Noisy MRI Reconstruction via MAP Estimation with an Implicit Deep-Denoiser Prior

ImMAP: diffusion-style reconstruction derived from Tweedie's formula and MRI physics

Nikola Janjušević[1] Amirhossein Khalilian-Gourtani[2] Yao Wang[3] Li Feng[1]

[1]Department of Radiology and [2]Department of Neurology, NYU Grossman School of Medicine
[3]Department of Electrical and Computer Engineering, NYU Tandon School of Engineering

NYU Langone Health

CAI²R Center for Advanced Imaging Innovation and Research

NYU TANDON SCHOOL OF ENGINEERING

Part 1

Introduction and Background

NYU Langone Health



image $\boldsymbol{x} \in \mathbb{C}^N$

coil maps $\boldsymbol{s}_c$,
2D-DFT $\boldsymbol{F}$

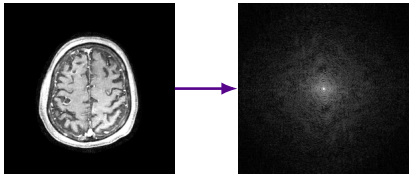
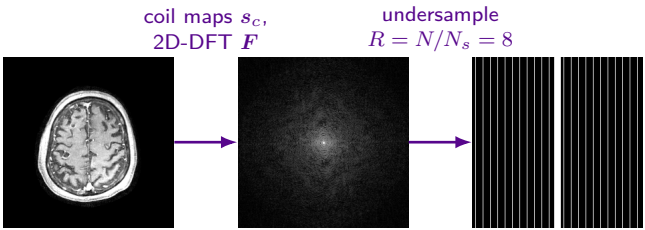
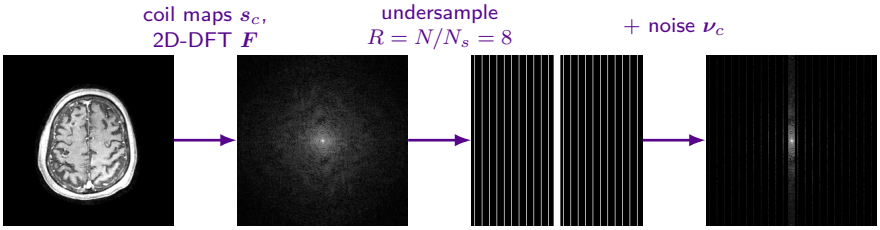


image $\boldsymbol{x} \in \mathbb{C}^N$ k-space $\boldsymbol{F}(\boldsymbol{s}_c \circ \boldsymbol{x})$


coil maps s_c,
2D-DFT F
undersample
$R = N/N_s = 8$
image $x \in \mathbb{C}^N$
k-space $F(s_c \circ x)$
sampling I_Ω

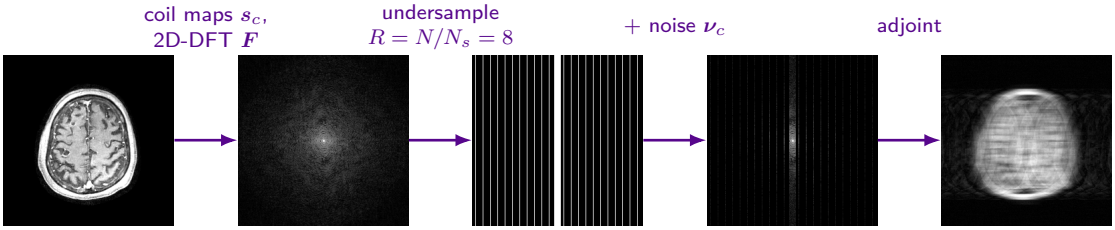
Accelerated multicoil MRI



$$y_c = I_\Omega F(s_c \circ x) + \nu_c, \quad c = 1, \dots, C, \qquad \nu[n] \sim \mathcal{N}(0, \Sigma_y) \ \forall n \quad \text{(noise-whitened; } \Sigma_y \in \mathbb{C}^{C \times C}\text{)}$$

per-coil model

Accelerated multicoil MRI



$$y_c = I_\Omega F(s_c \circ x) + \nu_c, \quad c = 1, \ldots, C, \qquad \nu[n] \sim \mathcal{N}(0, \Sigma_y) \ \forall\, n \quad \text{(noise-whitened; } \Sigma_y \in \mathbb{C}^{C \times C}\text{)}$$

per-coil model

$$y = Ax + \nu, \qquad A = I_\Omega F S : \ \mathbb{C}^N \to \mathbb{C}^{N_s C}$$

encoding operator

Classical regularized

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{y}\|^2_{\boldsymbol{\Sigma}_y^{-1}} + \lambda\mathcal{R}(\boldsymbol{x})$$

hand-crafted prior $\mathcal{R}$
SENSE, TV, wavelets

Reconstruction approaches

Classical regularized

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|^2_{\boldsymbol{\Sigma}_y^{-1}} + \lambda \mathcal{R}(\boldsymbol{x})$$

hand-crafted prior $\mathcal{R}$
SENSE, TV, wavelets

End-to-end deep networks

$$\hat{\boldsymbol{x}} = f_\theta(\boldsymbol{y}; \boldsymbol{A})$$

unrolled optimization,
supervised training
LPDSNet, E2E-VarNet

Reconstruction approaches

Classical regularized

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \; \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|^2_{\boldsymbol{\Sigma}_y^{-1}} + \lambda \mathcal{R}(\boldsymbol{x})$$

hand-crafted prior $\mathcal{R}$
SENSE, TV, wavelets

End-to-end deep networks

$$\hat{\boldsymbol{x}} = f_\theta(\boldsymbol{y}; \boldsymbol{A})$$

unrolled optimization,
supervised training
LPDSNet, E2E-VarNet

Zero-shot diffusion (DDGMs)

$$\nabla_{\boldsymbol{x}} \log p_\sigma(\boldsymbol{x}) \approx \frac{f(\boldsymbol{x}; \sigma) - \boldsymbol{x}}{\sigma^2}$$

score learned by a denoiser,
reverse-time SDE sampling
conditioned on $\boldsymbol{y}$: DDS, ΠGDM

Reconstruction approaches

Classical regularized

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{y}\|^2_{\boldsymbol{\Sigma}_y^{-1}} + \lambda\mathcal{R}(\boldsymbol{x})$$

hand-crafted prior $\mathcal{R}$
SENSE, TV, wavelets

End-to-end deep networks

$$\hat{\boldsymbol{x}} = f_\theta(\boldsymbol{y}; \boldsymbol{A})$$

unrolled optimization,
supervised training
LPDSNet, E2E-VarNet

Zero-shot diffusion (DDGMs)

$$\nabla_{\boldsymbol{x}} \log p_\sigma(\boldsymbol{x}) \approx \frac{f(\boldsymbol{x};\sigma) - \boldsymbol{x}}{\sigma^2}$$

score learned by a denoiser,
reverse-time SDE sampling
conditioned on $\boldsymbol{y}$: DDS, ΠGDM

DDGMs in MRI, open questions:

SDE / Markov-chain modeling decoupled from MRI physics
parameters sensitive to measurement noise
real scanner noise largely unexplored

This work

interpretable construction of deep priors, extended to diffusion





- Start from the stochastic ascent with the prior implicit in a denoiser
  Kadkhodaie & Simoncelli, NeurIPS 2021

