Optimizing computational efficiency and image quality in clinical breast ultrasound tomography through data downsampling

Victoria Nketia[1*], Trevor Mitcham[2], Maria Camila Gonzalez Garcia[2], Rehman Ali[2], Israel Owolabi[1], Melanie Singh[3], Neb Duric[1,2,3]

[1]University of Rochester, Department of Electrical and Computer Engineering, Rochester, NY, USA.

[2]University of Rochester, Department of Imaging Sciences, Rochester, NY, USA.

[3]University of Rochester, Department of Biomedical Engineering, Rochester, NY, USA.

ABSTRACT

Ultrasound tomography (UST) offers quantitative, high-resolution breast imaging, but widespread clinical adoption remains limited by long acquisition times, large raw data volumes, and high computational demands for image reconstruction. Some current systems employ a 1024-element ring transducer array, producing more than 2 GB of raw data per slice and nearly 200 GB for a full breast volume, resulting in considerable memory and storage requirements, hardware, and processing burdens. This study systematically investigates downsampling as a practical strategy to reduce data size, computational load, and reconstruction time while preserving diagnostic image quality. Using clinical UST data from a ring-array system, raw data were downsampled from 1024×1024 to 512×512 and 256×256, with additional transmit-only reductions applied by factors of 4, 8, 16, and 32. Images were reconstructed using a Full Waveform Inversion (FWI) algorithm, and the quality of the reconstructed images was evaluated across consistent anatomical Regions of Interest (ROIs). Quantitative metrics included Signal-to-Noise Ratio (SNR), Contrast-to-Noise Ratio (CNR), and Root Mean Square Error (RMSE). Results show that downsampling reduces reconstruction time by up to a factor of three while decreasing memory size. Although tumor SNR decreased with extreme downsampling, CNR remained largely comparable across many configurations, indicating preserved tumor detectability. The 64×256 configuration provided an optimal balance, offering strong CNR performance with a reconstruction 3 times faster compared to the 3.0 minutes of the full dataset. These findings demonstrate that strategic downsampling enables substantial efficiency gains with minimal loss of clinically relevant image quality, supporting the development of lower-cost, UST systems suitable for clinical use.

Keywords: Ultrasound tomography, downsampling, image quality, breast imaging

1. INTRODUCTION

Ultrasound tomography (UST) has emerged as a compelling three-dimensional breast imaging modality capable of producing quantitative maps of sound speed and acoustic attenuation, both of which provide clinically meaningful information about tissue composition. Quantitative sound speed measurements derived from UST have shown strong correlations with volumetric breast density, an established biomarker of breast cancer risk, demonstrating the modality's diagnostic and predictive potential[1], [2]. UST further offers whole-breast coverage, operator independence, and improved reproducibility compared to hand-held ultrasound, supporting its suitability for large-scale clinical deployment[3].

*vnketia@ur.rochester.edu

Medical Imaging 2026: Ultrasonic Imaging and Tomography, edited by Christian Boehm, Mohammad Mehrmohammadi, Shawn Liangzhong Xiang, Proc. of SPIE Vol. 13931, 139310L · 2026 Published by SPIE
1605-7422 · doi: 10.1117/12.3082370

Despite these advantages, widespread clinical adoption of UST remains constrained by significant data-handling and computational challenges. For systems that employ ring-shaped arrays with 1024 transmit (Tx) and 1024 receive (Rx) elements, generating a full dataset takes considerable storage and acquisition time.

In the most extreme case, or every transmit event, all 1024 receivers record a complete acoustic waveform, which is digitized into time samples to capture the entire acoustic propagation path through the breast.

As a result, each raw UST slice is stored as a three-dimensional dataset indexed by Time × Tx × Rx. Multiplying these dimensions can result in as much as 2 GB of data per slice.

A typical scan containing 80 to 100 slices, therefore, produces on the order of 160 to 200 GB of raw data per patient. This massive data volume imposes substantial demands on memory, storage bandwidth, acquisition time, and data transfer. Furthermore, advanced reconstruction algorithms such as Full Waveform Inversion (FWI) add additional computational burden due to their iterative and model-intensive nature[4], [5].

Numerous studies have attempted to reduce these computational demands. In earlier work, Duric et al. (2007) explored the use of sparse transducer arrays in ultrasound tomography through the development of the Computed Ultrasound Risk Evaluation (CURE) prototype. Their findings demonstrated that even with fewer transducer elements, it was possible to detect breast tumors[6]. Source-encoding strategies have been proposed to accelerate waveform inversion [7], while ray-based and refraction-corrected methods have been developed to improve efficiency in large-scale 3D reconstructions [3]. Hybrid transmission and reflection waveform tomography methods have also been shown to improve image quality by incorporating more information from the recorded wavefield [8]. More recently, data-driven and machine-learning-assisted reconstructions have been explored to reduce runtime and improve convergence in UST and related modalities[9], illustrating an increasing interest in computational efficiency across the field.

Despite these developments, the fundamental bottleneck remains the sheer size of the raw data, which scales quadratically with the number of transducer elements. System design studies emphasize that acquisition speed, memory constraints, and computational cost are among the main obstacles to bringing UST into routine clinical use. By reducing the number of Tx and Rx elements, a technique known as downsampling offers a straightforward strategy to mitigate these limitations. Because the total raw data size scales proportionally with the number of transmit-receive (Tx-Rx) pairs, reducing the array from 1024 × 1024 to 512 × 512 decreases the number of recorded waveforms by a factor of four, lowering the raw data per slice from roughly 2 GB to about 0.5 GB. A further reduction to 256 × 256 yields a sixteen times decrease, reducing each slice to approximately 0.125 GB. These reductions directly reduce memory requirements and shorten data-loading and reconstruction time, making downsampling a smart approach for improving the cost of clinical UST systems. However, systematic studies of how downsampling affects image quality, lesion detectability, and reconstruction accuracy in clinical UST data remain limited.

The purpose of this study is to address this gap by systematically evaluating transducer-element and transmit-element downsampling in clinical breast UST using FWI reconstruction. All data processing and analyses were performed in MATLAB[10]. Image quality was assessed using the Signal-to-Noise Ratio (SNR), Contrast-to-Noise Ratio (CNR), and Root Mean Square Error (RMSE), which enabled a comprehensive characterization of the trade-offs between computational efficiency and diagnostic value. By identifying downsampling strategies that preserve clinically relevant image quality while substantially reducing data volume and reconstruction time, this work contributes to the development of more efficient, cost-effective, and scalable UST systems aligned with future breast imaging [7]

2. METHODOLOGY

This study systematically evaluated how different downsampling strategies affect raw data volume, computational efficiency, and reconstructed image quality in clinical ultrasound tomography (UST). All data processing, downsampling operations, and quantitative analyses were conducted in MATLAB [10]. Image reconstruction for all configurations employed a Full Waveform Inversion (FWI) algorithm, an iterative physics-based method known for producing quantitative sound speed estimates in soft tissue [1], [5]. Reconstruction settings, including iteration count, regularization, and initial model, were held constant to ensure that all observed differences arose solely from the degree of data reduction.

2.1 Clinical data downsampling

Downsampling was performed to evaluate how reducing the number of transducer elements or transmit events influences raw data size, reconstruction efficiency, and diagnostic image quality in ultrasound tomography (UST). All preprocessing, downsampling procedures, and analyses were carried out in MATLAB [10]. Image reconstruction for every configuration employed a Full Waveform Inversion (FWI) algorithm, with all reconstruction parameters held constant to ensure that differences arose solely from the level of data reduction.

