MARCH 2023

TEXAS
The University of Texas at Austin

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

Asad Aali[1], Marius Arvinte[1,2], Sidharth Kumar[1], Yamin Ishraq Arefeen[1], and Jonathan I. Tamir[1]
[1]Chandra Family Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX, United States, [2]Intel Corporation, Hillsboro, OR, United States

Asad Aali
M.S. Student
Electrical & Computer Engineering
The University of Texas at Austin

UT Computational Sensing and Imaging Lab

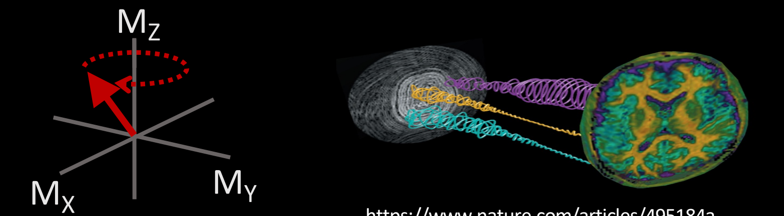
- Joint design of imaging system and software
- Particular focus on application to MRI
- Work with clinicians to translate work to hospital

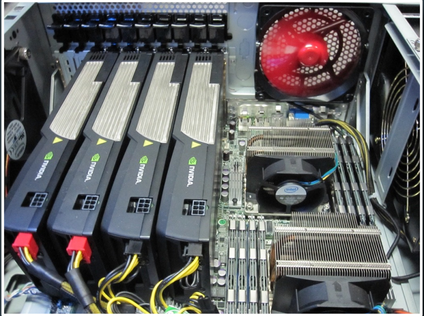
Jon Tamir, PhD
Assistant Professor, ECE, UT Austin
http://www.jtsense.com/