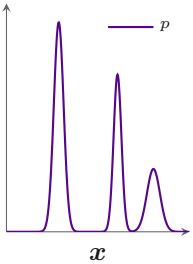


- Start from the stochastic ascent with the prior implicit in a denoiser
  Kadkhodaie & Simoncelli, NeurIPS 2021
- Extend it to measurement noise $\boldsymbol{\Sigma}_y$ and the large-scale MRI encoding operator $\boldsymbol{A}$



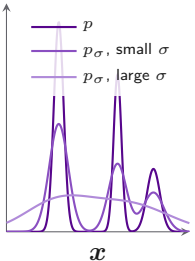
- Start from the stochastic ascent with the prior implicit in a denoiser
  Kadkhodaie & Simoncelli, NeurIPS 2021
- Extend it to measurement noise $\boldsymbol{\Sigma}_y$ and the large-scale MRI encoding operator $\boldsymbol{A}$
- Competitive with end-to-end and diffusion methods on simulated *and* real noise, with a single, insensitive hyperparameter

$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{n}, \quad \boldsymbol{x} \sim p, \quad \boldsymbol{n} \sim g_\sigma = \mathcal{N}(0, \sigma^2 \boldsymbol{I})$$ noisy image



$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{n}, \quad \boldsymbol{x} \sim p, \quad \boldsymbol{n} \sim g_\sigma = \mathcal{N}(0, \sigma^2 \boldsymbol{I}) \qquad \text{noisy image}$$

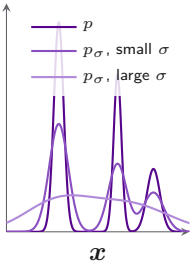
$$p_\sigma(\boldsymbol{z}) = \int g_\sigma(\boldsymbol{z} - \boldsymbol{x})\, p(\boldsymbol{x})\, d\boldsymbol{x} = (g_\sigma * p)(\boldsymbol{z}) \qquad \text{blurred prior}$$



$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{n}, \quad \boldsymbol{x} \sim p, \quad \boldsymbol{n} \sim g_\sigma = \mathcal{N}(0, \sigma^2 \boldsymbol{I}) \qquad \text{noisy image}$$

$$p_\sigma(\boldsymbol{z}) = \int g_\sigma(\boldsymbol{z} - \boldsymbol{x})\, p(\boldsymbol{x})\, d\boldsymbol{x} = (g_\sigma * p)(\boldsymbol{z}) \qquad \text{blurred prior}$$

$$\nabla_{\boldsymbol{z}}\, p_\sigma(\boldsymbol{z}) = \int \nabla_{\boldsymbol{z}}\, g_\sigma(\boldsymbol{z} - \boldsymbol{x})\, p(\boldsymbol{x})\, d\boldsymbol{x} \qquad \text{differentiate}$$

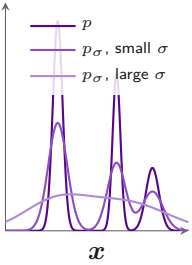


$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{n}, \quad \boldsymbol{x} \sim p, \quad \boldsymbol{n} \sim g_\sigma = \mathcal{N}(0, \sigma^2 \boldsymbol{I}) \qquad \text{noisy image}$$

$$p_\sigma(\boldsymbol{z}) = \int g_\sigma(\boldsymbol{z} - \boldsymbol{x})\, p(\boldsymbol{x})\, d\boldsymbol{x} = (g_\sigma * p)(\boldsymbol{z}) \qquad \text{blurred prior}$$

$$\nabla_{\boldsymbol{z}}\, p_\sigma(\boldsymbol{z}) = \int \nabla_{\boldsymbol{z}}\, g_\sigma(\boldsymbol{z} - \boldsymbol{x})\, p(\boldsymbol{x})\, d\boldsymbol{x} \qquad \text{differentiate}$$

$$= \tfrac{1}{\sigma^2} \int (\boldsymbol{x} - \boldsymbol{z})\, g_\sigma(\boldsymbol{z} - \boldsymbol{x})\, p(\boldsymbol{x})\, d\boldsymbol{x} \qquad \text{Gaussian pdf}$$



using $\nabla_{\boldsymbol{z}}\, g_\sigma(\boldsymbol{z} - \boldsymbol{x}) = \frac{\boldsymbol{x} - \boldsymbol{z}}{\sigma^2}\, g_\sigma(\boldsymbol{z} - \boldsymbol{x})$

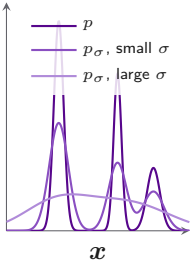
$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{n}, \quad \boldsymbol{x} \sim p, \quad \boldsymbol{n} \sim g_\sigma = \mathcal{N}(0, \sigma^2 \boldsymbol{I}) \qquad \text{noisy image}$$

$$p_\sigma(\boldsymbol{z}) = \int g_\sigma(\boldsymbol{z} - \boldsymbol{x})\, p(\boldsymbol{x})\, d\boldsymbol{x} = (g_\sigma * p)(\boldsymbol{z}) \qquad \text{blurred prior}$$

$$\nabla_{\boldsymbol{z}}\, p_\sigma(\boldsymbol{z}) = \int \nabla_{\boldsymbol{z}}\, g_\sigma(\boldsymbol{z} - \boldsymbol{x})\, p(\boldsymbol{x})\, d\boldsymbol{x} \qquad \text{differentiate}$$

$$= \tfrac{1}{\sigma^2} \int (\boldsymbol{x} - \boldsymbol{z})\, g_\sigma(\boldsymbol{z} - \boldsymbol{x})\, p(\boldsymbol{x})\, d\boldsymbol{x} \qquad \text{Gaussian pdf}$$

$$\sigma^2\, \nabla_{\boldsymbol{z}} \log p_\sigma(\boldsymbol{z}) = \int (\boldsymbol{x} - \boldsymbol{z})\, p(\boldsymbol{x} \,|\, \boldsymbol{z})\, d\boldsymbol{x} \qquad \text{log-derivative, Bayes}$$



using $\nabla_{\boldsymbol{z}}\, g_\sigma(\boldsymbol{z} - \boldsymbol{x}) = \frac{\boldsymbol{x} - \boldsymbol{z}}{\sigma^2}\, g_\sigma(\boldsymbol{z} - \boldsymbol{x})$ and $p(\boldsymbol{x} \,|\, \boldsymbol{z}) = g_\sigma(\boldsymbol{z} - \boldsymbol{x})\, p(\boldsymbol{x}) \,/\, p_\sigma(\boldsymbol{z})$

$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{n}, \quad \boldsymbol{x} \sim p, \quad \boldsymbol{n} \sim g_\sigma = \mathcal{N}(0, \sigma^2 \boldsymbol{I}) \qquad \text{noisy image}$$

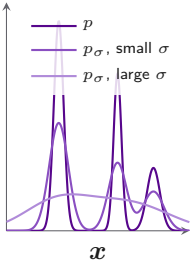
$$p_\sigma(\boldsymbol{z}) = \int g_\sigma(\boldsymbol{z} - \boldsymbol{x})\, p(\boldsymbol{x})\, d\boldsymbol{x} = (g_\sigma * p)(\boldsymbol{z}) \qquad \text{blurred prior}$$

$$\nabla_{\boldsymbol{z}}\, p_\sigma(\boldsymbol{z}) = \int \nabla_{\boldsymbol{z}}\, g_\sigma(\boldsymbol{z} - \boldsymbol{x})\, p(\boldsymbol{x})\, d\boldsymbol{x} \qquad \text{differentiate}$$

$$= \tfrac{1}{\sigma^2} \int (\boldsymbol{x} - \boldsymbol{z})\, g_\sigma(\boldsymbol{z} - \boldsymbol{x})\, p(\boldsymbol{x})\, d\boldsymbol{x} \qquad \text{Gaussian pdf}$$

$$\sigma^2\, \nabla_{\boldsymbol{z}} \log p_\sigma(\boldsymbol{z}) = \int (\boldsymbol{x} - \boldsymbol{z})\, p(\boldsymbol{x} \,|\, \boldsymbol{z})\, d\boldsymbol{x} \qquad \text{log-derivative, Bayes}$$

$$= \mathbb{E}[\boldsymbol{x} \,|\, \boldsymbol{z}] - \boldsymbol{z} \qquad \text{Tweedie}$$

