Solving Inverse Problems with Score-Based Generative Priors learned from Noisy Data

TEXAS
The University of Texas at Austin

ArXiv Link

UT CSI Lab

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

Introduction

- ❑ Generative models trained on clean data distribution have shown to outperform end-to-end supervised deep learning.
- ❑ A large collection of clean training data is prohibitively expensive to acquire.
- ❑ Our method approximately learns a generative model of the clean distribution from noisy data.
- ❑ We present SURE-Score: a novel loss function that leverages Stein's unbiased risk estimate (SURE) to jointly denoise the data and learn a score function

Forward Models

❑ Multi-Coil MRI

$$y_i = F_\alpha S_i x + n_i$$

Multi-coil MRI data are acquired in the frequency domain by placing multiple RF coils around the imaging anatomy

$x \in \mathbb{C}^N$	Vectorized Image	$F_\alpha \in \mathbb{C}^{\alpha N \times N}$	Fourier Sampling
$S_i \in \mathbb{C}^{N \times N}$	Coil Sensitivity Map	n_i	Gaussian Noise

❑ Multiple-Input Multiple-Output (MIMO) Channels

$$Y = HP + N$$

Point-to-point MIMO baseband communication scenario where transmitters and receivers equipped with N_t and N_r antennas

$P \in \mathbb{C}^{N_t \times N_p}$	Pilot Measurement
$H \in \mathbb{C}^{N_r \times N_t}$	Channel State Information
N	Gaussian Noise

Training SURE-Score

$$\mathcal{L}(\theta) = \alpha\left(\underbrace{\mathbb{E}_{\tilde{H},n_i}\left[\sigma_i^2\left\|s_\theta(\overbrace{\tilde{H}+\sigma_w^2 s_\theta(\tilde{H})}^{f_\theta(\tilde{H})}+n_i)+\frac{n_i}{\sigma_i^2}\right\|_2^2\right]}_{\text{Score Loss}}\right)+\underbrace{\left(\mathbb{E}_{\tilde{H},w}\left[\left\|\overbrace{\sigma_w^2 s_\theta(\tilde{H})}^{f_\theta(\tilde{H})-\tilde{H}}\right\|_2^2+2\sigma_w^2\operatorname{div}_{\tilde{H}}(\overbrace{\tilde{H}+\sigma_w^2 s_\theta(\tilde{H})}^{f_\theta(\tilde{H})})\right]\right)}_{\text{Modified SURE Loss}}$$

Where $\operatorname{div}_{\tilde{H}}(\tilde{H}+\sigma_w^2 s_\theta(\tilde{H})) = tr\left(J_{\tilde{H}+\sigma_w^2 s_\theta(\tilde{H})}\right)$

Where α is appropriate scaling applied to score model

$$\underbrace{\mathbb{E}_{H,w}\left[\left\|f_\theta(\tilde{H})-\tilde{H}\right\|_2^2+2\sigma_w^2\operatorname{div}_{\tilde{H}}\left(f_\theta(\tilde{H})\right)\right]}_{\text{Regular SURE Loss}}$$

