MAY 2024

TEXAS
The University of Texas at Austin

GSURE Denoising enables training of higher quality generative priors for accelerated Multi-Coil MRI Reconstruction

Authors: **Asad Aali**[1], Marius Arvinte[2], Sidharth Kumar[1], Yamin Arefeen[1], Jon Tamir[1]

[1]The University of Texas at Austin

[2]Intel Labs

Motivation

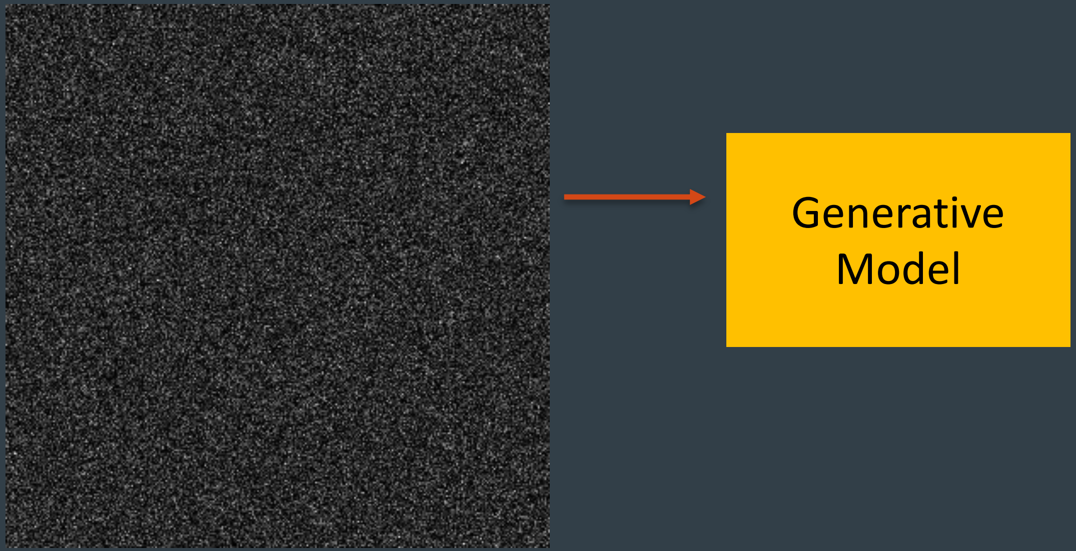
- Deep learning based accelerated MRI reconstruction with Generative Models
  - ✓ Develops a prior with large training databases
  - ✓ Decouples from the forward model

Training Dataset

Motivation

- Generative models are powerful image generators

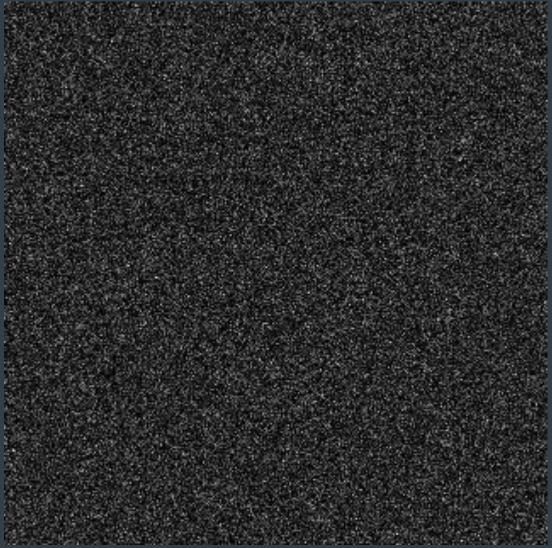
Sample from Gaussian Distribution



Motivation

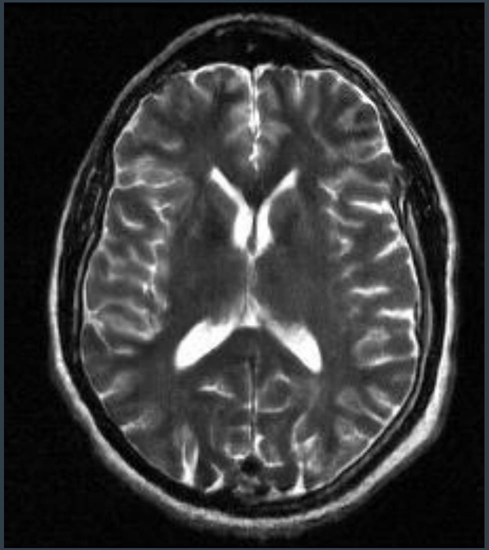
- Generative models are powerful image generators

Sample from Gaussian Distribution



Generative Model

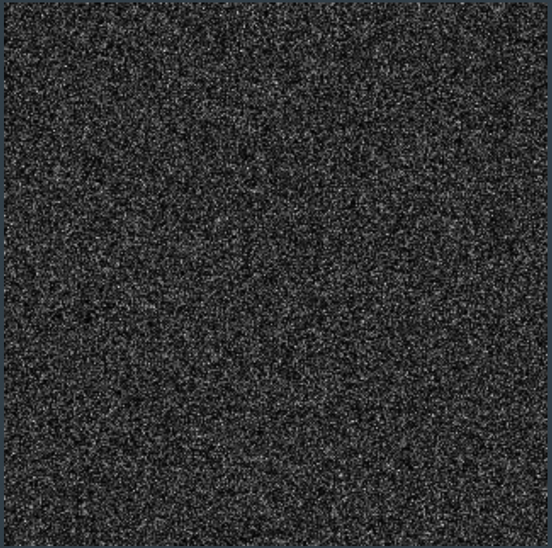
Sample from Image Distribution



Motivation

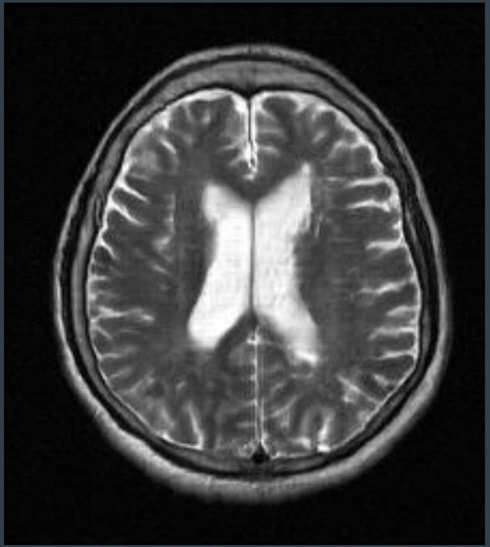
- Generative models are powerful image generators