https://github.com/utcsilab

Computational MRI

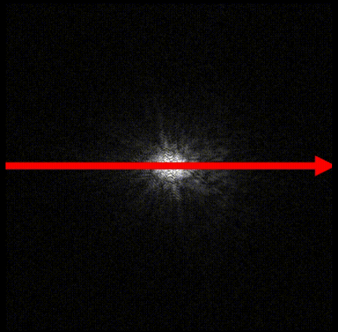
Imaging physics



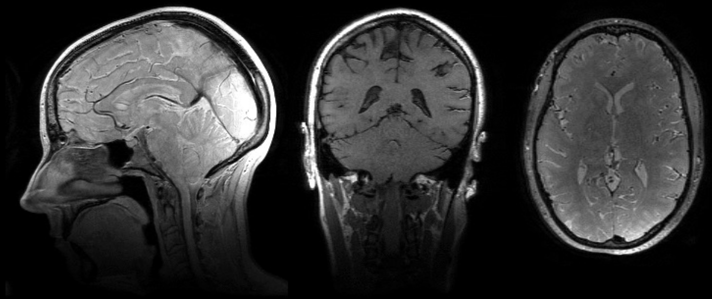


Prior knowledge

Acquisition

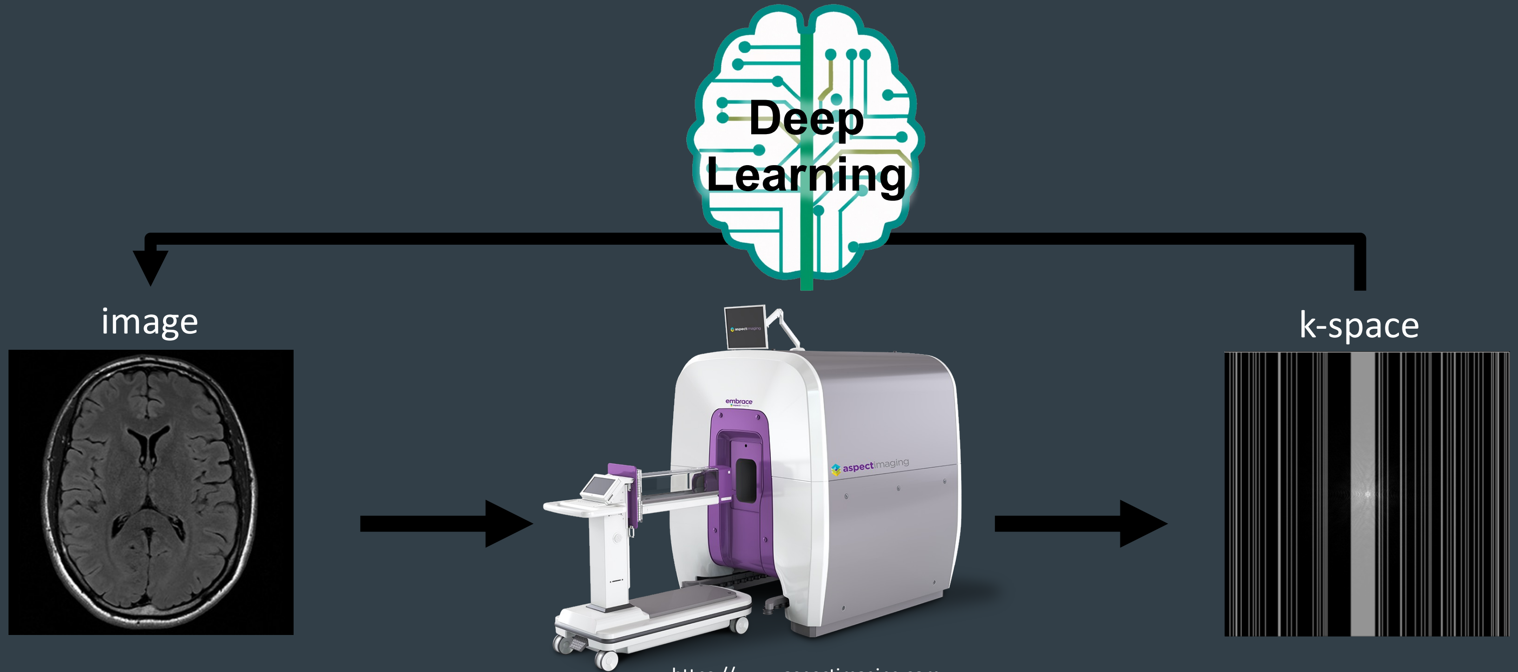


Reconstruction



Deep learning inversion for MRI

1. End-to-end supervised training
2. Distribution learning / generative modeling



https://www.aspectimaging.com

Generative models are powerful image generators

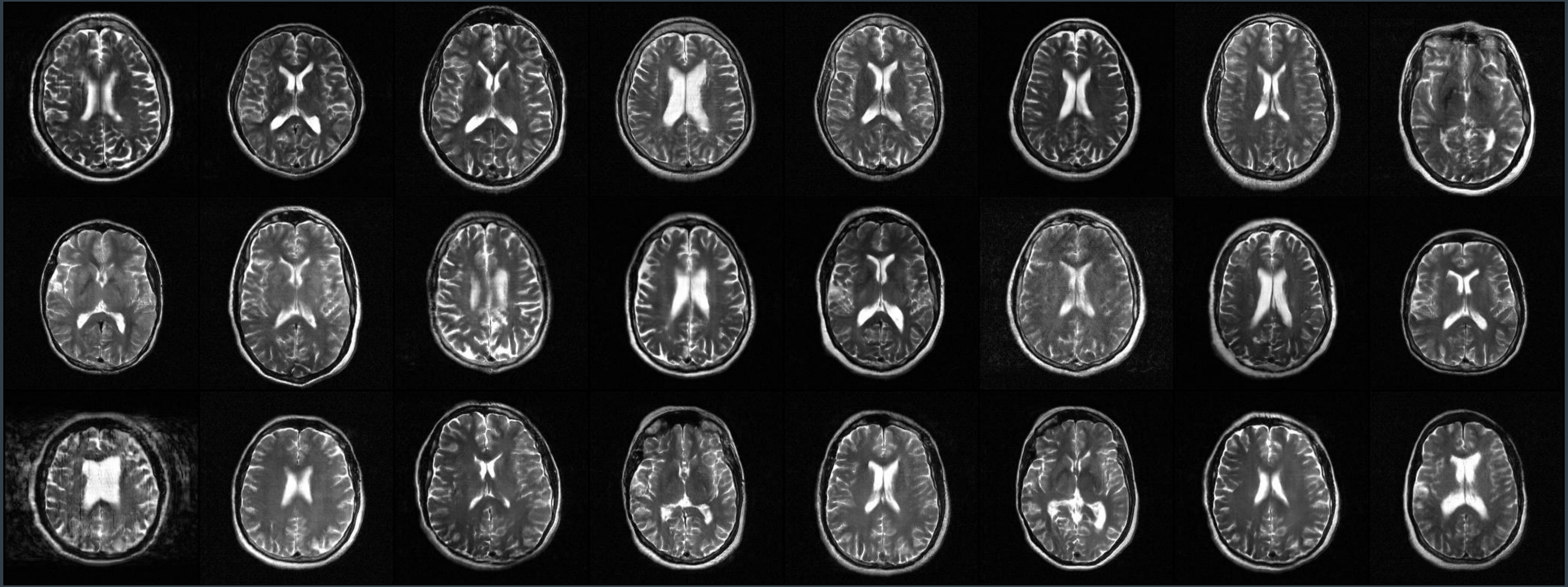
T Karras et al., CVPR 2020

Y Song et al., ICLR 2021

Generative models are powerful image generators

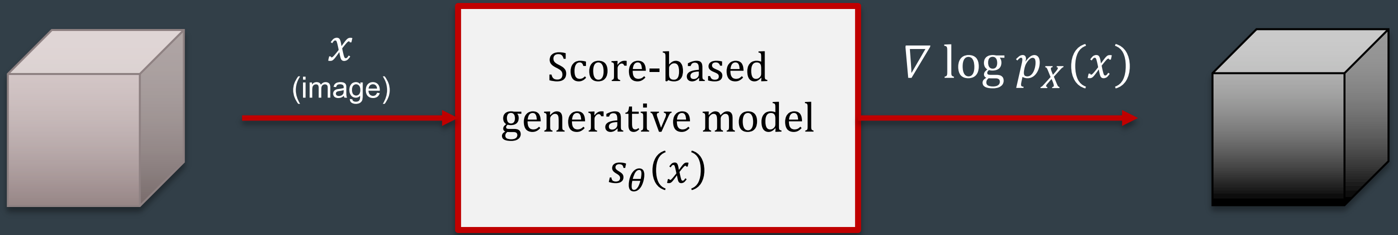
https://thiscatdoesnotexist.com/

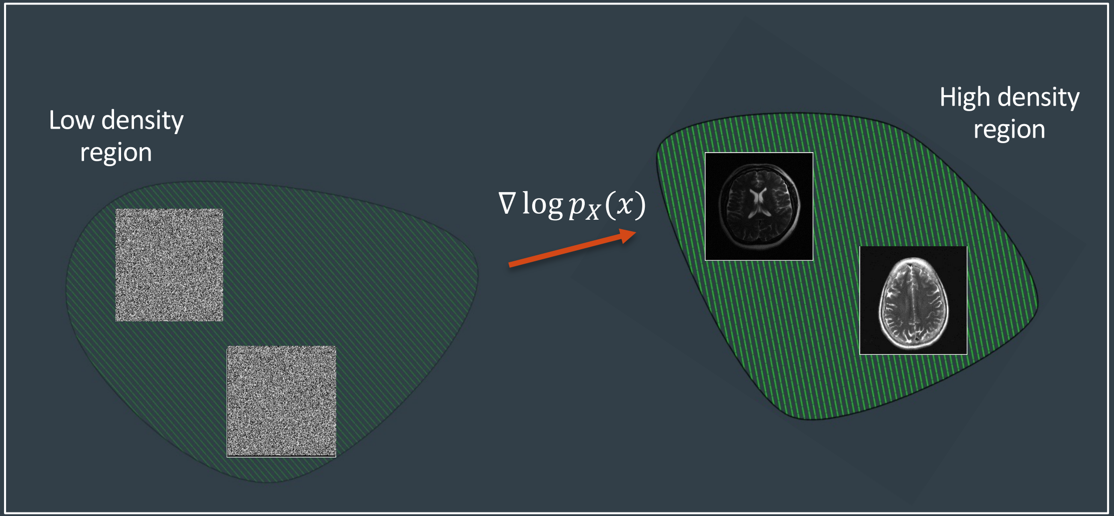
Generative models are powerful image generators