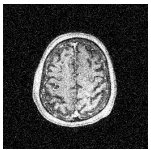


using $\nabla_{\boldsymbol{z}}\, g_\sigma(\boldsymbol{z} - \boldsymbol{x}) = \frac{\boldsymbol{x} - \boldsymbol{z}}{\sigma^2}\, g_\sigma(\boldsymbol{z} - \boldsymbol{x})$ and $p(\boldsymbol{x} \,|\, \boldsymbol{z}) = g_\sigma(\boldsymbol{z} - \boldsymbol{x})\, p(\boldsymbol{x}) \,/\, p_\sigma(\boldsymbol{z})$

the MMSE estimate $\mathbb{E}[\boldsymbol{x} \,|\, \boldsymbol{z}]$ is one gradient-ascent step on $\log p_\sigma$ from $\boldsymbol{z}$, with step-size σ^2

$$\sigma^2 \nabla_{\boldsymbol{z}} \log p_\sigma(\boldsymbol{z}) = \mathbb{E}[\boldsymbol{x} \mid \boldsymbol{z}] - \boldsymbol{z}$$

Tweedie



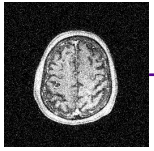
$\boldsymbol{z} \sim p_\sigma$

$$\sigma^2 \nabla_{\boldsymbol{z}} \log p_\sigma(\boldsymbol{z}) = \mathbb{E}[\boldsymbol{x} \mid \boldsymbol{z}] - \boldsymbol{z}$$

Tweedie

$$f = \arg\min_f \ \mathbb{E}\|\boldsymbol{x} - f(\boldsymbol{z};\sigma)\|_2^2$$

MMSE denoiser



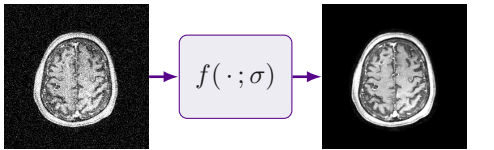
$\boldsymbol{z} \sim p_\sigma$

$f(\cdot\,;\sigma)$

$$\sigma^2 \nabla_{\boldsymbol{z}} \log p_\sigma(\boldsymbol{z}) = \mathbb{E}[\boldsymbol{x} \mid \boldsymbol{z}] - \boldsymbol{z}$$ Tweedie

$$f = \arg\min_f \ \mathbb{E}\|\boldsymbol{x} - f(\boldsymbol{z};\sigma)\|_2^2$$ MMSE denoiser

$$f(\boldsymbol{z};\sigma) \approx \mathbb{E}[\boldsymbol{x} \mid \boldsymbol{z}]$$ trained over σ

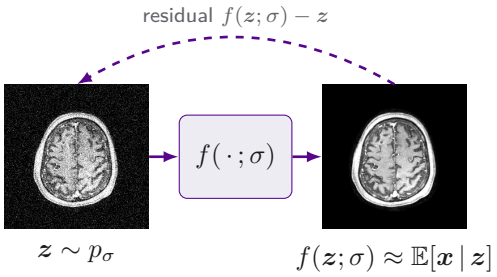


$$\sigma^2 \nabla_{\boldsymbol{z}} \log p_\sigma(\boldsymbol{z}) = \mathbb{E}[\boldsymbol{x} \mid \boldsymbol{z}] - \boldsymbol{z} \qquad \text{Tweedie}$$

$$f = \arg\min_f \ \mathbb{E}\|\boldsymbol{x} - f(\boldsymbol{z};\sigma)\|_2^2 \qquad \text{MMSE denoiser}$$

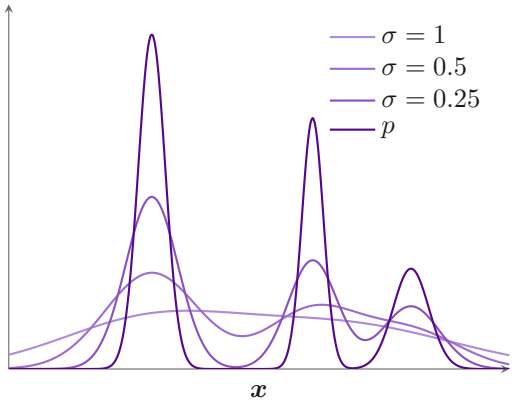
$$f(\boldsymbol{z};\sigma) \approx \mathbb{E}[\boldsymbol{x} \mid \boldsymbol{z}] \qquad \text{trained over } \sigma$$

$$\sigma^2 \nabla_{\boldsymbol{z}} \log p_\sigma(\boldsymbol{z}) \approx f(\boldsymbol{z};\sigma) - \boldsymbol{z} \qquad \text{score} \approx \text{residual}$$



the denoiser residual $f(\boldsymbol{z};\sigma) - \boldsymbol{z}$ is one gradient-ascent step on $\log p_\sigma$ with step-size σ^2

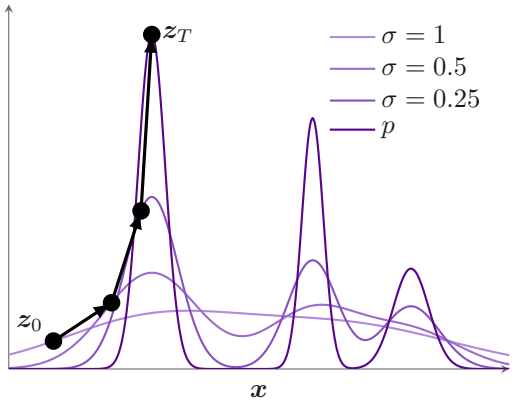
Key insight: $\sigma^2 \nabla_{\boldsymbol{z}} \log p_\sigma(\boldsymbol{z})$, the score of the *noisy* distribution, is an approximate ascent direction on the *true* distribution p.