Sample from Gaussian Distribution



Generative Model

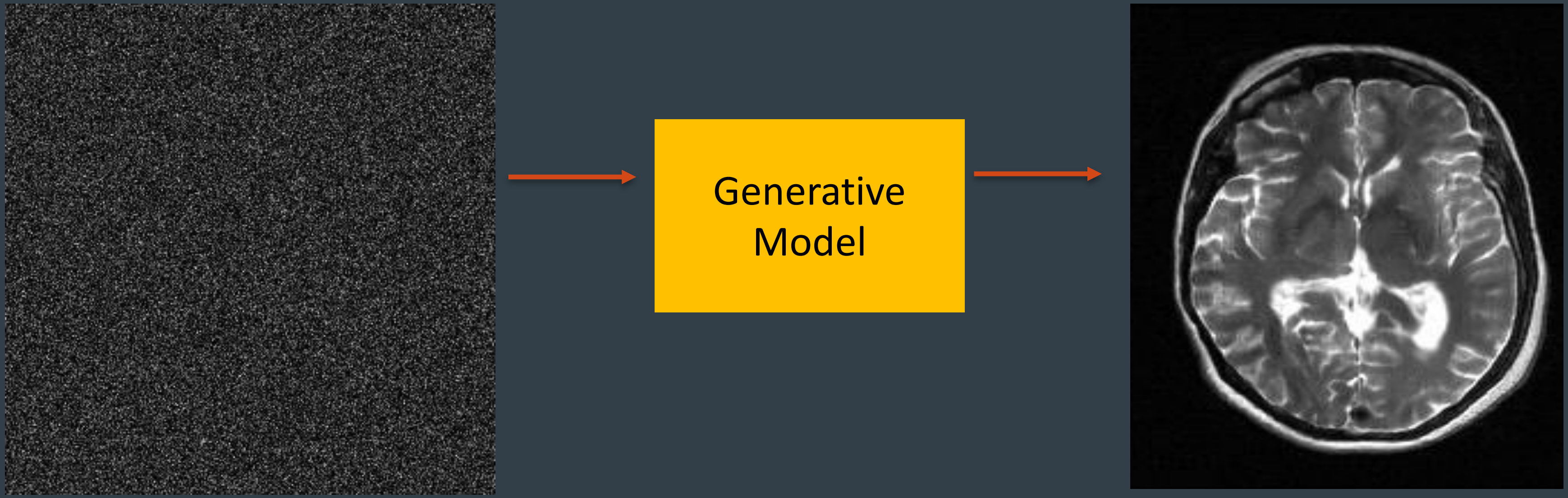
Sample from Image Distribution



Motivation

- Generative models are powerful image generators

Sample from Gaussian Distribution **Sample from Image Distribution**



Motivation

- Generative models are powerful image generators

Sample from Gaussian Distribution

$$\sim \mathcal{N}(z)$$

Image Prior

$\rightarrow$ Generative Model $\rightarrow$

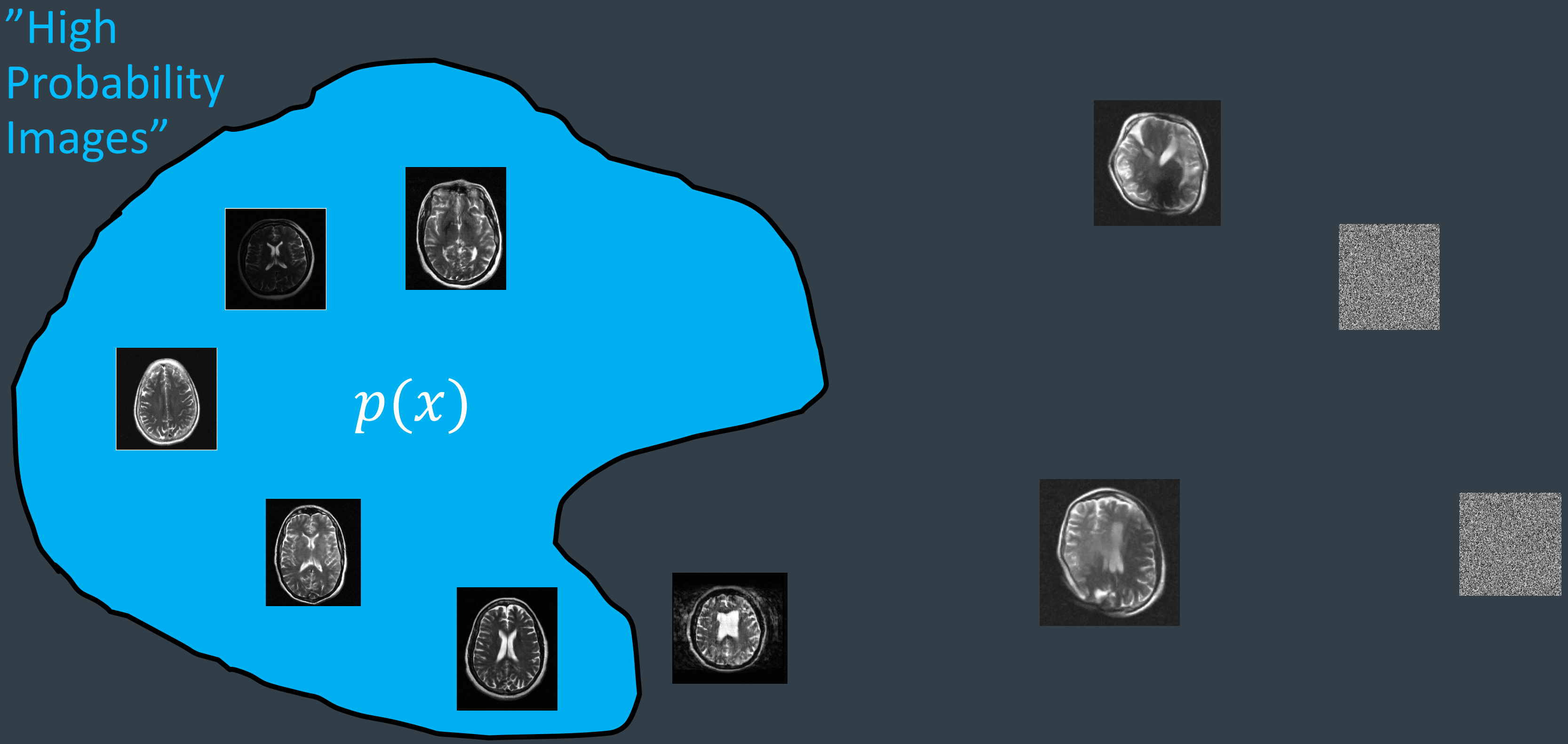
Sample from Image Distribution

$$\sim p(x)$$

Image Prior

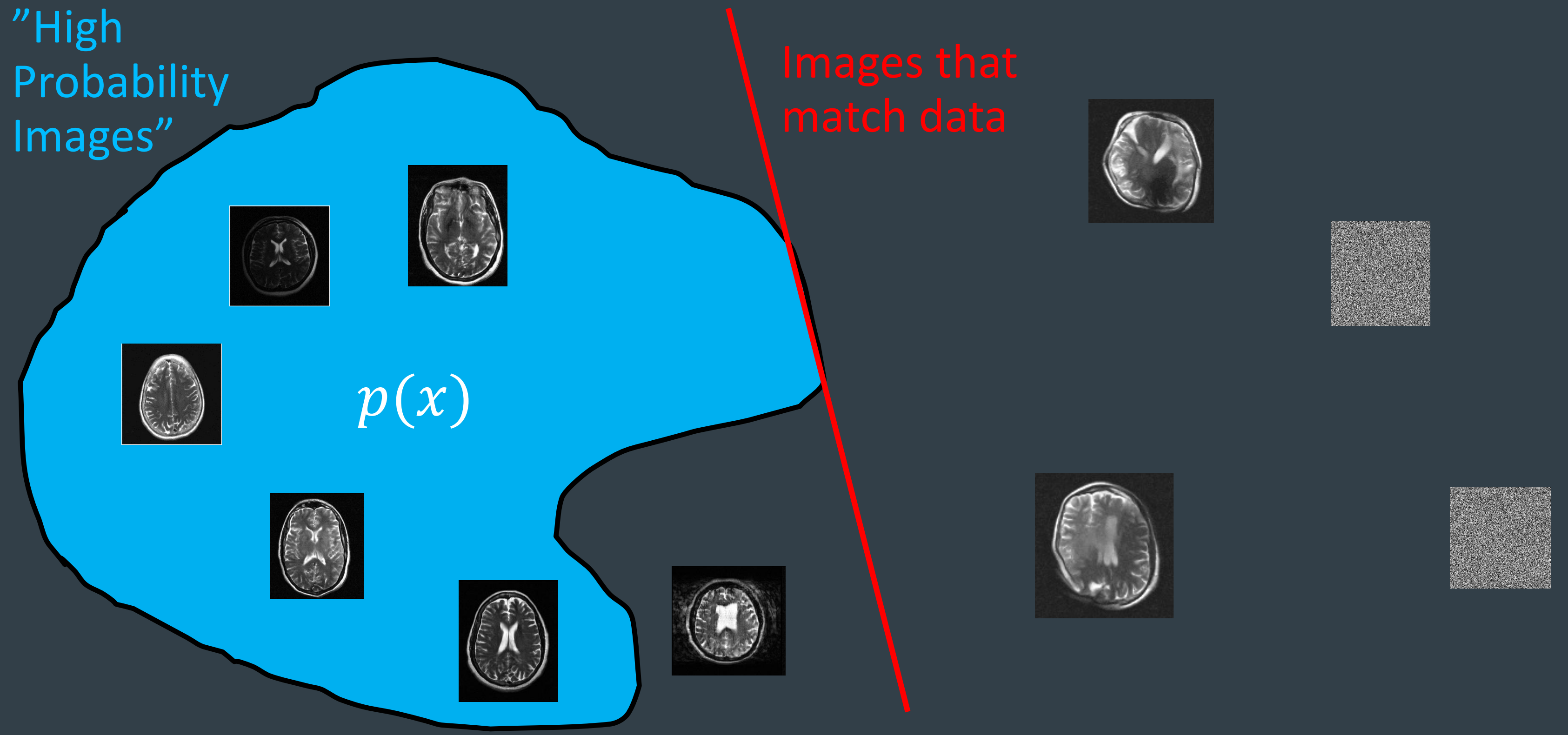
Motivation

- Using Generative Models to guide accelerated MRI reconstructions.



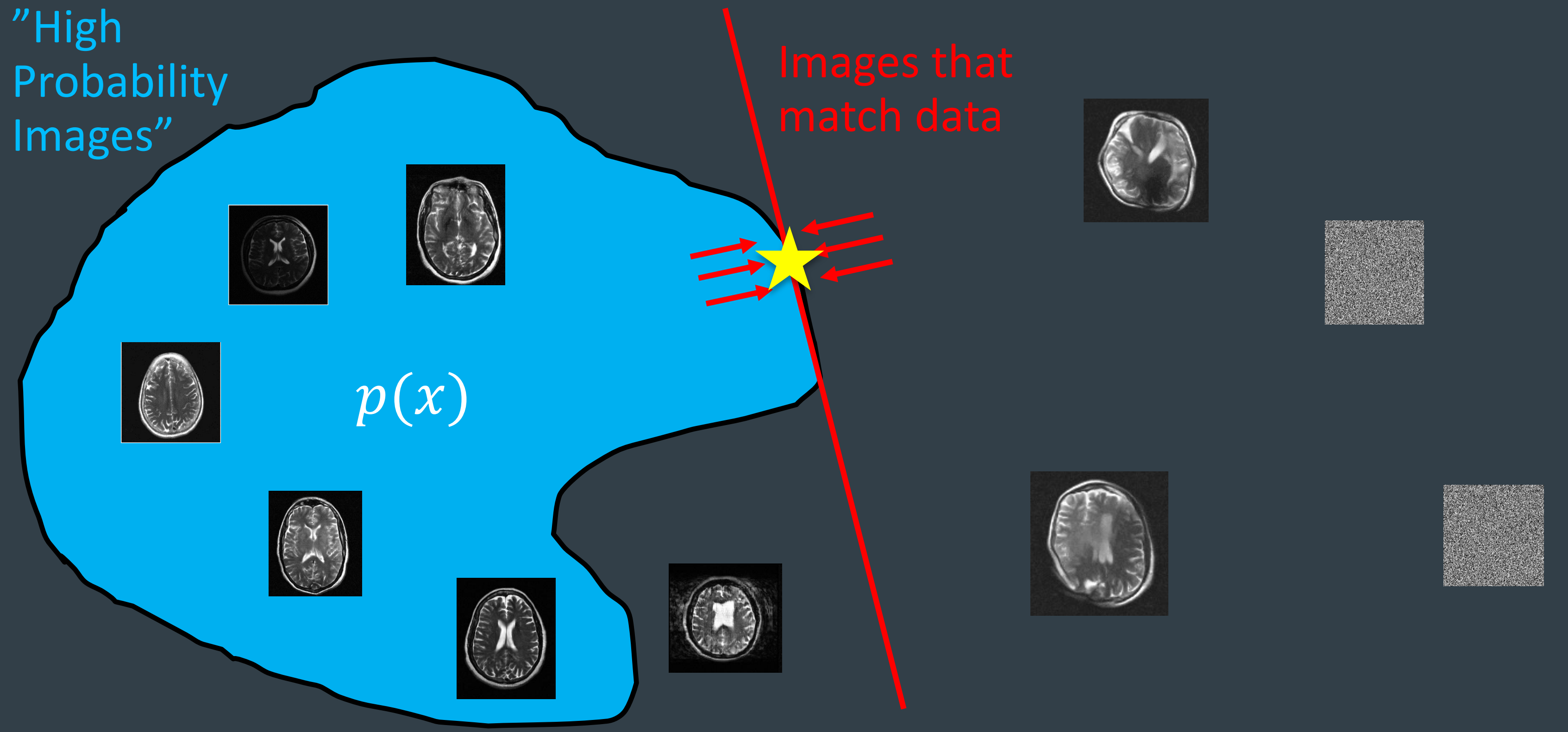
Motivation

- Using Generative Models to guide accelerated MRI reconstructions.



Motivation

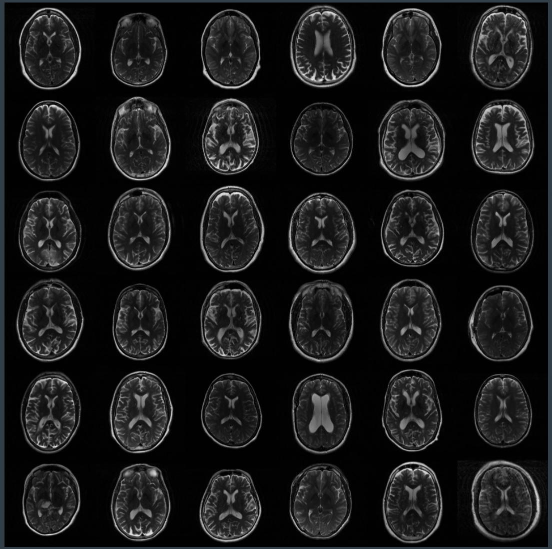
- Using Generative Models to guide accelerated MRI reconstructions.



Motivation

- Generative Models rely on large amounts of ***high-quality data***.

Training Dataset

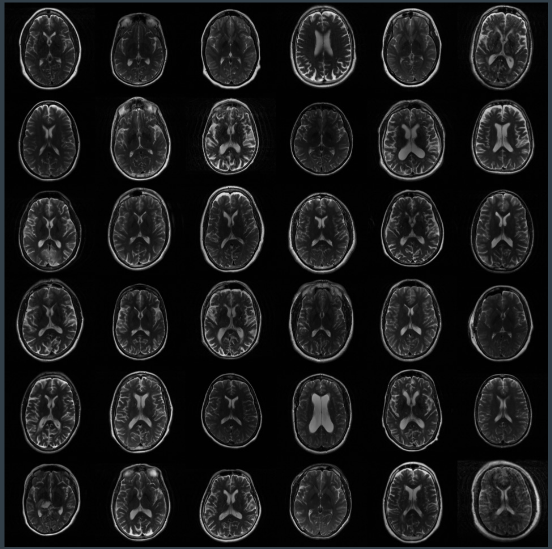


[1]FastMRI (https://fastmri.med.nyu.edu/), [2]Wang, *MRM* (2024)

Motivation

- Generative Models rely on large amounts of ***high-quality data***.
- MRI data are ***inherently noisy***[1,2].

Training Dataset

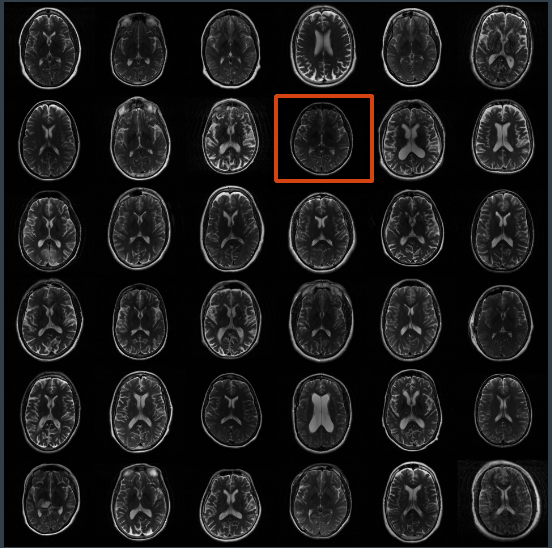


[1]FastMRI (https://fastmri.med.nyu.edu/), [2]Wang, *MRM* (2024)

Motivation

- Generative Models rely on large amounts of ***high-quality data***.
- MRI data are ***inherently noisy***[1,2].

Training Dataset

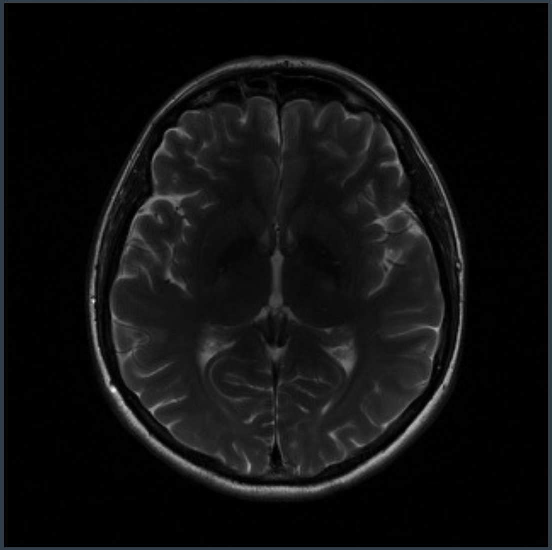


[1]FastMRI (https://fastmri.med.nyu.edu/), [2]Wang, *MRM* (2024)

Motivation

- Generative Models rely on large amounts of ***high-quality data***.
- MRI data are ***inherently noisy***[1,2].

Training Dataset

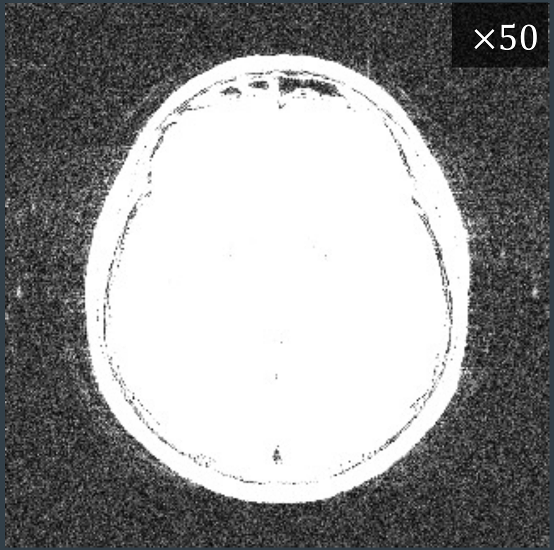


[1]FastMRI (https://fastmri.med.nyu.edu/), [2]Wang, *MRM* (2024)

Motivation

- Generative Models rely on large amounts of ***high-quality data***.
- MRI data are ***inherently noisy***[1,2].

Training Dataset

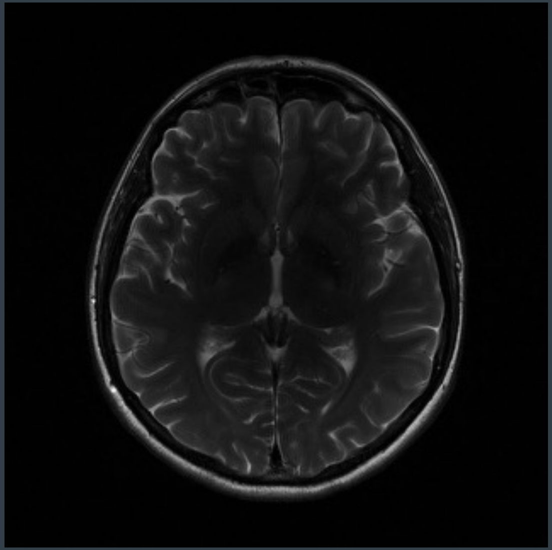


[1]FastMRI (https://fastmri.med.nyu.edu/), [2]Wang, *MRM* (2024)

Motivation

- Generative Models rely on large amounts of ***high-quality data***.
- MRI data are ***inherently noisy***[1,2].