Generative model trained on FastMRI data

Score-based generative models



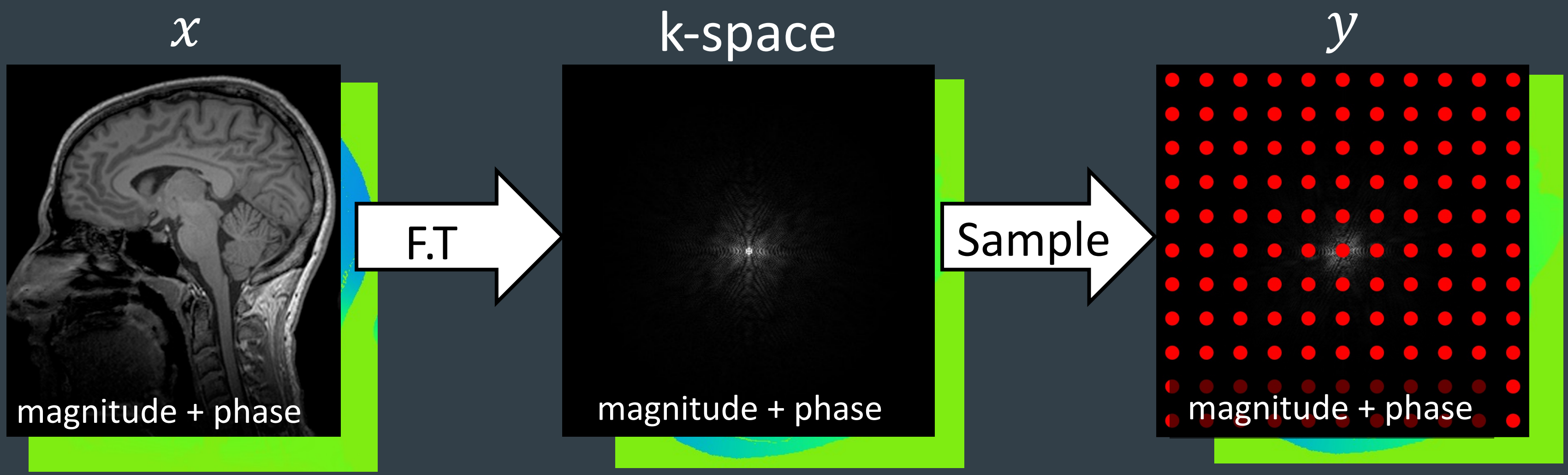


A Hyvarinen, JMLR 2005, Y Song et al., UAI 2018, Vincent et al., MIT Press 2011, Y Song et al., NeurIPS 2019. P Dhariwal et al,. NeurIPS 2021.

MRI: Problem Formulation

Signal is the Fourier transform of the image

$$y = Ax + noise$$



Back to basics: statistical interpretation

$$y = Ax + n$$

Goal: estimate the image from noisy measurements

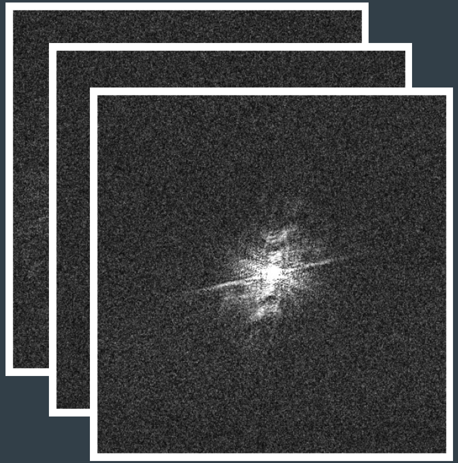
- **Likelihood function:** described by the imaging physics $p(y|x)$
- Prior knowledge: probability distribution for the image: $p(x)$
- MAP estimator: maximize the posterior $\max_{x} p(x|y)$
- Another option: sample from the posterior $x \sim p(x|y)$
- Another option: conditional expectation $E\,[x|y]$

MRI Samples are inherently noisy

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

$$y = Ax + noise$$

y



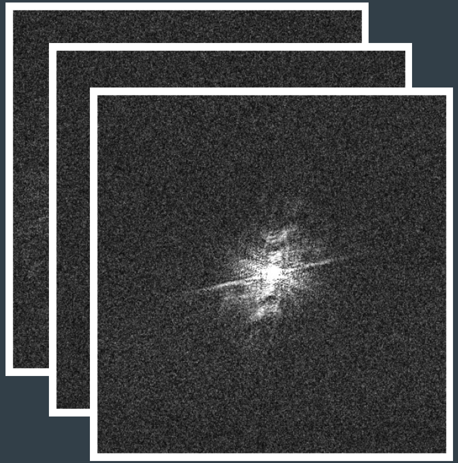
Original K-Space

MRI Samples are inherently noisy

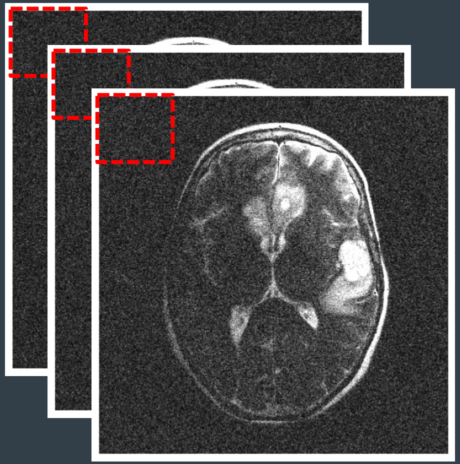
Goal is to learn the **clean distribution** using *noisy* data (i.i.d Gaussian, with known power σ_w^2).

