Full Waveform Inversion Using Generative Adversarial Networks for Breast Sound Speed Reconstruction in 2D Ultrasound Computed Tomography

Gangwon Jeong

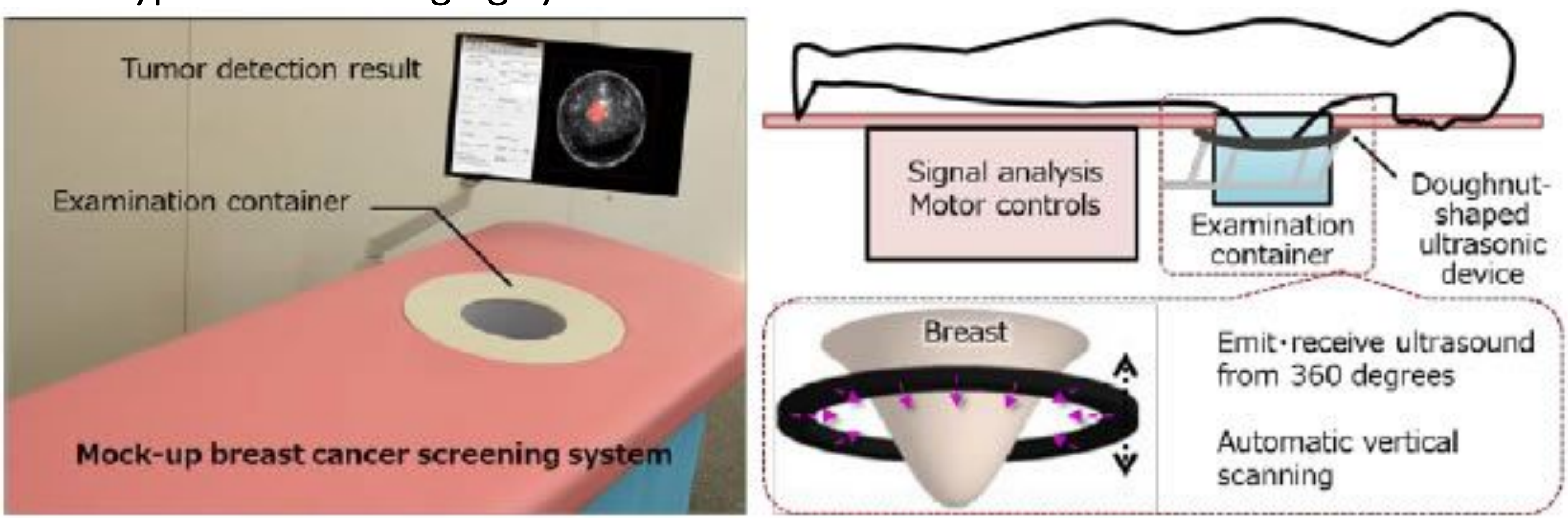
Outline

1. Introduction
   - Ultrasound computed tomography (USCT)
   - Regularization
2. Background
   - Waveform inversion for USCT image reconstruction
   - Generative adversarial nets (GAN)
3. Method
   - GAN-based waveform inversion
4. Numerical study
5. Conclusions and future work

Ultrasound computed tomography (USCT)

- USCT is a non-invasive ***quantitative*** imaging technique for detecting breast cancer
  - Non-ionizing radiation
  - Breast-compression-free
  - High contrast profile for early detection

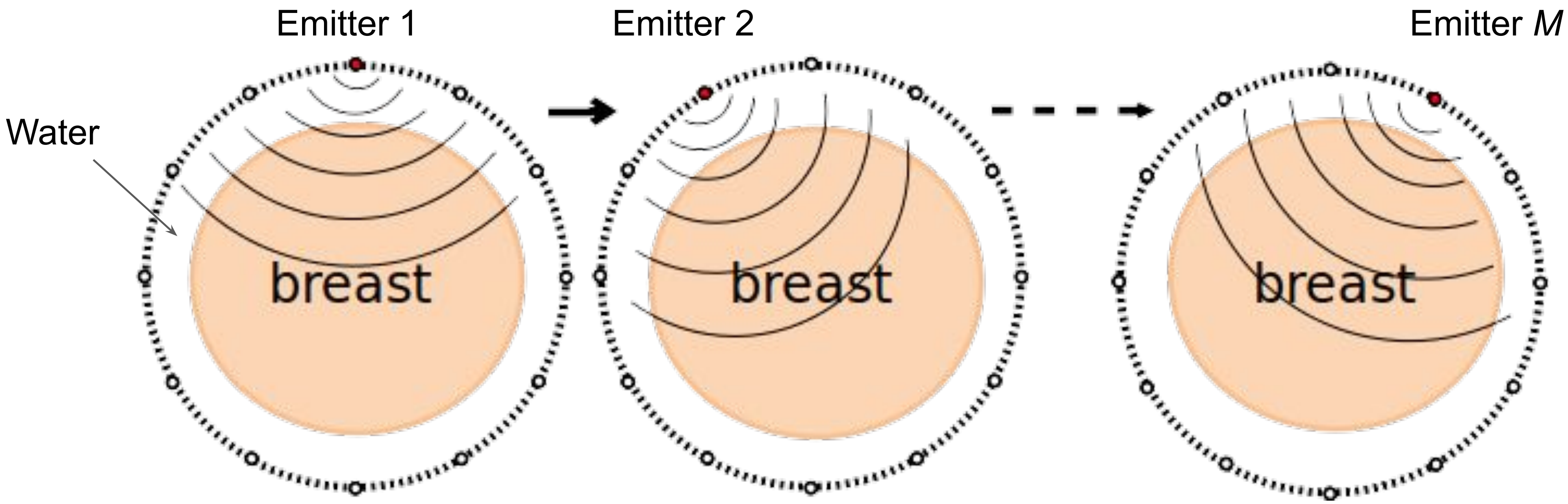
- A typical USCT imaging system



Breast cancer screening system mock-up and prototype configuration [from Hitach website]

Ultrasound computed tomography (USCT)

- A typical USCT imaging system [cont'd]