2.2 Computational platform and time measurements

All reconstruction time benchmarks were performed on a dedicated Dell workstation to ensure consistent and reproducible performance measurements. The system was equipped with an Intel® Xeon® Gold 6226R CPU operating at 2.90 GHz, with 16 CPU cores, and 96 GB of system memory, running a 64-bit Linux operating system. No GPU acceleration was used for this study; all Full Waveform Inversion (FWI) reconstructions were executed entirely on the CPU. Reconstruction times reported in this work correspond to the total wall-clock time required to reconstruct a single slice for each downsampling configuration. While absolute reconstruction times may vary with hardware configuration, the relative trends observed across downsampling levels are independent of the computational platform.

2.3 Transducer-Element downsampling (Tx-Rx downsampling)

Transducer-element downsampling refers to the uniform reduction of both transmit (Tx) and receive (Rx) channels. The full UST dataset contains 1024 transmitters and 1024 receivers, forming a 1024 × 1024 acquisition matrix. For each transmit event, all receivers record 2112-time samples, giving a total raw dataset size of:

$$N_{full\ data} = 2112 \times 1024 \times 1024 = 2.21 \times 10^9 \text{ samples per slice.}$$

With the system's effective sample packing, this corresponds to approximately 2 GB per slice. To evaluate reduced-aperture system designs, the array was uniformly downsampled to 512 × 512 (factor of 2) and 256 × 256 (factor of 4). Since the number of recorded samples is proportional to the number of Tx-Rx pairs, the downsampled datasets scale as:

$N_{512} = 2112 \times 512 \times 512\,,\ N_{256} = 2112\ \times 56 \times 256$

To express these relative to the full dataset:

$$\frac{N_{512}}{N_{full\ data}} = \frac{512\ \times 512}{1024 \times 1024} = \frac{1}{4},$$

$$\frac{N_{256}}{N_{full\ data}} = \frac{256\ \times 256}{1024 \times 1024} = \frac{1}{16}$$

Thus, the 512 × 512 configuration contains ¼ of the full dataset, and the 256 × 256 configuration contains 1/16. With effective storage formatting, these correspond to approximately 0.5 GB and 0.125 GB per slice, respectively. These reductions directly decrease memory usage, data-loading time, and FWI computational cost. Because FWI requires simulating a wavefield for each transmitter, reducing both Tx and Rx elements lowers the number of wavefields simulated per iteration and reduces the computational grid size. However, reducing array aperture also reduces spatial sampling density, which may influence resolution and lesion detectability. Evaluating this trade-off is one of the goals of this study.

2.4 Transmit-Only downsampling

Transmit-only downsampling reduces the number of transmit events while keeping the full receive aperture intact. This strategy is clinically relevant because acquisition time is directly proportional to the number of transmitters, while full Rx sampling is preserved to maintain spatial information.

For the full dataset, transmit-only downsampling factors of 8 and 32 were applied, producing 128 × 1024 and 32 × 1024 datasets. The total number of samples follows:

$$N = 2112 \times N_{Tx} \times N_{RX}$$

Thus, relative to the full dataset:

$$N_{128\times1024} = \frac{1}{8} N_{full\ data}, \quad N_{32\times1024} = \frac{1}{32} N_{full\ data}$$

Because each transmit event produces a complete set of receive waveforms, reducing the number of transmitters proportionally reduces the number of wavefield simulations required during FWI, thereby reducing reconstruction time.

Transmit-only downsampling was also applied to the already downsampled Tx-Rx configurations (512 × 512 and 256 × 256). This generated hybrid configurations such as 128 × 512, 64 × 512, 32 × 512, 64 × 256, and 32 × 256. These configurations emulate practical system designs that minimize hardware cost and acquisition time while maintaining sufficient spatial sampling for tomographic imaging. Across all configurations, the full dataset configuration 1024 × 1024 reconstruction was used as the reference for quantitative comparisons.

2.5 Image reconstruction

All datasets (full and downsampled) were reconstructed using a consistent Full Waveform Inversion (FWI) framework. FWI iteratively updates the spatial distribution of sound speed by minimizing the difference between measured acoustic wavefields and those simulated through the current forward model. For each transmit event, a forward simulation and an adjoint simulation are required, making reconstruction cost directly proportional to the number of transmitters and the spatial extent of the computational grid.

Because Tx-Rx downsampling reduces both the number of transmit events and the number of recorded receiver channels, it decreases the number of forward/adjoint simulations and lowers memory requirements for storing intermediate wavefields. Similarly, transmit-only downsampling reduces the number of forward simulations without altering the receive aperture, resulting in shorter iteration times but maintaining dense Rx sampling. Thus, downsampling provides a direct and predictable reduction in FWI computational burden. To ensure fair comparison across configurations, all reconstructions used the same number of iterations, identical model parameterization, matched boundary conditions, and consistent regularization settings.

Reconstruction time was recorded for each dataset to quantify the computational efficiency gained through data reduction. The full-resolution 1024 × 1024 reconstruction was used as the reference for image quality and computational comparisons.

2.6 Quantitative image quality assessment and trade-off determination

Quantitative assessment of image quality was performed to evaluate how different downsampling strategies influenced tumor detectability, tissue uniformity, system noise performance, and overall reconstruction fidelity. Four anatomically meaningful Regions of Interest (ROIs) were defined consistently across all images, as shown in Figure 1.

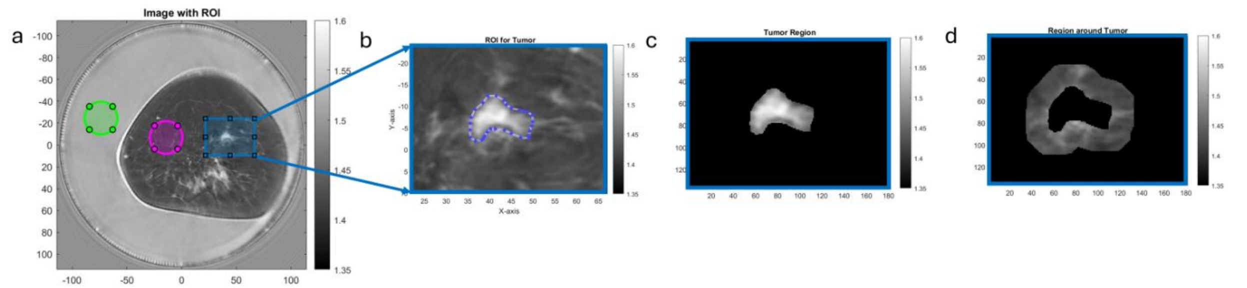

Figure 1: Regions of interest (ROIs) used for quantitative image quality assessment. (a) Reconstructed image showing ROIs for the region surrounding the tumor(blue), background breast tissue (pink), and water region (green). (b) Zoomed view of the tumor area and ROI. (c) Masked tumor ROI (d) Masked region around the tumor

These ROIs were manually verified for consistency and extracted using MATLAB to ensure that all quantitative comparisons were performed using identical anatomical locations across reconstruction configurations.

To comprehensively characterize image quality, three families of metrics were computed: Signal-to-Noise Ratio (SNR), Contrast-to-Noise Ratio (CNR), and Root Mean Square Error (RMSE). Together, these metrics provided complementary information about signal clarity, contrast, uniformity, and reconstruction accuracy.