$$y = Ax + noise$$

$$y \longrightarrow F^{\mathrm{H}}y$$



Original K-Space

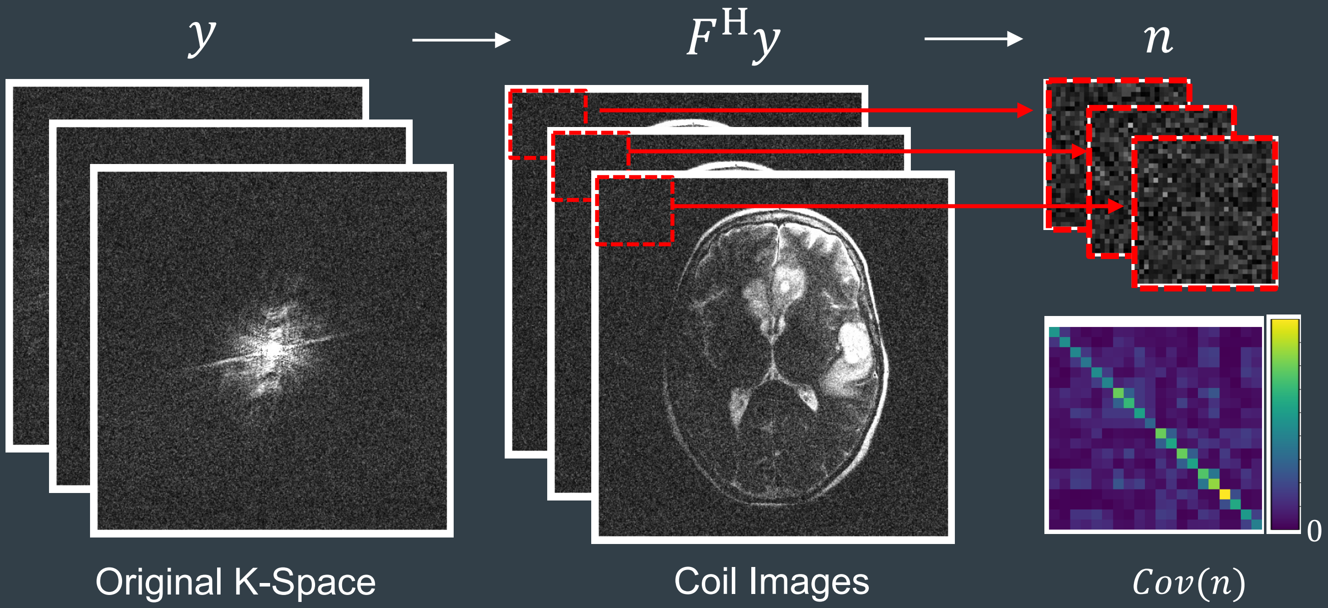


Coil Images

MRI Samples are inherently noisy

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

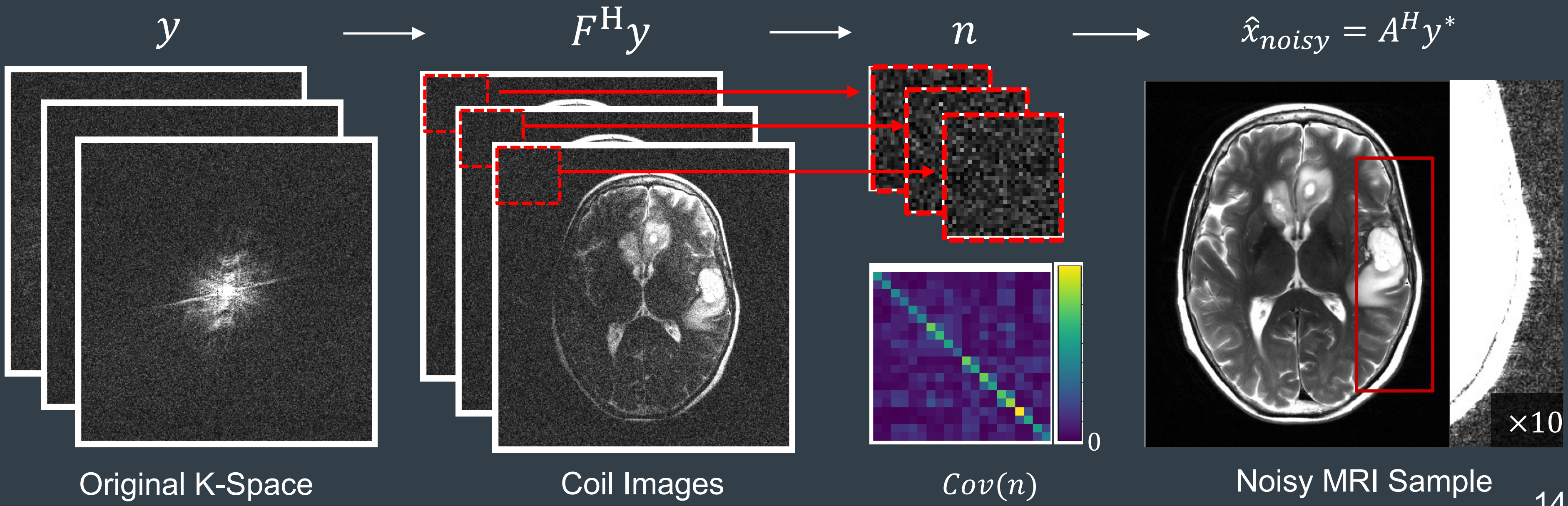
$$y = Ax + noise$$



MRI Samples are inherently noisy

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

$$y = Ax + noise$$



Learning from Noisy Data

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

$$y = Ax + noise$$

- Case 1: Assume A is Identity -> **SURE-Score**

$$y = x + noise$$

Learning from Noisy Data

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

$$y = Ax + noise$$

- Case 1: Assume A is Identity -> **SURE-Score**

$$y = x + noise$$

- Case 2: Assume A is a Linear Forward Operator (Fully-Sampled) -> **GSURE-Score**

$$y = FSx + noise$$

Learning from Noisy Data

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

$$y = Ax + noise$$

- Case 1: Assume A is Identity -> **SURE-Score**

$$y = x + noise$$

- Case 2: Assume A is a Linear Forward Operator (Fully-Sampled) -> **GSURE-Score**

$$y = FSx + noise$$

- Case 3: Assume A is a Linear Forward Operator (Under-Sampled) -> **GSURE + Ambient Diffusion ???**

$$y = PFSx + noise$$

*In ambient diffusion, we assume kspace is noise-free.

Case 1: Denoising with SURE -> $y = x + n$

$$\mathbb{E}_w\left[\frac{1}{n}\|x - f_\theta(y)\|_2^2\right] = \underbrace{\mathbb{E}_w\left[\frac{1}{n}\|y - f_\theta(y)\|_2^2\right] - \sigma_w^2 + \frac{2\sigma_w^2}{n}\,\mathrm{div}_y\big(f_\theta(y)\big)}_{\text{Does not depend on the clean sample } x}$$

- We can train a denoiser that removes Gaussian noise at a single noise level σ_w^2 using the loss [1]:

$$\mathcal{L}(\theta) = \mathbb{E}_{x,w}\left[\|y - f_\theta(y)\|_2^2 + 2\sigma_w^2 \mathrm{div}_y\big(f_\theta(y)\big)\right]$$

[1] Soltanayev, Shakarim, and Se Young Chun. "Training deep learning based denoisers without ground truth data." *Advances in neural information processing systems* 31 (2018).

Case 2:
Denoising with Generalized-SURE

$$y = FSx + n$$

Generalized-SURE Equation

- An unbiased estimate of the MSE

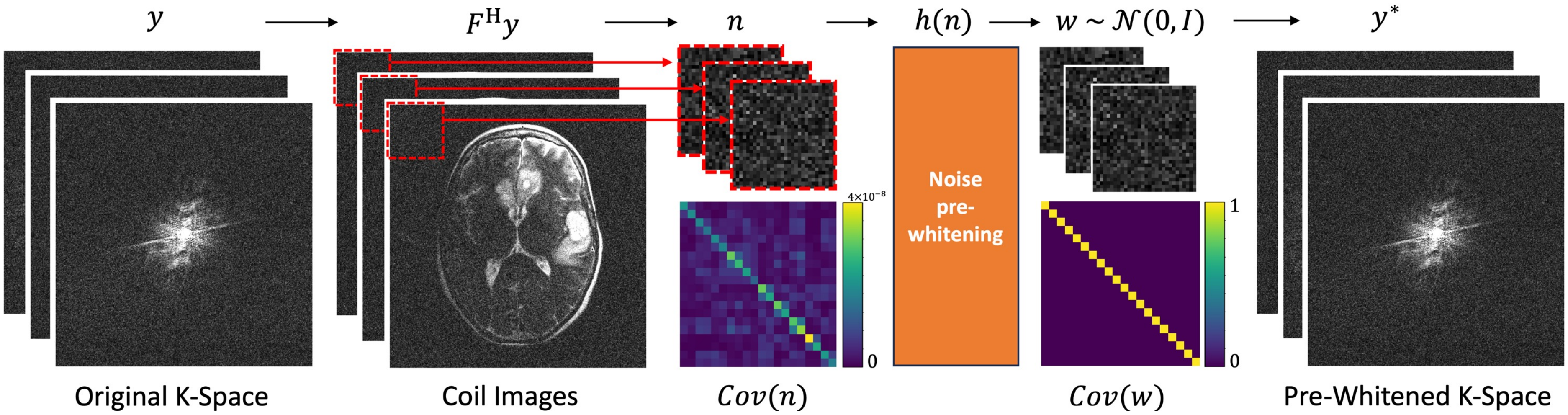
$$s(g) = \|x\|^2 + \left\|g_\phi(u)\right\|^2 + 2\left(Tr\left(\frac{\partial g_\phi(u)}{\partial u}\right) + g_\phi^T(u)\frac{\partial \ln q(u)}{\partial u}\right)$$