Key insight: $\sigma^2 \nabla_{\boldsymbol{z}} \log p_\sigma(\boldsymbol{z})$, the score of the *noisy* distribution, is an approximate ascent direction on the *true* distribution p.

$$\boldsymbol{z}_{t+1} = \boldsymbol{z}_t + h_t\Big(f(\boldsymbol{z}_t) - \boldsymbol{z}_t\Big) + \boldsymbol{\epsilon}_t$$

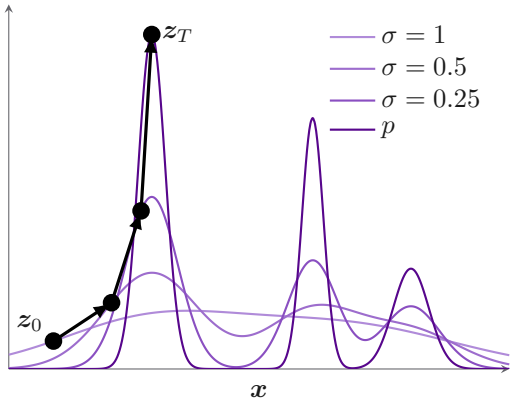
$$\boldsymbol{z}_0 \sim \mathcal{N}(0, \boldsymbol{I}), \quad h_t \in [0,1],$$
$$\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \gamma_t^2 \boldsymbol{I})$$



Key insight: $\sigma^2 \nabla_{\boldsymbol{z}} \log p_\sigma(\boldsymbol{z})$, the score of the *noisy* distribution, is an approximate ascent direction on the *true* distribution p.

$$\boldsymbol{z}_{t+1} = \boldsymbol{z}_t + h_t\Big(f(\boldsymbol{z}_t) - \boldsymbol{z}_t\Big) + \boldsymbol{\epsilon}_t$$

$$\boldsymbol{z}_0 \sim \mathcal{N}(0, \boldsymbol{I}), \quad h_t \in [0,1],$$
$$\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \gamma_t^2 \boldsymbol{I})$$



coarse-to-fine ascent on $\log p$:
each fractional step reduces σ
and sharpens the next ascent direction

$$\arg\max_{\boldsymbol{x}} \ \log p(\boldsymbol{x} \mid \boldsymbol{y}) \ = \ \arg\max_{\boldsymbol{x}} \ \log p(\boldsymbol{x}) + \log p(\boldsymbol{y} \mid \boldsymbol{x}) \qquad \text{Bayes' rule}$$

$$\arg\max_{\boldsymbol{x}} \ \log p(\boldsymbol{x} \,|\, \boldsymbol{y}) \ = \ \arg\max_{\boldsymbol{x}} \ \log p(\boldsymbol{x}) + \log p(\boldsymbol{y} \,|\, \boldsymbol{x}) \qquad \text{Bayes' rule}$$

The measurement-conditioned sampler of Kadkhodaie & Simoncelli assumes

- noise-free measurements $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$

MRI: $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\nu}, \ \ \boldsymbol{\nu} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_y)$

$$\arg\max_{\boldsymbol{x}} \ \log p(\boldsymbol{x} \,|\, \boldsymbol{y}) \ = \ \arg\max_{\boldsymbol{x}} \ \log p(\boldsymbol{x}) + \log p(\boldsymbol{y} \,|\, \boldsymbol{x}) \qquad \text{Bayes' rule}$$

The measurement-conditioned sampler of Kadkhodaie & Simoncelli assumes

- noise-free measurements $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ — MRI: $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\nu}, \ \ \boldsymbol{\nu} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_y)$
- an easily SVD-decomposable $\boldsymbol{A}$ — MRI: $\boldsymbol{A} = \boldsymbol{I}_\Omega \boldsymbol{F} \boldsymbol{S}$, large-scale, multicoil

$$\arg\max_{\boldsymbol{x}} \ \log p(\boldsymbol{x} \,|\, \boldsymbol{y}) \ = \ \arg\max_{\boldsymbol{x}} \ \log p(\boldsymbol{x}) + \log p(\boldsymbol{y} \,|\, \boldsymbol{x}) \qquad \text{Bayes' rule}$$

The measurement-conditioned sampler of Kadkhodaie & Simoncelli assumes

- noise-free measurements $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ — MRI: $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\nu}, \ \ \boldsymbol{\nu} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_y)$
- an easily SVD-decomposable $\boldsymbol{A}$ — MRI: $\boldsymbol{A} = \boldsymbol{I}_\Omega \boldsymbol{F} \boldsymbol{S}$, large-scale, multicoil

ImMAP: a measurement-conditioned stochastic ascent that accounts for $\boldsymbol{\Sigma}_y$ and a general MRI encoding operator $\boldsymbol{A}$

Part 2

Proposed Method: ImMAP

NYU Langone Health

$$\boldsymbol{x}^{\diamond} = \arg\max_{\boldsymbol{x}} \; p(\boldsymbol{x} \,|\, \boldsymbol{y}) \;=\; \arg\max_{\boldsymbol{x}} \; \log p(\boldsymbol{x}) + \log p(\boldsymbol{y} \,|\, \boldsymbol{x})$$

MAP objective

$$\boldsymbol{x}^{\diamond} = \arg\max_{\boldsymbol{x}} \; p(\boldsymbol{x} \,|\, \boldsymbol{y}) \;=\; \arg\max_{\boldsymbol{x}} \; \log p(\boldsymbol{x}) + \log p(\boldsymbol{y} \,|\, \boldsymbol{x}) \qquad \text{MAP objective}$$

$$\boldsymbol{z}_{t+1} = \boldsymbol{z}_t + h_t \sigma_t^2 \Big(\underbrace{\nabla_{\boldsymbol{z}_t} \log p_{\sigma_t}(\boldsymbol{z}_t)}_{\text{prior score}} + \underbrace{\nabla_{\boldsymbol{z}_t} \log p(\boldsymbol{y} \,|\, \boldsymbol{z}_t)}_{\text{likelihood score}} \Big) + \boldsymbol{\epsilon}_t, \qquad \boldsymbol{z}_t \sim p_{\sigma_t} \qquad \text{stochastic ascent}$$

$$\boldsymbol{x}^{\diamond} = \arg\max_{\boldsymbol{x}} \; p(\boldsymbol{x} \,|\, \boldsymbol{y}) \;=\; \arg\max_{\boldsymbol{x}} \; \log p(\boldsymbol{x}) + \log p(\boldsymbol{y} \,|\, \boldsymbol{x})$$

MAP objective

$$\boldsymbol{z}_{t+1} = \boldsymbol{z}_t + h_t \sigma_t^2 \Big(\underbrace{\nabla_{\boldsymbol{z}_t} \log p_{\sigma_t}(\boldsymbol{z}_t)}_{\text{prior score}} + \underbrace{\nabla_{\boldsymbol{z}_t} \log p(\boldsymbol{y} \,|\, \boldsymbol{z}_t)}_{\text{likelihood score}} \Big) + \boldsymbol{\epsilon}_t, \qquad \boldsymbol{z}_t \sim p_{\sigma_t}$$

stochastic ascent

Prior score: Tweedie's formula

$$\sigma_t^2 \, \nabla_{\boldsymbol{z}_t} \log p_{\sigma_t}(\boldsymbol{z}_t) \approx f(\boldsymbol{z}_t; \sigma_t) - \boldsymbol{z}_t$$

$$\boldsymbol{x}^{\diamond} = \arg\max_{\boldsymbol{x}} \; p(\boldsymbol{x} \,|\, \boldsymbol{y}) \;=\; \arg\max_{\boldsymbol{x}} \; \log p(\boldsymbol{x}) + \log p(\boldsymbol{y} \,|\, \boldsymbol{x})$$

MAP objective

$$\boldsymbol{z}_{t+1} = \boldsymbol{z}_t + h_t \sigma_t^2 \Big(\underbrace{\nabla_{\boldsymbol{z}_t} \log p_{\sigma_t}(\boldsymbol{z}_t)}_{\text{prior score}} + \underbrace{\nabla_{\boldsymbol{z}_t} \log p(\boldsymbol{y} \,|\, \boldsymbol{z}_t)}_{\text{likelihood score}} \Big) + \boldsymbol{\epsilon}_t, \qquad \boldsymbol{z}_t \sim p_{\sigma_t}$$

stochastic ascent

Prior score: Tweedie's formula

$$\sigma_t^2 \, \nabla_{\boldsymbol{z}_t} \log p_{\sigma_t}(\boldsymbol{z}_t) \approx f(\boldsymbol{z}_t; \sigma_t) - \boldsymbol{z}_t$$