Training Dataset

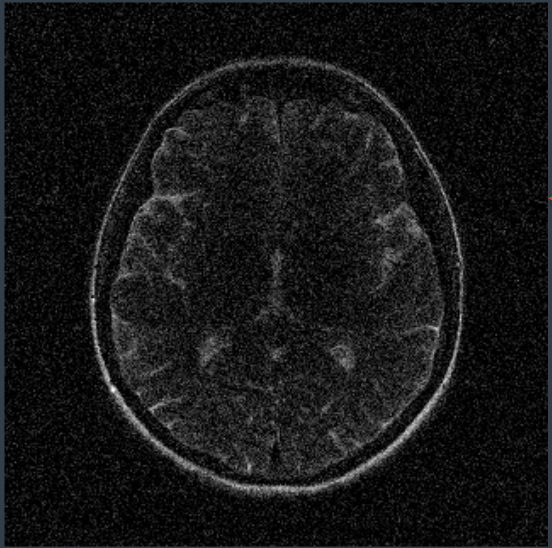


[1]FastMRI (https://fastmri.med.nyu.edu/), [2]Wang, *MRM* (2024)

Motivation

- Training generative models with noisy datasets leads to a poor prior.

Training Dataset



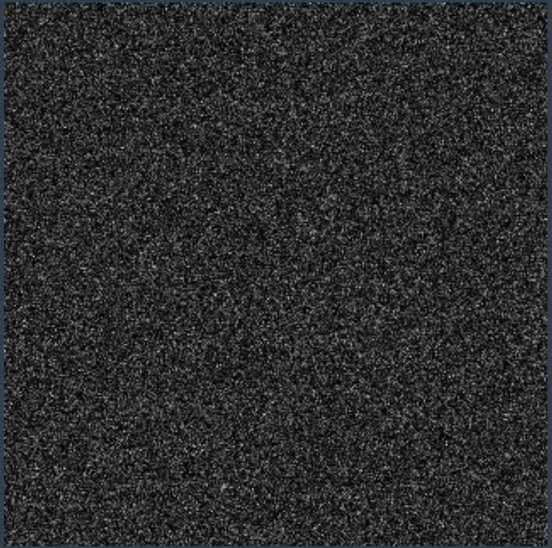
training

Generative Model

Motivation

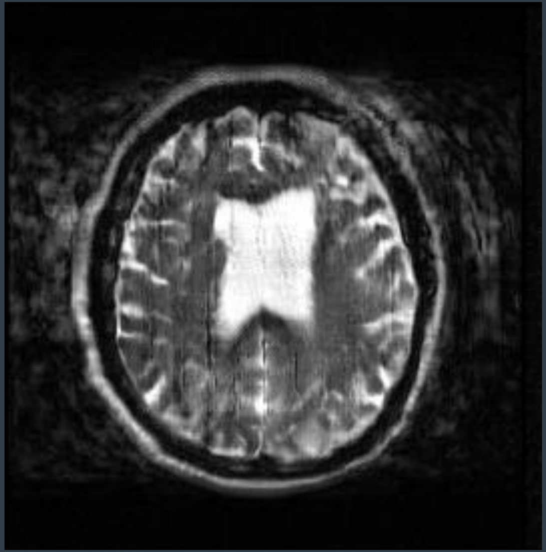
- Training generative models with noisy datasets leads to a poor prior.

Sample from Gaussian Distribution



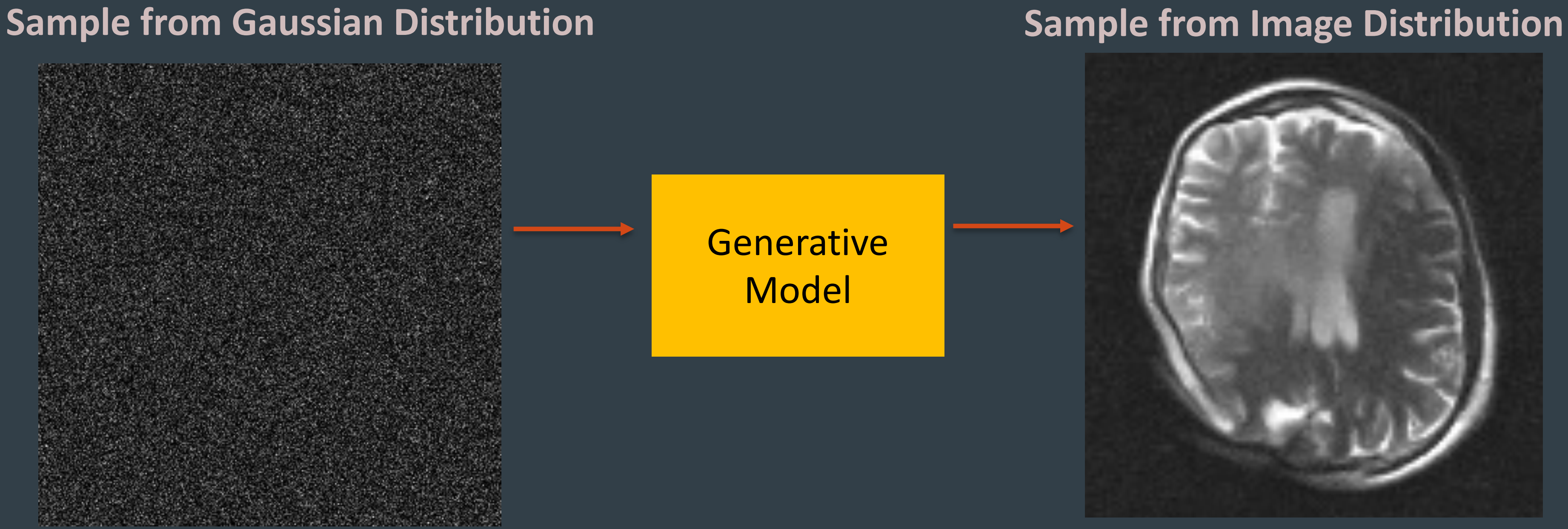
Generative Model

Sample from Image Distribution



Motivation

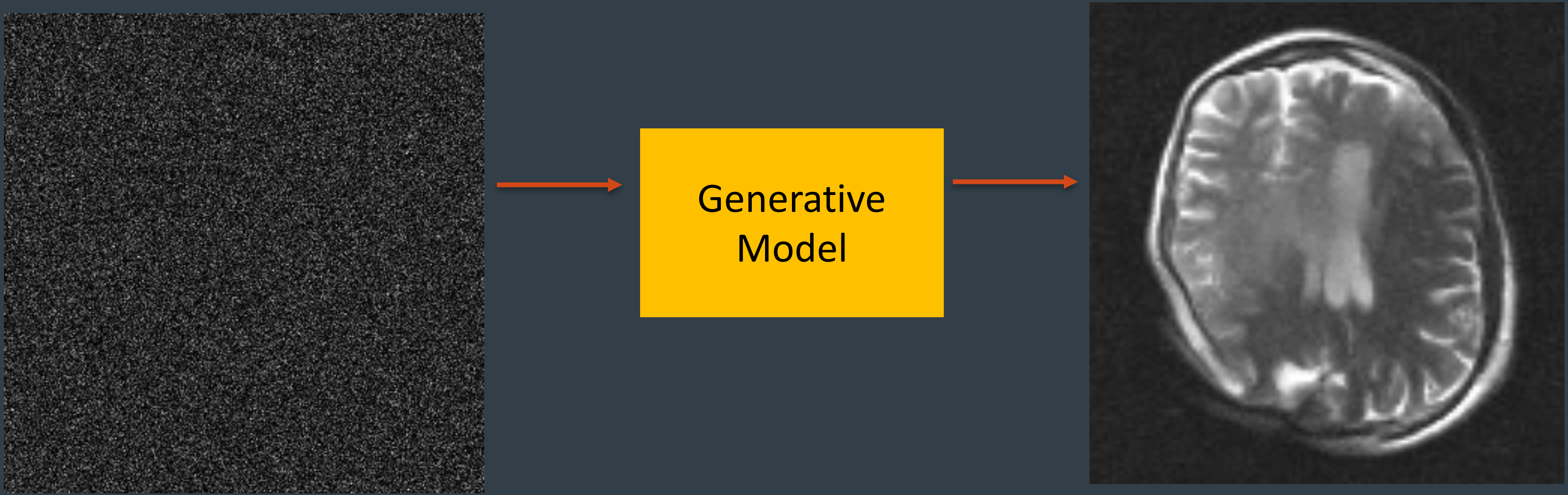
- Training generative models with noisy datasets leads to a poor prior.



Motivation

- Training generative models with noisy datasets leads to a poor prior.
- **Reconstruction performance depends on accuracy of priors**

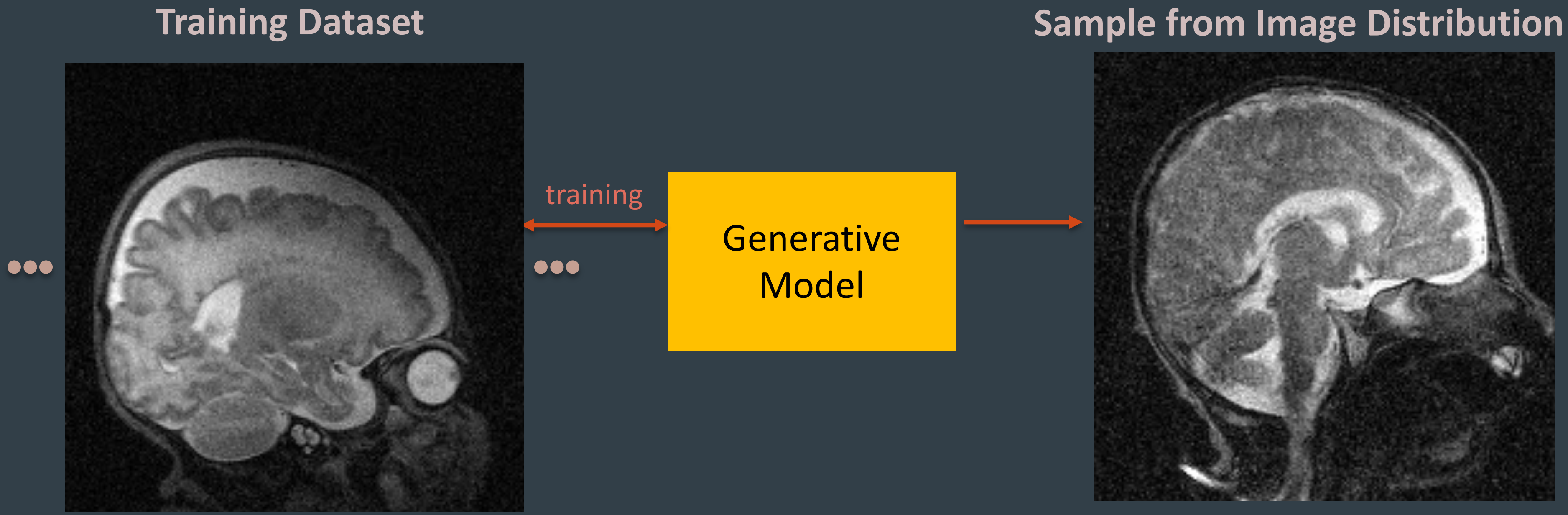
Sample from Gaussian Distribution **Sample from Image Distribution**



Motivation

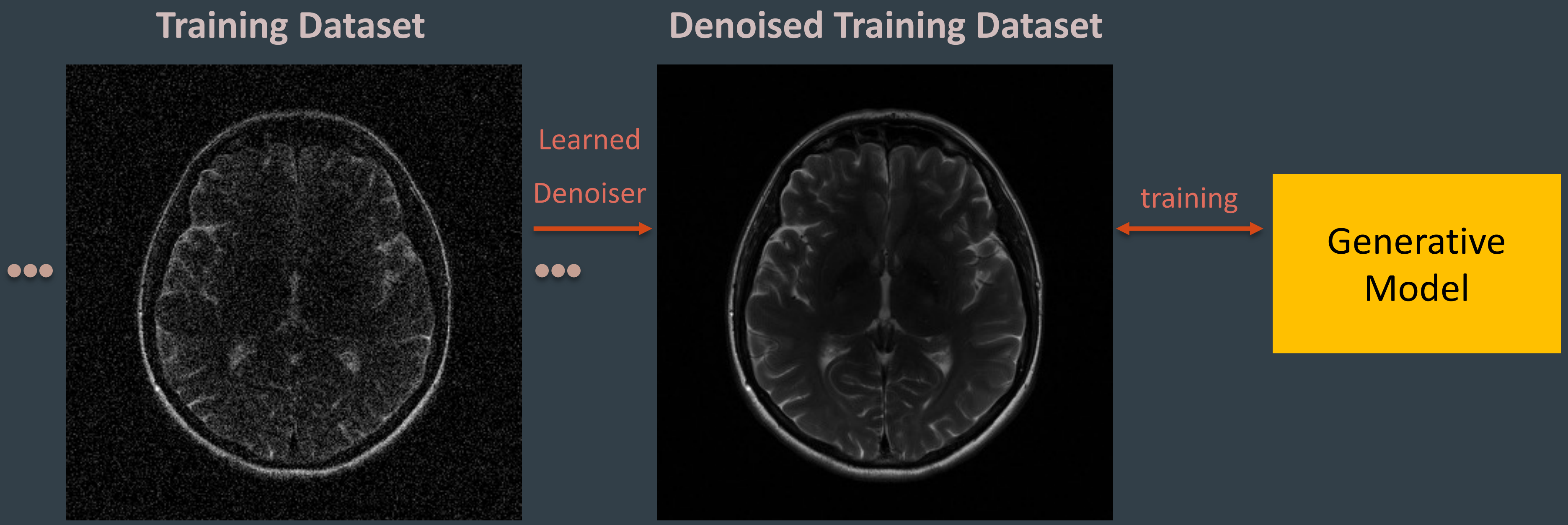
Application in real world datasets: low field neo-natal MRI

- Training generative models with noisy datasets leads to a poor prior.
- **Reconstruction performance depends on accuracy of priors**



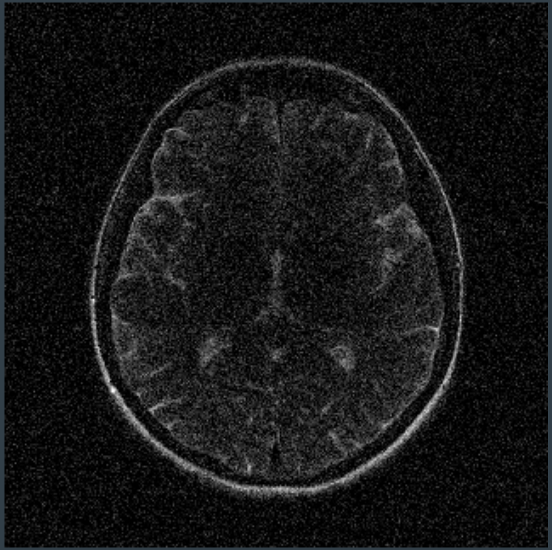
Purpose

- Use a learned denoiser to denoise the training dataset before training the generative model



Purpose

Training a learned denoiser without access clean training samples?