2.6.1 Signal-to-Noise Ratio (SNR)

SNR was evaluated in three distinct regions to capture different aspects of noise behavior:

1. Tumor SNR quantified how clearly the tumor stands out against its immediate surroundings. The signal was defined as the mean intensity of the tumor ROI (μ_{tumor}), while noise was defined as the standard deviation in the region surrounding the tumor ($\sigma_{region\ around\ tumor}$):

$$SNR_{tumor} = \frac{\mu_{tumor}}{\sigma_{region\ around\ tumor}} \tag{1}$$

2. Water SNR evaluated system noise and background uniformity in the homogeneous water bath. The signal component was defined as the mean pixel intensity within the water region itself, μ_{water}, identified by the green circular ROI in Figure 1(a). The noise component was defined as the standard deviation of pixel intensities within the same water region,σ_{water}.

$$SNR_{water} = \frac{\mu_{water}}{\sigma_{water}} \tag{2}$$

3. Background SNR assessed noise performance within the breast parenchyma. The signal component was defined as the mean pixel intensity within the background breast tissue region, $\mu_{background}$, as indicated by the pink circular ROI in Figure 1(a). The noise component was defined as the standard deviation of pixel intensities within this same anatomical background tissue region, $\sigma_{background}$. This metric characterizes tissue variability and reflects anatomical noise that influences interpretation.

$$SNR_{background} = \frac{\mu_{background}}{\sigma_{background}} \tag{3}$$

Collectively, the three SNR metrics provided a detailed assessment of local tumor detectability, global system noise, and background tissue consistency.

2.6.2 Contrast-to-Noise Ratio (CNR)

CNR was computed to quantify tumor conspicuity relative to surrounding tissue. It was defined using the absolute mean intensity difference between the tumor and the region surrounding the tumor, normalized by the combined noise from both areas. CNR is a clinically meaningful metric because it directly reflects how distinguishable a tumor is from the surrounding parenchyma.

$$CNR = \frac{|\mu_{tumor} - \mu_{region\ around\ tumor}|}{\sqrt{\sigma_{tumor}^2 + \sigma_{region\ around\ tumor}^2}} \tag{4}$$

$$\sigma_{tumor}: standard\ deviation\ of\ the\ tumor$$

$$\sigma_{region\ around\ tumor}: standard\ deviation\ of\ region\ around\ tumor\ [11]$$

2.6.3 Root mean square error (RMSE)

To evaluate reconstruction fidelity, RMSE was computed between each downsampled reconstruction and the full-resolution reference image (1024 × 1024). Pixel-wise intensity differences were calculated as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(I_{reference} - I_{downsampled}\right)^2} \tag{5}$$

where n is the total number of valid pixel pairs being compared. $I_{reference}$ represent pixels from the image with the full dataset (1024 × 1024 elements) and $I_{downsampled}$ is the pixel intensity from the downsampled image

2.6.4 Trade-off determination

To determine the most effective downsampling strategy, all image quality metrics: SNR, CNR, and RMSE, were evaluated collectively across the different reconstruction configurations. A configuration was considered optimal when it maintained acceptable tumor visibility, reflected by high tumor SNR and CNR, while also preserving the uniformity of background tissue and the water bath through stable background and water SNR values. In addition, optimal configurations exhibited low reconstruction error relative to the full data reference image, as indicated by a low RMSE, and simultaneously achieved substantial reductions in data size and reconstruction time. This integrated, multi-metric evaluation framework enabled the objective selection of downsampling parameters that balance diagnostic image quality with computational and hardware efficiency, ensuring suitability for a cost-effective and clinically practical UST system.

3. RESULTS

This section presents the effects of transducer-element downsampling and transmit-only downsampling on reconstruction quality, computational performance, and quantitative image metrics. All reconstructions were compared with the 1024 × 1024 reference image.

3.1 Visual assessment of reconstructed images

Figure 2 presents reconstructed breast sound-speed images obtained using the full dataset and progressively more intense transmit-only downsampling configurations, along with corresponding zoomed-in views of the tumor region.

As shown in Figure 2(a), the full dataset reconstruction (1024 × 1024 elements) exhibits well-defined anatomical boundaries, homogeneous background tissue, and clear visualization of the tumor with high local contrast. When the number of transmit elements is reduced by a factor of two to 512 × 1024 (Figure 2(b)), overall image quality remains largely comparable to the full dataset, with preserved anatomical structure and minimal smoothing effects. Further transmit downsampling to 128 × 1024 (Figure 2(c)) introduces mild blurring of fine tissue features; however, the tumor remains clearly identifiable in both the full-field image and the zoomed-in region.

More intense transmit-only downsampling to 32 × 1024 (Figure 2(d)) results in noticeable smoothing of internal breast structures and reduced definition of tissue boundaries, with increased background noise and reduced tumor conspicuity in the zoomed-in view. These visual trends indicate that moderate transmit-only downsampling preserves clinically relevant image features, whereas severe reductions significantly degrade image quality and lesion definition.

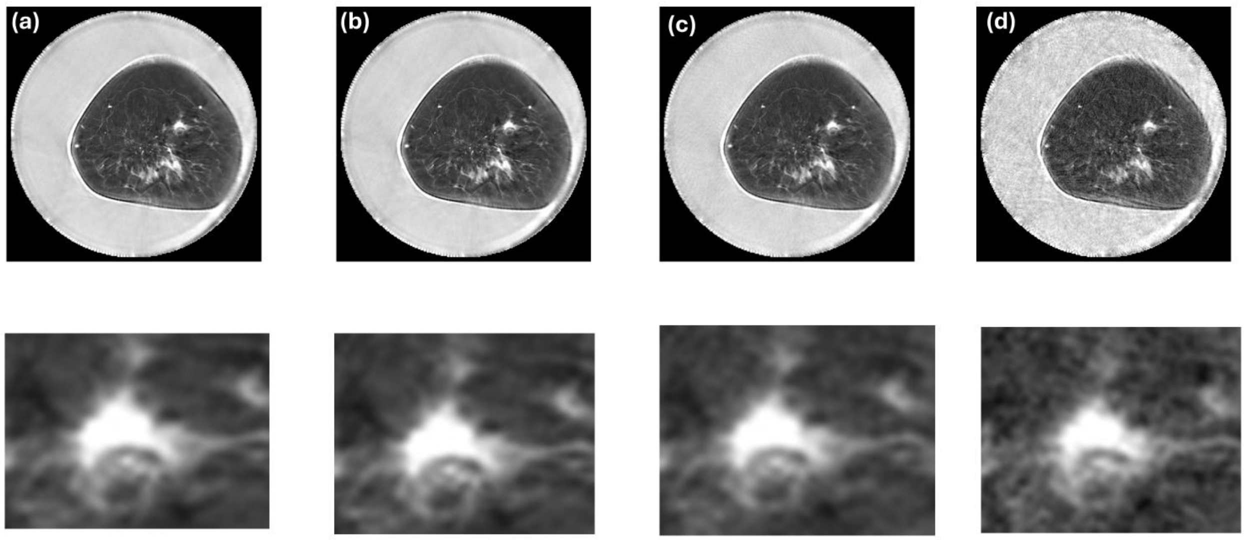

Figure 2: Reconstructed breast images under varying levels of transmit downsampling and their corresponding zoomed-in tumor regions. (a) Full dataset (1024 × 1024 elements); (b) full dataset with a transmit downsampling factor of 2, resulting in 512 × 1024 elements; (c) full dataset with a transmit downsampling factor of 8, resulting in 128 × 1024 elements; and (4) full dataset with a transmit downsampling factor of 32, resulting in 32 × 1024 elements. For each case, the corresponding zoomed-in tumor region is shown directly below the reconstructed image.

3.2 Visual assessment of reconstructed images from transducer-element and transmit-only downsampling

Figure 3 presents reconstructed breast images obtained using combined transducer-element and transmit-only downsampling strategies, along with corresponding zoomed-in views of the tumor region.