SURE-based denoising with s_θ Denoising score matching with s_θ

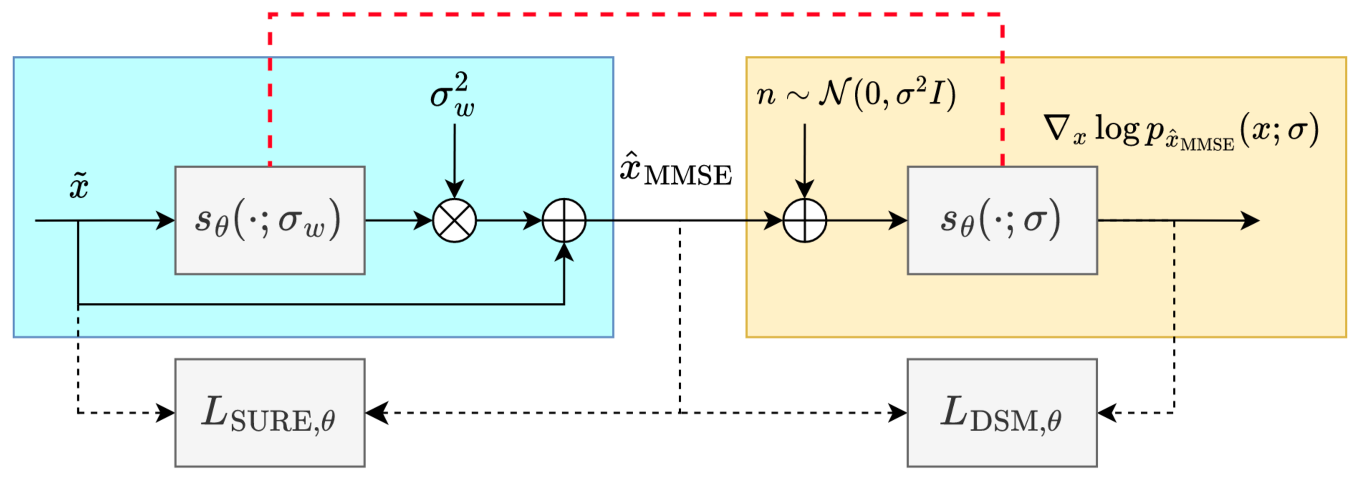

Fig. 1. Flow of SURE-Score during training. The same deep neural network s_θ is used first for denoising and subsequently for denoising score matching.

Generic PyTorch Training Pipeline

Denoising

- ❑ Using *Tweedie's rule* and training score models with: (i) Noisy Data (*Naive*), (ii) SURE-Score, and (iii) Noise-Free (*Supervised*) data
- ❑ Following table lists NRMSE ($\mu \pm \sigma$) of 100 validation Multi-Coil MRI slices.

Denoising Performance (NRMSE)

SNR^w	0 dB	10 dB
Naive	2.48 ± 0.24	0.70 ± 0.07
Supervised	0.21 ± 0.01	0.14 ± 0.01
SURE-Score	0.23 ± 0.01	0.16 ± 0.01

Takeaways:

- ❑ Denoising performance of SURE-Score nearly matches supervised learning
- ❑ Shows consistency between the SURE and score-matching objective.

Annealed Langevin Dynamics

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

For prior sampling $p_{\hat{x}_{MMSe}}$, we set A = 0

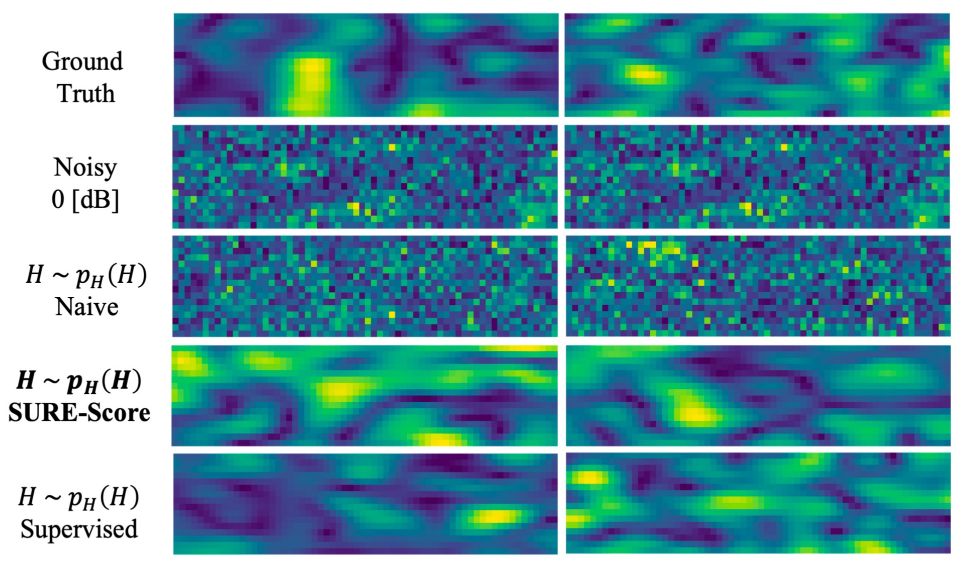

Fig. 2. Prior sampling for three methods: Naive, SURE-Score at SNR^w 0 dB, and Supervised. Each column is different realization of a CDL-C channel.

