

---

# WISE: full-Waveform variational Inference via Subsurface Extensions

---

Ziyi Yin<sup>1,\*</sup>, Rafael Orozco<sup>1,\*</sup>, Mathias Louboutin<sup>2</sup>, Felix J. Herrmann<sup>1</sup>

<sup>1</sup> Georgia Institute of Technology, <sup>2</sup> Devito Codes Ltd, \* First two authors contributed equally

## Abstract

We introduce a probabilistic technique for full-waveform inversion, employing variational inference and conditional normalizing flows to quantify uncertainty in migration-velocity models and its impact on imaging. Our approach integrates generative artificial intelligence with physics-informed common-image gathers, reducing reliance on accurate initial velocity models. Considered case studies demonstrate its efficacy producing realizations of migration-velocity models conditioned by the data. These models are used to quantify amplitude and positioning effects during subsequent imaging.

## 1 Introduction

Full-waveform inversion (FWI) plays a pivotal role in exploration, primarily focusing on estimating Earth’s subsurface properties from observed seismic data. The inherent complexity of FWI stems from its nonlinearity, further complicated by ill-posedness and computational intensiveness of the wave modeling. To address these challenges, we introduce a computationally cost-effective probabilistic framework that generates multiple migration-velocity models conditioned on observed seismic data. By combining deep learning with physics, our approach harnesses advancements in variational inference [VI, 11] and generative artificial intelligence [AI, 6, 13, 28]. We achieve this by forming common-image gathers (CIGs), followed by training conditional normalizing flows (CNFs) that quantify uncertainties in migration-velocity models.

Our paper is organized as follows. First, we delineate the FWI problem and its inherent challenges. Subsequently, we explore VI to quantify FWI’s uncertainty. To reduce VI’s computational costs, we introduce *physics-informed summary statistics* and justify the use of CIGs as these statistics. Our framework’s capabilities are validated through two case studies, which include studying the effects of uncertainty in the generated migration-velocity models on migration.

## 2 Methodology

We present a Bayesian inference approach to FWI by briefly introducing FWI and VI used as a framework for uncertainty quantification (UQ).

### 2.1 Full-waveform inversion

Estimation of unknown migration-velocity models,  $\mathbf{x}$ , from noisy seismic data,  $\mathbf{y}$  involves inverting nonlinear forward operator,  $\mathcal{F}$ , which links  $\mathbf{x}$  to  $\mathbf{y}$  via  $\mathbf{y} = \mathcal{F}(\mathbf{x}) + \epsilon$  with  $\epsilon$  measurement noise. Source/receiver signatures are assumed known and absorbed into  $\mathcal{F}$ . Solving this nonlinear inverse problem is challenging because of the noise, the non-convexity of the objective function, and the non-trivial null-space of the modeling [37]. As a result, multiple migration-velocity models fit the data, necessitating a Bayesian framework for UQ.

## 2.2 Full-waveform inference

Rather than seeking a single migration-velocity model, our goal is to invert for a range of models compatible with the data, termed “full-waveform inference”. From a Bayesian perspective, this involves determining the posterior distribution of migration-velocity models given the data,  $p(\mathbf{x}|\mathbf{y})$ . We focus on amortized VI, which exchanges the computational cost of posterior sampling for neural network training [4, 26, 29, 31, 44, 45]. Specifically, we employ amortized VI, which incurs offline computational training cost but enables cheap online posterior inference on many datasets  $\mathbf{y}$  [15]. Next, we discuss how to use CNFs for amortized VI.

## 2.3 Amortized variational inference with conditional normalizing flows

During VI, the posterior distribution  $p(\mathbf{x}|\mathbf{y})$  is approximated by the surrogate,  $p_{\theta}(\mathbf{x}|\mathbf{y})$ , with learnable parameters,  $\theta$ . Given the sample pairs  $\{(\mathbf{x}^{(i)}, \mathbf{y}^{(i)})\}_{i=1}^N$ , CNFs are suitable to act as surrogates for the posterior because of their low-cost training and rapid sampling [18, 27]. Their training involves minimization of the Kullback-Leibler divergence between the true and surrogate posterior distribution. In practice, this requires access to  $N$  training pairs of migration-velocity model and observed data to minimize the following objective:

$$\underset{\theta}{\text{minimize}} \quad \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{2} \|f_{\theta}(\mathbf{x}^{(i)}; \mathbf{y}^{(i)})\|_2^2 - \log |\det \mathbf{J}_{f_{\theta}}| \right). \quad (1)$$

Here,  $f_{\theta}$  is the CNF with network parameters,  $\theta$ , and Jacobian,  $\mathbf{J}_{f_{\theta}}$ . It transforms each velocity model,  $\mathbf{x}^{(i)}$ , into white noise (as indicated by the  $\ell_2$ -norm), conditioned on the observation,  $\mathbf{y}^{(i)}$ . After training, the inverse of CNF turns random realizations of the standard Gaussian distribution into posterior samples (migration-velocity models) conditioned on any seismic observation that is in the same statistical distribution as the training data.