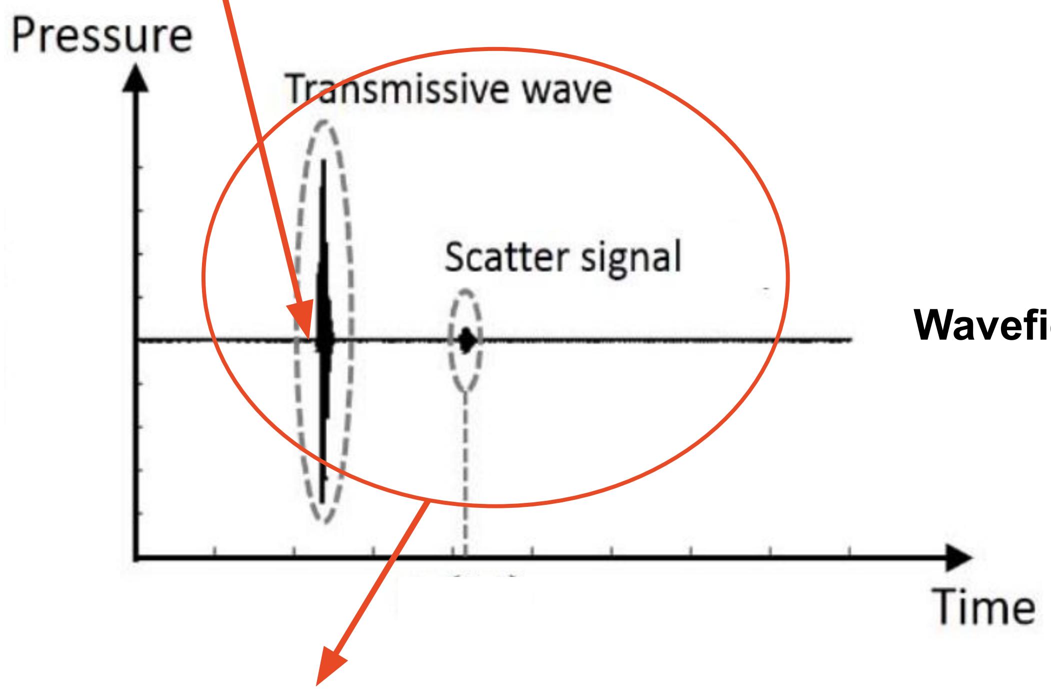


  - A transducer (emitter) generates an acoustic pulse (1~20MHz) to insonify a breast
  - All transducers (receivers) measure the resultant wavefield data
  - One dataset contains *M* single shots

- The **Goal of USCT** is to reconstruct the acoustic properties (e.g., sound speed) from the wavefield dataset

Ultrasound computed tomography (USCT)

- Most USCT image reconstruction methods use simplified approximate models of wave propagation
  - E.g., Time-of-flight tomography: scattering events are not considered (Poor quality)



Wavefield trace at a receiver from a source

- Waveform inversion methods provide more accurate reconstructions
  - Inversion of the acoustic wave equation to the acoustic properties
  - High computational cost of solving wave equation
  - **Enhanced resolution** with multiple scattering events

Challenges of waveform inversion methods

1. The solution may not be unique due to incomplete measurement data

2. Noise in data may result in a large change in the solutions

- The uncertainty in data lead to the uncertainty in the solution

3. Many possible solutions may exist, which minimize the closeness between the measured and modeled data

- The objective functions is usually nonconvex

4. Simulation error may result in inaccurate solutions

- Mismatch between 3D measurement and 2D simulation
- Mismatch of transducer locations between measurement and simulation
- Inaccurate acoustic properties in simulation such as density and attenuation

Regularization methods

- **Regularization** helps promoting “desirable” properties in the accomplished image
  - Regularizer is specified based on a prior knowledge of the sought-after objects
  - Examples of classic regularization methods
    - Tikhonov regularization: promote smoothness
    - Total variation regularization: preserve edges
    - $l1$ regularization in a wavelet basis: sparsify

- **Issue:** Classic regularizers may be suboptimal
  - Tikhonov regularization: over-smooth
  - Total variation regularization: low contrast, over-blocky
  - $l1$ regularization in a wavelet basis: chosen wavelet may not be optimal

Regularization methods

- **Alternative:** Data-driven regularizers can be formulated from realistic medical image set via machine learning approaches
    - Deep neural network (DNN)-based function can represent the realistic medical images
    - Examples of data-driven regularization
        - Sparse dictionary learning [Singh et al. (2016)]
            - Image patches are restricted to a sparse representation using learnt dictionary
        - Generative model [Bora *et al.* (2017)] ⇒ **To be elaborated**
            - The solution space is restricted to be near the range of generative model

- The aim of this study is to investigate the feasibility of a **Generative model-based data-driven regularizer** learnt on realistic breast sound speed maps for the waveform inversion methods in 2D USCT problem

Full-waveform-based imaging model for USCT

- The acoustic wave equation for the time domain approaches

$$\underbrace{\nabla^2 p_j(\mathbf{r},t)}_{\text{Pressure at j-th view}} - \frac{1}{\underbrace{c^2(\mathbf{r})}_{\text{Sound speed}}}\partial_{tt} p_j(\mathbf{r},t) = \underbrace{s_j(\mathbf{r},t)}_{\text{Source at j-th view}}$$

$$p_j(\mathbf{r},0) = 0,\ \ \partial_t p_j(\mathbf{r},0) = 0$$

- Numerical wave equation solver (solved by pseudospectral k-space method, Mast *et al.* (2001))

$$\underbrace{\mathbf{g}_j}_{\text{Estimated pressure at j-th view}} = \underbrace{\mathbf{M}}_{\text{Sampling operator}}\underbrace{\mathbf{H}(\mathbf{c})}_{\text{Wave solver depending on } c}\mathbf{s}_j$$

  - 2D spatiotemporal domain
  - D-D imaging equation
  - Matrix-free computation

Standard waveform inversion problem

- The formulation of a minimization problem

$$\hat{\mathbf{c}} = \underset{\mathbf{c}}{\operatorname{argmin}}\{\underbrace{F(\mathbf{c})}_{\text{Data fidelity}} + \alpha \underbrace{R(\mathbf{c})}_{\text{Regularization}}\}$$

- Data fidelity term

$$F(\mathbf{c}) = \frac{1}{2}\sum_{j=1}^{M} \|\underbrace{\mathbf{g}_j}_{\substack{\text{Noisy measurement}\\ \text{at j-th view}}} - \mathbf{MH}(\mathbf{c})\mathbf{s}_j\|^2$$

  - Nonlinear and Nonconvex w.r.t **c**
    - Good initial guess of ***c*** is required

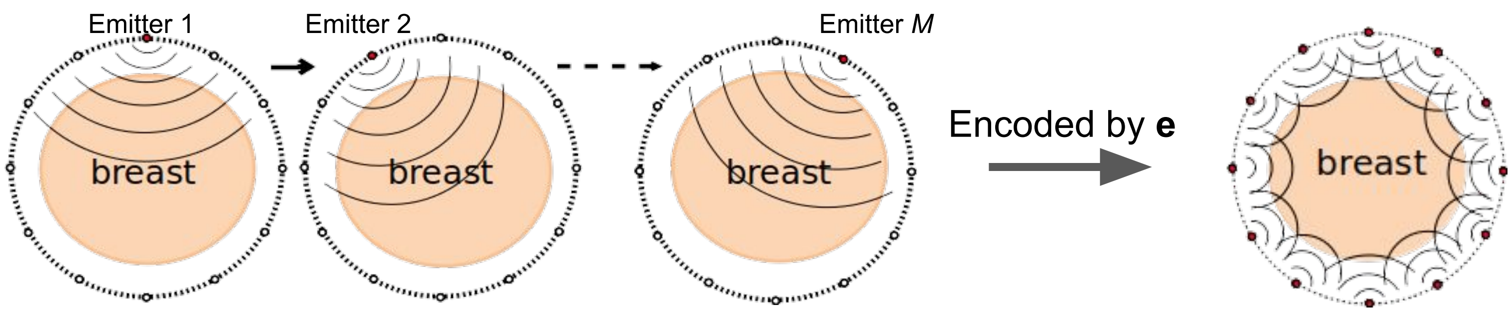
- The selected regularization term
  - Total variation (TV): smoothing away noise-corruption and preserving edges
  - **Pros:** well-established, simple implementation
  - **Cons:** underestimates the tissue variations, unwanted image degradation

Waveform inversion with source encoding

- The computation of solving wave equation is highly burdensome
  ⇒ Randomization of the data fidelity term has been proposed [Wang *et al.* (2017)]
- Choose random vector ***e*** to encode the sources and measurements

$$\mathbf{s}^{\mathbf{e}} = \sum_{j=1}^{M} [\mathbf{e}]_j \mathbf{s}_j \qquad \text{and} \qquad \underline{\mathbf{g}}^{\mathbf{e}} = \sum_{j=1}^{M} [\mathbf{e}]_j \underline{\mathbf{g}}_j$$