Posterior Reconstruction – Multi-Coil MRI

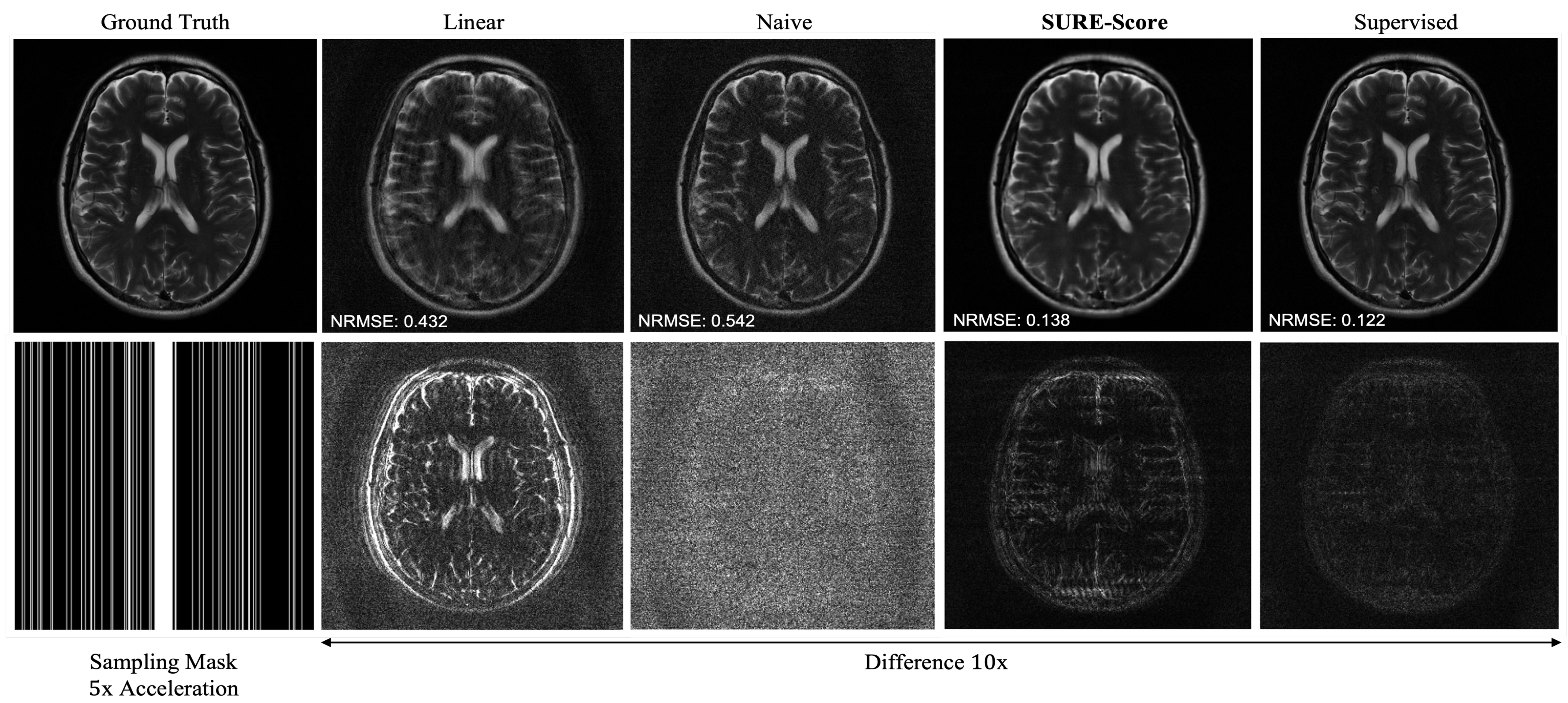

Fig. 3. Multi-coil MRI reconstruction at acceleration factor of 5×. From left to right: fully sampled ground truth, linear reconstruction, posterior sampling after naively training on noisy data at SNR^w, posterior sampling after training with SURE-Score, and posterior sampling after training with noise-free data. The bottom row shows the sampling pattern and difference images for each method, respectively.