## 2.4 Physics-informed summary statistics

While CNFs are capable of approximating the posterior distribution, training the CNFs on pairs  $(\mathbf{x}, \mathbf{y})$  presents challenges when changes in the acquisition occur or when physical principles simplifying the mapping between model and data are lacking, both of which lead to increasing training costs. To tackle these challenges, Radev et al. [25] introduced fixed reduced-size *summary statistics* that encapsulate observed data and inform the posterior distribution. Building on this concept, Orozco et al. [23] uses the gradient as the set of *physics-informed summary statistics*, partially reversing the forward map and therefore accelerating CNF training. For linear inverse problems with Gaussian noise, these statistics are unbiased — maintaining the same posterior distribution, whether conditioned on original shot data or on the gradient. Based on this principle, Siahkoohi et al. [35] used reverse-time migration (RTM), given by the action of the adjoint of the linearized Born modeling, to summarize data and quantify imaging uncertainties for a fixed accurate migration-velocity model.

We aim to extend this approach to the nonlinear FWI problems. While RTM transfers information from the data to the image domain, its performance diminishes for incorrect migration velocities. Hou and Symes [9] showed that least-squares migration can perfectly fit the data for correct migration-velocity models, but this fit fails for inaccurate velocity models. This highlights a fundamental limitation in cases where the velocity model is inaccurate and RTM does not correctly summarize the original shot data, which leads to a biased posterior. For an inaccurate initial FWI-velocity model  $\mathbf{x}_0$ ,  $p(\mathbf{x}|\mathbf{y}) \neq p(\mathbf{x}|\nabla \mathcal{F}(\mathbf{x}_0)^{\top} \mathbf{y})$  with  $\nabla \mathcal{F}$  Born modeling and  $\top$  the adjoint. To avoid this problem, more robust *physics-informed summary statistics* are needed to preserve information.

## 2.5 Common-image gathers as summary statistics

Migration-velocity analysis has a rich history in the literature [36]. Following Hou and Symes [9], we employ relatively artifact-free subsurface-offset extended Born modeling to calculate summary statistics. Because it is closer to an isometry—i.e, the adjoint of extended Born modeling is closer to its inverse [14, 41] and therefore preserves information — its adjoint can nullify residuals even when the FWI-velocity model is incorrect as shown by Hou and Symes [9]. Geng et al. [5] further demonstrate that neural networks can be used to map CIGs to velocity models. Both these findings

shed important light on the role of CIGs during VI because CIGs preserve more information than the gradient, which leads to less biased physics-informed summary statistics when given an inaccurate initial FWI-velocity model. Formally, this means  $p(\mathbf{x}|\mathbf{y}) \approx p(\mathbf{x}|\overline{\nabla\mathcal{F}}(\mathbf{x}_0)^\top \mathbf{y})$ , where  $\overline{\nabla\mathcal{F}}$  is extended Born modeling. Leveraging this mathematical observation, we propose WISE, short for full-Waveform variational Inference via Subsurface Extensions. The core of this technique is to train CNFs with pairs of velocity models,  $\mathbf{x}$ , and CIGs,  $\overline{\nabla\mathcal{F}}(\mathbf{x}_0)^\top \mathbf{y}$ , guided by the objective of Equation 1. Our case studies will demonstrate that even with inaccurate initial FWI-velocity models, CIGs encapsulate more information, enabling the trained CNFs to generate accurate migration-velocity models consistent with the observed shot data.

### 3 Synthetic case studies

Our study evaluates the performance of WISE through synthetic case studies on 2D slices of the Compass dataset [10], known for its “velocity kickback” challenge for FWI algorithms. For a poor initial FWI-velocity model, we aim to compare the quality of posterior samples informed by RTM alone versus those informed by CIGs to verify the superior information content of CIGs. We also illustrate how uncertainty in migration-velocity models can be converted into uncertainties in amplitude and positioning of imaged reflectors.