  - Elements of ***e*** from i.i.d *Rademacher* distribution ([-1,1])



All *M* shots are simultaneously excited

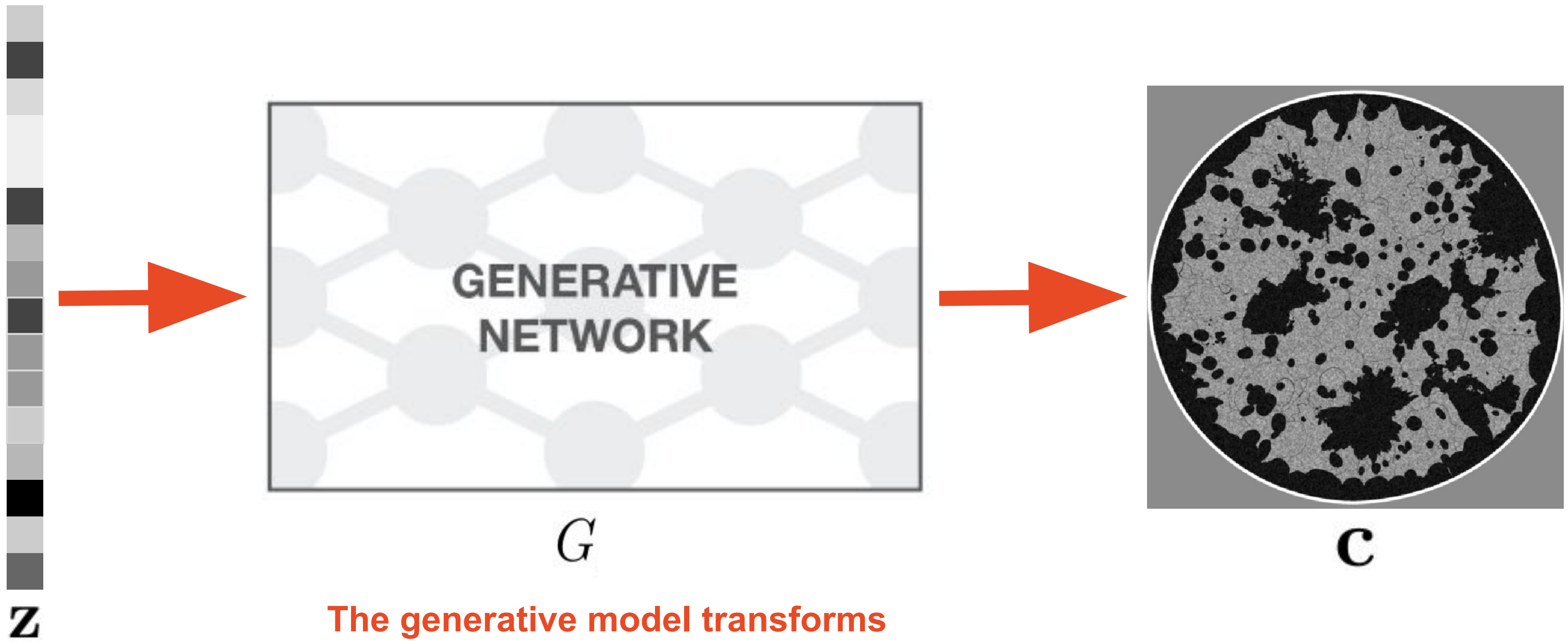
- Stochastic data fidelity term

$$F_s(\mathbf{c}) = \frac{1}{2}\underbrace{\mathbb{E}_{\mathbf{e}}}_{\text{Expectation over e}}\{\|\underline{\mathbf{g}}^{\mathbf{e}} - \mathbf{M}\mathbf{H}(\mathbf{c})\mathbf{s}^{\mathbf{e}}\|^2\} = \underbrace{F(\mathbf{c})}_{\text{Deterministic data fidelity}}$$

  - Stochastic gradient descent (SGD) algorithm is used
    - One realization of ***e*** at each iteration

Generative adversarial nets (GANs)

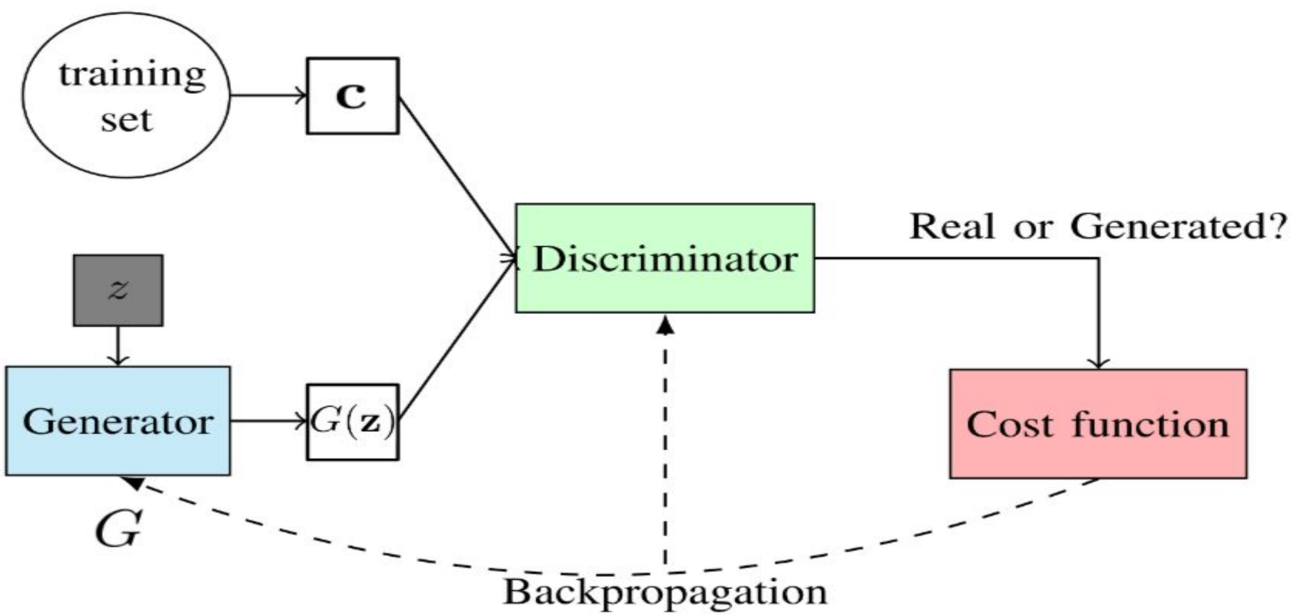
- Generative models based on DNN to represent realistic breast sound speed maps