Purpose

Investigate the **feasibility** and **effectiveness** of self-supervised **denoising** MRI samples as a pre-processing step to learning **generative priors** for accelerated MRI reconstruction

Problem Formulation

Problem Formulation

Goal is to learn the **clean distribution** using *noisy* data (i.i.d Gaussian, with known power σ_w^2).

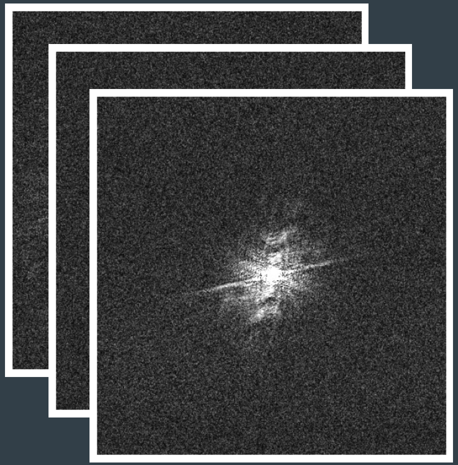
$$y = Ax + noise$$

Problem Formulation

Goal is to learn the **clean distribution** using *noisy* data (i.i.d Gaussian, with known power σ_w^2).

$$y = Ax + noise$$

y

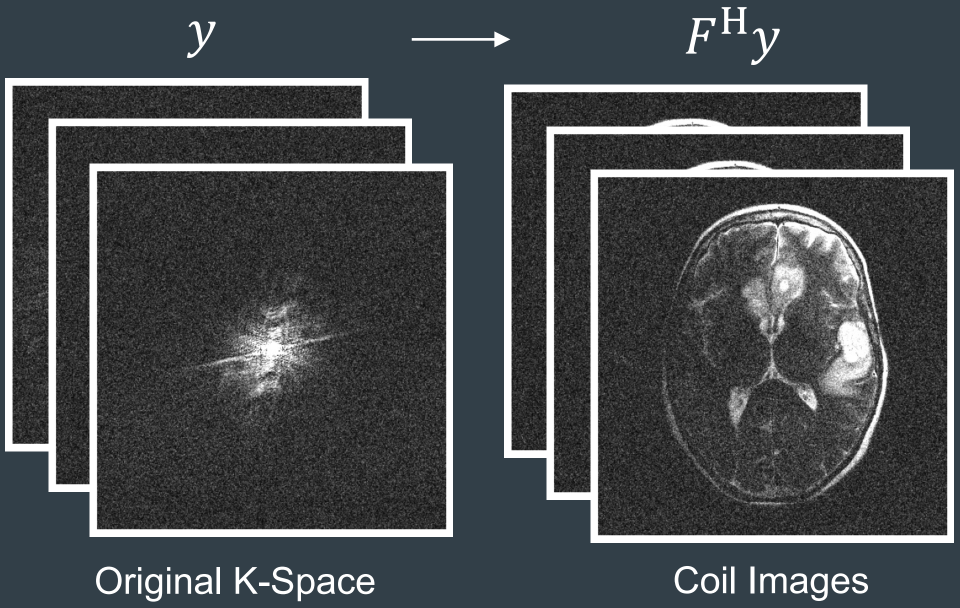


Original K-Space

Problem Formulation

Goal is to learn the **clean distribution** using *noisy* data (i.i.d Gaussian, with known power σ_w^2).

$$y = Ax + noise$$

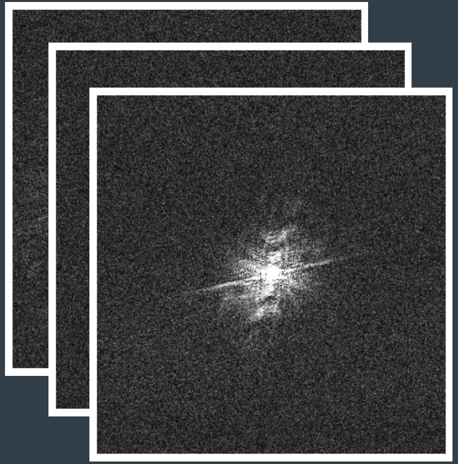


Problem Formulation

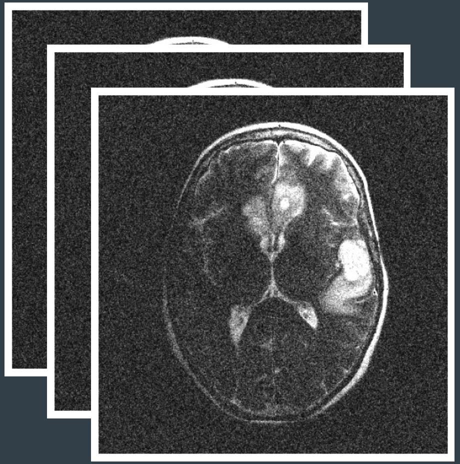
Goal is to learn the **clean distribution** using *noisy* data (i.i.d Gaussian, with known power σ_w^2).

$$y = Ax + noise$$

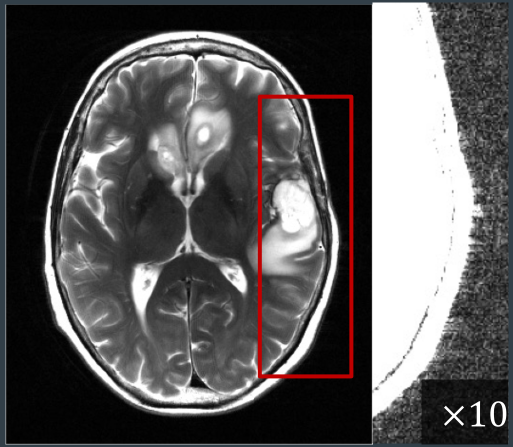
$y \longrightarrow F^{\mathrm{H}}y \longrightarrow \boxed{\hat{x}_{noisy} = A^H y}$



Original K-Space



Coil Images

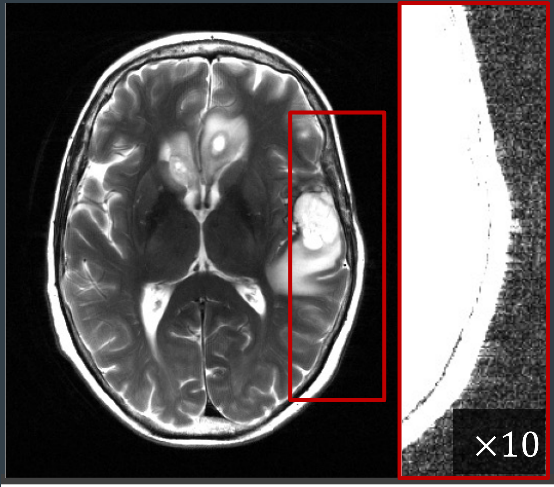


Noisy MRI Sample

Proposed Methods

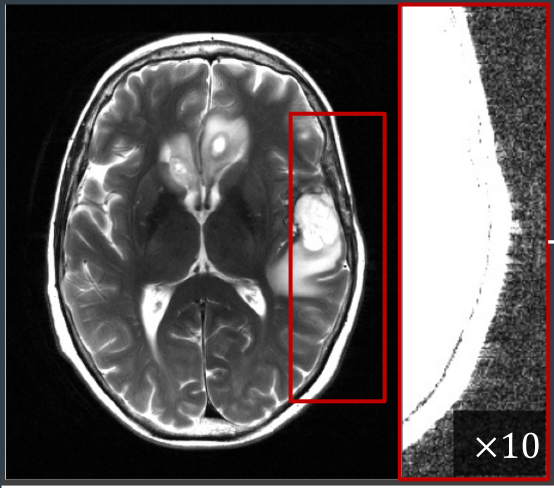
Proposed Methods

$\hat{x}_{noisy}$



Proposed Methods

$\hat{x}_{noisy}$

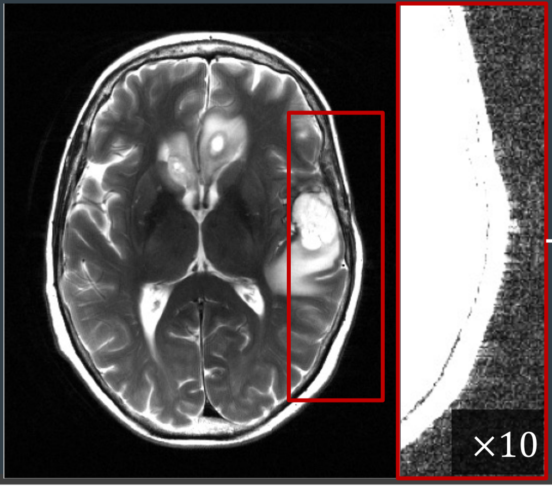


Denoiser Network
g_ϕ

$g_\phi(\hat{x}_{noisy})$

Proposed Methods

$\hat{x}_{noisy}$

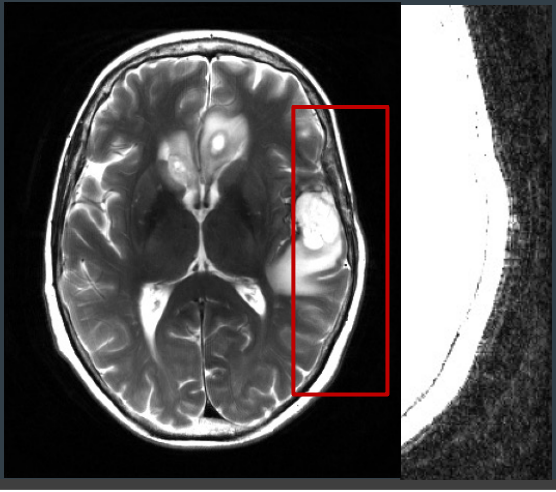


Denoiser Network

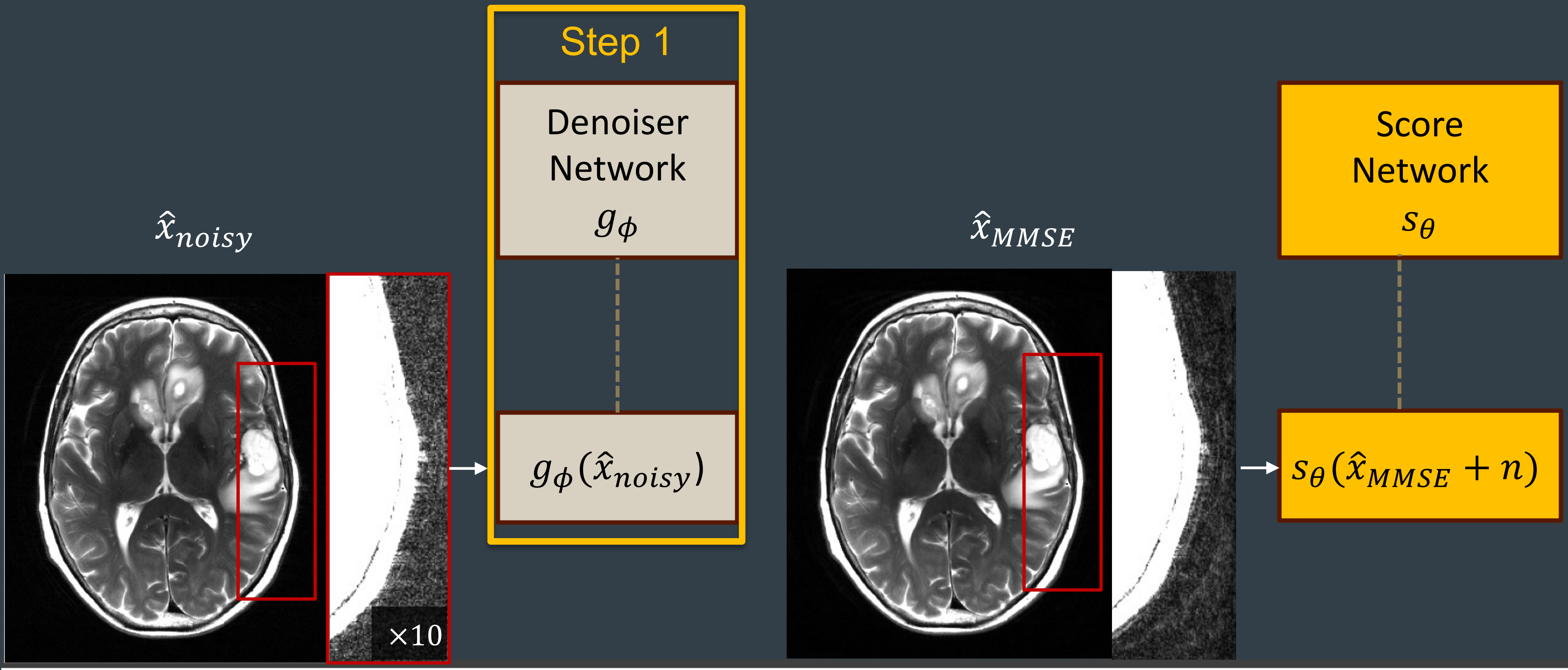
g_ϕ

$g_\phi(\hat{x}_{noisy})$

$\hat{x}_{MMSE}$



Proposed Methods



Self-Supervised Denoising

Goal is to learn the **clean distribution** using *noisy* data (i.i.d Gaussian, with known power σ_w^2).