**Dataset generation and network training.** We take 800 2D slices of the Compass model of 6.4 km by 3.2 km, with 512 equally spaced sources towed at 12.5m depth and 64 ocean-bottom nodes (OBNs) located at jittered sampled horizontal positions [7, 8]. This sampling scheme utilizes compressive sensing techniques to improve acquisition productivity in various situations [21, 38, 39, 43]. The surface is assumed absorbing. Using a 15Hz central frequency Ricker wavelet with energy below 3Hz removed for realism, acoustic data is simulated with Devito [17, 19] and JUDI.jl [40]. Uncorrelated band-limited Gaussian noise is added (S/N 12dB). The arithmetic mean over all velocity models is used as the 1D initial FWI-velocity model (shown in Figure 1(b)). 51 horizontal subsurface offsets ranging from  $-500\text{m}$  to  $+500\text{m}$  are used to compute CIGs (shown in Figure 1(e)). Each offset is input to the network as a separate channel. We use the conditional glow network structure [24] for the CNFs because of its capability to generate superior natural [12] and seismic [18] images.

**Results.** After CNF training, our method’s performance is evaluated on an unseen 2D Compass slice shown in Figure 1(a). When RTM is used to summarize the data, the conditional mean estimate (Figure 1(c)) does not capture the shape of the unconformity. Thanks to the CIGs, WISE captures more information and as a result produces a more accurate conditional mean (Figure 1(d)). For the 50 test samples, the structural similarity index measure (SSIM) with CIGs yields a mean of 0.63, outperforming RTM-based statistics with a mean SSIM of 0.52.

**Quality control.** To verify the inferred migration-velocity model as the conditional mean of the posterior, CIGs calculated for the initial FWI-velocity model (Figure 1(b)), plotted in Figure 1(e), are juxtaposed against CIGs calculated for the inferred migration-velocity model (Figure 1(d)), plotted in Figure 1(f). Significant improvement in near-offset focused energy is observed in the CIGs for the inferred migration-velocity model. A similar focusing behavior is noted for the posterior samples themselves, as shown in the ancillary material.

**Uncertainty quantification and downstream imaging.** While access to the posterior represents an important step towards grasping uncertainty, understanding its impact on imaging with (30Hz) RTMs is more relevant because it concerns uncertainty in the final product. For this purpose, we display the posterior velocity samples in Figure 2(a) and the point-wise standard deviation in Figure 2(b). These deviations increase with depth and correlate with complex geology where the RTM-based inference struggled. To understand how this uncertainty propagates to imaged reflectors, forward uncertainty is assessed by carrying out RTMs for different posterior samples with results shown in Figure 2(c) and the standard deviations plotted in Figure 2(d). These amplitude deviations are different because mapping migration-velocities to RTMs is highly nonlinear, leading to large areas of intense amplitude variation and dimming at the edges caused by the Born modeling’s null-space. While these amplitude sensitivities are useful, deviations in the migration velocities also leads to differences in reflector positioning. Vertical shifts between the envelope of the reference image (central image in Figure 2(c)) and the envelopes of RTMs for different posterior samples are calculated with a local cross-correlation technique and included in Figure 2(d) where blue/red areas correspond to up/down shifts. As expected,