As shown in Figure 3(a), uniform transducer-element downsampling to 512 × 512 elements preserves overall anatomical structure, background homogeneity, and clear tumor visibility, with image quality remaining comparable to the full-resolution reference. When transmit-only downsampling is applied to this configuration with a factor of four, resulting in a 128 × 512 dataset (Figure 3(b)), a modest reduction in fine structural detail is observed; however, the tumor remains clearly identifiable in both the full-field image and the zoomed-in region.

More extreme transmit-only downsampling to 32 × 512 elements (Figure 3(c)) produces noticeable image smoothing and reduced definition of internal breast structures. Boundary features appear less distinct, and tumor contrast is moderately reduced in the zoomed-in view. Despite these degradations, the tumor remains visible, indicating that moderate combined downsampling retains diagnostic content, whereas extreme reductions increasingly compromise spatial resolution and angular sampling.

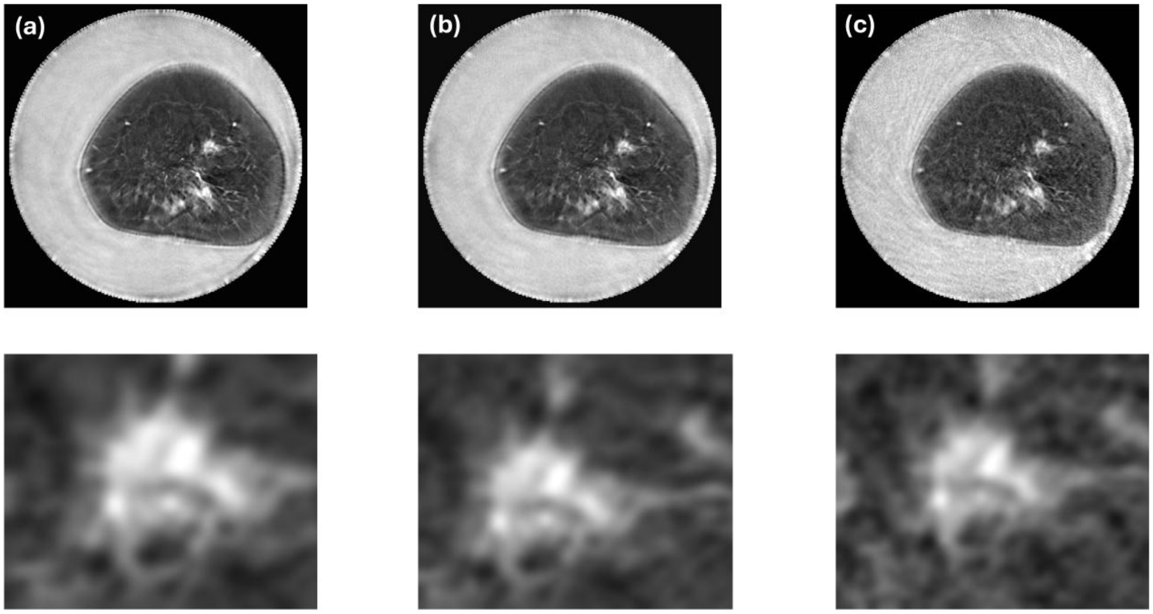

Figure 3 Reconstructed breast images under **combined transmit and receive (transducer element) downsampling** and their corresponding zoomed-in tumor regions. (a) Full dataset downsampled by a factor of 2, resulting in 512 × 512 elements; (b) full dataset downsampled by a factor of 2 with a transmit downsampling factor of 4, resulting in 128 × 512 elements; and (c) full dataset downsampled by a factor of 2 with a transmit downsampling factor of 16, resulting in 32 × 512 elements. For each case, the corresponding zoomed-in tumor region is shown **directly below** the reconstructed image.

3.3 Visual assessment of reconstructed images under extreme combined downsampling

Figure 4 illustrates reconstructed breast images obtained using more extreme combined transducer-element and transmit-only downsampling configurations, along with corresponding zoomed-in views of the tumor region. These cases represent the lower limit of spatial and angular sampling explored in this study and highlight the effects of significant data reduction on image quality.

As shown in Figure 4(a), reducing the array to 256 × 256 elements results in noticeable smoothing of fine anatomical features and reduced sharpness of tissue boundaries compared to higher-resolution configurations; however, the tumor remains identifiable in both the full-field image and the zoomed-in view. Applying additional transmit-only downsampling with a factor of four to this configuration (64 × 256, Figure 4(b)) further degrades spatial detail and contrast, with increased background texture and diminished definition of internal structures. At the most extreme configuration of 32 × 256 elements (Figure 4(c)), image degradation becomes pronounced, with substantial loss of structural detail, increased noise, and reduced tumor conspicuity in the zoomed-in region. These visual observations confirm the compounded impact of simultaneous spatial and angular undersampling on reconstruction quality.

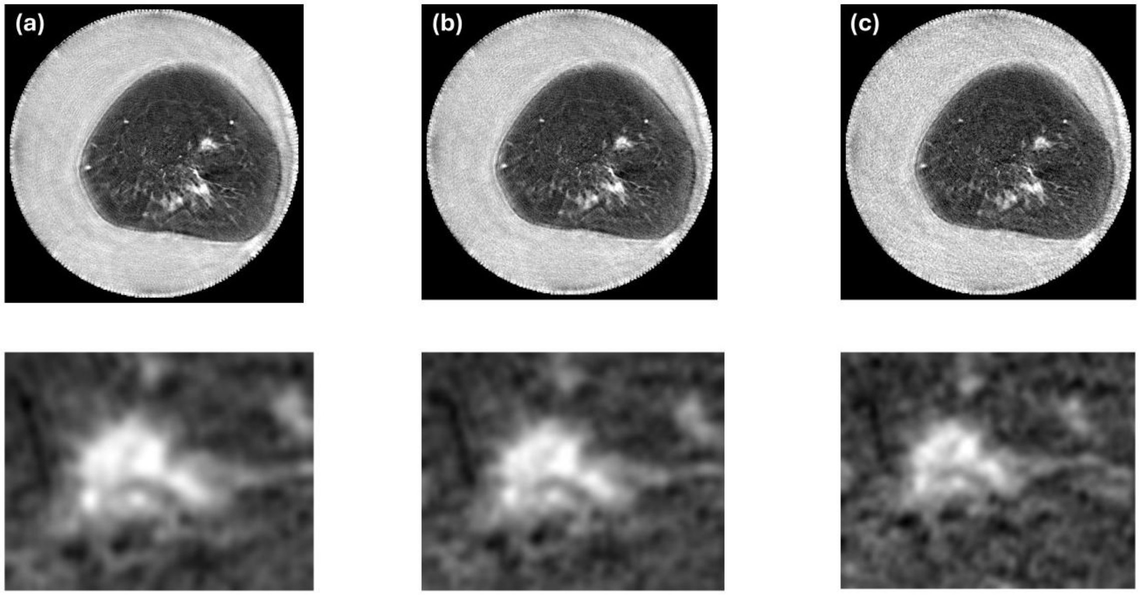

Figure 4 Reconstructed breast images under combined transmit and receive (transducer element) downsampling and their corresponding zoomed-in tumor regions. (a) Full dataset downsampled by a factor of 4, resulting in 256 × 256 elements; (b) full dataset downsampled by a factor of 4 with a transmit downsampling factor of 4, resulting in 64 × 256 elements; and (c) full dataset downsampled by a factor of 4 with a transmit downsampling factor of 8 applied during reconstruction, resulting in 32 × 256 elements. For each case, the corresponding zoomed-in tumor region is shown directly below the reconstructed image.