A is a Linear Forward Operator (Fully-Sampled) -> **GSURE**[1,2,3]

$$y = FSx + noise$$

[1]Soltanayev, *NeurIPS,* 2018, [2]Eldar, *IEEE Transactions on Signal Processing,* 2008, [3]Kawar, *TMLR,* 2023

Generalized SURE (GSURE) Basics

- GSURE[1]: Self-supervised denoising technique, only need access to:
  - $\hat{x}_{noisy}$ → **Noisy Samples**
  - Noise **Covariance Matrix**

- An unbiased estimate of the MSE

$$E[\mathrm{L}_{GSURE}] = E\left\|g_\phi(\hat{x}_{noisy}) - x\right\|$$

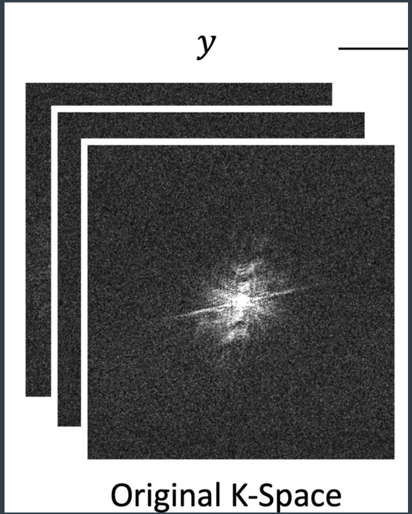
[1]Yonina, *IEEE Transactions on Signal Processing, 2008*

Pre-Processing Noisy Dataset

- GSURE loss requires the **noise covariance** matrix
- **Pre-whitening** (noise covariance $= I$) makes computation relatively straight-forward

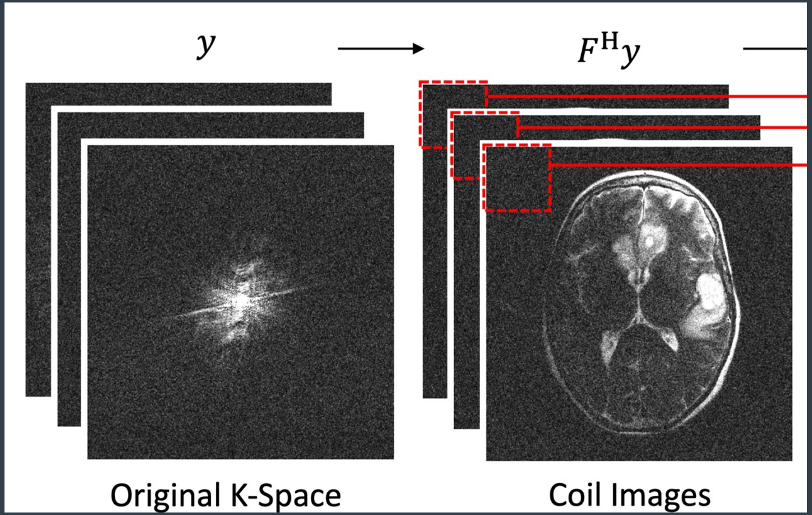
Pre-Processing Noisy Dataset

- GSURE loss requires the **noise covariance** matrix
- **Pre-whitening** (noise covariance $= I$) makes computation relatively straight-forward



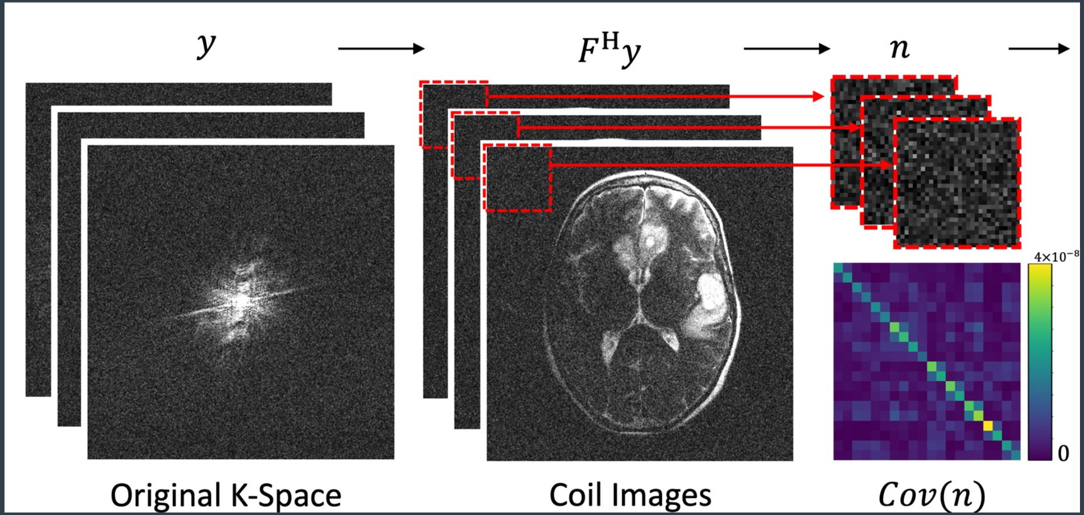
Pre-Processing Noisy Dataset

- GSURE loss requires the **noise covariance** matrix
- **Pre-whitening** (noise covariance $= I$) makes computation relatively straight-forward



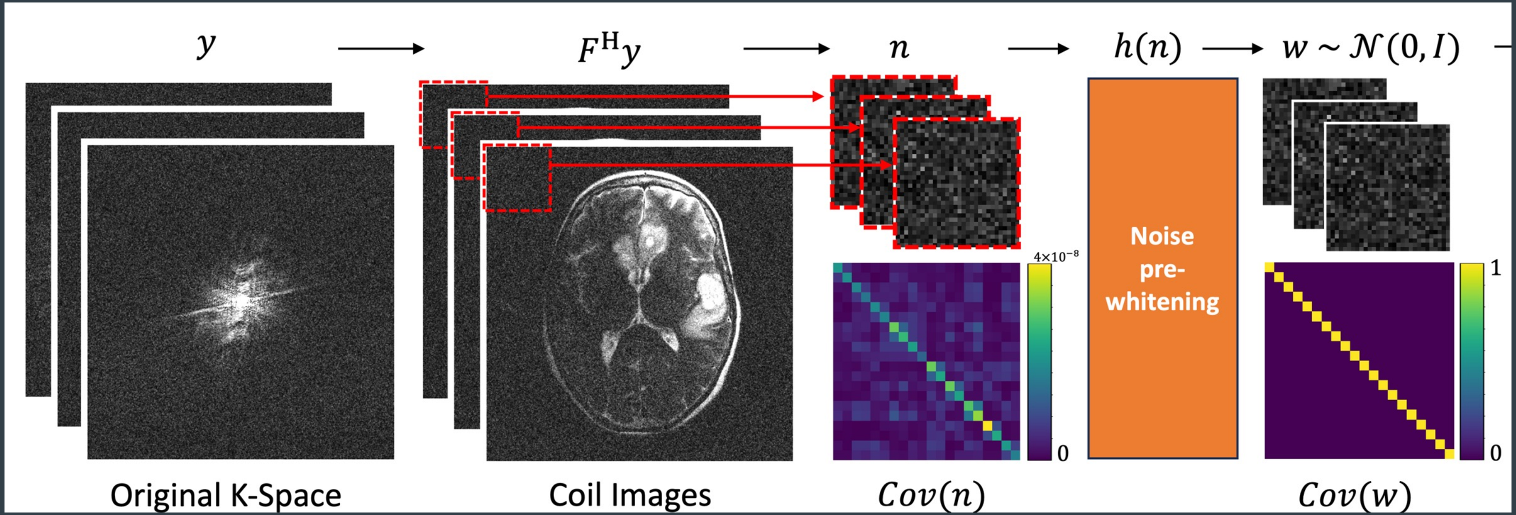
Pre-Processing Noisy Dataset

- GSURE loss requires the **noise covariance** matrix
- **Pre-whitening** (noise covariance $= I$) makes computation relatively straight-forward



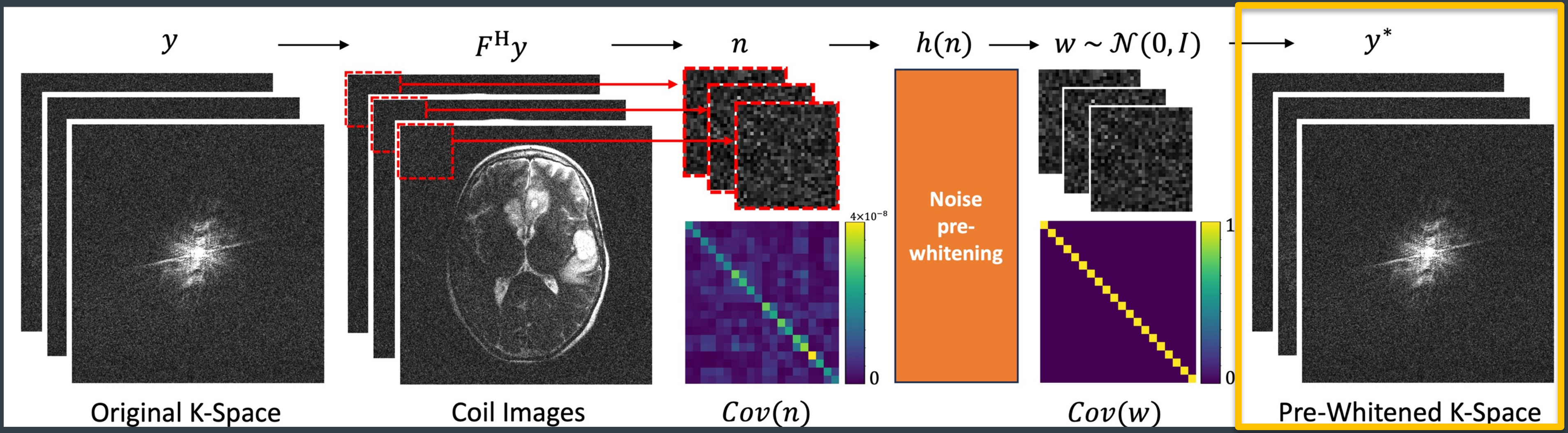
Pre-Processing Noisy Dataset

- GSURE loss requires the **noise covariance** matrix
- **Pre-whitening** (noise covariance $= I$) makes computation relatively straight-forward



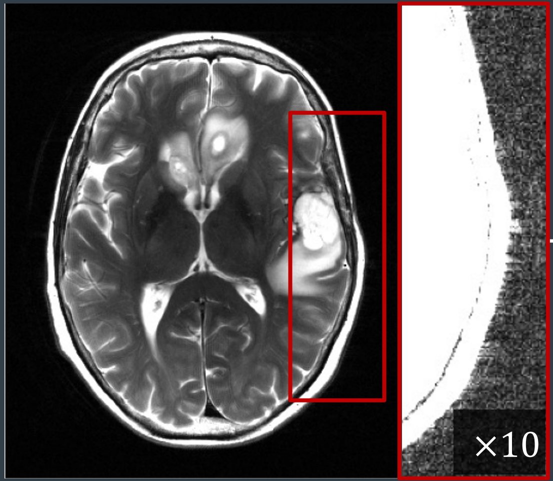
Pre-Processing Noisy Dataset

- GSURE loss requires the **noise covariance** matrix
- **Pre-whitening** (noise covariance $= I$) makes computation relatively straight-forward