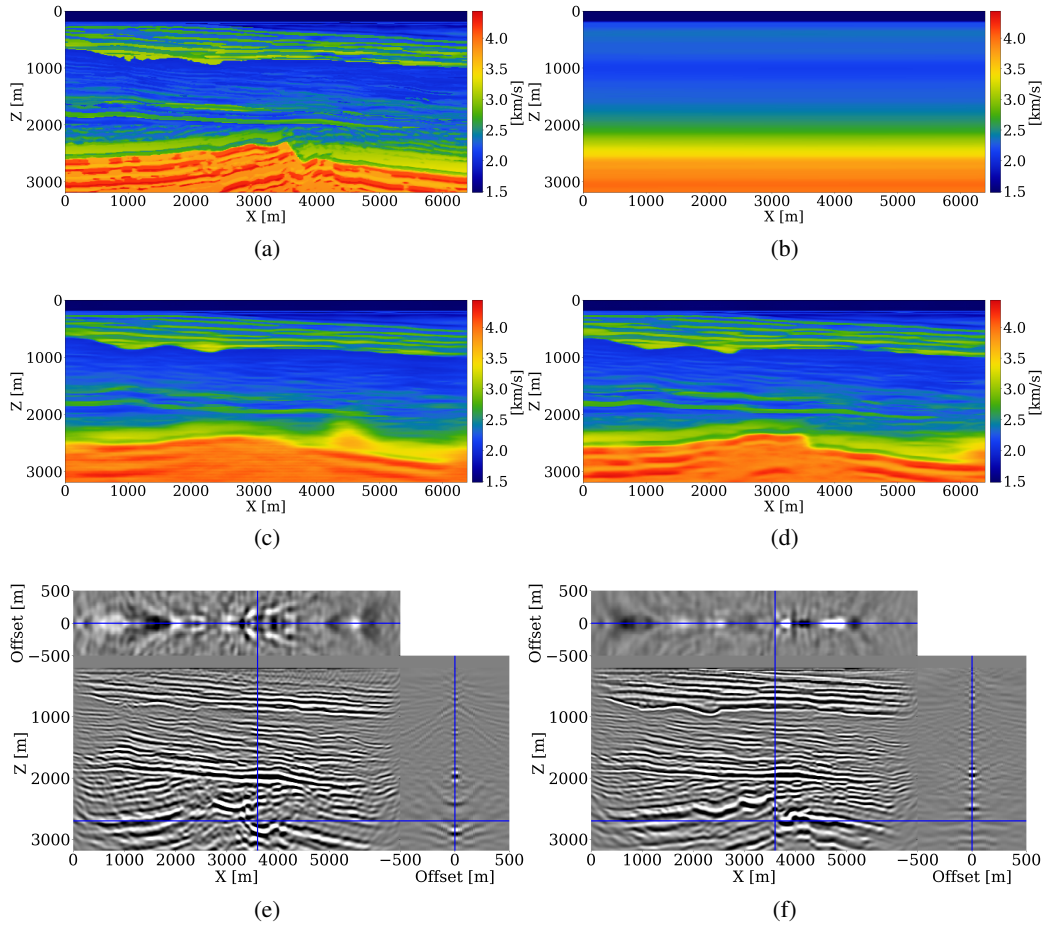


Figure 1: (a) an unseen ground-truth velocity model; (b) 1D initial FWI-velocity model; (c) conditional mean estimate for RTM as summary statistics (SSIM = 0.48); (d) conditional mean estimate from WISE (SSIM = 0.56); (e) CIGs calculated by the initial FWI-velocity model given by (b); (f) CIGs calculated by (d).

these shifts are most notable in the deeper regions and at the edges where velocity variations are the largest.

## 4 Discussion

Once the offline costs of computing 800 CIGs and network training are covered, WISE enables generation of velocity models for unseen seismic data at the low computational cost of a single set of CIGs for a poor initial FWI-velocity model. The Open FWI [2] case study in the [ancillary material](#) demonstrates WISE’s capability to produce realistic posterior samples and conditional means for a broad range of unseen velocity models. In the case of the Compass model, the initial FWI-velocity model was poor. Still, CIGs obtained from a single 1D initial model capture relevant information from the non-zero offsets. From this information, the network learns to produce migration-velocity models that focus CIGs. WISE also produced two types of uncertainty, namely (i) inverse uncertainty in migration-velocity model estimation, which arises from both the non-trivial null-space of FWI and the measurement noise, and (ii) forward uncertainty where uncertainty in migration-velocity models is propagated to uncertainty in amplitude and positioning of imaged reflectors.

Opportunities for future research remain. One area concerns dealing with the “amortization gap” where CNFs tend to maximize performance across multiple datasets rather than excelling at a single observation [20]. While we discovered that training CNFs on a diverse set of samples enhances