Likelihood score: ?

$$\nabla_{\boldsymbol{z}_t} \log p(\boldsymbol{y} \,|\, \boldsymbol{z}_t), \qquad \boldsymbol{z}_t \sim \mathcal{N}(\boldsymbol{x}, \, \sigma_t^2 \boldsymbol{I})$$

$$f(\boldsymbol{z}_t) \sim \mathcal{N}\big(\mathbb{E}[\boldsymbol{x} \mid \boldsymbol{z}_t],\ \boldsymbol{V}_t\big) \qquad \text{Laplace approx.}$$

$$\boldsymbol{V}_t = \tfrac{\sigma_t^2}{1+\sigma_t^2}\,\boldsymbol{I} \qquad \text{modeling choice}$$

Likelihood score via a Laplace approximation

$$f(\boldsymbol{z}_t) \sim \mathcal{N}\big(\mathbb{E}[\boldsymbol{x} \,|\, \boldsymbol{z}_t],\ \boldsymbol{V}_t\big) \qquad \text{Laplace approx.}$$

$$\boldsymbol{V}_t = \tfrac{\sigma_t^2}{1+\sigma_t^2}\,\boldsymbol{I} \qquad \text{modeling choice}$$

$$p(\boldsymbol{y} \,|\, \boldsymbol{z}_t) = \mathcal{N}\big(\boldsymbol{y} \,\big|\, \boldsymbol{A}\ \mathbb{E}[\boldsymbol{x} \,|\, \boldsymbol{z}_t],\ \boldsymbol{\Sigma}_t\big) \qquad \text{likelihood}$$

$$\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}_y + \boldsymbol{A}\boldsymbol{V}_t\boldsymbol{A}^H \qquad \text{noise + prior variance}$$

Likelihood score via a Laplace approximation

$$f(\boldsymbol{z}_t) \sim \mathcal{N}\big(\mathbb{E}[\boldsymbol{x} \,|\, \boldsymbol{z}_t],\ \boldsymbol{V}_t\big) \qquad \text{Laplace approx.}$$

$$\boldsymbol{V}_t = \tfrac{\sigma_t^2}{1+\sigma_t^2}\,\boldsymbol{I} \qquad \text{modeling choice}$$

$$p(\boldsymbol{y} \,|\, \boldsymbol{z}_t) = \mathcal{N}\big(\boldsymbol{y} \,\big|\, \boldsymbol{A}\ \mathbb{E}[\boldsymbol{x} \,|\, \boldsymbol{z}_t],\ \boldsymbol{\Sigma}_t\big) \qquad \text{likelihood}$$

$$\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}_y + \boldsymbol{A}\boldsymbol{V}_t\boldsymbol{A}^H \qquad \text{noise + prior variance}$$

$$\begin{aligned}\nabla_{\boldsymbol{z}_t} \log p(\boldsymbol{y} \,|\, \boldsymbol{z}_t) &= -\nabla_{\boldsymbol{z}_t}\, \tfrac{1}{2} \|\boldsymbol{\Sigma}_t^{-1/2}\big(\boldsymbol{A}\ \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}_t] - \boldsymbol{y}\big)\|_2^2 \\ &= -\nabla_{\boldsymbol{z}_t}\, \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}_t]^H\ \boldsymbol{A}^H\, \boldsymbol{\Sigma}_t^{-1}\big(\boldsymbol{A}\ \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}_t] - \boldsymbol{y}\big)\end{aligned}$$

likelihood score

$$f(\boldsymbol{z}_t) \sim \mathcal{N}\big(\mathbb{E}[\boldsymbol{x} \mid \boldsymbol{z}_t],\ \boldsymbol{V}_t\big) \qquad \text{Laplace approx.}$$

$$\boldsymbol{V}_t = \tfrac{\sigma_t^2}{1+\sigma_t^2}\,\boldsymbol{I} \qquad \text{modeling choice}$$

$$p(\boldsymbol{y} \mid \boldsymbol{z}_t) = \mathcal{N}\big(\boldsymbol{y} \,\big|\, \boldsymbol{A}\ \mathbb{E}[\boldsymbol{x} \mid \boldsymbol{z}_t],\ \boldsymbol{\Sigma}_t\big) \qquad \text{likelihood}$$

$$\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}_y + \boldsymbol{A}\boldsymbol{V}_t\boldsymbol{A}^H \qquad \text{noise + prior variance}$$

$$\begin{aligned}\nabla_{\boldsymbol{z}_t} \log p(\boldsymbol{y} \mid \boldsymbol{z}_t) &= -\nabla_{\boldsymbol{z}_t}\, \tfrac{1}{2}\|\boldsymbol{\Sigma}_t^{-1/2}\big(\boldsymbol{A}\ \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}_t] - \boldsymbol{y}\big)\|_2^2 \\ &= -\nabla_{\boldsymbol{z}_t}\, \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}_t]^H\ \boldsymbol{A}^H\, \boldsymbol{\Sigma}_t^{-1}\big(\boldsymbol{A}\ \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}_t] - \boldsymbol{y}\big)\end{aligned}$$

likelihood score

Implementation

- $\mathbb{E}[\boldsymbol{x}|\boldsymbol{z}_t] \approx f(\boldsymbol{z}_t)$
- $\nabla_{\boldsymbol{z}_t}\, \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}_t]^H \boldsymbol{v} \approx \boldsymbol{J}_t^H \boldsymbol{v}$: VJP of f via autodiff
- $\boldsymbol{\Sigma}_t^{-1}\boldsymbol{u}$: symmetric system, solved with CG

Likelihood score via a Laplace approximation

$$f(\boldsymbol{z}_t) \sim \mathcal{N}\big(\mathbb{E}[\boldsymbol{x} \mid \boldsymbol{z}_t],\ \boldsymbol{V}_t\big) \qquad \text{Laplace approx.}$$

$$\boldsymbol{V}_t = \tfrac{\sigma_t^2}{1+\sigma_t^2}\,\boldsymbol{I} \qquad \text{modeling choice}$$

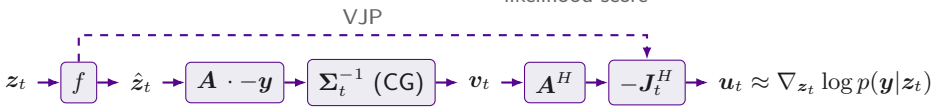
$$p(\boldsymbol{y} \mid \boldsymbol{z}_t) = \mathcal{N}\big(\boldsymbol{y} \,\big|\, \boldsymbol{A}\ \mathbb{E}[\boldsymbol{x} \mid \boldsymbol{z}_t],\ \boldsymbol{\Sigma}_t\big) \qquad \text{likelihood}$$

$$\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}_y + \boldsymbol{A}\boldsymbol{V}_t\boldsymbol{A}^H \qquad \text{noise + prior variance}$$

$$\begin{aligned}\nabla_{\boldsymbol{z}_t} \log p(\boldsymbol{y} \mid \boldsymbol{z}_t) &= -\nabla_{\boldsymbol{z}_t}\, \tfrac{1}{2} \|\boldsymbol{\Sigma}_t^{-1/2}\big(\boldsymbol{A}\ \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}_t] - \boldsymbol{y}\big)\|_2^2 \\ &= -\nabla_{\boldsymbol{z}_t}\, \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}_t]^H\ \boldsymbol{A}^H\, \boldsymbol{\Sigma}_t^{-1}\big(\boldsymbol{A}\ \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}_t] - \boldsymbol{y}\big)\end{aligned} \qquad \text{likelihood score}$$