Posterior Reconstruction – MIMO Channels

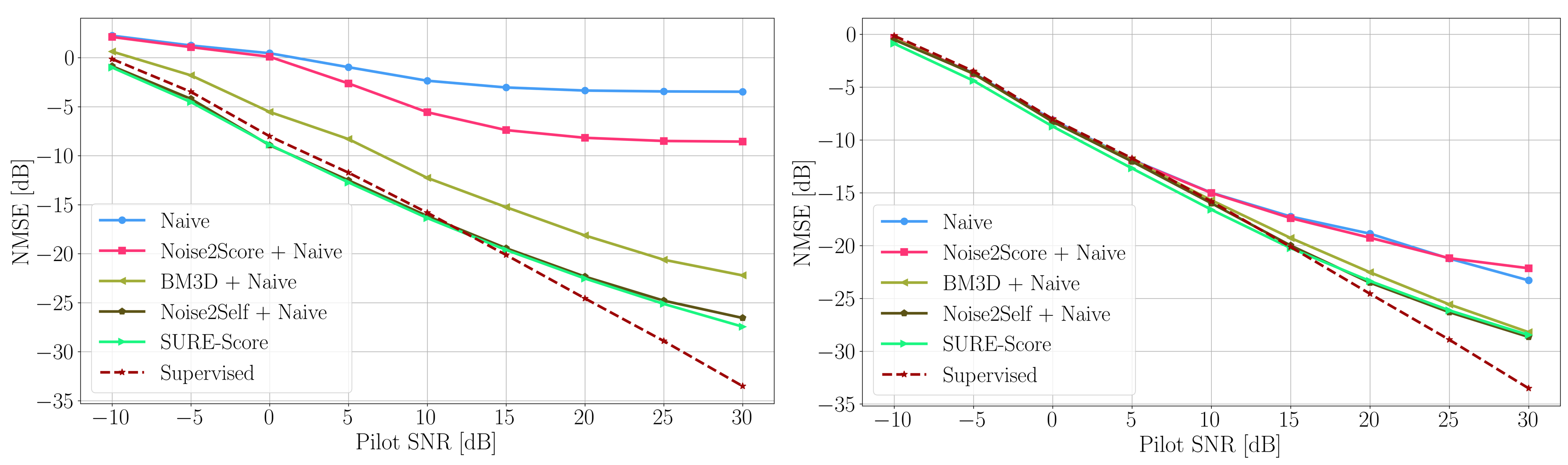

Fig. 3. Channel estimation performance at α = 0.6 (38 pilots) using score models trained on CDL-C channels at SNR^w: 0 dB (left) and 10 dB (right).

Key Takeaways:

- ❑ **SURE-Score** performs close to **optimal** with respect to supervised DSM except at higher pilot SNR
- ❑ Naive training plateaus in estimation performance because of ***overfitting***
- ❑ Noise2Score and BM3D suffer at lower SNR^w and improve at higher SNR
- ❑ Performance gap at high pilot SNR likely due to performance limits of MMSE denoiser and finite training data

Generalized SURE-Score

❑ Goal: Learn the score directly from noisy measurements y

$$y = Ax + n$$

- Where n is a zero-mean Gaussian random vector
- A is full-rank

Methodology:

❑ Utilize extended SURE principle to obtain unbiased MSE estimate for exponential family noise corruption

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

❑ Use a single-network to jointly denoise the data and learn score-function

Discussion and Conclusion

- ❑ Self-supervised techniques can **match** supervised techniques in denoising and inverse problem performance
- ❑ Runtime per iteration increases due to additional pass through the network
- ❑ Choosing hyper-parameters without access to ground truth data is an open challenge
- ❑ **Next Steps:** Our work currently assumes white Gaussian noise corruption but could be extended to arbitrary exponential families

Selected References

1. Y. Song and S. Ermon, "Generative modeling by estimating gradients of the data distribution," Advances in neural information processing systems, vol. 32, 2019.
2. A. Jalal, M. Arvinte, G. Daras, E. Price, A. G. Dimakis, and J. Tamir, "Robust compressed sensing mri with deep generative priors," Advances in Neural Information Processing Systems, vol. 34, pp. 14938–14954, 2021.
3. M.Arvinte and J.I.Tamir,"Mimo channel estimation using score-based generative models," IEEE Transactions on Wireless Communications, 2022.
4. C. A. Metzler, A. Mousavi, R. Heckel, and R. G. Baraniuk, "Unsu- pervised learning with stein's unbiased risk estimator," arXiv preprint arXiv:1805.10531, 2018.