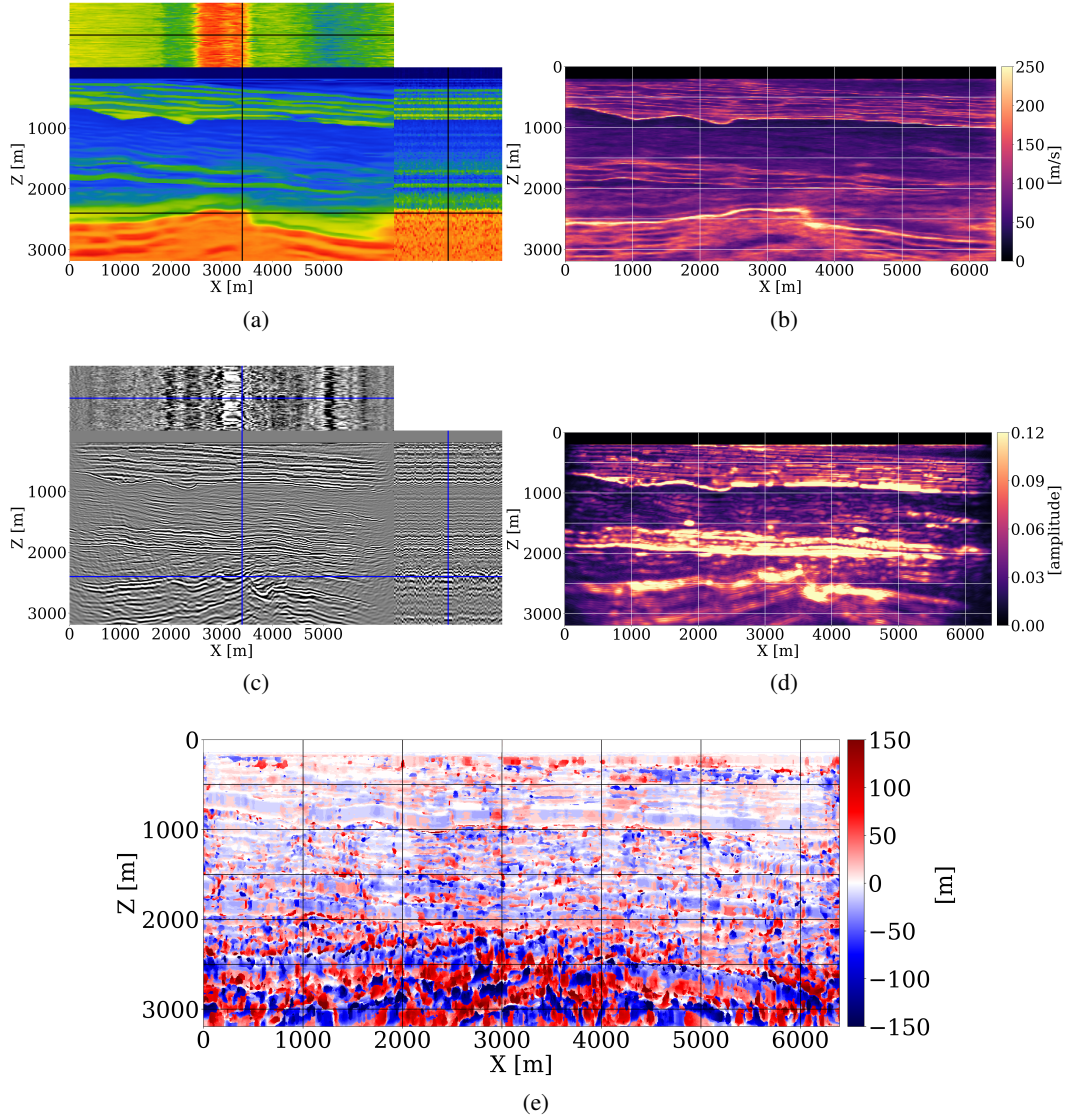


Figure 2: Variability in velocity models and imaged reflectors. (a) Posterior velocity samples from WISE visualized similar to CIGs by plotting the conditional mean (Figure 1(d)) in the central image. Above it shows the posterior sample traces at  $Z = 2.4$  km. On the right shows the traces at  $X = 3.4$  km. (b) Point-wise standard deviation of the posterior velocity samples. (c) Samples of imaged reflectors, where the central image displays imaged reflectors using the conditional mean estimate. The layout of the traces remains the same as (a). (d) Point-wise standard deviation of the imaged reflectors. (e) Point-wise maximum depth shift.