Where:

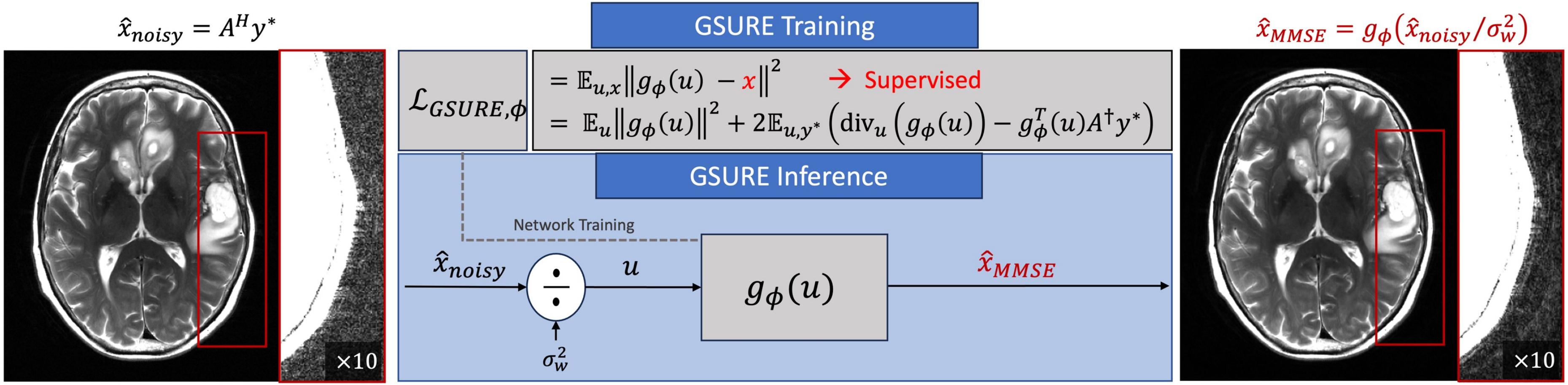
- g_ϕ is the denoiser network
- $u = \frac{A^H y}{\sigma_w^2}$
- $\frac{\partial \ln q(u)}{\partial u} = -A^+ y$
- In the i.i.d. Gaussian case, this equation becomes SURE (previously shown)

$$L_{SURE} = \|x\|^2 + 2\sigma_w^2 \frac{\partial g_\phi(y)}{\partial y} - 2g_\phi^T(y)y$$