Implementation

- $\mathbb{E}[\boldsymbol{x}|\boldsymbol{z}_t] \approx f(\boldsymbol{z}_t)$
- $\nabla_{\boldsymbol{z}_t}\, \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}_t]^H \boldsymbol{v} \approx \boldsymbol{J}_t^H \boldsymbol{v}$: VJP of f via autodiff
- $\boldsymbol{\Sigma}_t^{-1}\boldsymbol{u}$: symmetric system, solved with CG



The ImMAP algorithm

Input: $\boldsymbol{y} \sim \mathcal{CN}(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{\Sigma}_y)$, $\boldsymbol{\Sigma}_y$, $\boldsymbol{A}$, f; β=0.01, σ_L=0.01, h_0=0.01

Output: $\boldsymbol{x}^\diamond = \arg\max_{\boldsymbol{x}} \log p(\boldsymbol{x}) + \log p(\boldsymbol{y} \,|\, \boldsymbol{x})$

1 Draw $\boldsymbol{z}_0 \sim \mathcal{CN}(0, \boldsymbol{I})$; $t = 1$, $\sigma_t^2 = 1$

2 **while** $\sigma_t > \sigma_L$ **do**

11 $t \leftarrow t + 1$

12 **return** $\boldsymbol{x}^\diamond = \boldsymbol{z}_t$

Inputs

- $\boldsymbol{A}$: MRI encoding operator (coil maps, DFT, sampling)
- $\boldsymbol{\Sigma}_y$: coil-noise covariance, estimated from multicoil data
  Rep2Rep: Janjušević et al., MRM 2025, arXiv:2504.17698
- f: noise-adaptive denoiser

The ImMAP algorithm

Input: $\boldsymbol{y} \sim \mathcal{CN}(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{\Sigma}_y)$, $\boldsymbol{\Sigma}_y$, $\boldsymbol{A}$, f; β=0.01, σ_L=0.01, h_0=0.01

Output: $\boldsymbol{x}^\diamond = \arg\max_{\boldsymbol{x}} \log p(\boldsymbol{x}) + \log p(\boldsymbol{y} \mid \boldsymbol{x})$

Draw $\boldsymbol{z}_0 \sim \mathcal{CN}(0, \boldsymbol{I})$; $t = 1$, $\sigma_t^2 = 1$

while $\sigma_t > \sigma_L$ **do**

- $\hat{\boldsymbol{z}}_t,\ \boldsymbol{J}_t^H \leftarrow \text{pullback}\big(f(\cdot\,; \sigma_t),\ \boldsymbol{z}_t\big)$ — denoise + VJP
- $\sigma_t^2 \leftarrow \|\hat{\boldsymbol{z}}_t - \boldsymbol{z}_t\|_2^2 / N$ — noise-level estimate
- $t \leftarrow t + 1$

return $\boldsymbol{x}^\diamond = \boldsymbol{z}_t$

Prior score (Tweedie)

$$\sigma_t^2 \nabla \log p_{\sigma_t}(\boldsymbol{z}_t) \approx \hat{\boldsymbol{z}}_t - \boldsymbol{z}_t$$

σ_t re-estimated from the residual every iteration

The ImMAP algorithm

Input: $\boldsymbol{y} \sim \mathcal{CN}(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{\Sigma}_y)$, $\boldsymbol{\Sigma}_y$, $\boldsymbol{A}$, f; β=0.01, σ_L=0.01, h_0=0.01

Output: $\boldsymbol{x}^\diamond = \arg\max_{\boldsymbol{x}} \log p(\boldsymbol{x}) + \log p(\boldsymbol{y} \mid \boldsymbol{x})$

Draw $\boldsymbol{z}_0 \sim \mathcal{CN}(0, \boldsymbol{I})$; $t = 1$, $\sigma_t^2 = 1$

while $\sigma_t > \sigma_L$ **do**

- $\hat{\boldsymbol{z}}_t,\ \boldsymbol{J}_t^H \leftarrow \text{pullback}\big(f(\cdot\,; \sigma_t),\ \boldsymbol{z}_t\big)$ — denoise + VJP
- $\sigma_t^2 \leftarrow \|\hat{\boldsymbol{z}}_t - \boldsymbol{z}_t\|_2^2 / N$ — noise-level estimate
- $\boldsymbol{\Sigma}_t \leftarrow \boldsymbol{\Sigma}_y + \frac{\sigma_t^2}{1+\sigma_t^2}\boldsymbol{A}\boldsymbol{A}^H$ — Laplace approximation
- $\boldsymbol{v}_t \leftarrow \boldsymbol{\Sigma}_t^{-1}\big(\boldsymbol{A}\hat{\boldsymbol{z}}_t - \boldsymbol{y}\big)$ — conjugate gradient
- $\boldsymbol{u}_t \leftarrow -\boldsymbol{J}_t^H \boldsymbol{A}^H \boldsymbol{v}_t$ — likelihood score
- $t \leftarrow t + 1$

return $\boldsymbol{x}^\diamond = \boldsymbol{z}_t$

Likelihood score (Laplace)

$$\boldsymbol{v}_t = \boldsymbol{\Sigma}_t^{-1}\big(\boldsymbol{A}\hat{\boldsymbol{z}}_t - \boldsymbol{y}\big)$$

$$\boldsymbol{u}_t = -\boldsymbol{J}_t^H \boldsymbol{A}^H \boldsymbol{v}_t \approx \nabla_{\boldsymbol{z}_t} \log p(\boldsymbol{y} \mid \boldsymbol{z}_t)$$

one VJP and one CG solve per iteration

The ImMAP algorithm

Input: $\boldsymbol{y} \sim \mathcal{CN}(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{\Sigma}_y)$, $\boldsymbol{\Sigma}_y$, $\boldsymbol{A}$, f; β=0.01, σ_L=0.01, h_0=0.01

Output: $\boldsymbol{x}^\diamond = \arg\max_{\boldsymbol{x}} \log p(\boldsymbol{x}) + \log p(\boldsymbol{y} \,|\, \boldsymbol{x})$

Draw $\boldsymbol{z}_0 \sim \mathcal{CN}(0, \boldsymbol{I})$; $t = 1$, $\sigma_t^2 = 1$

while $\sigma_t > \sigma_L$ **do**

- $\hat{\boldsymbol{z}}_t,\ \boldsymbol{J}_t^H \leftarrow \text{pullback}\big(f(\cdot\,; \sigma_t),\ \boldsymbol{z}_t\big)$ — denoise + VJP
- $\sigma_t^2 \leftarrow \|\hat{\boldsymbol{z}}_t - \boldsymbol{z}_t\|_2^2 / N$ — noise-level estimate
- $\boldsymbol{\Sigma}_t \leftarrow \boldsymbol{\Sigma}_y + \frac{\sigma_t^2}{1+\sigma_t^2}\,\boldsymbol{A}\boldsymbol{A}^H$ — Laplace approximation
- $\boldsymbol{v}_t \leftarrow \boldsymbol{\Sigma}_t^{-1}\big(\boldsymbol{A}\hat{\boldsymbol{z}}_t - \boldsymbol{y}\big)$ — conjugate gradient
- $\boldsymbol{u}_t \leftarrow -\boldsymbol{J}_t^H \boldsymbol{A}^H \boldsymbol{v}_t$ — likelihood score
- $h_t \leftarrow h_0 \frac{t}{1+h_0(t-1)}$ — step-size
- $\gamma_t^2 \leftarrow \sigma_t^2\big((1-\beta h_t)^2 - (1-h_t)^2\big)$, $\boldsymbol{\epsilon}_t \sim \mathcal{CN}(0, \gamma_t^2 \boldsymbol{I})$
- $t \leftarrow t + 1$

return $\boldsymbol{x}^\diamond = \boldsymbol{z}_t$

Step-size and noise injection

$h_t \nearrow 1$; $\beta \in (0, 1]$ trades noise injection ($\beta \to 0$) against determinism ($\beta = 1$)

