

Supplementary Materials

November 26, 2024

1 Validation on real clinical data

To validate the alignment between OpenWaves data and real-world data, we used a BFNO model trained on the OpenWaves dataset as a forward surrogate to reconstruct a public clinical dataset of human breast USCT [Ali et al. \(2024\)](#). The clinical dataset was acquired at the Karmanos Cancer Institute (KCI) under IRB No. 040912M1F.

The USCT instrument for data collection employed a 22 cm-diameter ring transducer array with 1024 elements and a pulse center frequency of 2.5 MHz. The system recorded time-series channel data from all 1024 receivers for each individual emitter on the ring, producing a full 1024×1024 matrix. The received channel data was windowed for each transmission, and the discrete-time Fourier transform (DTFT) was applied to isolate the frequencies used in waveform inversion. Figure 1 presents the high-resolution reconstruction results of the dataset using a 2D FDFD solver. This result, reported in the original paper ([Ali et al., 2024](#)), was obtained with data from 20 frequencies and 1024 sensors, and is reasonably regarded as the ground truth for this experiment.

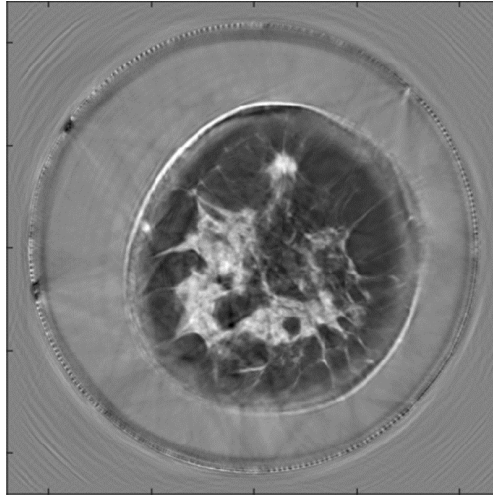


Figure 1: **Clinical breast phantom ground truth.** The figure shows the reconstruction results of a breast phantom using FDFD solver with data from 20 frequencies in the range of 0.3 MHz to 1.25 MHz, as presented in [Ali et al. \(2024\)](#). This reconstruction is considered the ground truth.

To balance cost and time constraints, we selected two forward models and one direct inversion model as representatives to reconstruct clinical data. As shown in Fig. 2, the BFNO model successfully reconstructed a clear malignant tumor together with other internal tissue structures, including skin, fat, and glands. These results demonstrate a strong correspondence between the anatomical structures in the OpenWaves dataset and real breast tissues. In contrast, the FNO model failed to capture many high-frequency details, indicating its limitations in learning high-frequency features. Moreover, the direct inversion model, which utilized only three-frequency data, was unable to reconstruct the correct phantom structures, underscoring the lack of generalization capability in direct inversion methods compared to FWI-based approaches.

We attribute this difference to two factors: First, forward models leverage adjoint updates that integrate physical principles, thereby preserving generalization capabilities. Second, while forward

models focus on modeling $p(y|x)$, direct inversion methods require learning $p(x|y) \propto p(y|x)p(x)$, which involves the simultaneous modeling of two complex distributions and is more susceptible to the influence of the data distribution $p(x)$.