Pre processing Data for GSURE Denoising

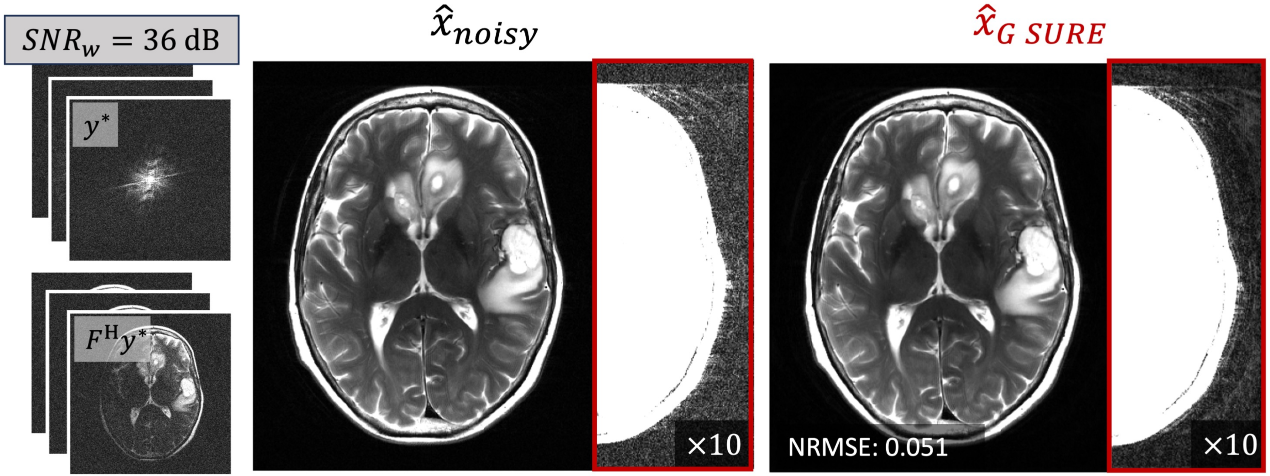


GSURE Training and inference



GSURE Denoising – T2 Brain Scans

Original FastMRI



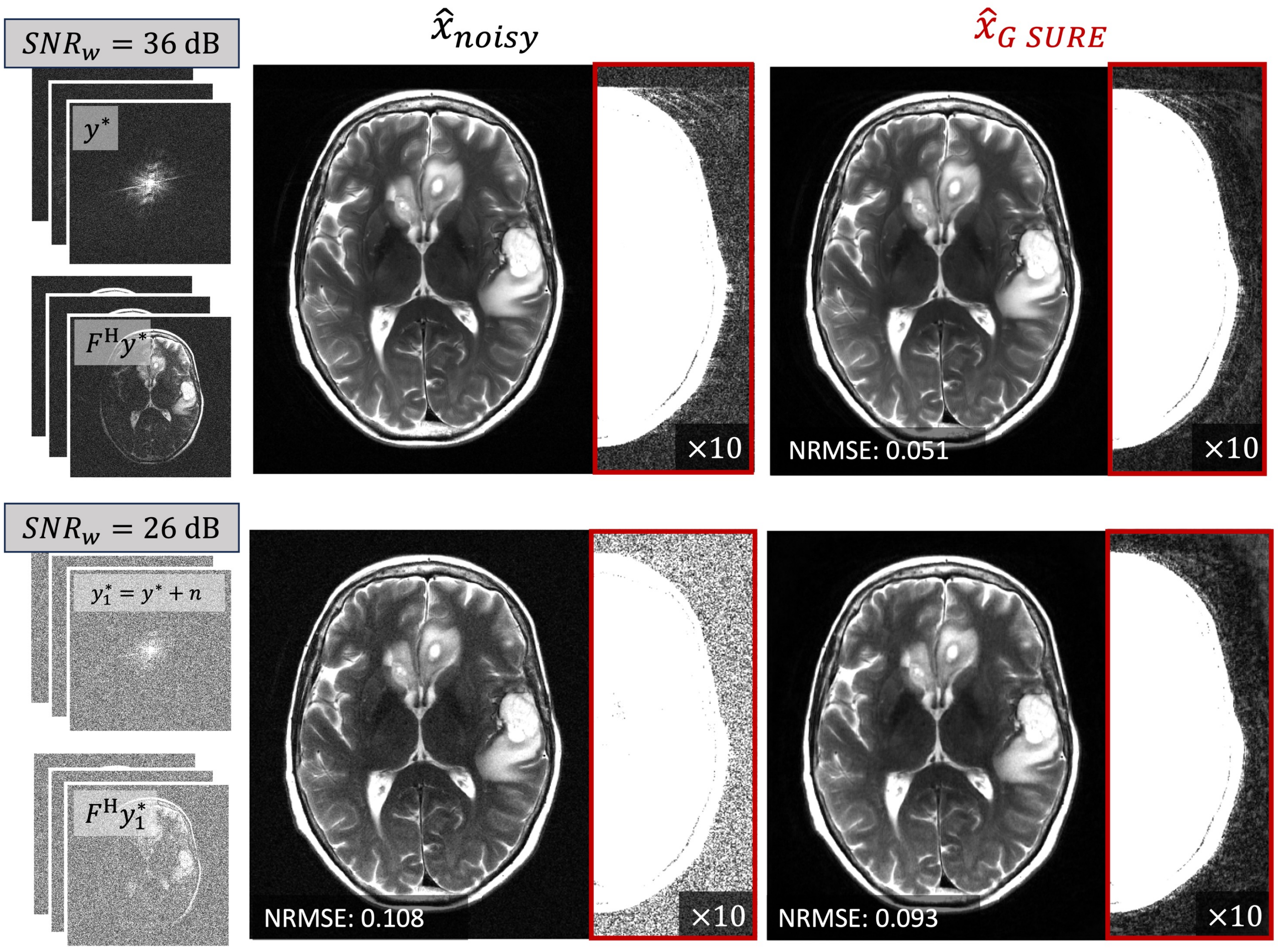
GSURE Denoising – T2 Brain Scans

Original FastMRI

Original FastMRI

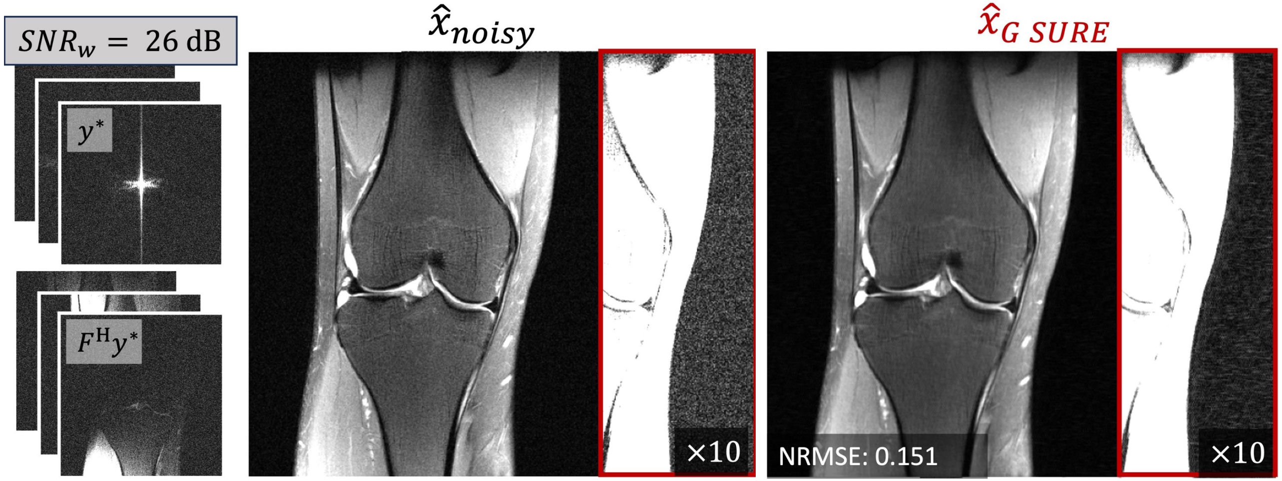
+

Additive Gaussian Noise



GSURE Denoising – Knee Scans

Original FastMRI



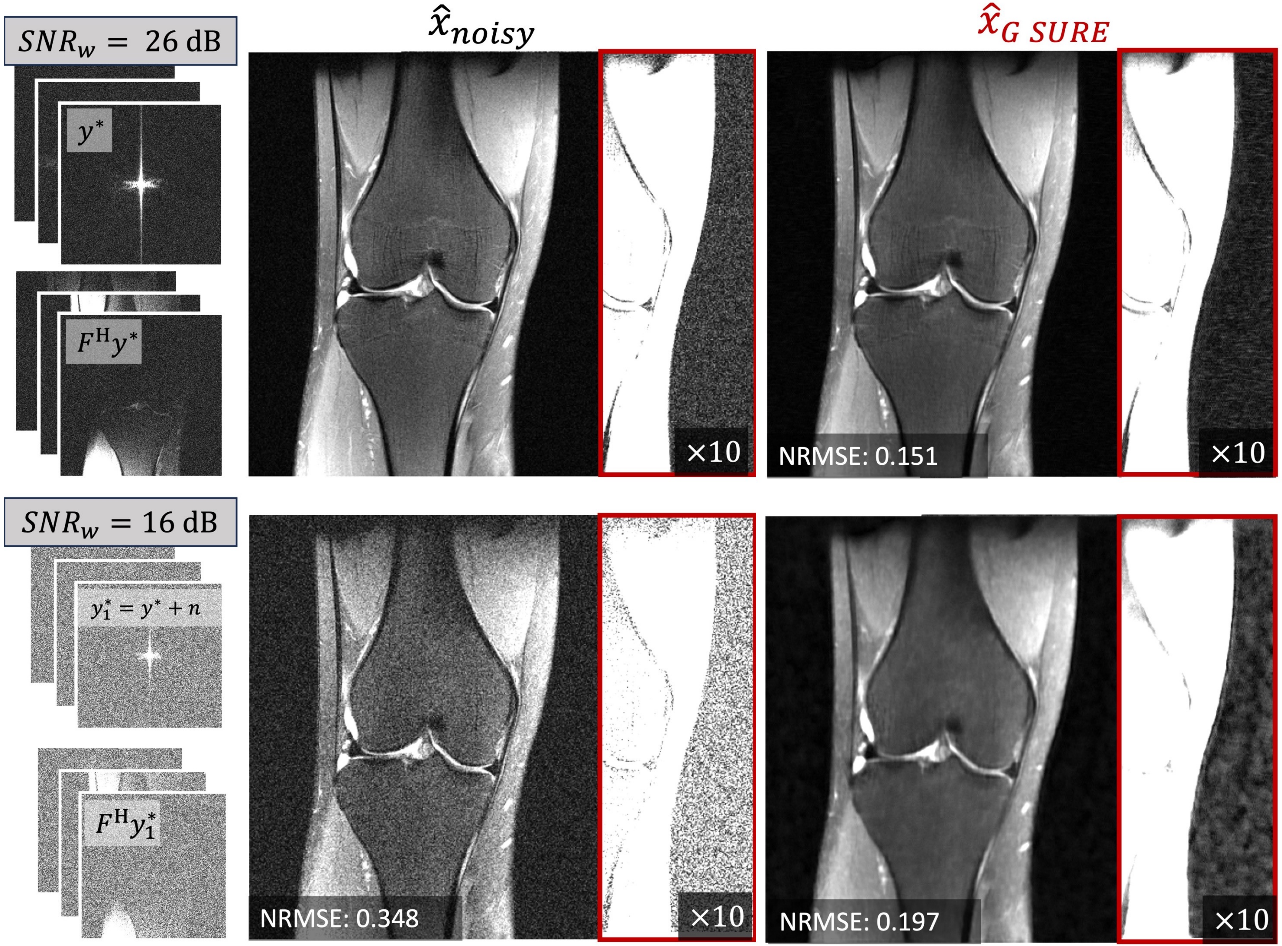
GSURE Denoising – Knee Scans