The ImMAP algorithm

Input: $\boldsymbol{y} \sim \mathcal{CN}(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{\Sigma}_y)$, $\boldsymbol{\Sigma}_y$, $\boldsymbol{A}$, f; β=0.01, σ_L=0.01, h_0=0.01

Output: $\boldsymbol{x}^\diamond = \arg\max_{\boldsymbol{x}} \log p(\boldsymbol{x}) + \log p(\boldsymbol{y} \mid \boldsymbol{x})$

Draw $\boldsymbol{z}_0 \sim \mathcal{CN}(0, \boldsymbol{I})$; $t = 1$, $\sigma_t^2 = 1$

while $\sigma_t > \sigma_L$ **do**

- $\hat{\boldsymbol{z}}_t,\ \boldsymbol{J}_t^H \leftarrow \text{pullback}\big(f(\cdot\,; \sigma_t),\ \boldsymbol{z}_t\big)$ — denoise + VJP
- $\sigma_t^2 \leftarrow \|\hat{\boldsymbol{z}}_t - \boldsymbol{z}_t\|_2^2 / N$ — noise-level estimate
- $\boldsymbol{\Sigma}_t \leftarrow \boldsymbol{\Sigma}_y + \frac{\sigma_t^2}{1+\sigma_t^2}\,\boldsymbol{A}\boldsymbol{A}^H$ — Laplace approximation
- $\boldsymbol{v}_t \leftarrow \boldsymbol{\Sigma}_t^{-1}\big(\boldsymbol{A}\hat{\boldsymbol{z}}_t - \boldsymbol{y}\big)$ — conjugate gradient
- $\boldsymbol{u}_t \leftarrow -\boldsymbol{J}_t^H \boldsymbol{A}^H \boldsymbol{v}_t$ — likelihood score
- $h_t \leftarrow h_0 \frac{t}{1+h_0(t-1)}$ — step-size
- $\gamma_t^2 \leftarrow \sigma_t^2\big((1-\beta h_t)^2 - (1-h_t)^2\big)$, $\boldsymbol{\epsilon}_t \sim \mathcal{CN}(0, \gamma_t^2 \boldsymbol{I})$
- $\boldsymbol{z}_{t+1} \leftarrow \boldsymbol{z}_t + h_t\big(\hat{\boldsymbol{z}}_t - \boldsymbol{z}_t + \sigma_t^2 \boldsymbol{u}_t\big) + \boldsymbol{\epsilon}_t$ — stochastic ascent
- $t \leftarrow t + 1$

return $\boldsymbol{x}^\diamond = \boldsymbol{z}_t$

Update

$$\boldsymbol{z}_{t+1} = \boldsymbol{z}_t + h_t \sigma_t^2 \big(\nabla \log p_{\sigma_t}(\boldsymbol{z}_t) + \nabla \log p(\boldsymbol{y} \mid \boldsymbol{z}_t)\big) + \boldsymbol{\epsilon}_t$$

stop when the estimated noise-level falls below σ_L

$$h_t = h_0 \frac{t}{1 + h_0(t-1)} \nearrow 1$$ step-size

Noise injection and the implicit schedule

one hyperparameter $\beta \in (0, 1]$

$$h_t = h_0 \frac{t}{1 + h_0(t-1)} \nearrow 1 \qquad \text{step-size}$$

$$\gamma_t^2 = \sigma_t^2 \big((1 - \beta h_t)^2 - (1 - h_t)^2 \big) \qquad \text{noise injection}$$

$$h_t = h_0 \frac{t}{1 + h_0(t-1)} \nearrow 1 \qquad \text{step-size}$$

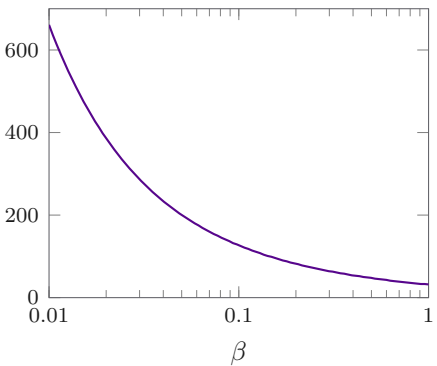
$$\gamma_t^2 = \sigma_t^2\big((1-\beta h_t)^2 - (1-h_t)^2\big) \qquad \text{noise injection}$$

$$\sigma_{t+1}^2 = (1-h_t)^2\sigma_t^2 + \gamma_t^2 = (1-\beta h_t)^2\,\sigma_t^2 \qquad \text{schedule}$$

(last line assumes an ideal denoiser)

- $\beta = 1$: no noise injection, deterministic
- $\beta \to 0$: maximal noise injection, slower schedule
- in practice σ_t is *estimated* from the residual: the schedule is not followed strictly

iterations until $\sigma_t < \sigma_L$ (ideal denoiser)



Relation to SDE-based diffusion models

Denoising diffusion generative models

$$d\boldsymbol{x}_t = h(\boldsymbol{x}_t, t)\, dt + l(t)\, d\boldsymbol{W}_t, \qquad \boldsymbol{x}_0 \sim p_0$$

reverse-time solution $\Rightarrow$ sample from p_0 or $p_0(\boldsymbol{x} \,|\, \boldsymbol{y})$

modeling choices: h, l, T, discretization
all unrelated to the imaging inverse problem

Relation to SDE-based diffusion models

Denoising diffusion generative models

$$d\boldsymbol{x}_t = h(\boldsymbol{x}_t, t)\, dt + l(t)\, d\boldsymbol{W}_t, \qquad \boldsymbol{x}_0 \sim p_0$$

reverse-time solution $\Rightarrow$ sample from p_0 or $p_0(\boldsymbol{x} \,|\, \boldsymbol{y})$

modeling choices: h, l, T, discretization
all unrelated to the imaging inverse problem

ImMAP

$$\boldsymbol{z}_{t+1} = \boldsymbol{z}_t + h_t \sigma_t^2 \big(\nabla \log p_{\sigma_t}(\boldsymbol{z}_t) + \nabla \log p(\boldsymbol{y} \,|\, \boldsymbol{z}_t) \big) + \boldsymbol{\epsilon}_t$$

assumptions: Tweedie's formula + MRI physics $(\boldsymbol{A}, \boldsymbol{\Sigma}_y)$

hyperparameter: β (and stopping level σ_L)

Relation to SDE-based diffusion models

Denoising diffusion generative models

$$d\boldsymbol{x}_t = h(\boldsymbol{x}_t, t)\, dt + l(t)\, d\boldsymbol{W}_t, \qquad \boldsymbol{x}_0 \sim p_0$$

reverse-time solution $\Rightarrow$ sample from p_0 or $p_0(\boldsymbol{x} \,|\, \boldsymbol{y})$

modeling choices: h, l, T, discretization
all unrelated to the imaging inverse problem