3.4 Reconstruction time and computational performance

Reconstruction times for all downsampling configurations are summarized in Table 1. As expected, computational time decreased with increasing levels of data reduction; however, the observed improvements were not proportional to the reduction in raw data size. The full dataset required approximately 3.0 minutes per slice, whereas the 32 × 512 configuration was reconstructed in 1.06 minutes, corresponding to an improvement of nearly a factor of three rather than the order-of-magnitude reduction suggested by the decrease in data volume.

This behavior reflects the structure of the Full Waveform Inversion (FWI) algorithm, in which only a subset of computational operations scales with the amount of acquired data. Reducing the number of transmit events lowers the number of forward and adjoint wavefield simulations performed per iteration and reducing the number of receive channels decreases data input and memory size. However, the dominant computational cost in FWI arises from numerical wave propagation on a fixed spatial grid, which must be executed for each iteration regardless of data size. Additional steps, such as gradient accumulation, model updates, and regularization, also remain largely independent of the number of transmit-receive measurements. Consequently, these data-independent operations limit the achievable reduction in total reconstruction time, even under aggressive downsampling.

As a result, aggressive downsampling yields diminishing returns in reconstruction speed, even as data volume continues to decrease. While extreme downsampling configurations produced the shortest reconstruction times, the overall speedup was limited by these data-independent computational costs and was accompanied by substantial degradation in image quality. These findings indicate that downsampling can provide meaningful, but not unlimited, improvements in computational efficiency and highlight the importance of balancing data reduction with the fixed computational overhead inherent to FWI-based reconstruction.

3.5 Quantitative image quality metrics

Quantitative image quality metrics, including signal-to-noise ratio (SNR), contrast-to-noise ratio (CNR), and root mean square error (RMSE), are summarized in Table 1 for all downsampling configurations. RMSE increased monotonically with increasing levels of downsampling, indicating progressively greater deviation from the full-resolution reference image. Moderate data reductions, such as the 512 × 512 and 128 × 1024 configurations, resulted in only minor increases in RMSE, whereas extremely downsampling, for example, the 32 × 256 configuration, produced substantially larger errors.

Tumor SNR and CNR remained relatively stable across moderate downsampling levels, indicating preserved tumor conspicuity and diagnostic quality. In contrast, both metrics declined rapidly under extreme downsampling, reflecting reduced tumor visibility and loss of contrast relative to surrounding tissue. Background SNR and water SNR exhibited only modest variation for moderate configurations, suggesting that system noise performance and background uniformity were largely maintained until more severe spatial and angular undersampling were introduced.

Table 1. Quantitative image quality metrics for various downsampling factors

Element Configuration	Reconstruction time (min)	SNR_{tumor}	SNR_{water}	$SNR_{background}$	CNR	RMSE
1024 x 1024	3.00	80.41	799.58	154.67	1.65	0.0000
512 x1024	2.07	79.62	624.13	149.77	1.66	0.0015
128 x1024	1.70	79.00	484.97	142.81	1.68	0.0040
32 x1024	1.16	71.77	193.79	114.72	1.67	0.0107
512 x 512	1.95	73.59	399.33	133.43	1.78	0.0146
128 x 512	1.17	70.69	362.80	130.53	1.82	0.0154
32 x512	1.06	70.22	206.54	120.42	1.78	0.0162
256 x 256	1.65	70.29	178.40	106.60	1.83	0.0172
64 x 256	1.08	69.43	156.15	93.27	1.85	0.0180
32 x256	1.07	66.77	123.54	85.18	1.77	0.0194

3.5.1 Identification of image quality inflection using SNR trends under transmit-only downsampling

To quantitatively assess the impact of transmit-only downsampling on image quality, the signal-to-noise ratio (SNR) was evaluated across a range of transmit–receive element configurations while maintaining a fixed receive aperture of 1024 elements. Tumor SNR, water SNR, and background SNR were computed for each configuration to characterize lesion detectability, system noise performance, and tissue uniformity, respectively. By examining how these metrics change as the number of transmit elements is progressively reduced from 1024 to 32, saturation or inflection points in image quality can be identified, where further reductions in the transmit aperture result in marginal changes in SNR and limited additional impact on diagnostic performance.

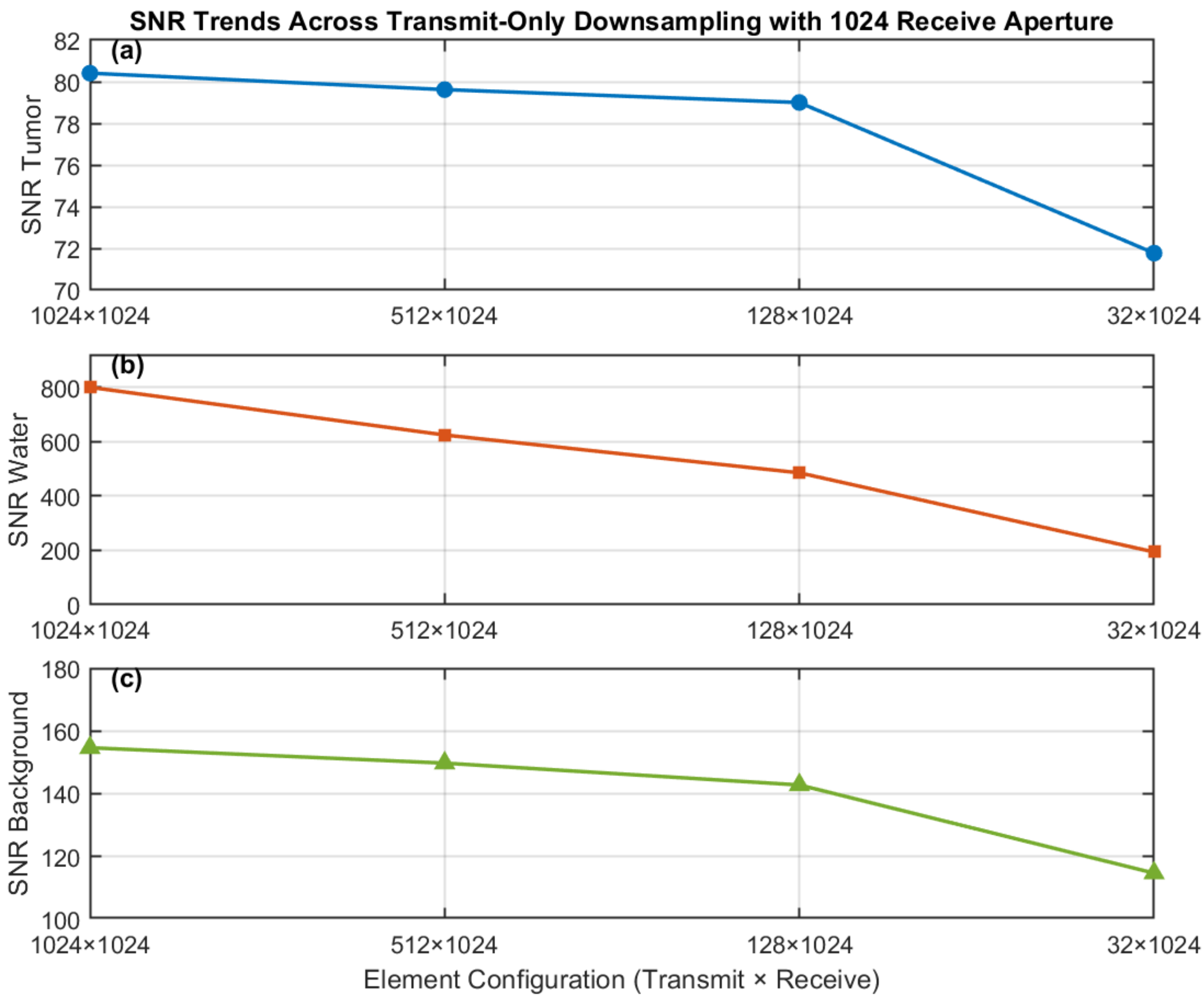

Figure 5: Signal-to-noise ratio (SNR) as a function of transmit-only downsampling with a fixed receive aperture (Rx = 1024). Element configurations range from 1024 to 32 × 1024. Panels show (a) tumor SNR, (b) water SNR, and (c) background SNR, illustrating how image quality changes as the number of transmit elements is reduced.