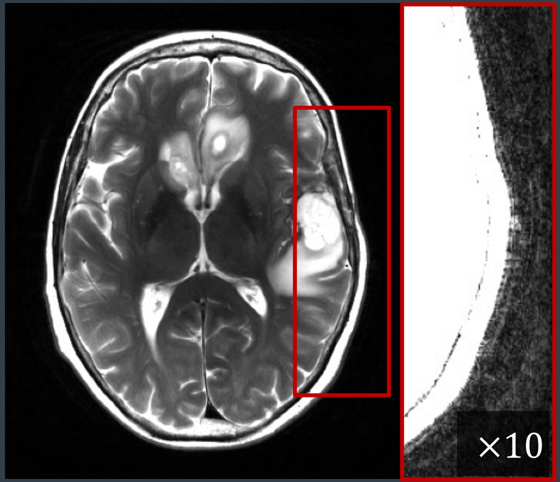
GSURE Denoising - Summary

$\hat{x}_{noisy}$



$g_\phi(\hat{x}_{noisy})$

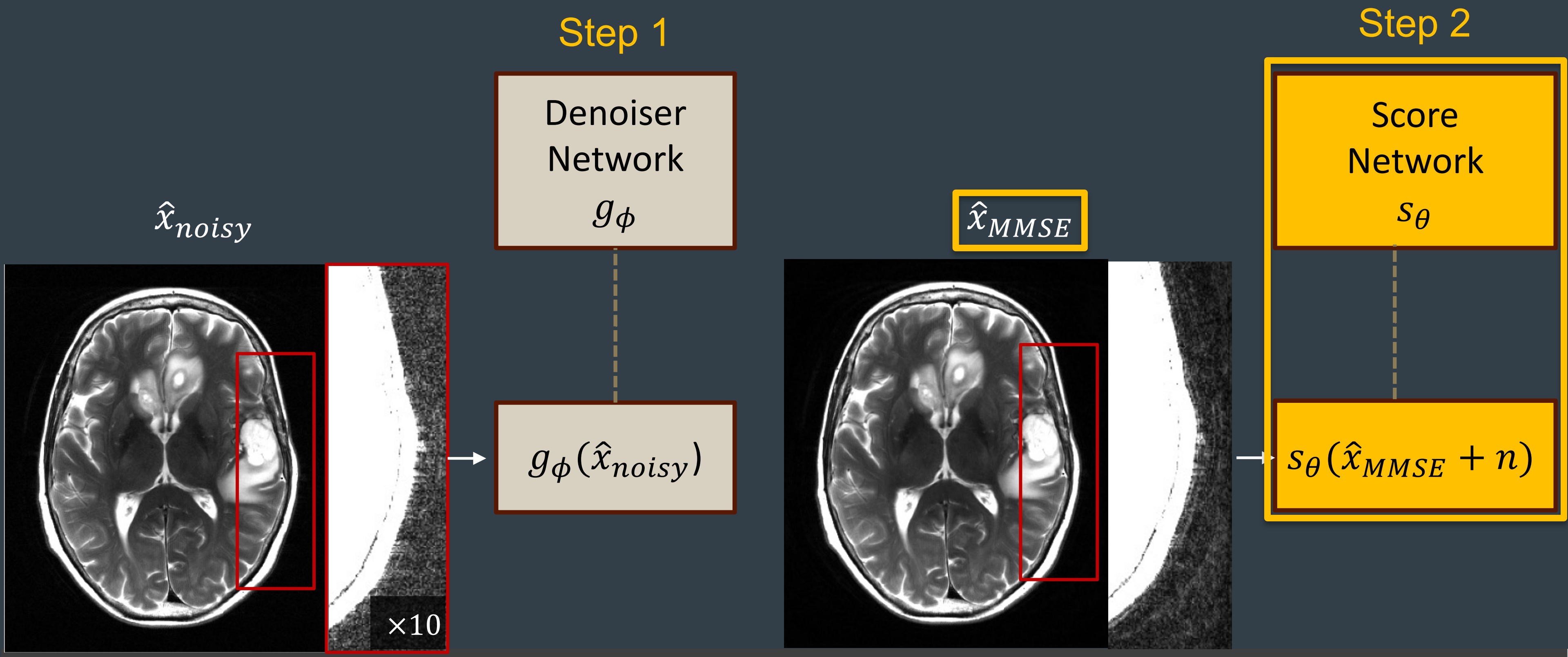
$\hat{x}_{MMSE}$



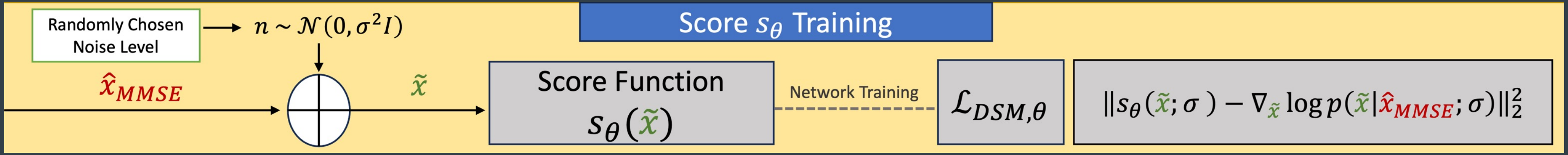
$$L_{GSURE} = \left\|g_\phi(u)\right\|^2 + 2\left(Tr\left(\frac{\partial g_\phi(u)}{\partial u}\right) + g_\phi^T(u)\frac{\partial \ln q(u)}{\partial u}\right)$$

GSURE Training

Proposed Methods



Score-based generative models



Hyvarinen, *JMLR* 2005, Song, *UAI,* 2018, Vincent, *MIT Press* 2011, Song, *NeurIPS* 2019, Dhariwal, *NeurIPS* 2021

Score-based generative models

Score s_θ Training

Randomly Chosen Noise Level → $n \sim \mathcal{N}(0, \sigma^2 I)$

$\hat{x}_{MMSE}$ ⊕ n → $\tilde{x}$ → Score Function $s_\theta(\tilde{x})$

Network Training

$$\mathcal{L}_{DSM,\theta} \quad \| s_\theta(\tilde{x}; \sigma) - \nabla_{\tilde{x}} \log p(\tilde{x} | \hat{x}_{MMSE}; \sigma) \|_2^2$$

Inference - Annealed Langevin Dynamics

$x_0 \sim \mathcal{N}(0, I)$

$$x_{t+1} = x_t + \alpha_t \left(\frac{A^H (y - A x_t)}{\sigma_n^2 + \gamma_t^2} + s_\theta(x_t; \sigma_t) \right) + \sqrt{2\beta\alpha_t}\,\eta_t$$

$\eta_t \sim \mathcal{N}(0, I)$: Annealing Noise
α_t: Step-Size at t
β: Noise Controlling Parameter
γ_t: Annealing Factor

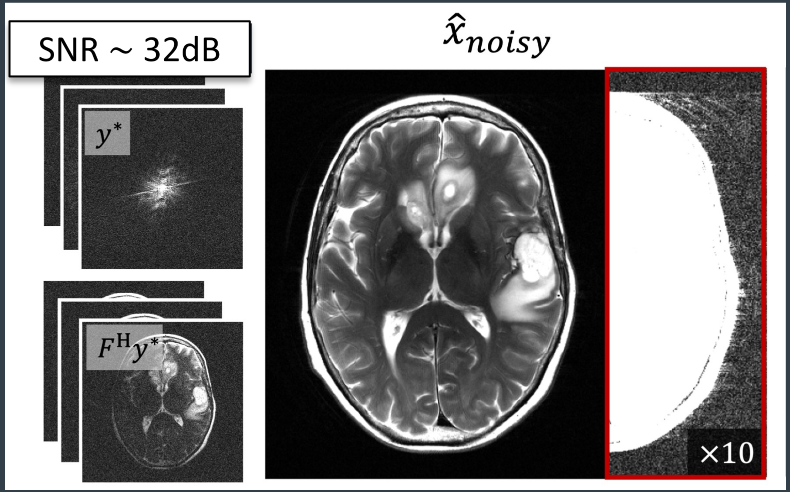
Hyvarinen, *JMLR* 2005, Song, *UAI,* 2018, Vincent, *MIT Press* 2011, Song, *NeurIPS* 2019, Dhariwal, *NeurIPS* 2021

Experiments

1. Evaluation of **Self-Supervised Denoising** (GSURE)
2. **Prior sampling** performance of score models trained on Noisy vs. GSURE Denoised data
3. **Posterior Reconstruction** performance of score models trained on Noisy vs. GSURE Denoised data