Input random vector drawn from a simple distribution

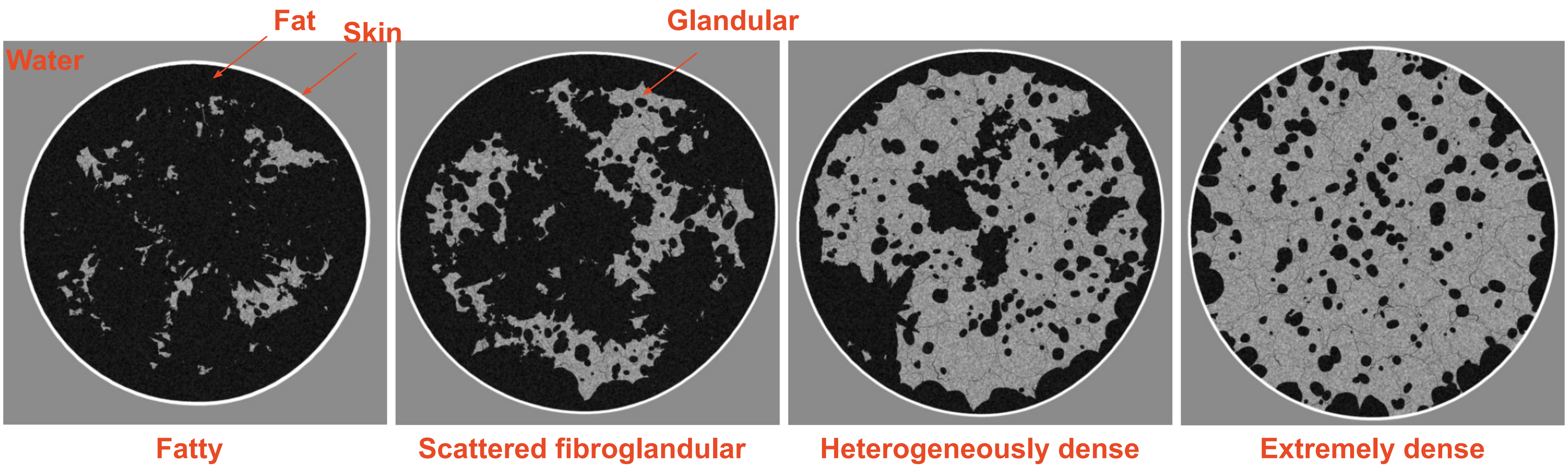
The generative model transforms the simple random vector into a more complex one

Output sound speed maps



Generation of training dataset for GAN

- FDA's numerical breast phantoms (NBPs) [Badano *et al.* (2018)]
  - Generation of 3D random numerical phantoms for mammographic studies
    - Breast size, shape, location, density, and extent of different tissues
  - Extension for USCT study [Li *et al.* (2021)]
    - Adjustment of breast shape consistent with a prone imaging position
    - Stochastic assignment of tissue specific acoustic properties
    - Modeling of acoustic heterogeneity within fatty and glandular tissues
- Example of 2D coronal slices of NBPs (**'real' samples**, pixel size=0.2mm)



* Dynamic range: [1.42, 1.58] mm/μs

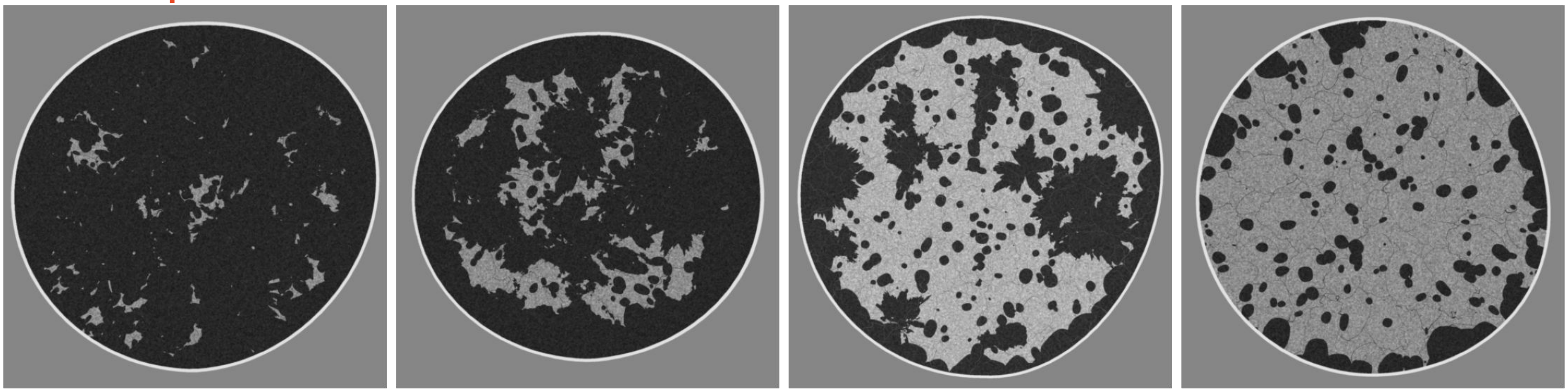
StyleGAN2-based generative model

- One of state-of-art GAN architecture (*StyleGAN2*) was used [Kerras *et al.* (2020)] to represent high-quality NBPs
  - Originally developed to model high-quality human face images
  - High-capacity generator with successive CNN blocks
- *StyleGAN2*-based generator for high-quality NBPs
  - Input random vectors
    - Style vectors $\mathcal{W}$ control overall style of breasts
    - Noise vectors $\mathcal{N}$ control minor changes of features
  - Training dataset: 33,000 2D coronal slices from 1,800 3D NBPs
  - Generated breast sound speed map: $\mathbf{c} = G(\mathcal{W}, \mathcal{N})$

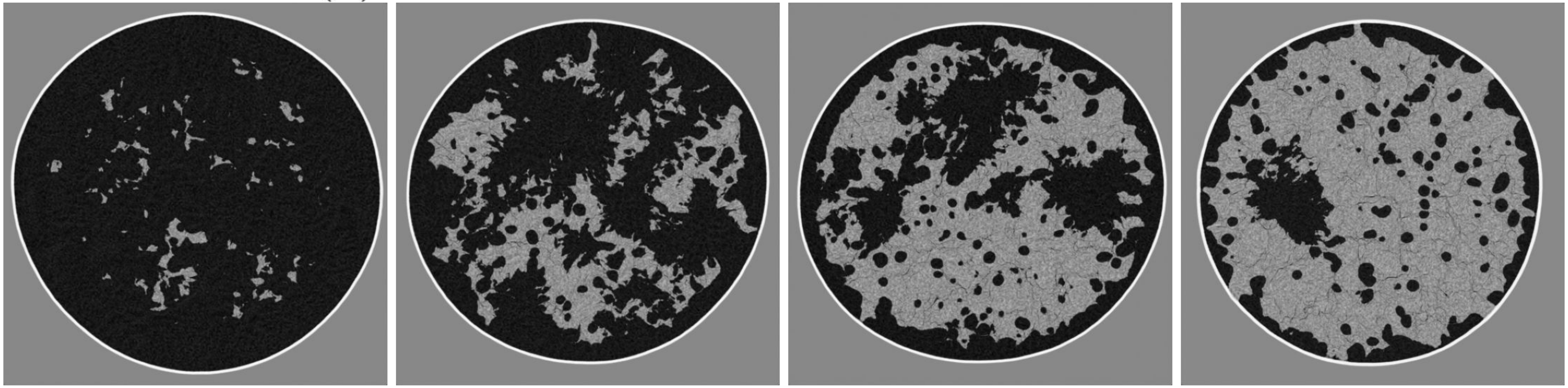
Visual comparison between 'real' and 'fake' samples

- Examples of 'real' and 'fake' breast sound speed maps, [1.412, 1.58] $mm/\mu s$