generalization, applying AI techniques to unseen, out-of-distribution samples remains a challenge. However, our WISE framework is compatible with several fine-tuning approaches. To improve single-observation performance, particularly for out-of-distribution samples, computationally more expensive latent space corrections [35] can be employed that incorporate the physics. Recent studies also have indicated that trained CNFs can act as preconditioners or regularizers for physics-based, non-amortized inference [29, 34]. These correction methods can enhance the fit of posterior samples to observed data, as shown in Siahkoohi et al. [35], or enable the generation of more focused CIGs through migration velocity analysis. Moreover, velocity continuation methods [3] could be used including recent advances in neural operators [33]. These could offset the cost of running RTMs for each posterior sample, thus accelerating forward uncertainty propagation. While we observed that providing more offsets can enhance the quality of the inference, we recognize the resulting increase in CIG computation costs and CNF memory consumption. This necessitates cost-effective frameworks for determining optimal offset numbers or sampling strategies for CIGs. In this context, recent work on using CNFs for Bayesian optimal experimental design [22] seamlessly integrates as an advancement to the WISE framework. Considering low-rank approximations of CIGs [41] may reduce computational demands. Additionally, exploring other conditional generative models like diffusion models [1] may be worthwhile. Our case studies have yet to account for inverse uncertainty due to modeling errors, such as attenuation effects, multiples, or residual shear wave energy, which could be addressed through Bayesian model misspecification techniques [30]. Recent advances suggest that transfer learning could correct these modeling inaccuracies [32, 42], a solution that our approach is amenable to.

Incurring computational cost on an NVIDIA A100 GPU can be broken down as follows: generating training pairs requires generation of 64 OBN datasets and corresponding CIGs for 800 models, totaling approximately 80 hours of runtime. After generating the training set, training the CNF takes around 16 hours. With these initial runtime investments, the cost for a single inference involves only a single CIG computation, which takes about 6 minutes. For context, running a single FWI starting from the velocity model included in Figure 1(b) requires 12.5 data passes taking roughly 50 minutes to complete (the final result is shown in the ancillary material). Traditional UQ methods require the compute equivalent to hundreds of FWI runs [44], but here we estimate at least 50 FWI runs. Based on these numbers, the computational savings from employing CNF surrogates offset the upfront costs after inference on approximately 3 datasets. We emphasize that as long as the statistics of the underlying geology remains similar, our amortized network can be applied to different observed datasets in the complete basin without retraining. Furthermore, the parallel execution of training pair generation on clusters can significantly reduce initial computational time. Although our study primarily demonstrates a proof of concept on a realistic 2D experiment, the WISE software tool chain is designed for large-scale 3D problems. CNFs, favored for their memory efficiency through invertibility [24], are well-suited for 3D problems. In addition, memory consumption of CIG computation can be reduced significantly with random trace estimation techniques [16]. Since our work requires training samples of Earth models, we envision these samples coming from legacy proxy models and future work will explore automatic workflows for generating these from field observations.

## 5 Conclusions

We present WISE, full-Waveform variational Inference via Subsurface Extensions, for computationally efficient uncertainty quantification of FWI. This framework underscores the potential of generative AI to address FWI challenges, paving the way for a new seismic inversion and imaging paradigm that is uncertainty-aware. By having common-image gathers act as information-preserving summary statistics, a principled approach to UQ is achieved where generative AI is successfully combined with wave physics. Because WISE automatically produces distributions for migration-velocity models conditioned by the data, it moves well beyond traditional velocity model building. It was shown that this distributional information can be employed to quantify uncertainties in the migration-velocity models that can be used to better understand amplitude and positioning uncertainty in migration.

## 6 Acknowledgements

This research was carried out with the support of Georgia Research Alliance and partners of the ML4Seismic Center. The authors would like to thank Charles Jones (Osokey) for the constructive discussion.

## 7 Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT to improve readability and language. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