3.5.2 SNR trends under combined transmit and receive downsampling with a 512-element receive aperture

Following the transmit-only analysis with a full receive aperture, the effect of combined transmit and receive downsampling on image quality was evaluated using a reduced receive aperture of 512 elements. Tumor SNR, water SNR, and background SNR were computed for each transmit-receive configuration to assess tumor detectability, system noise behavior, and tissue uniformity under reduced spatial sampling. This analysis enables examination of how simultaneous reductions in transmit and receive apertures influence SNR trends and helps identify thresholds beyond which combined spatial and angular undersampling leads to noticeable degradation in image quality.

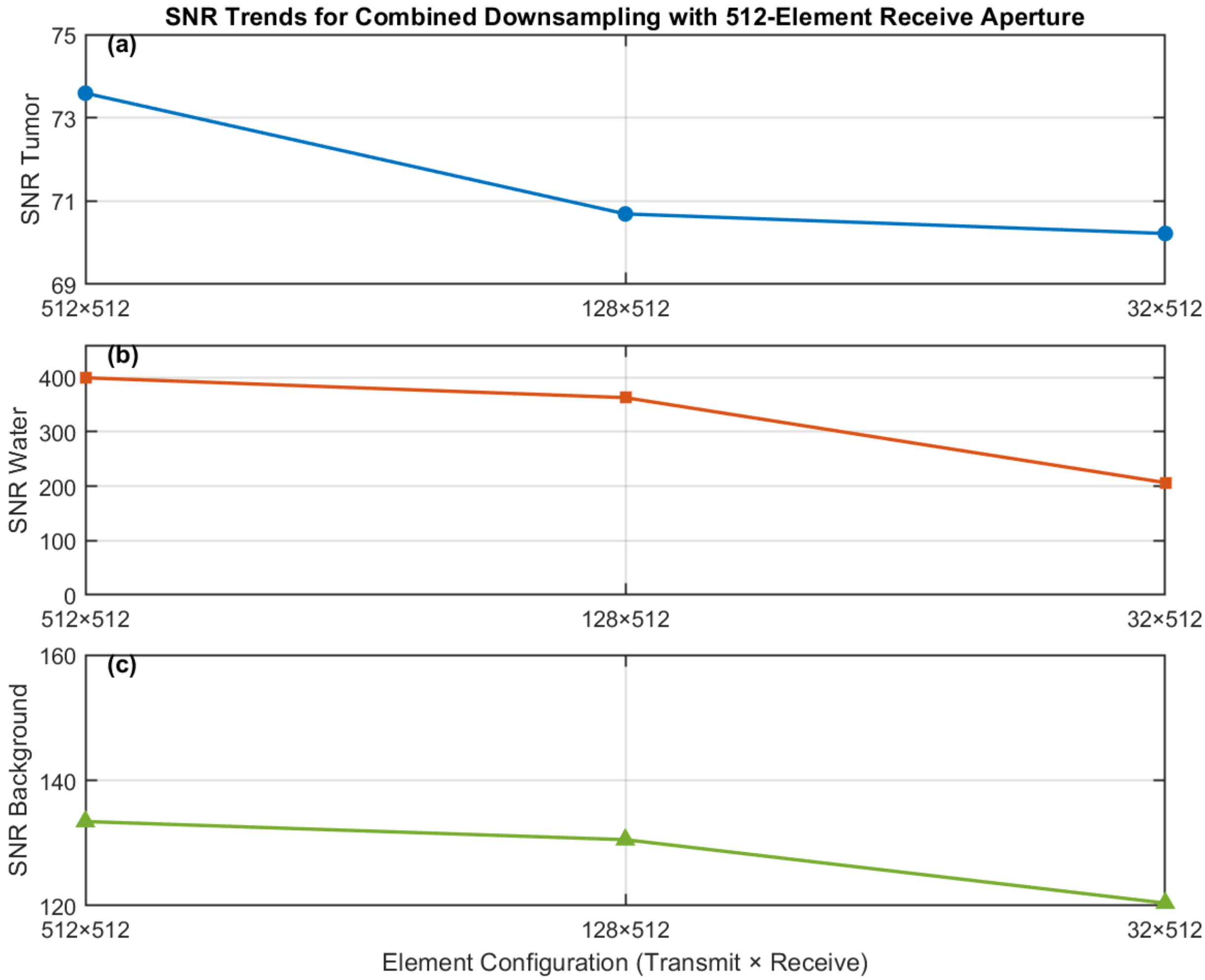

Figure 6: Signal-to-noise ratio (SNR) as a function of combined transmit and receive downsampling with a fixed receive aperture (Rx = 512). Element configurations range from 512 × 512 to 32 × 512 (transmit × receive). Panels show (a) tumor SNR, (b) water SNR, and (c) background SNR, illustrating how image quality changes as both transmit and receive apertures are reduced.

3.5.3 SNR trends under combined transmit and receive downsampling with a 256-element receive aperture

To investigate the impact of progressive downsampling, SNR trends were examined with the receive aperture further reduced to 256 elements while systematically decreasing the number of transmit elements. This analysis characterizes image quality behavior under reduced spatial sampling and enables assessment of tumor, water, and background SNR at lower acquisition densities.