'real' samples

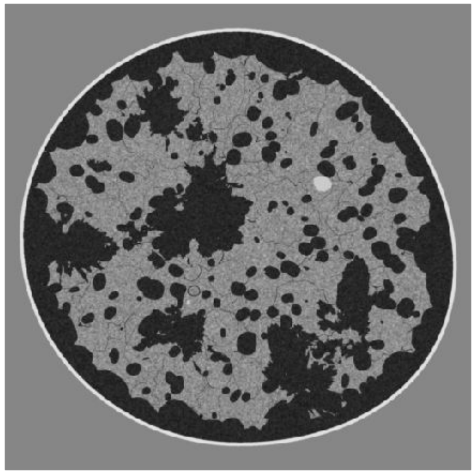


'fake' samples $G(\mathbf{z})$



Evaluation of GAN

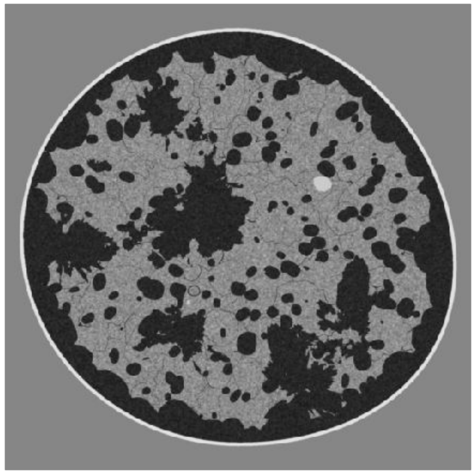
- Representation ability of the learnt generative model



Projection onto latent space

$$\longrightarrow \mathcal{W}^*, \mathcal{N}^* \xrightarrow{G}$$

Mapping from latent space



A breast sound speed map sampled from ‘real’ objects distribution

$G(\mathcal{W}^*, \mathcal{N}^*)$ SSIM: 0.99997

- The formulation of the projection problem

$$\mathcal{W}^*, \mathcal{N}^* = \underset{\mathcal{W}, \mathcal{N}}{\operatorname{argmin}} \|\mathbf{c} - G(\mathcal{W}, \mathcal{N})\|_2^2$$

- SSIM values evaluated on 150 samples of testing dataset
  - Mean(SSIM): 9.9996E-1
  - Stdev(SSIM): 9.21162E-6

Waveform inversion using GAN-based regularizer

- Data fidelity term with the generative model

$$F_s(G(\mathcal{W},\mathcal{N})) = \frac{1}{2}\mathbb{E}_{\mathbf{e}}\{\|\underline{\mathbf{g}}^{\mathbf{e}} - \mathbf{MH}(G(\mathcal{W},\mathcal{N}))\mathbf{s}^{\mathbf{e}}\|^2\}$$

  - The breast sound speed map ***c*** is replaced with *G*(*W, N*)
  - Highly nonlinear and nonconex

- The formulation of a minimization problem

$$\hat{\mathcal{W}},\hat{\mathcal{N}} = \underset{\mathcal{W},\mathcal{N}}{\operatorname{argmin}}\, F_s(G(\mathcal{W},\mathcal{N}))$$

  - The final estimate $\hat{\mathbf{c}} = G(\hat{\mathcal{W}},\hat{\mathcal{N}})$

- The optimization process
  - SGD update formula based on the chain rule

$$\mathcal{W}_{k+1} = \mathcal{W}_k - \beta_k\{\nabla_{\mathbf{c}} f(\mathbf{c}_k,\mathbf{e}_k)\nabla_{\mathcal{W}} G(\mathcal{W}_k,\mathcal{N}_k)\}$$

$$\mathcal{N}_{k+1} = \mathcal{N}_k - \beta_k\{\nabla_{\mathbf{c}} f(\mathbf{c}_k,\mathbf{e}_k)\nabla_{\mathcal{N}} G(\mathcal{W}_k,\mathcal{N}_k)\}$$

Numerical study

- Sparse view USCT data, forward modeling, and reconstruction

Simulation of pressure measurement data

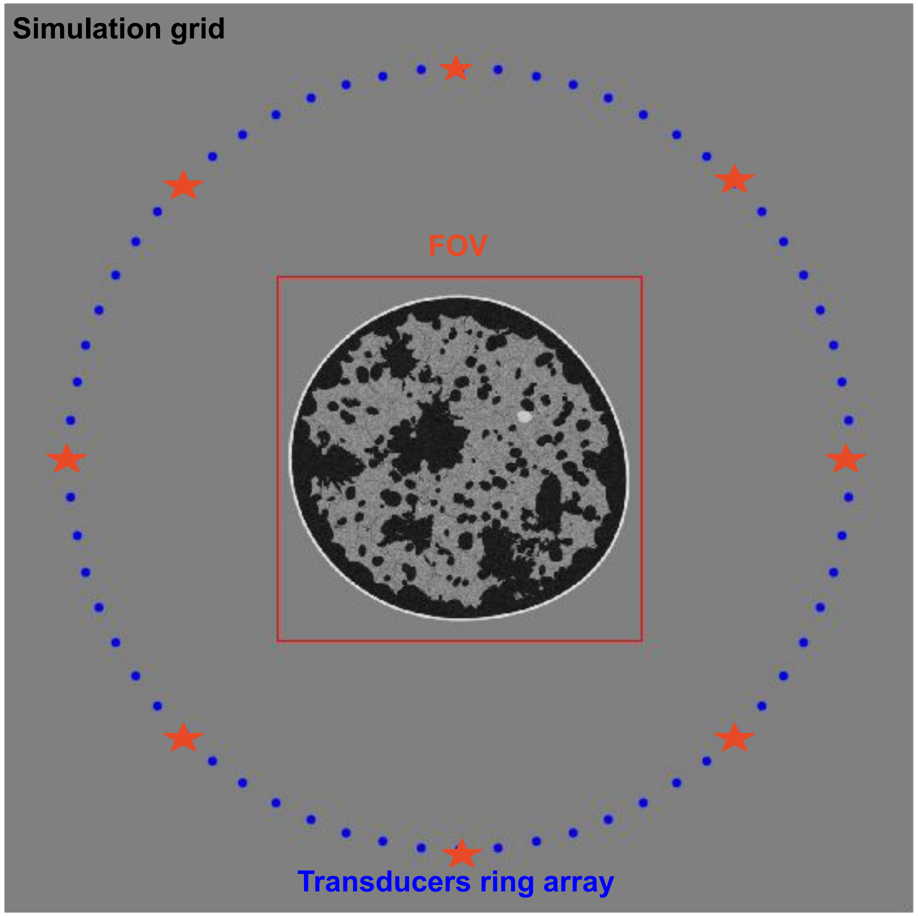
- Grid points: 1280 X 1280
- Pixel size: 0.2 mm (256 X 256 mm)
- Number of time points: 4250
- Time step: 0.04 μs
- No measurement noise

Transducers

- Pointwise type
- Number of emitters and receivers: **8, 64**
  ⇒ less than common USCT setting
- Radius of transducer array: 110 mm