## References

- [1] Baldassari, L., A. Siahkoohi, J. Garnier, K. Solna, and M. V. de Hoop, 2024, Conditional score-based diffusion models for bayesian inference in infinite dimensions: *Advances in Neural Information Processing Systems*, **36**.
- [2] Deng, C., S. Feng, H. Wang, X. Zhang, P. Jin, Y. Feng, Q. Zeng, Y. Chen, and Y. Lin, 2022, Openfwi: Large-scale multi-structural benchmark datasets for full waveform inversion: *Advances in Neural Information Processing Systems*, **35**, 6007–6020.
- [3] Fomel, S., 2003, Time-migration velocity analysis by velocity continuation: *Geophysics*, **68**, 1662–1672.
- [4] Gahlot, A. P., H. T. Erdinc, R. Orozco, Z. Yin, and F. J. Herrmann, 2023, Inference of CO2 flow patterns—a feasibility study: *NeurIPS 2023 Workshop - Tackling Climate Change with Machine Learning*.
- [5] Geng, Z., Z. Zhao, Y. Shi, X. Wu, S. Fomel, and M. Sen, 2022, Deep learning for velocity model building with common-image gather volumes: *Geophysical Journal International*, **228**, 1054–1070.
- [6] Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, 2014, Generative adversarial nets: *Advances in neural information processing systems*, **27**.
- [7] Hennenfent, G., and F. J. Herrmann, 2008, Simply denoise: Wavefield reconstruction via jittered undersampling: *Geophysics*, **73**, V19–V28.
- [8] Herrmann, F. J., 2010, Randomized sampling and sparsity: Getting more information from fewer samples: *Geophysics*, **75**, WB173–WB187.
- [9] Hou, J., and W. W. Symes, 2016, Accelerating extended least-squares migration with weighted conjugate gradient iteration: *Geophysics*, **81**, S165–S179.
- [10] Jones, C., J. Edgar, J. Selvage, and H. Crook, 2012, Building complex synthetic models to evaluate acquisition geometries and velocity inversion technologies: *74th EAGE Conference and Exhibition incorporating EUROPEC 2012, European Association of Geoscientists & Engineers*, cp–293.
- [11] Jordan, M. I., Z. Ghahramani, T. S. Jaakkola, and L. K. Saul, 1999, An introduction to variational methods for graphical models: *Machine learning*, **37**, 183–233.
- [12] Kingma, D. P., and P. Dhariwal, 2018, Glow: Generative flow with invertible 1x1 convolutions: *Advances in neural information processing systems*, **31**.
- [13] Kingma, D. P., and M. Welling, 2013, Auto-encoding variational bayes: *arXiv preprint arXiv:1312.6114*.
- [14] Kroode, F. t., 2023, An omnidirectional seismic image extension: *Inverse Problems*, **39**, 035003.
- [15] Kruse, J., G. Detommaso, U. Köthe, and R. Scheichl, 2021, Hint: Hierarchical invertible neural transport for density estimation and bayesian inference: *Proceedings of the AAAI Conference on Artificial Intelligence*, 8191–8199.
- [16] Louboutin, M., and F. J. Herrmann, 2024, Wave-based inversion at scale on graphical processing units with randomized trace estimation: *Geophysical Prospecting*, **72**, 353–366.
- [17] Louboutin, M., M. Lange, F. Luporini, N. Kukreja, P. A. Witte, F. J. Herrmann, P. Velesko, and G. J. Gorman, 2019, Devito (v3.1.0): an embedded domain-specific language for finite differences and geophysical exploration: *Geoscientific Model Development*, **12**, 1165–1187.