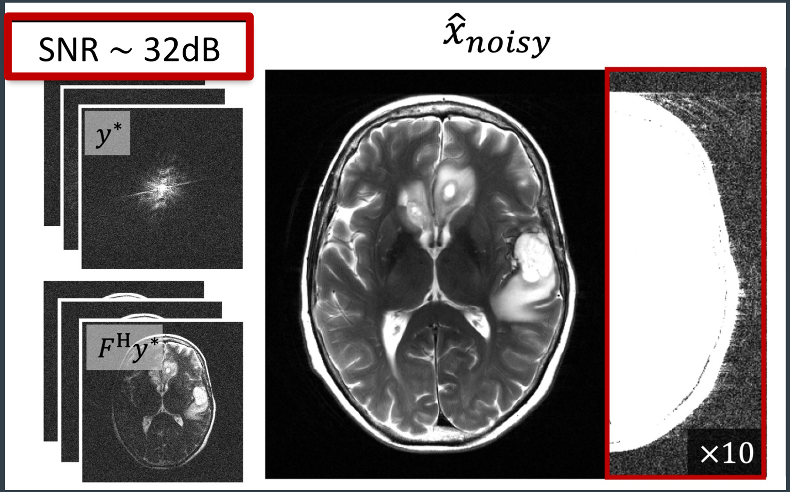
T2 Brain Scans

Original FastMRI



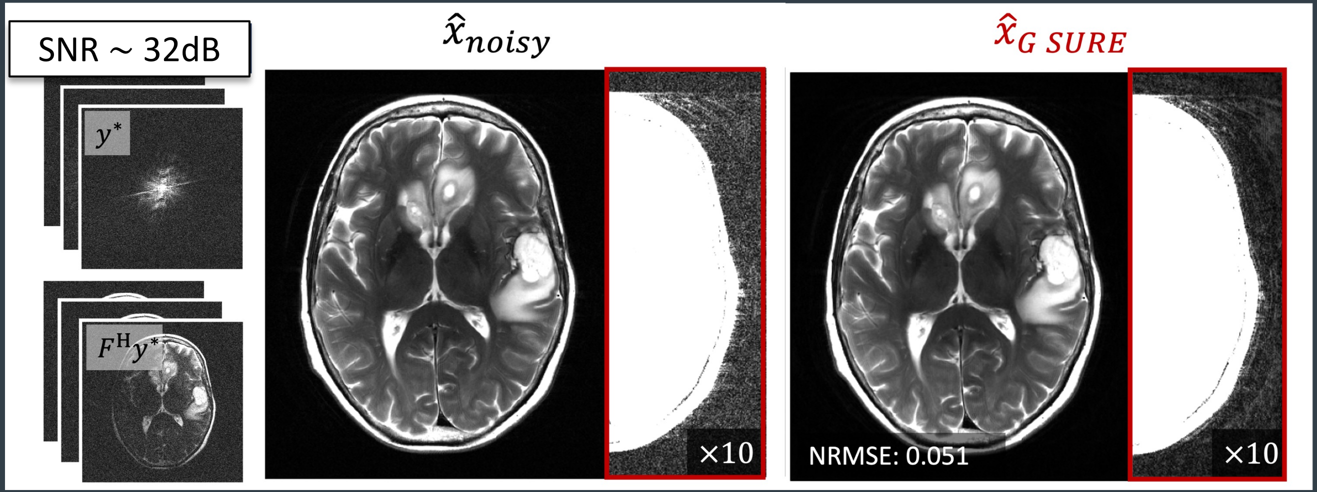
T2 Brain Scans

Original FastMRI



T2 Brain Scans

Original FastMRI



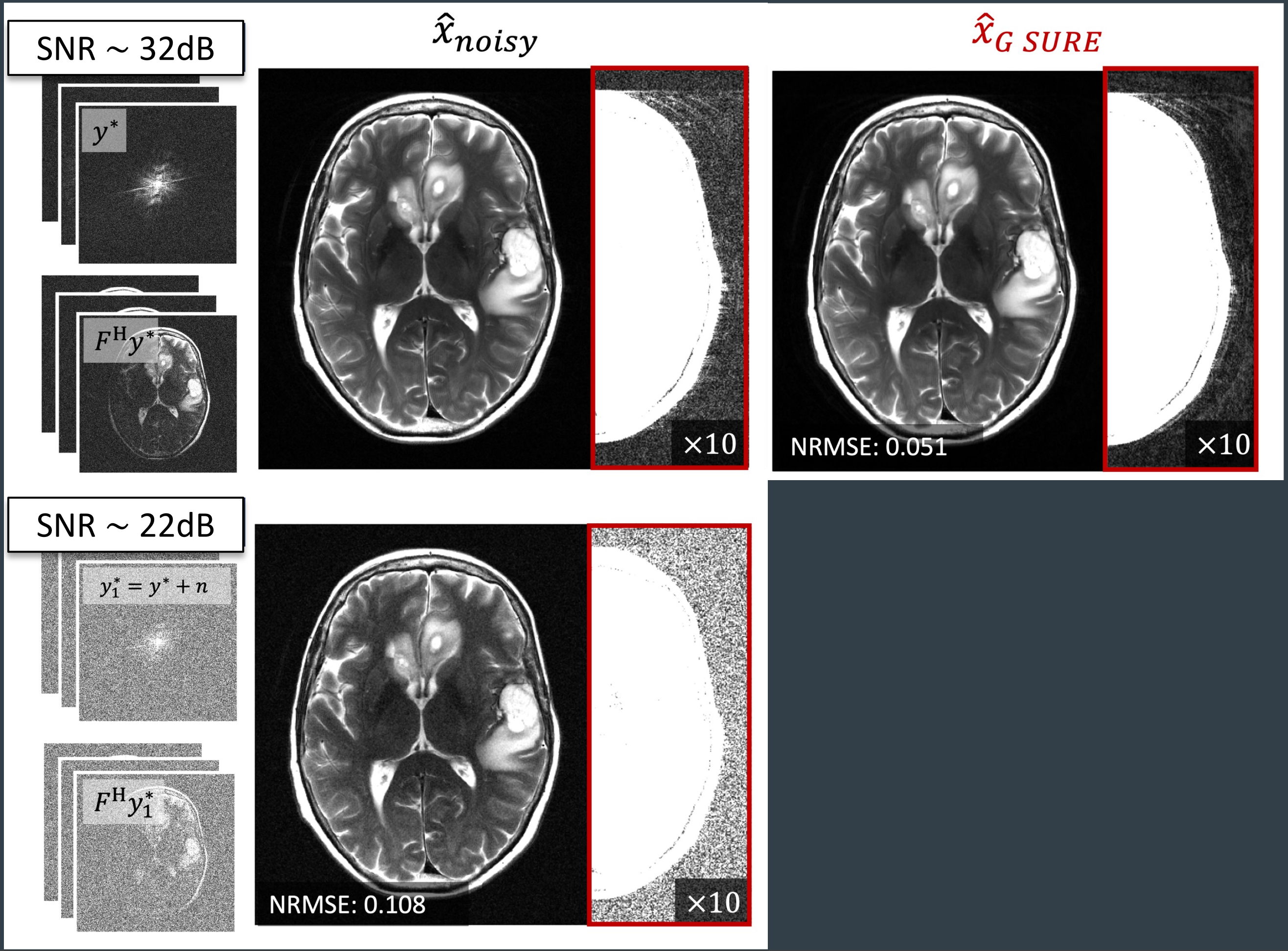
T2 Brain Scans

Original FastMRI

Original FastMRI

+

Additive Gaussian Noise



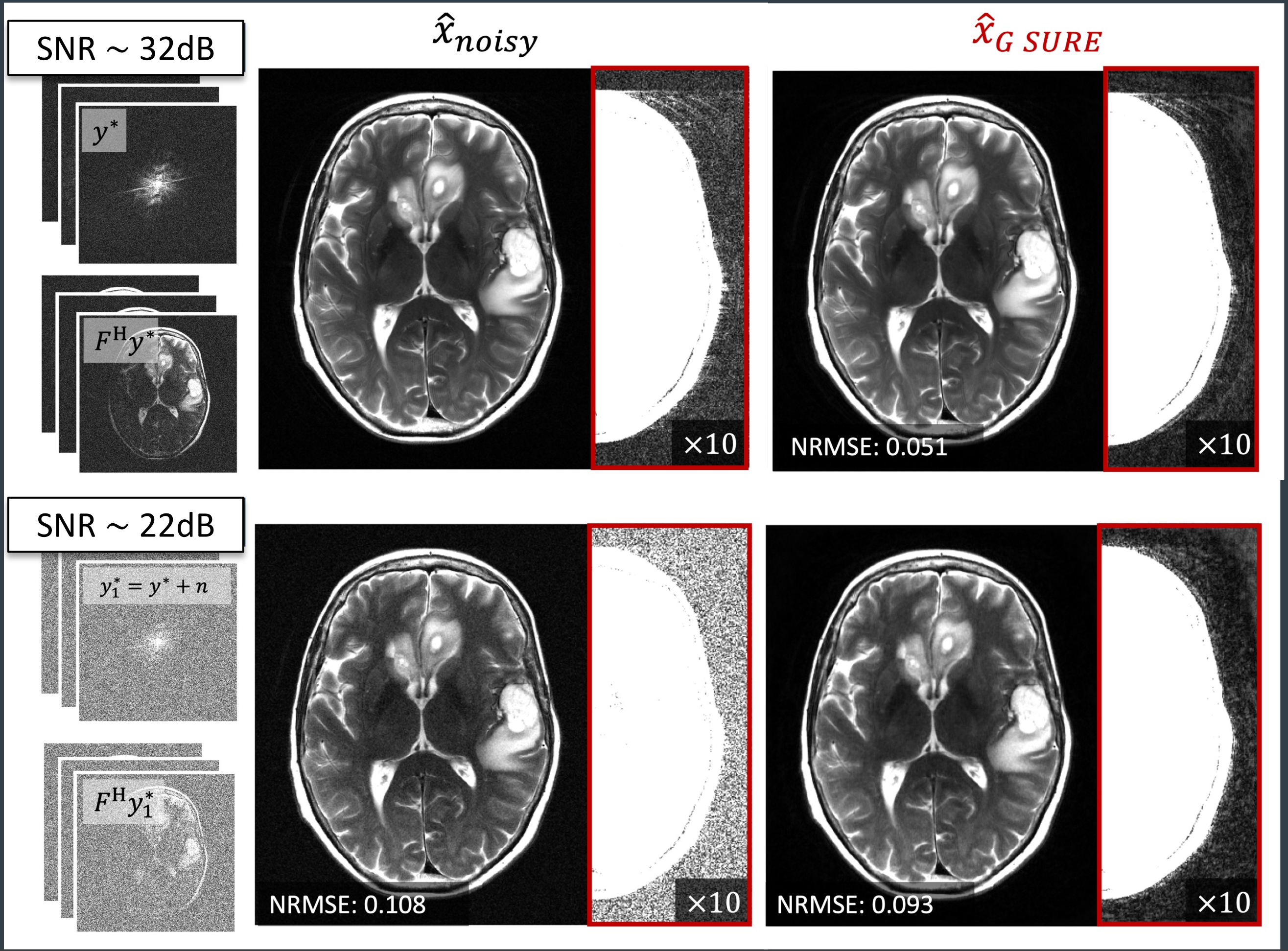
T2 Brain Scans

Original FastMRI

Original FastMRI

+

Additive Gaussian Noise



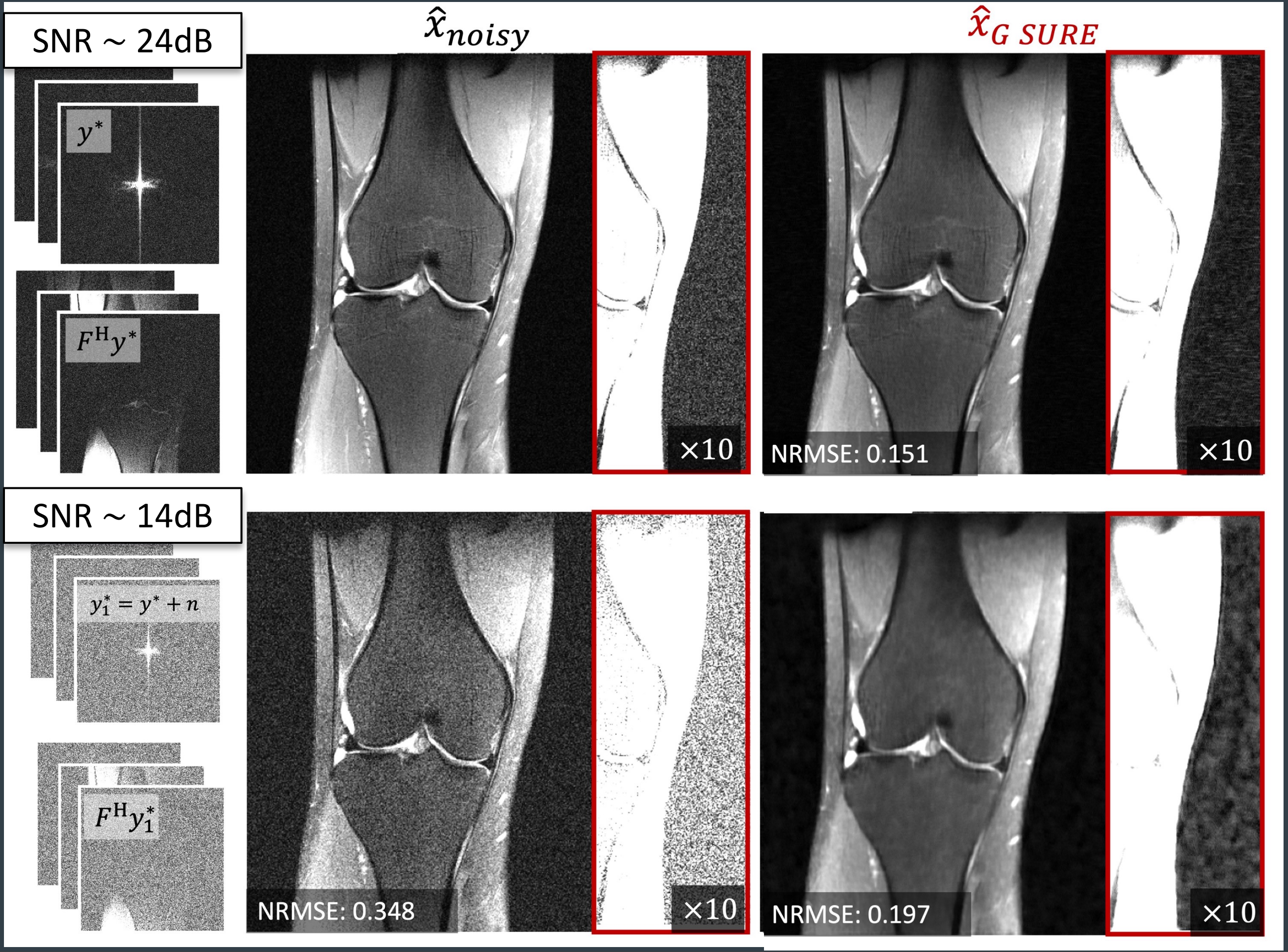
Knee Scans

Original FastMRI

Original FastMRI

+

Additive Gaussian Noise

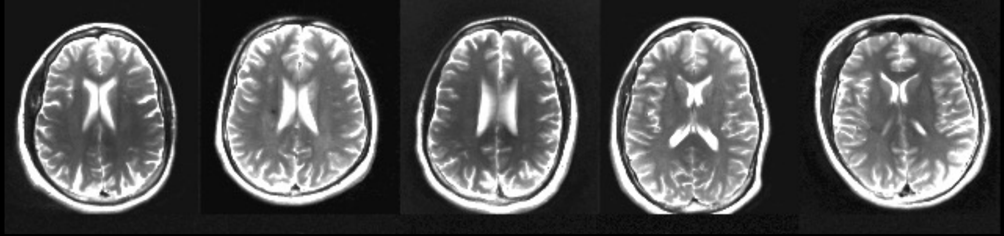


Results

1. Evaluation of **Self-Supervised Denoising** (GSURE)
2. **Prior sampling** performance of score models trained on Noisy vs. GSURE Denoised data
3. **Posterior Reconstruction** performance of score models trained on Noisy vs. GSURE Denoised data