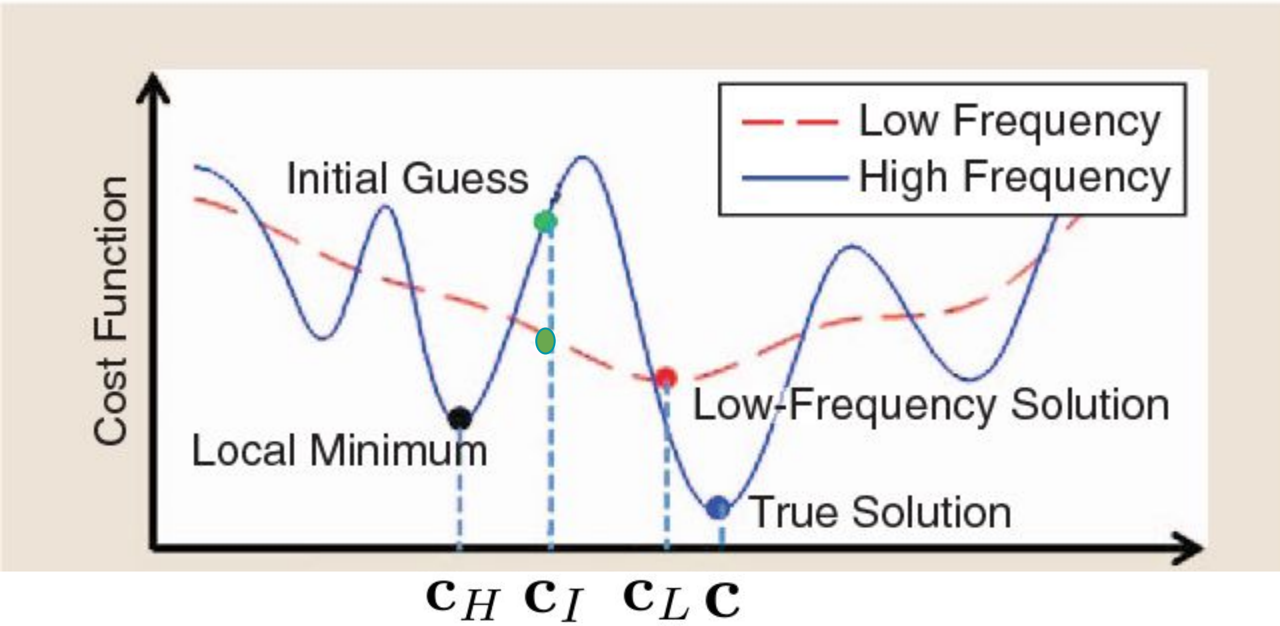
Reconstruction

- Same discretization parameters as those used to generate data (**Inverse crime setting**)
- FOV: 512 X 512 (102.4 X 102.4 mm)



Inversion strategy

- Nonlinearty and nonconvexity of the cost function [Hu et al. (2017)]



  - Due to cycle-skipping phenomena
  - <u>Even worse due to the nonlinearty of the generative model</u>

- Inversion strategy

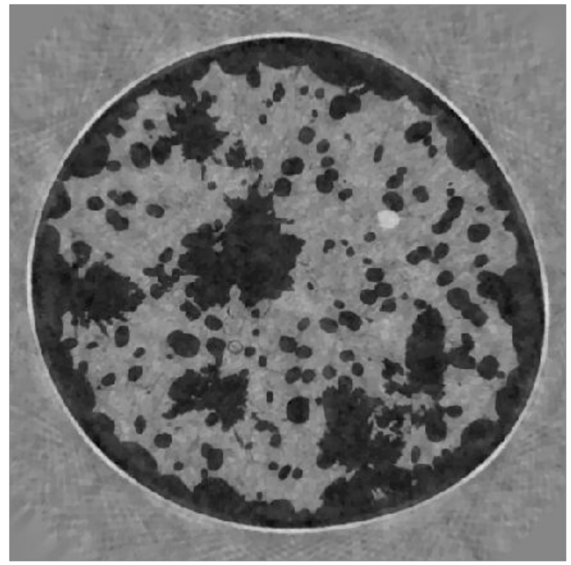
Step 1. Low-frequency waveform inversion

Step 2. Full-frequency waveform inversion

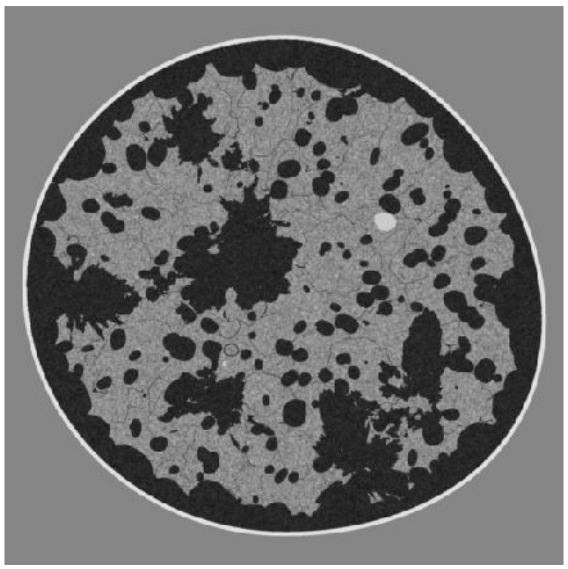
Numerical study

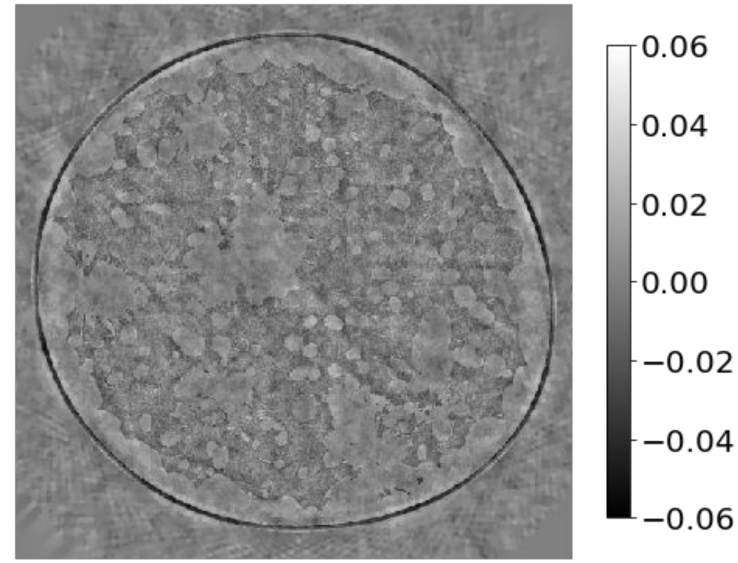
- Waveform inversion results from sparse views, [1.42, 1.58] (mm/μs)