Original FastMRI

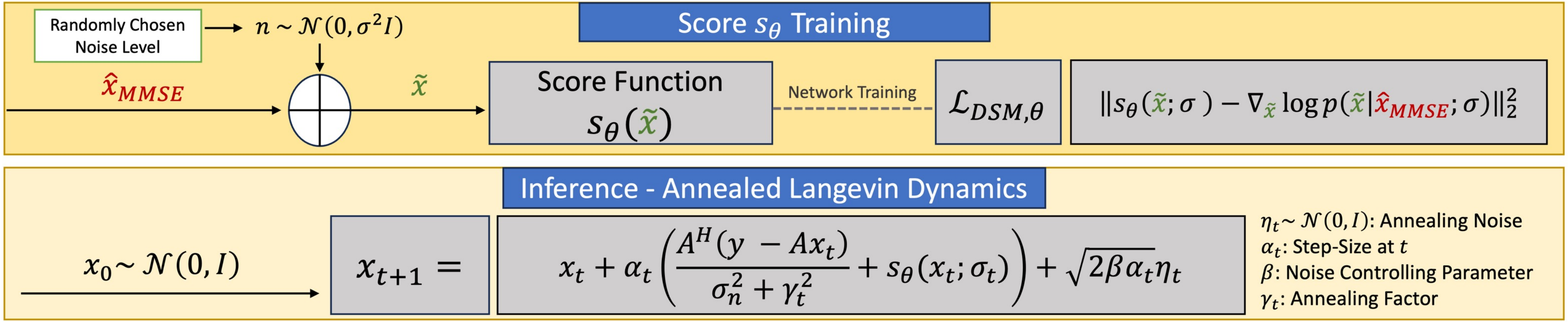
Original FastMRI

+

Additive Gaussian Noise



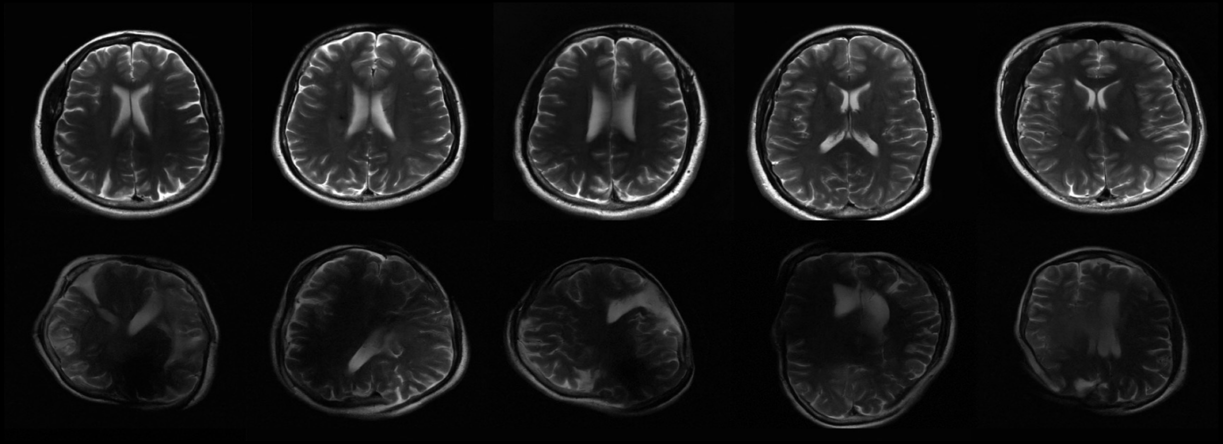
Score Model Training + Inference



Prior Sampling – $p(x)$

Naive Score
~ 32dB

Naive Score
~ 22dB



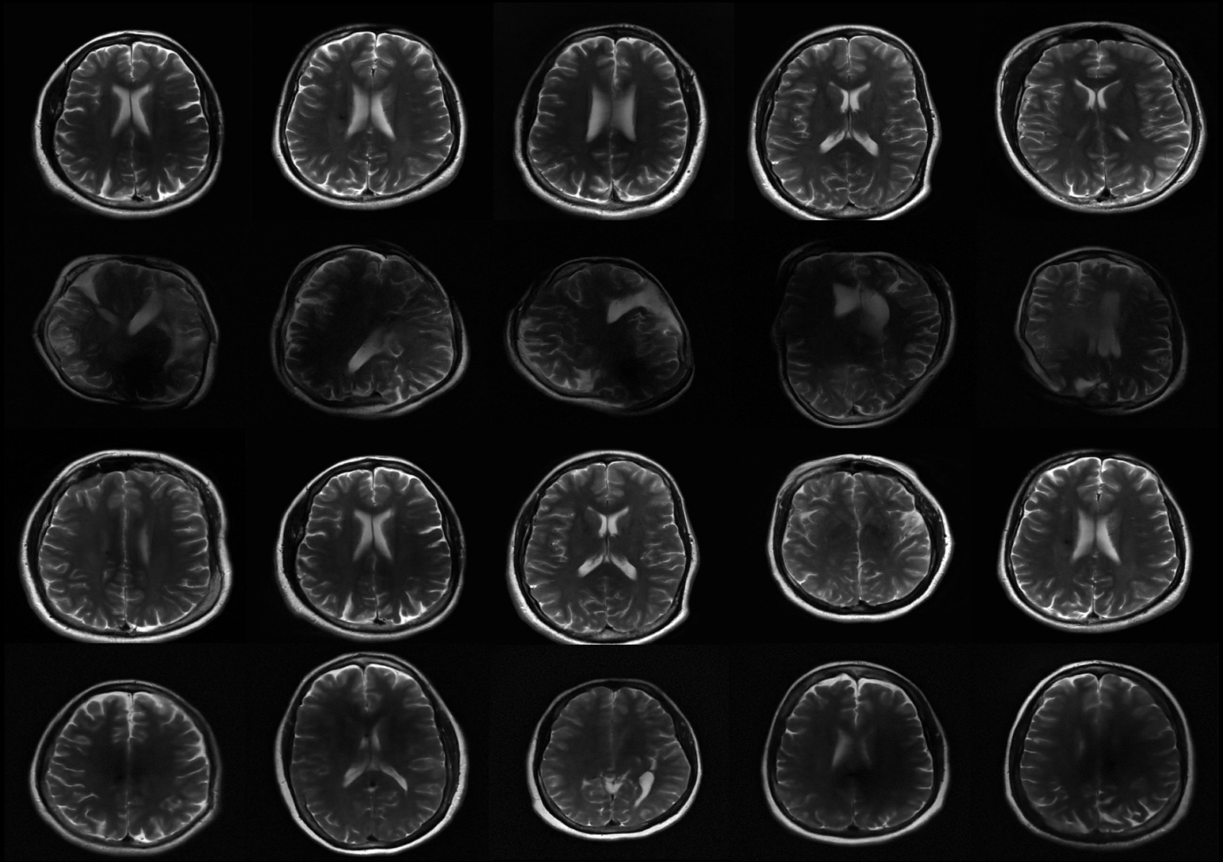
Prior Sampling – $p(x)$

Naive Score
~ 32dB

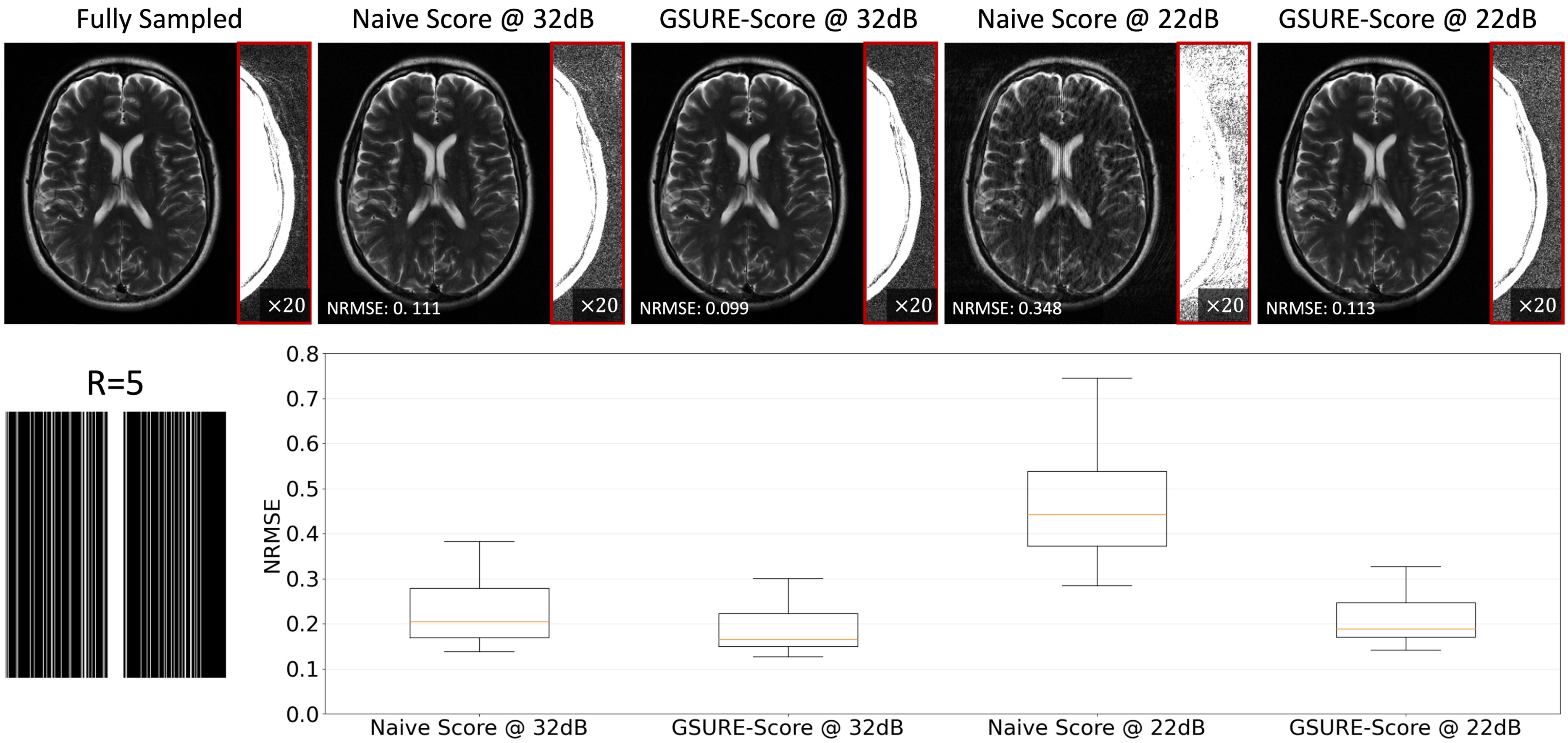
Naive Score
~ 22dB

GSURE-Score
~ 32dB

GSURE-Score
~ 22dB



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



Conclusions

1. Self-supervised techniques like GSURE can successfully remove noise

2. Denoising as a pre-processing step, severely improves the quality of generative priors

3. Priors trained on denoised FastMRI are better inverse problem solvers than naive training

Thank you!

Asad Aali

asad.aali@utexas.edu

https://www.linkedin.com/in/asadaali/

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

M.S. ECE Student

The University of Texas at Austin