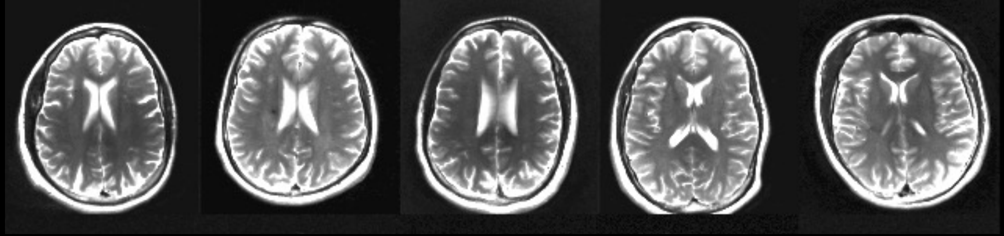
Prior Sampling
$p(x)$

Naive Score
~ 32dB

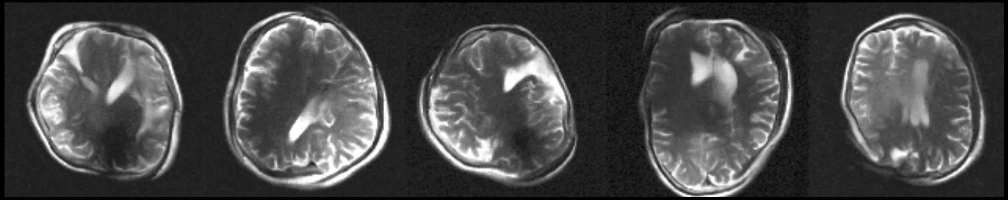


Prior Sampling
$p(x)$

Naive Score
~ 32dB



Naive Score
~ 22dB



Prior Sampling

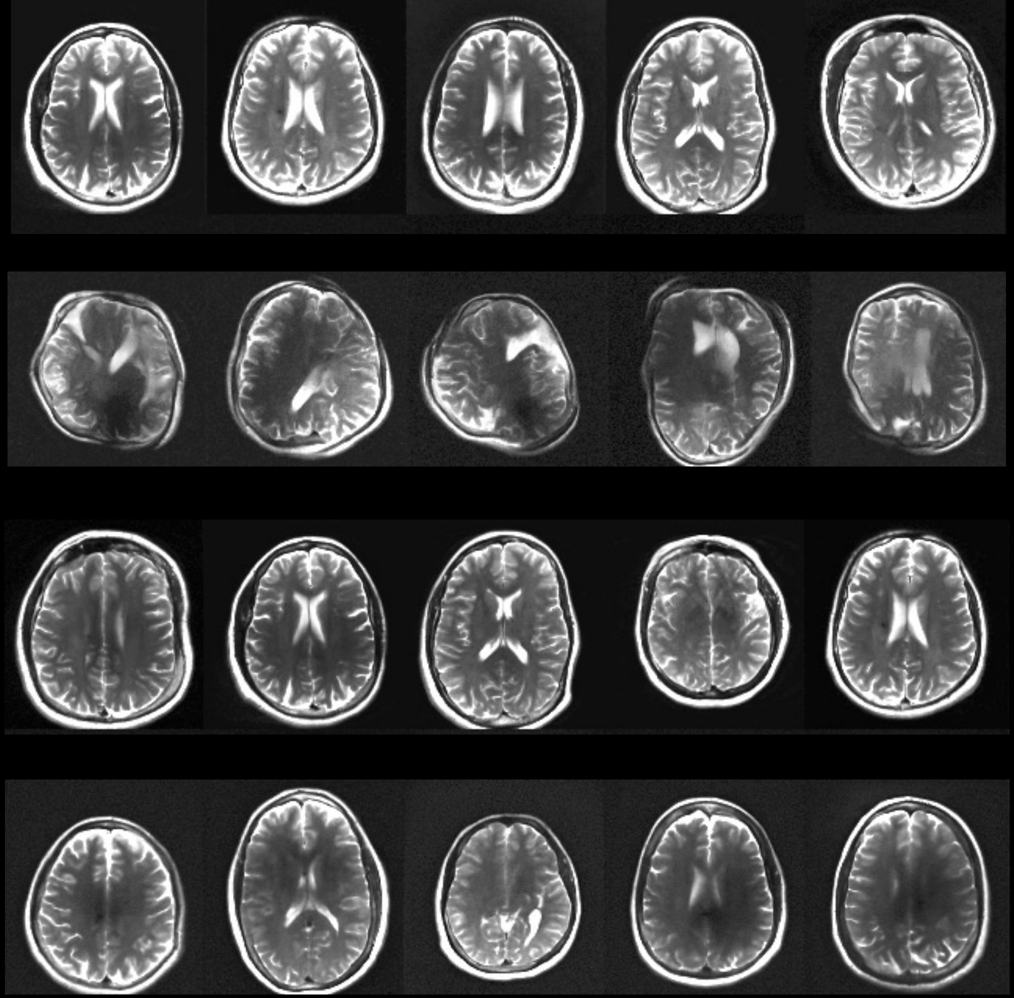
$p(x)$

Naive Score
~ 32dB

Naive Score
~ 22dB

GSURE-Score
~ 32dB

GSURE-Score
~ 22dB

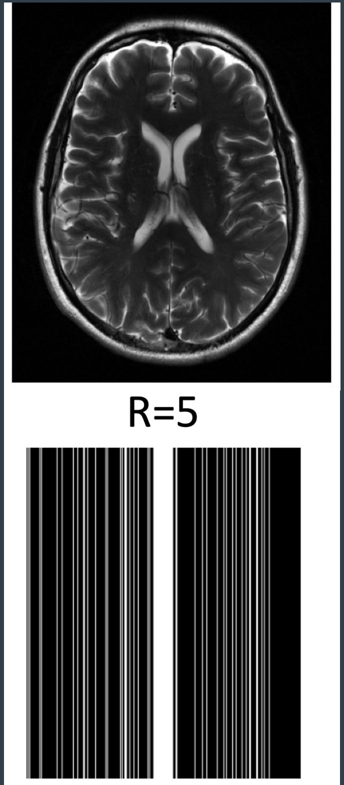


Results

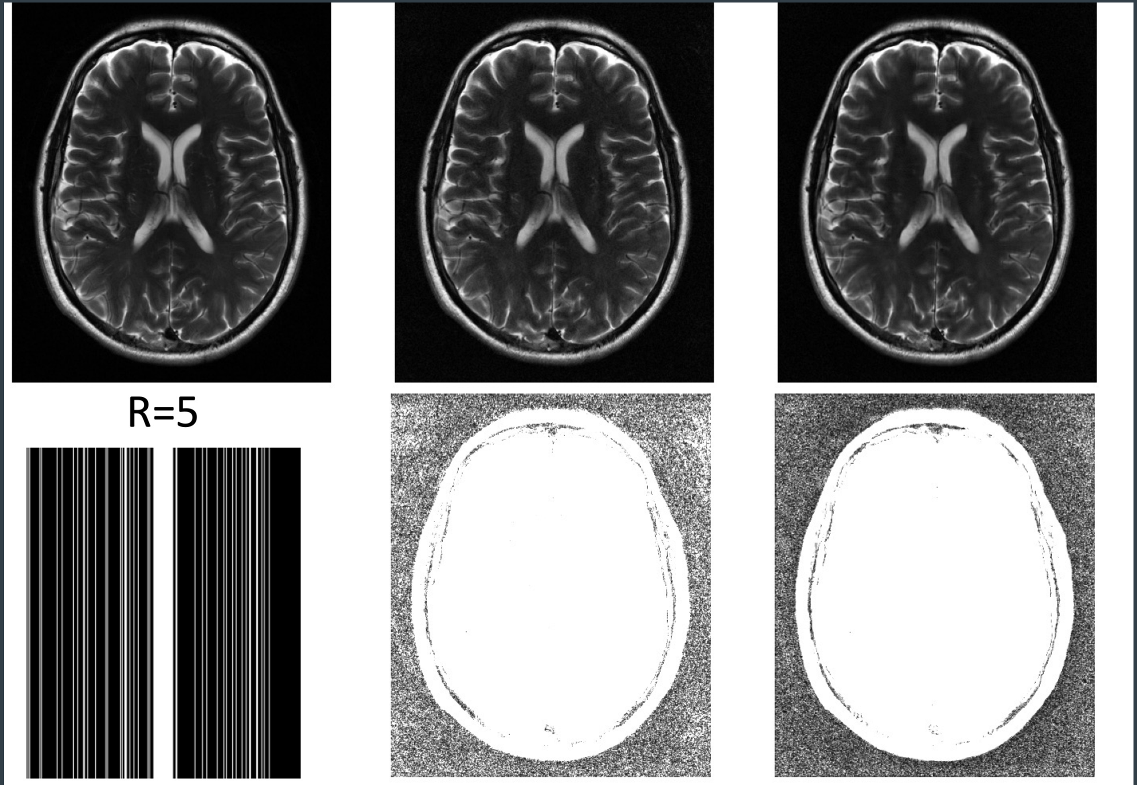
1. Evaluation of **Self-Supervised Denoising** (GSURE)
2. **Prior sampling** performance of score models trained on Noisy vs. GSURE Denoised data
3. **Posterior Reconstruction** performance of score models trained on Noisy vs. GSURE Denoised data

Posterior Sampling $x \sim p(x|y)$

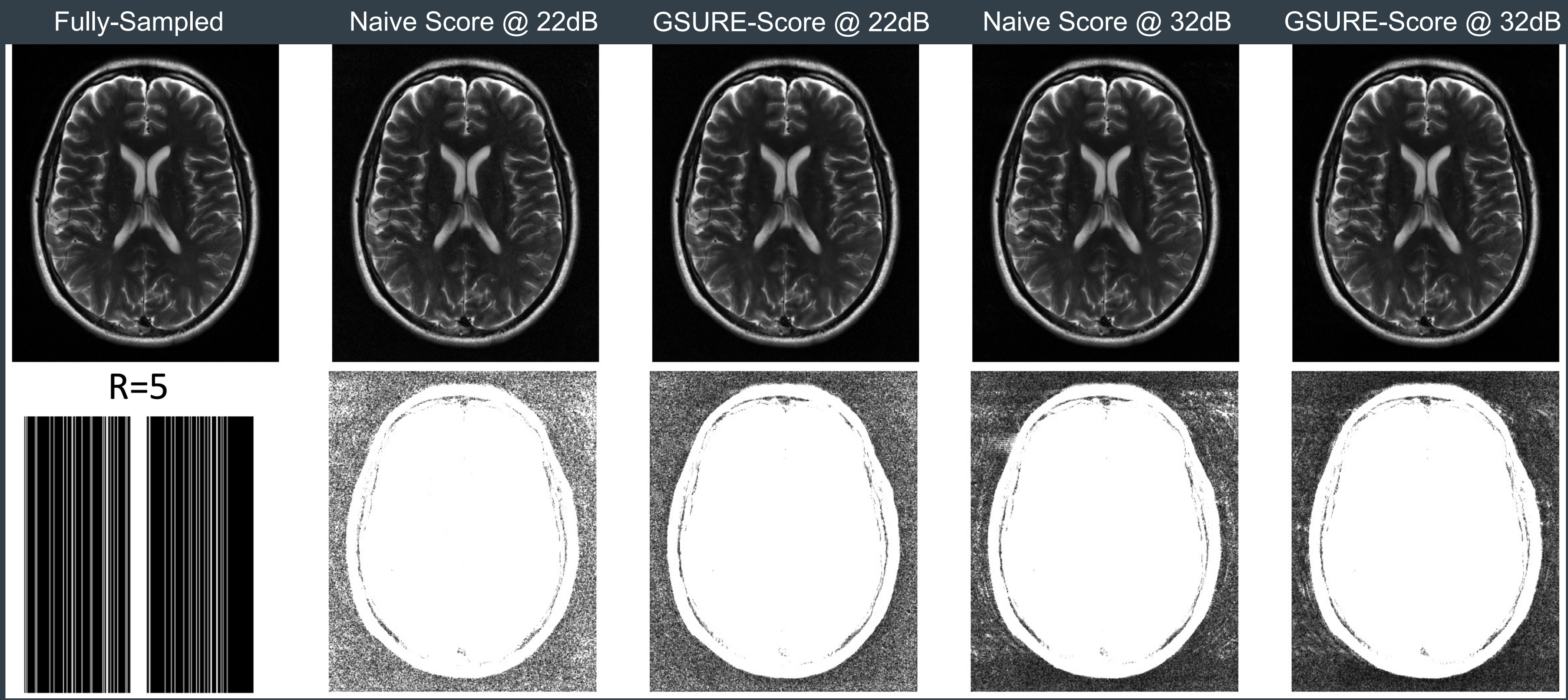
Fully-Sampled | Naive Score @ 22dB | GSURE-Score @ 22dB | Naive Score @ 32dB | GSURE-Score @ 32dB



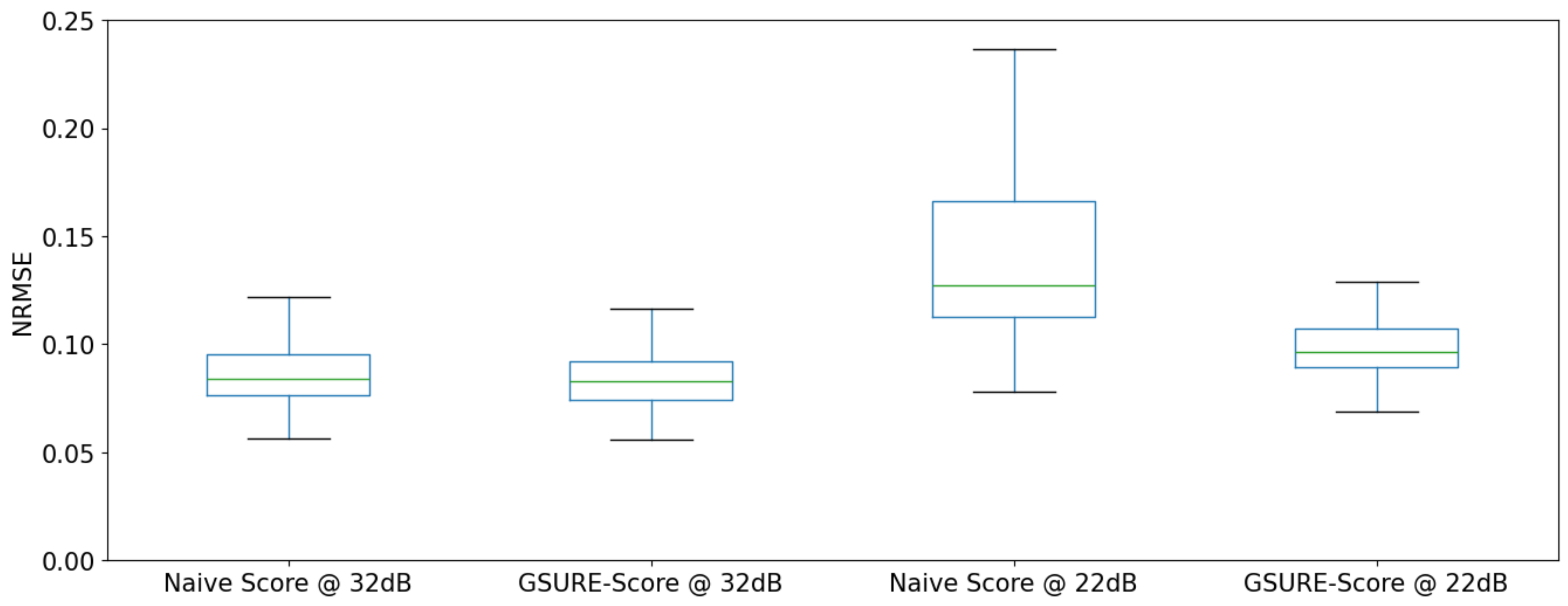
Posterior Sampling $x \sim p(x|y)$



Posterior Sampling $x \sim p(x|y)$



Posterior Sampling $x \sim p(x|y)$



Discussion and Conclusion

1. Self-supervised denoising techniques can successfully **remove noise**
2. Denoising as a pre-processing step helps train more **accurate priors**
3. The benefit of denoising is more visible in **lower SNR** settings
4. Priors trained on denoised FastMRI are **better inverse problem solvers** than naive training
5. We see application in the low-field setting (neo-natal MRI)

Future Works

A is a Linear Forward Operator (**Fully-Sampled**) -> GSURE[1,2,3]

$$y = FSx + noise$$

Assume A is a **Low-Rank** Forward Operator[4]

$$y = \boldsymbol{P}FSx + noise$$

[1]Soltanayev, *NeurIPS,* 2018, [2]Eldar, *IEEE Transactions on Signal Processing,* 2008, [3]Kawar, *TMLR,* 2023, [4]Aali, AmbientDPS, *Arxiv*, 2024

Thank you!

Asad Aali

Email: asad.aali@utexas.edu

Website: https://asad-aali.github.io/

MS ECE Student, UT Austin

Source Code: https://github.com/utcsilab/GsureScore-Diffusion.git