Standard waveform inversion with TV



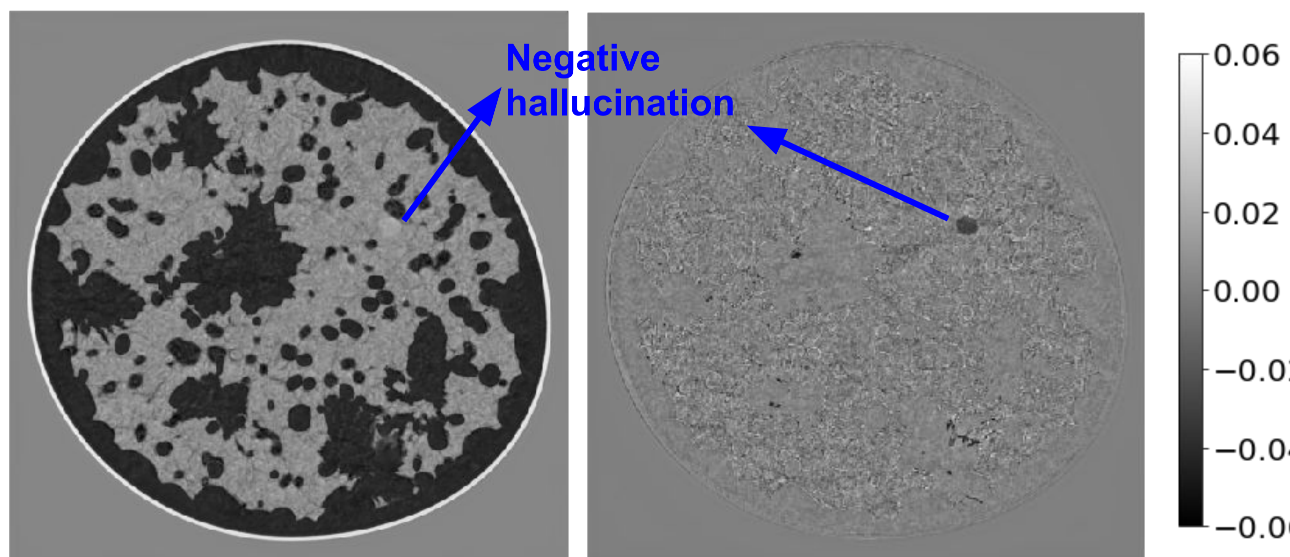
True sound speed map





Estimate w/ TV regularization
SSIM: 0.7108

GAN-based waveform inversion



Estimate w/ GAN-based regularization
SSIM:0.8293

Conclusions and future work

1. Waveform inversion is a method to reconstruct quantitatively accurate breast sound speed map

2. However, classic regularization method may not be sufficient to compensate ill-posedness

3. Generative models can successfully model the distribution of realistic breast sound speed maps

4. This preliminary study suggests that the waveform inversion method with the data-driven regularizer may be a better alternative to the classic regularizer (e.g., TV) when the measurement data is extremely sparse.

5. However, there may be possibility of bad hallucinations for performing detection tasks.

Future work: Systematic analyses for other sparse view settings, noisy measurements, and model mismatch

References

[1] Badano, A., Graff, C. G., Badal, A., Sharma, D., Zeng, R., Samuelson, F. W., ... & Myers, K. J. (2018). Evaluation of digital breast tomosynthesis as replacement of full-field digital mammography using an in silico imaging trial. *JAMA network open*, 1(7), e185474-e185474.

[2] Bora, A., Jalal, A., Price, E., & Dimakis, A. G. (2017, July). Compressed sensing using generative models. *In International Conference on Machine Learning* (pp. 537-546). PMLR.

[3] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems,* 27.

[4] Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and improving the image quality of stylegan. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8110-8119).

[5] Laloy, E., Linde, N., Ruffino, C., Hérault, R., Gasso, G., & Jacques, D. (2019). Gradient-based deterministic inversion of geophysical data with generative adversarial networks: Is it feasible?. *Computers & Geosciences,* 133, 104333.

[6] Li, F., Villa, U., Park, S., & Anastasio, M. A. (2021). Three-dimensional stochastic numerical breast phantoms for enabling virtual imaging trials of ultrasound computed tomography. *arXiv preprint arXiv*:2106.02744.

[7] Mast, T. D., Souriau, L. P., Liu, D. L., Tabei, M., Nachman, A. I., & Waag, R. C. (2001). A k-space method for large-scale models of wave propagation in tissue. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 48(2), 341-354.

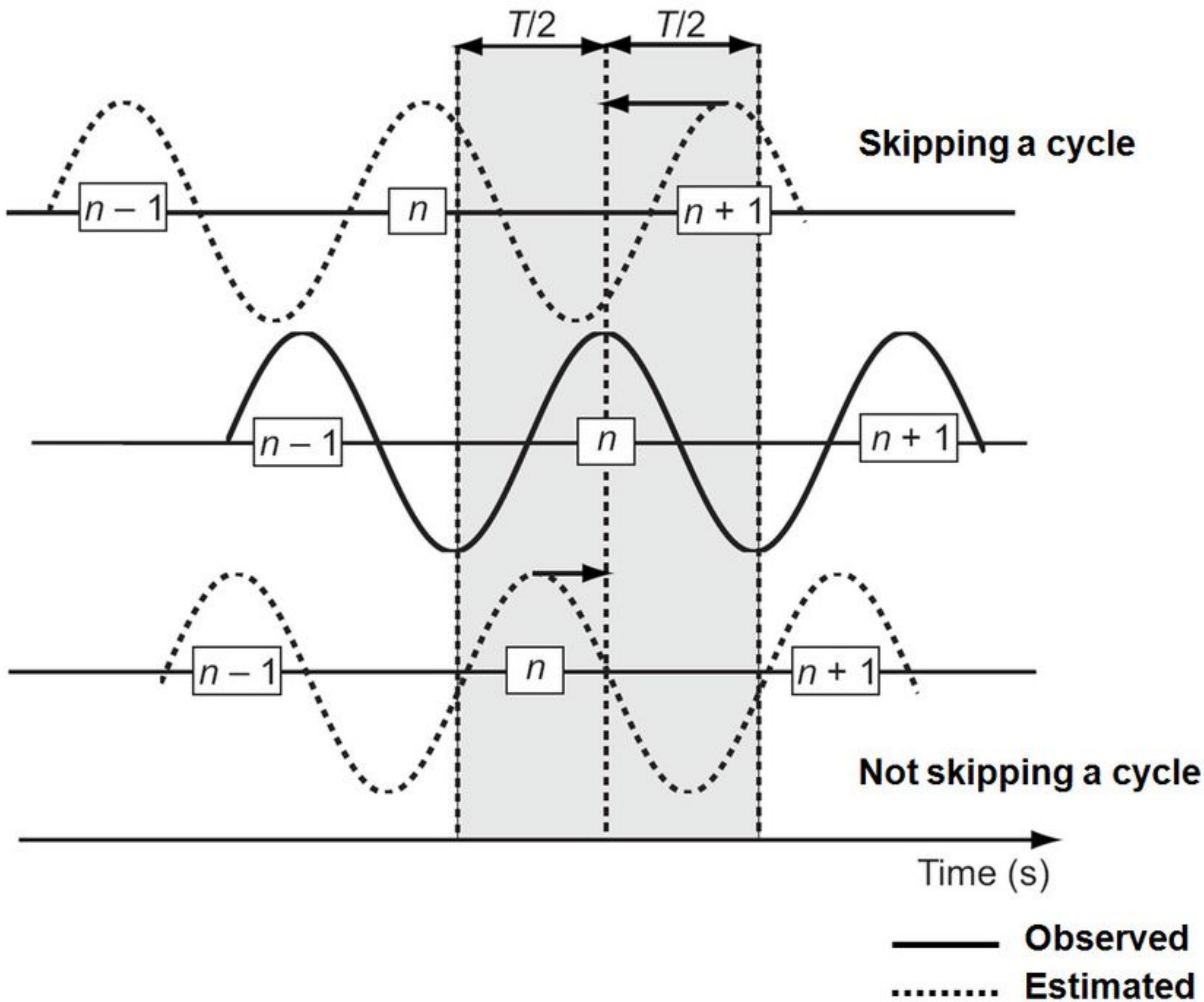
[8] Mosser, L., Dubrule, O., & Blunt, M. J. (2020). Stochastic seismic waveform inversion using generative adversarial networks as a geological prior. *Mathematical Geosciences*, 52(1), 53-79.

[9] Singh, S., Singhal, V., & Majumdar, A. (2016). Deep blind compressed sensing. *arXiv preprint arXiv:1612.07453*.