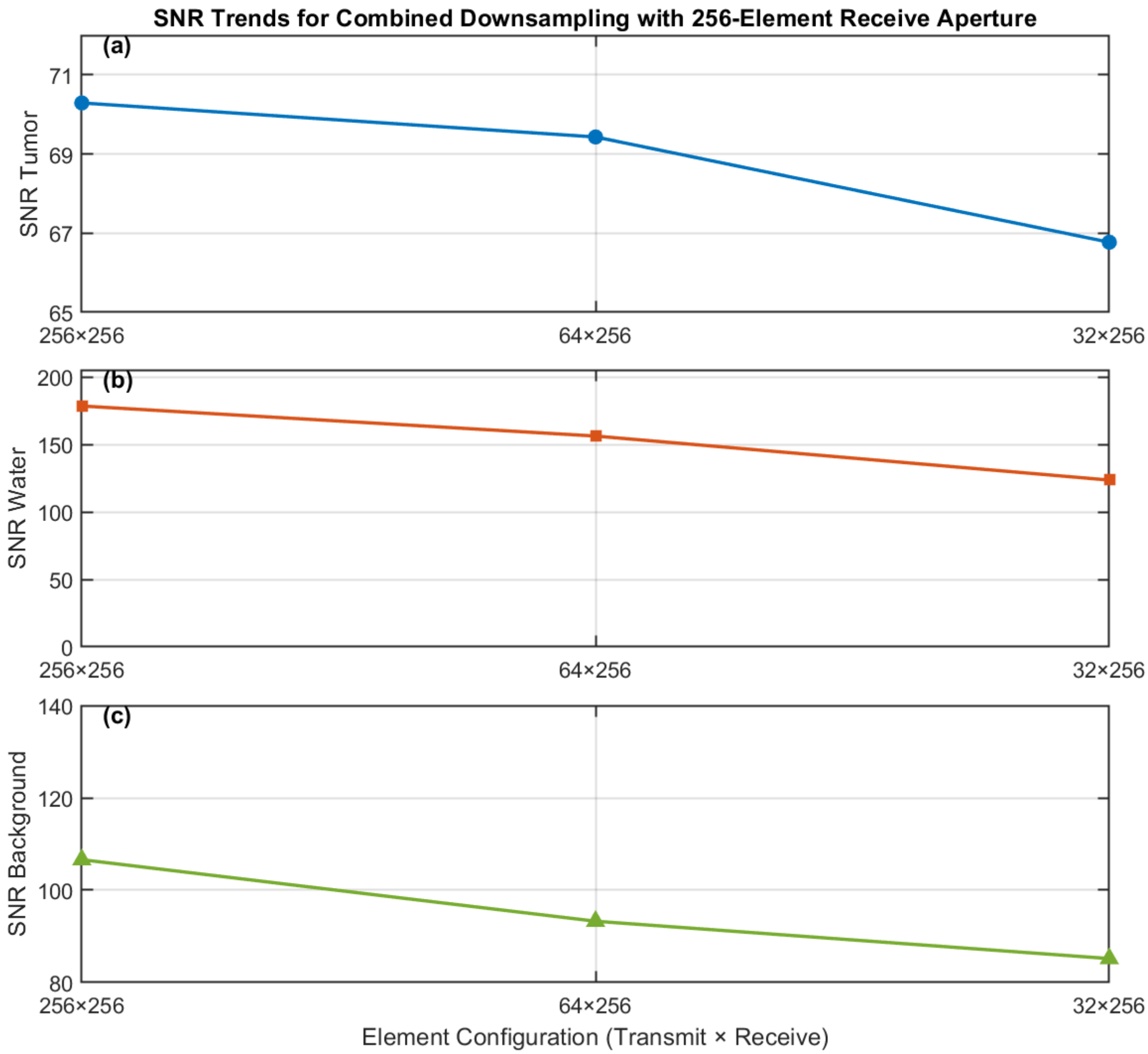

Figure 7: Signal-to-noise ratio (SNR) as a function of combined transmit and receive downsampling with a fixed receive aperture (Rx = 256). Element configurations range from 256 × 256 to 32 × 256 (transmit × receive). Panels show (a) tumor SNR, (b) water SNR, and (c) background SNR, illustrating how image quality changes as both transmit and receive apertures are reduced.

3.5.4 Quantitative error analysis

To further quantify the effect of downsampling on reconstruction fidelity, the root-mean-square error (RMSE) was evaluated across all transmit-only and combined transmit–receive downsampling configurations. RMSE provides a global measure of deviation from the reference full-data reconstruction and is particularly sensitive to cumulative errors introduced by reduced spatial and angular sampling. By examining RMSE trends under fixed receive apertures, this analysis complements the SNR and CNR metrics by capturing overall reconstruction accuracy as downsampling becomes more aggressive. Figure 8 summarizes the RMSE behavior for transmit-only downsampling with a fixed receive aperture of 1024 elements, as well as combined transmit-receive downsampling with receive apertures fixed at 512 and 256 elements.

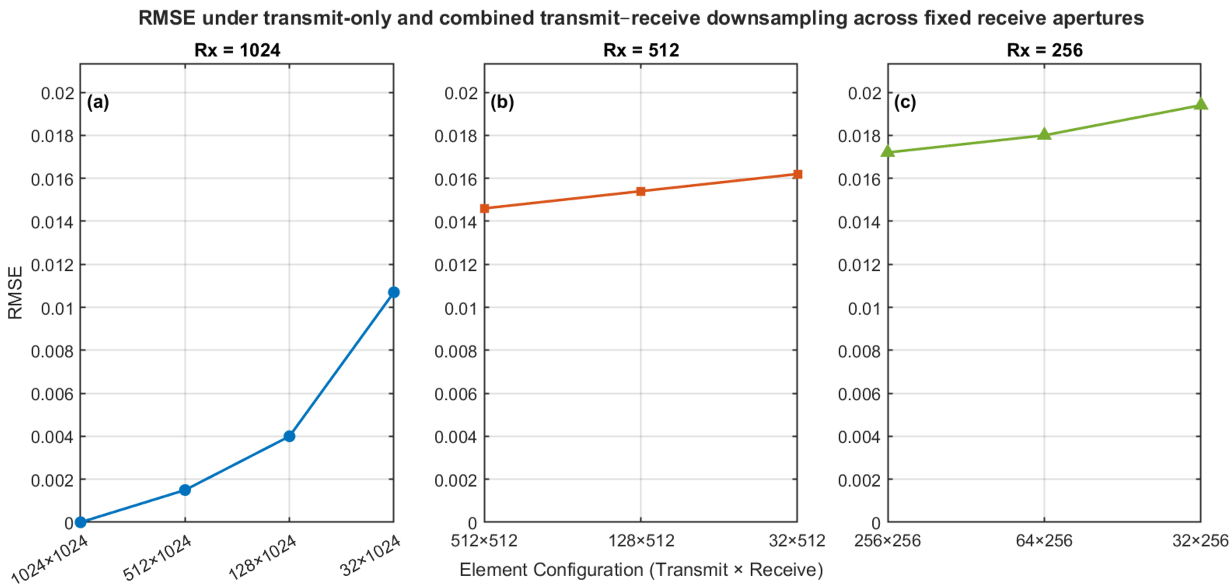

Figure 8 Root-mean-square error (RMSE) as a function of element configuration under transmit-only and combined transmit-receive downsampling across fixed receive apertures. Panels show RMSE trends for (a) transmit-only downsampling with Rx = 1024, (b) combined transmit–receive downsampling with Rx = 512, and (c) combined transmit–receive downsampling with Rx = 256.

4. DISCUSSION

This study systematically evaluated the effects of transducer-element downsampling and transmit-only downsampling on image quality and computational performance in clinical ultrasound tomography (UST). The results demonstrate that substantial reductions in raw data volume and reconstruction time can be achieved without a proportional loss in diagnostically relevant image quality, particularly when moderate downsampling strategies are employed. By integrating qualitative visual assessment, as shown in Figures 2-4, with quantitative image quality metrics as shown in Figures 5-8 and Table 1, this work provides a comprehensive characterization of the trade-offs between computational efficiency and reconstruction quality in full-waveform inversion-based UST.

Transmit-only downsampling with a fixed receive aperture of 1024 elements exhibited a clear saturation-like behavior in image quality metrics, as shown in Figure 5. Tumor, background, and water SNR values remained relatively stable as the number of transmit elements was reduced from 1024 to approximately 128, indicating a plateau region in which additional transmit diversity yields diminishing improvements in image quality. This behavior suggests that, beyond a certain level of angular sampling, the inversion process becomes increasingly constrained by factors other than transmit aperture density, including modeling inaccuracies, noise measurement, and the fundamental limitations of the inverse problem. A more pronounced decline in SNR was observed when the number of transmit elements was reduced to 32, with the most significant degradation occurring at this most aggressive downsampling level. These results indicate that the transmit aperture becomes a limiting factor only under severe downsampling, when angular coverage and wavefield diversity are insufficient to support accurate reconstruction. Within the context of the dataset and reconstruction framework considered, transmit-only downsampling to approximately 128 elements therefore provide a near-optimal balance between image quality preservation and computational efficiency.

Moderate transducer-element downsampling further supports this conclusion. Visual inspection of Figure 3 and quantitative metrics summarized in Table 1 show that reducing the array to 512 × 512 elements preserves essential anatomical features and tumor visibility, with only modest increases in RMSE and largely stable tumor SNR and CNR values relative to the full dataset. This indicates that the spatial sampling density at this level remains above the threshold required for reliable full-waveform inversion. These findings are consistent with prior studies reporting that, in practical UST systems, effective resolution is often limited more by modeling approximations, acoustic heterogeneity, and noise than by transducer count alone [1,5,7]. Consequently, moderate reductions in both transmit and receive apertures do not substantially degrade clinically relevant image features.