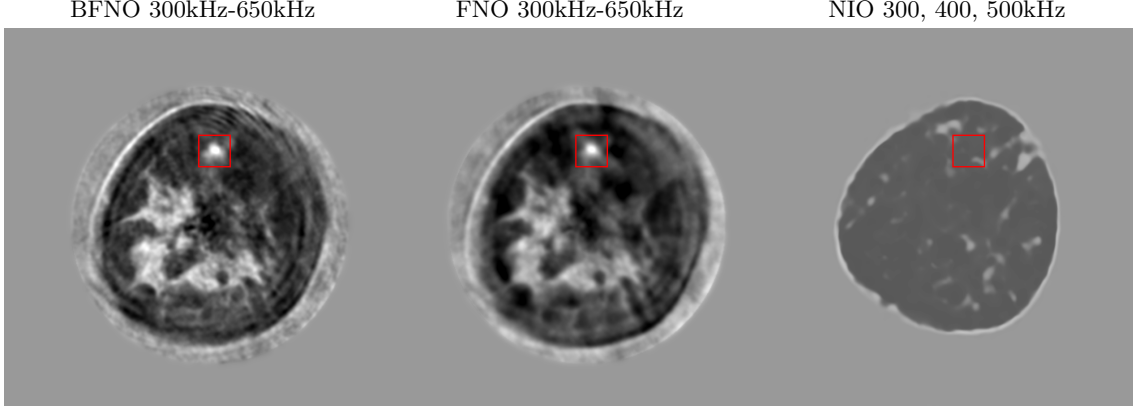


Figure 2: **Ablation study on clinical data reconstruction from different models.** The figure presents the reconstruction results on the clinical USCT dataset of a human breast (Ali et al., 2024), using forward models and a direct inversion model trained on OpenWaves. (a): BFNO with 8 frequencies and 256 transmitters; (b): FNO with 8 frequencies and 256 transmitters; (c): NIO inversion. The malignant tumor is highlighted in the red box.

Next, we conducted an ablation study using the BFNO forward model to evaluate the impact of different data configurations. As shown in Figure 3, our findings reveal that reducing the number of sensors significantly degrades the quality of reconstructed images. Specifically, when only 64 sensors are used, the tissue details become noticeably blurred. Similarly, reducing the number of frequencies results in the loss of fine details and introduces artifacts. These observations align with the well-established behavior of traditional FWI reconstruction methods.

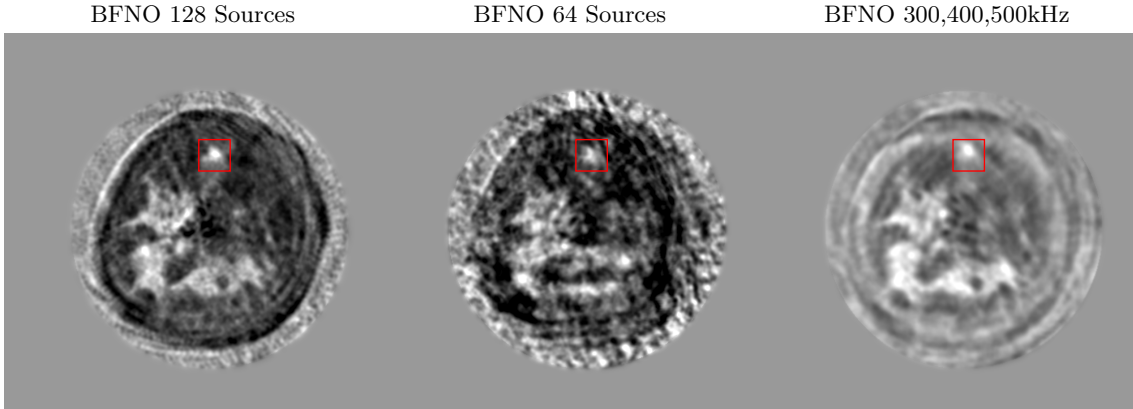


Figure 3: **Ablation study on clinical data reconstruction from different experimental settings.** The figure presents the reconstruction results on the clinical USCT dataset of a human breast (Ali et al., 2024) using BFNO. Different observational data settings were employed to investigate the impact of frequency count and source number on the inversion. (a): 8 frequencies and 128 transmitters; (b): 8 frequencies and 64 transmitters; (c): 3 frequencies and 128 transmitters. The malignant tumor is highlighted in the red box.

References

Rehman Ali, Trevor M. Mitcham, Thurston Brevett, Òscar Calderón Agudo, Cristina Durán Martínez, Cuiping Li, Marvin M. Dooley, and Nebojsa Duric. 2-d slicewise waveform inversion of sound speed and acoustic attenuation for ring array ultrasound tomography based on a block lu solver. *IEEE Transactions on Medical Imaging*, pp. 1–1, 2024. doi: 10.1109/TMI.2024.3383816.