res.: Machine Learning and Computation,* **1**(3), e2024JH000275. .

Russell, B. & Hampson, D., 2005. Comparison of poststack seismic inversion methods, *SEG Technical Program Expanded Abstracts 1991,* 876–878.

Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.,* **40**(3), 3–1-3-29. .

Scales, J.A. & Tenorio, L., 2001. Prior information and uncertainty in inverse problems, *Geophysics,* **66**(2), 389–397. .

Sen, M.K. & Stoffa, P.L., 1996. Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion, *Geophys. Prospect.,* **44**(2), 313–350. .

Shen, J., Ruiz, A., Agudo, A. & Moreno-Noguer, F., 2021. Stochastic neural radiance fields: quantifying uncertainty in implicit 3D representations,arXiv:2109.02123 [cs].

Shen, J., Agudo, A., Moreno-Noguer, F. & Ruiz, A., 2022. Conditional-flow nerf: Accurate 3D modelling with reliable uncertainty quantification.arXiv:2203.10192 [cs]

Siahkoohi, A., Rizzuti, G., Witte, P.A. & Herrmann, F.J., 2020. Faster uncertainty quantification for inverse problems with conditional normalizing flows,arXiv:2007.07985.

Siahkoohi, A., Rizzuti, G., Orozco, R. & Herrmann, F.J., 2023. Reliable amortized variational inference with physics-based latent distribution correction, *Geophysics,* **88**(3), R297–R322. .

Sitzmann, V., Martel, J.N.P., Bergman, A.W., Lindell, D.B. & Wetzstein, G., 2020. Implicit neural representations with periodic activation functions,arXiv:2006.09661 [cs]

Steeples, D.W. & Miller, R.D., 1988. Seismic reflection methods applied to engineering, environmental, and ground-water problems, ISSN: 2214-4609. in *Geotechnical and Environmental Geophysics: Volume I, Review and Tutorial,* p. cp, European Association of Geoscientists and Engineers.

Strutz, D. & Curtis, A., 2024. Variational Bayesian experimental design for geophysical applications: seismic source location, amplitude versus offset inversion, and estimating CO2 saturations in a subsurface reservoir, *Geophys. J. Int.,* **236**(3), 1309–1331. .

Sun, J., Innanen, K., Zhang, T. & Trad, D., 2023. Implicit seismic full waveform inversion with deep neural representation, *J. geophys. Res.: Solid Earth,* **128**(3), e2022JB025964.

Talarico, E.C.E.S., Grana, D., Passos De Figueiredo, L. & Pesco, S., 2020. Uncertainty quantification in seismic facies inversion, *Geophysics,* **85**(4), M43–M56. .

Tancik, M. *et al.*, 2020. Fourier features let networks learn high frequency functions in low dimensional domains,arXiv:2006.10739 [cs].

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics,* **49**(8), 1259–1266. ,

Tarantola, A. & Valette, B., 1981. Inverse problems = quest for information, *J. Geophys.,* **50**(1), 159–170.

Telford, W.M., Geldart, L.P. & Sheriff, R.E., 1990. *Applied Geophysics,* Cambridge Univ. Press.

Ulrych, T.J., Sacchi, M.D. & Woodbury, A., 2001. A Bayes tour of inversion: a tutorial, *Geophysics,* **66**(1), 55–69. .

Yilmaz, O., 2001. *Seismic Data Analysis,* Society of Exploration Geophysicists.

Zhang, C., Butepage, J., Kjellstrom, H. & Mandt, S., 2018. Advances in variational inference,arXiv:1711.05597 [cs].

Zhang, X., Nawaz, M.A., Zhao, X. & Curtis, A., 2021. Chapter Two - An introduction to variational inference in geophysical inverse problems, in *Advances in Geophysics, Vol. 62 of Inversion of Geophysical Data,* pp. 73–140, ed, Schmelzbach, Cedric, Elsevier.

Zhao, X., Curtis, A. & Zhang, X., 2022. Bayesian seismic tomography using normalizing flows, *Geophys. J. Int.,* **228**(1), 213–239. .

Zoeppritz, K., 1919. Uber reflexion und durchgang seismischer wellen durch unstetigkeitsflächen, *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische Klasse,* 66–84.**1919**