In contrast, more severe combined downsampling strategies resulted in clear and predictable degradation in image quality. Reducing the array to 256 × 256 elements or operating at progressively reduced sampling densities produced increased image smoothing, reduced tumor contrast, and elevated RMSE values, as evident in Figures 4, 7, and 8. The RMSE trends in Figure 8 show a monotonic increase as downsampling becomes more aggressive across all fixed receive apertures, with particularly pronounced error growth under transmit-only downsampling at Rx = 1024 and under combined downsampling at Rx = 256. These results quantitatively confirm the visual and SNR-based observations, indicating that reconstruction fidelity degrades as both spatial and angular sampling are reduced. These effects arise from the combined loss of spatial and angular sampling, which limits the diversity of the recorded wavefield and reduces the information content available to the inversion algorithm. As a result, reconstruction accuracy declines, particularly in regions characterized by strong sound-speed gradients or fine-scale anatomical structure.

Transmit-only downsampling demonstrated intermediate and clinically relevant behavior relative to combined downsampling. As shown in Figures 2 and 5, reductions in the transmit aperture by factors of four to eight maintained acceptable tumor visibility while substantially reducing reconstruction time, provided that a dense receive aperture was preserved. This observation suggests that the transmit dimension can tolerate more aggressive data reduction than the receive dimension in UST acquisitions. From a physical perspective, this behavior is consistent with tomographic imaging principles, in which dense receive sampling can partially compensate for reduced transmit diversity as long as overall angular coverage remains sufficient. However, the combined downsampling results in Figures 6-8 indicate that simultaneous reductions in both transmit and receive apertures lead to more rapid degradation in image quality, underscoring the compounded effects of spatial and angular undersampling.

Overall, the multi-metric evaluation framework employed in this study, integrating visual assessment with SNR, CNR, RMSE, and reconstruction time, proved effective for identifying downsampling configurations that balance diagnostic performance with computational efficiency. The results suggest that carefully selected downsampling strategies can substantially reduce memory requirements, hardware complexity, acquisition time, and reconstruction cost, thereby supporting the development of more accessible and cost-effective UST systems capable of near–real-time imaging in clinical workflows.

Several limitations should be acknowledged. The analysis was conducted using a single clinical case, which limits the generalizability of the conclusions. Breast composition, tumor morphology, and acoustic heterogeneity vary widely across patients, and further evaluation across diverse clinical datasets is required to assess the robustness of the identified downsampling strategies. In addition, the impact of downsampling on the detectability of tumors with varying sizes, contrasts, and spatial locations was not explicitly examined. Future work will incorporate simulation studies and additional clinical cases with controlled lesion characteristics to further evaluate diagnostic sensitivity and establish generalizable design guidelines for downsampled UST systems.

5. CONCLUSIONS

This study provides a systematic and quantitative evaluation of data downsampling strategies in clinical ultrasound tomography (UST) using full-waveform inversion (FWI), with explicit consideration of their impact on image quality, data volume, and computational performance. By jointly analyzing transmit-only and combined transmit-receive downsampling across multiple aperture configurations, this work establishes practical operating configurations that achieve substantial reductions in raw data volume and reconstruction time without proportional degradation in diagnostically relevant image features. For optimally selected configurations, reconstruction time reductions of up to approximately three times were achieved while preserving tumor visibility and overall image quality.

A key contribution of this work is the identification of saturation-like behavior in image quality metrics under transmit-only downsampling with dense receive sampling, revealing a configuration in which further increases in transmit diversity yield only marginal improvements in image quality. In contrast, progressively reduced spatial and angular sampling through combined transmit-receive downsampling resulted in predictable degradation of image quality, delineating clear thresholds beyond which reconstruction quality is compromised. These findings provide quantitative guidance for balancing computational efficiency and image quality, advancing beyond practical or experimental downsampling approaches commonly reported in prior studies.

Collectively, the results demonstrate that carefully selected downsampling strategies can substantially reduce acquisition time, memory requirements, hardware, and computational cost in UST systems without compromising clinically meaningful image features. This has direct implications for the design of next-generation UST system, enabling more accessible imaging workflows suitable for clinical deployment.

Future work will extend this framework through larger multi-patient clinical studies and controlled simulation experiments incorporating lesions of varying sizes, contrasts, and spatial distributions. These will further assess the robustness and generalizability of the identified operating configurations and support the development of principled, application-specific design guidelines for downsampled UST systems.

REFERENCES

[1] M. Sak, P. Littrup, R. Brem, and N. Duric, "Whole breast sound speed measurement from US tomography correlates strongly with volumetric breast density from mammography," *J Breast Imaging*, vol. 2, no. 5, pp. 443–451, 2020, doi: 10.1093/jbi/wbaa052.

[2] N. Duric *et al.*, "Breast density measurements with ultrasound tomography: A comparison with film and digital mammography," *Med Phys*, vol. 40, no. 1, 2013, doi: 10.1118/1.4772057.

[3] A. Javaherian, F. Lucka, and B. T. Cox, "Refraction-corrected ray-based inversion for three-dimensional ultrasound tomography of the breast," *Inverse Probl*, vol. 36, no. 12, Dec. 2020, doi: 10.1088/1361-6420/abc0fc.

[4] R. Ali *et al.*, "2-D Slicewise Waveform Inversion of Sound Speed and Acoustic Attenuation for Ring Array Ultrasound Tomography Based on a Block LU Solver," *IEEE Trans Med Imaging*, vol. 43, no. 8, pp. 2988–3000, 2024, doi: 10.1109/TMI.2024.3383816.

[5] P. J. Littrup, M. Mehrmohammadi, and N. Duric, "Breast Tomographic Ultrasound: The Spectrum from Current Dense Breast Cancer Screenings to Future Theranostic Treatments," Apr. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/tomography10040044.

[6] N. Duric *et al.*, "Detection of breast cancer with ultrasound tomography: First results with the Computed Ultrasound Risk Evaluation (CURE) prototype," *Med Phys*, vol. 34, no. 2, pp. 773–785, 2007, doi: 10.1118/1.2432161.

[7] K. Wang, T. Matthews, F. Anis, C. Li, N. Duric, and M. Anastasio, "Waveform inversion with source encoding for breast sound speed reconstruction in ultrasound computed tomography," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 62, no. 3, pp. 475–493, 2015, doi: 10.1109/TUFFC.2014.006788.

[8] L. Huang *et al.*, "Breast ultrasound waveform tomography: using both transmission and reflection data, and numerical virtual point sources," in *Medical Imaging 2014: Ultrasonic Imaging and Tomography*, SPIE, Mar. 2014, p. 90400T. doi: 10.1117/12.2043136.

[9] L. Lozenski, H. Wang, B. Wohlberg, U. Villa, and Y. Lin, "Data driven methods for ultrasound computed tomography," SPIE-Intl Soc Optical Eng, Apr. 2023, p. 27. doi: 10.1117/12.2654442.

[10] The MathWorks Inc., "MATLAB," 2024, *The MathWorks Inc., Natick, Massachusetts, USA* : R2024a.

[11] Y. Wang, C. Zheng, H. Peng, and Q. Chen, "An adaptive beamforming method for ultrasound imaging based on the mean-to-standard-deviation factor," *Ultrasonics*, vol. 90, pp. 32–41, Nov. 2018, doi: 10.1016/j.ultras.2018.06.006.