- [18] Louboutin, M., Z. Yin, R. Orozco, T. J. Grady, A. Siahkoochi, G. Rizzuti, P. A. Witte, O. Møyner, G. J. Gorman, and F. J. Herrmann, 2023, Learned multiphysics inversion with differentiable programming and machine learning: *The Leading Edge*, **42**, 474–486.
- [19] Luporini, F., M. Louboutin, M. Lange, N. Kukreja, P. Witte, J. Hückelheim, C. Yount, P. H. J. Kelly, F. J. Herrmann, and G. J. Gorman, 2020, Architecture and performance of devito, a system for automated stencil computation: *ACM Trans. Math. Softw.*, **46**.
- [20] Marino, J., Y. Yue, and S. Mandt, 2018, Iterative amortized inference: *International Conference on Machine Learning*, PMLR, 3403–3412.
- [21] Oghenekohwo, F., H. Wason, E. Esser, and F. J. Herrmann, 2017, Low-cost time-lapse seismic with distributed compressive sensing—part 1: Exploiting common information among the vintages: *Geophysics*, **82**, P1–P13.
- [22] Orozco, R., F. J. Herrmann, and P. Chen, 2024, Probabilistic bayesian optimal experimental design using conditional normalizing flows: *arXiv preprint arXiv:2402.18337*.
- [23] Orozco, R., A. Siahkoochi, G. Rizzuti, T. van Leeuwen, and F. J. Herrmann, 2023a, Adjoint operators enable fast and amortized machine learning based bayesian uncertainty quantification: *Medical Imaging 2023: Image Processing*, SPIE, 357–367.
- [24] Orozco, R., P. Witte, M. Louboutin, A. Siahkoochi, G. Rizzuti, B. Peters, and F. J. Herrmann, 2023b, Invertiblenetworks. jl: A julia package for scalable normalizing flows: *arXiv preprint arXiv:2312.13480*.
- [25] Radev, S. T., U. K. Mertens, A. Voss, L. Ardizzone, and U. Kothe, 2022, Bayesflow: Learning complex stochastic models with invertible neural networks: *IEEE Transactions on Neural Networks and Learning Systems*, **33**, 1452–1466.
- [26] Ren, Y., P. A. Witte, A. Siahkoochi, M. Louboutin, Z. Yin, and F. J. Herrmann, 2021, Seismic velocity inversion and uncertainty quantification using conditional normalizing flows: Presented at the AGU Fall Meeting 2021, AGU.
- [27] Rezende, D., and S. Mohamed, 2015, Variational inference with normalizing flows: *International conference on machine learning*, PMLR, 1530–1538.
- [28] Rezende, D. J., S. Mohamed, and D. Wierstra, 2014, Stochastic backpropagation and approximate inference in deep generative models: *International conference on machine learning*, PMLR, 1278–1286.
- [29] Rizzuti, G., A. Siahkoochi, P. A. Witte, and F. J. Herrmann, 2020, Parameterizing uncertainty by deep invertible networks: An application to reservoir characterization: *SEG Technical Program Expanded Abstracts 2020*.
- [30] Schmitt, M., P.-C. Bürkner, U. Köthe, and S. T. Radev, 2023, Detecting model misspecification in amortized bayesian inference with neural networks: *DAGM German Conference on Pattern Recognition*, Springer, 541–557.
- [31] Siahkoochi, A., and F. J. Herrmann, 2021, Learning by example: Fast reliability-aware seismic imaging with normalizing flows: *First International Meeting for Applied Geoscience & Energy Expanded Abstracts*.
- [32] Siahkoochi, A., M. Louboutin, and F. J. Herrmann, 2019, The importance of transfer learning in seismic modeling and imaging: *Geophysics*, **84**, A47–A52.
- [33] ———, 2022, Velocity continuation with fourier neural operators for accelerated uncertainty quantification: *SEG International Exposition and Annual Meeting*, SEG, D011S092R004.
- [34] Siahkoochi, A., G. Rizzuti, M. Louboutin, P. Witte, and F. Herrmann, 2021, Preconditioned training of normalizing flows for variational inference in inverse problems: Presented at the Third Symposium on Advances in Approximate Bayesian Inference.
- [35] Siahkoochi, A., G. Rizzuti, R. Orozco, and F. J. Herrmann, 2023, Reliable amortized variational inference with physics-based latent distribution correction: *Geophysics*, **88**, R297–R322.
- [36] Symes, W. W., 2008, Migration velocity analysis and waveform inversion: *Geophysical prospecting*, **56**, 765–790.
- [37] Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266.
- [38] Wason, H., and F. J. Herrmann, 2013, Time-jittered ocean bottom seismic acquisition: *SEG International Exposition and Annual Meeting*, SEG, SEG–2013.
- [39] Wason, H., F. Oghenekohwo, and F. J. Herrmann, 2017, Low-cost time-lapse seismic with distributed compressive sensing—part 2: Impact on repeatability: *Geophysics*, **82**, P15–P30.
- [40] Witte, P. A., M. Louboutin, N. Kukreja, F. Luporini, M. Lange, G. J. Gorman, and F. J. Herrmann, 2019, A large-scale framework for symbolic implementations of seismic inversion algorithms in julia: *Geophysics*, **84**, F57–F71.



- [41] Yang, M., M. Graff, R. Kumar, and F. J. Herrmann, 2021, Low-rank representation of omnidirectional subsurface extended image volumes: *Geophysics*, **86**, S165–S183.
- [42] Yao, J., L. Guasch, and M. Warner, 2023, Neural networks as a tool for domain translation of geophysical data: *Geophysics*, **88**, V267–V275.
- [43] Yin, Z., H. T. Erdinc, A. P. Gahlot, M. Louboutin, and F. J. Herrmann, 2023, Derisking geologic carbon storage from high-resolution time-lapse seismic to explainable leakage detection: *The Leading Edge*, **42**, 69–76.
- [44] Zhang, X., A. Lomas, M. Zhou, Y. Zheng, and A. Curtis, 2023, 3-d bayesian variational full waveform inversion: *Geophysical Journal International*, **234**, 546–561.
- [45] Zhang, X., M. A. Nawaz, X. Zhao, and A. Curtis, 2021, An introduction to variational inference in geophysical inverse problems, *in* *Advances in Geophysics*: Elsevier, **62**, 73–140.