[10] Wang, K., Matthews, T., Anis, F., Li, C., Duric, N., & Anastasio, M. A. (2015). Waveform inversion with source encoding for breast sound speed reconstruction in ultrasound computed tomography. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control,* 62(3), 475-493.

Thank you

Cycle-skipping phenomena

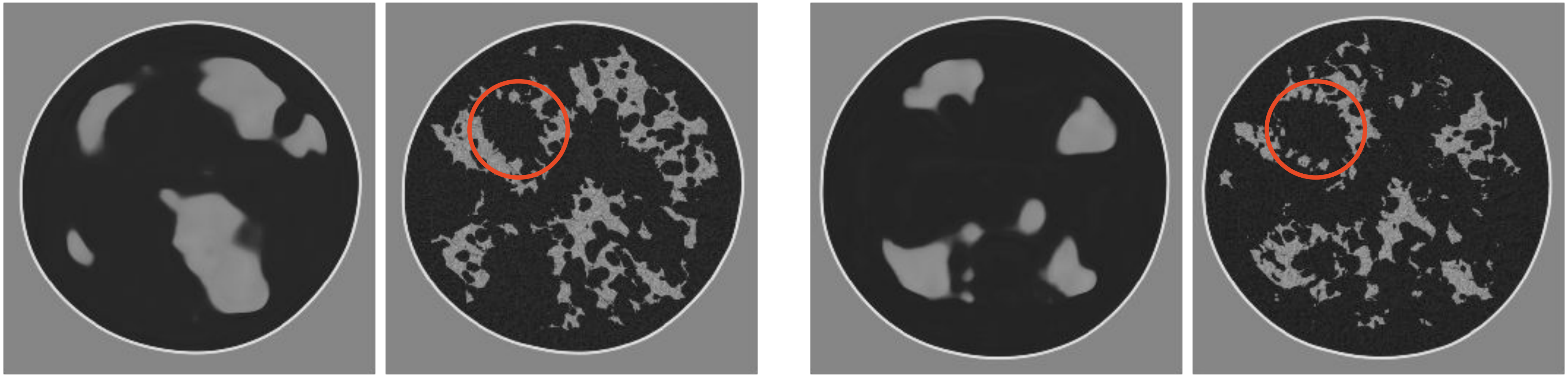


Generated NBPs using StyleGAN2

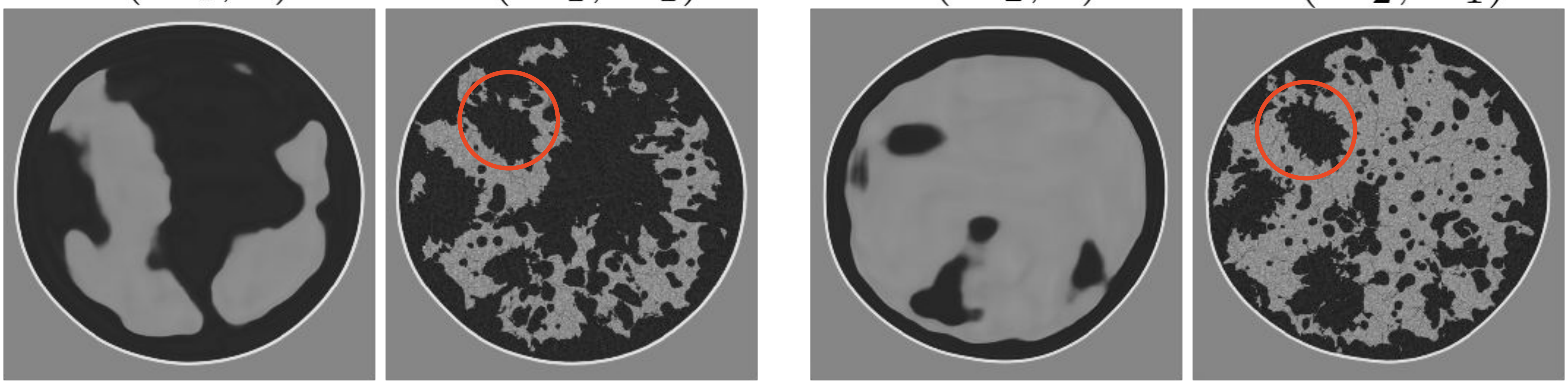
$\mathcal{W}$: style latent vectors

$\mathcal{N}$: noise latent vectors

- Different style vectors with the same noise vectors



$G(\mathcal{W}_1, \mathbf{0})$ $G(\mathcal{W}_1, \mathcal{N}_1)$ $G(\mathcal{W}_2, \mathbf{0})$ $G(\mathcal{W}_2, \mathcal{N}_1)$



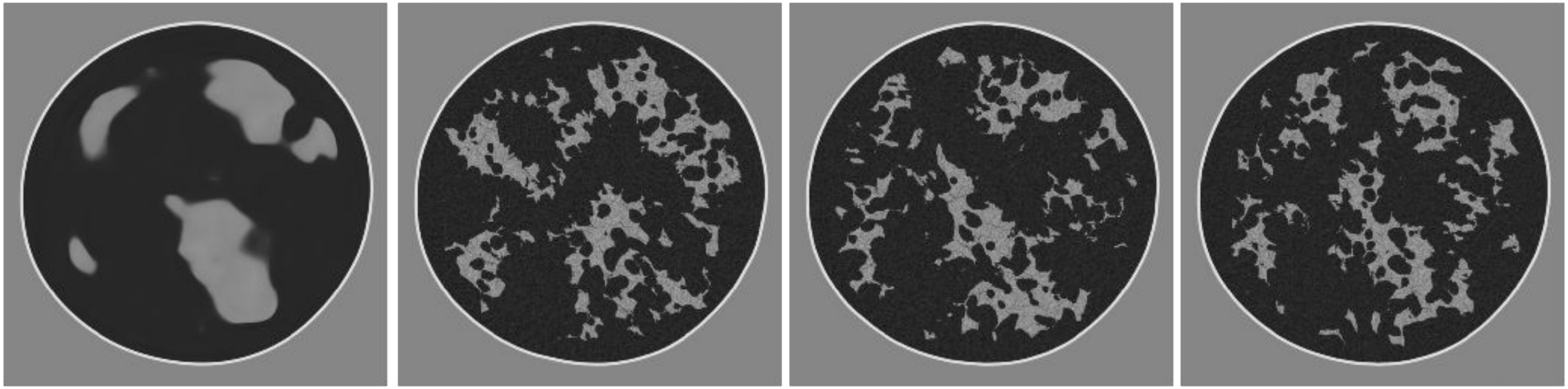
$G(\mathcal{W}_3, \mathbf{0})$ $G(\mathcal{W}_3, \mathcal{N}_1)$ $G(\mathcal{W}_4, \mathbf{0})$ $G(\mathcal{W}_4, \mathcal{N}_1)$

Generated NBPs using StyleGAN2

$\mathcal{W}$: style latent vectors

$\mathcal{N}$: noise latent vectors

- The same style vectors with the different noise vectors



$G(\mathcal{W}_1, \mathbf{0})$ $G(\mathcal{W}_1, \mathcal{N}_1)$ $G(\mathcal{W}_1, \mathcal{N}_2)$ $G(\mathcal{W}_1, \mathcal{N}_3)$

- Style vectors determine
  - The breast size, shape, location, concentration of density, …
- Noise vectors determine
  - The details of fat, glandular, ligaments, vein, heterogeneous texture, ...