ImMAP

$$\boldsymbol{z}_{t+1} = \boldsymbol{z}_t + h_t \sigma_t^2 \big(\nabla \log p_{\sigma_t}(\boldsymbol{z}_t) + \nabla \log p(\boldsymbol{y} \,|\, \boldsymbol{z}_t)\big) + \boldsymbol{\epsilon}_t$$

assumptions: Tweedie's formula + MRI physics $(\boldsymbol{A}, \boldsymbol{\Sigma}_y)$

hyperparameter: β (and stopping level σ_L)

Decomposed Diffusion Sampling (DDS) Chung et al., ICLR 2024

few-step (M) conjugate gradient on a Krylov subspace assumed tangent to the noisy image manifold; hyperparameters M, η, and γ (data-consistency vs. denoiser trade-off)

$\Rightarrow$ γ must be re-tuned per noise level

Part 3

Experiments

NYU Langone Health

Dataset and two noise settings

66 training volumes / 16 held-out test volumes (∼60 slices each) · $8\times$ Cartesian, 4% autocalibration · ESPIRiT coil maps

Dataset and two noise settings

66 training volumes / 16 held-out test volumes (∼60 slices each) · $8\times$ Cartesian, 4% autocalibration · ESPIRiT coil maps

1. Simulated noise
$\sigma_y \in \{0.01, 0.02, 0.05\}$

$$\boldsymbol{x}' = \boldsymbol{S}^H \boldsymbol{F}^H \boldsymbol{y} \qquad \text{coil-combined reference}$$

$$\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{x}' + \boldsymbol{\nu}', \qquad \boldsymbol{\nu}' \sim \mathcal{N}(0, \sigma_y^2 \boldsymbol{I}) \qquad \text{AWGN}$$

reference $\boldsymbol{x}'$: no irreducible error

Dataset and two noise settings

66 training volumes / 16 held-out test volumes (∼60 slices each) · $8\times$ Cartesian, 4% autocalibration · ESPIRiT coil maps

1. Simulated noise
$\sigma_y \in \{0.01, 0.02, 0.05\}$

$$\boldsymbol{x}' = \boldsymbol{S}^H \boldsymbol{F}^H \boldsymbol{y}$$ coil-combined reference

$$\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{x}' + \boldsymbol{\nu}', \qquad \boldsymbol{\nu}' \sim \mathcal{N}(0, \sigma_y^2 \boldsymbol{I})$$ AWGN

reference $\boldsymbol{x}'$: no irreducible error

2. Measurement noise "σ_y = meas."

$$\boldsymbol{y}' = \boldsymbol{I}_\Omega \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{I}_\Omega \boldsymbol{\nu}$$ scanner noise

$\hat{\boldsymbol{\Sigma}}_y$ estimated in the multicoil image domain

noise covariance

reference $\boldsymbol{x}'$ contains $\boldsymbol{\nu}$: irreducible error

$\hat{\boldsymbol{\Sigma}}_y$: Rep2Rep wavelet-filter method, MRM 2025

Training and baselines

	Training data	Notes
Denoiser f (LPDSNet arch., $\boldsymbol{A} = \boldsymbol{I}$)	fully-sampled recons. + AWGN, $\sigma_t \in [0, 0.5]$	one noise-adaptive model, shared by ImMAP and DDS
LPDSNet, end-to-end ($\times 2$)	$8\times$ simulated noise / measurement noise	same architecture: 30 layers, 169 channels, 7×7
DDS Chung et al., ICLR 2024	zero-shot	$M = 5$, $\eta = 0.8$, 83 iterations, γ grid-searched per noise level
ImMAP (ours)	zero-shot	$\beta = 0.01$, $\sigma_L = 0.01$, $h_0 = 0.01$

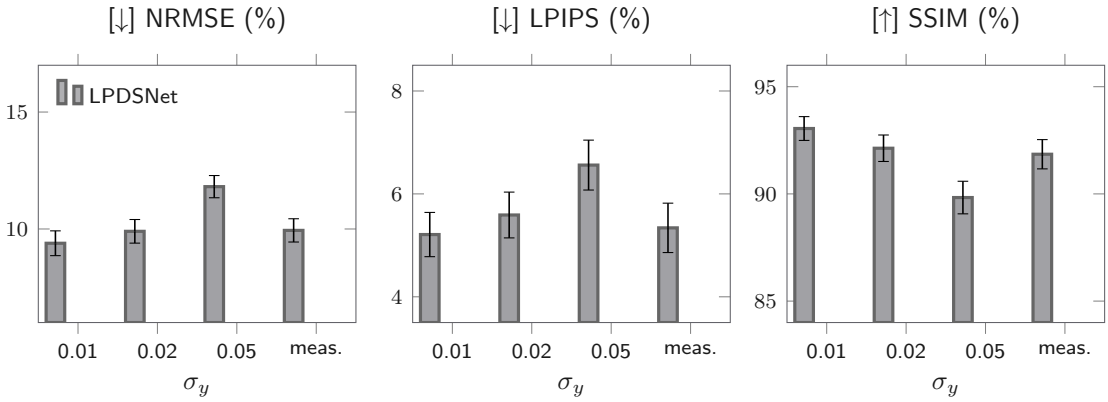
Training and baselines

	Training data	Notes
Denoiser f (LPDSNet arch., $\boldsymbol{A} = \boldsymbol{I}$)	fully-sampled recons. + AWGN, $\sigma_t \in [0, 0.5]$	one noise-adaptive model, shared by ImMAP and DDS
LPDSNet, end-to-end ($\times 2$)	$8\times$ simulated noise / measurement noise	same architecture: 30 layers, 169 channels, 7×7
DDS Chung et al., ICLR 2024	zero-shot	$M = 5$, $\eta = 0.8$, 83 iterations, γ grid-searched per noise level
ImMAP (ours)	zero-shot	$\beta = 0.01$, $\sigma_L = 0.01$, $h_0 = 0.01$

metrics: NRMSE $\downarrow$, LPIPS $\downarrow$, SSIM $\uparrow$ (all in %); mean $\pm$ s.e.m. over 16 test volumes

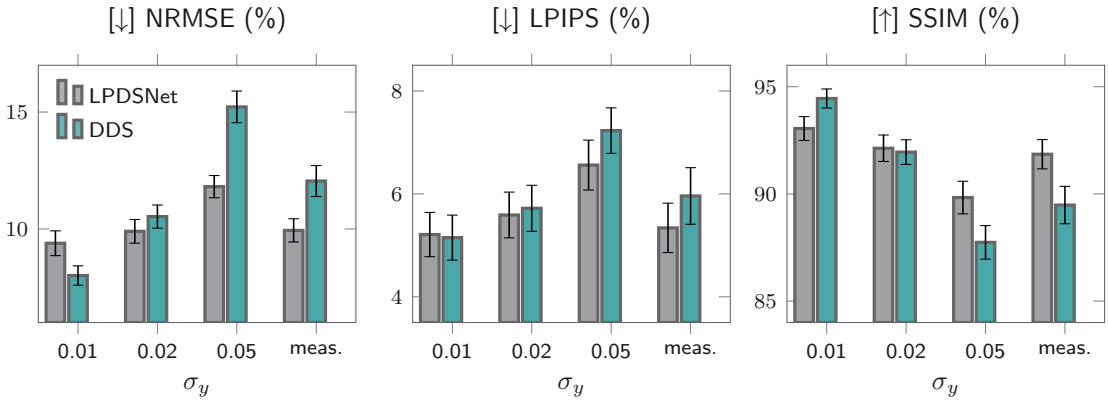
Quantitative results

8×, T1w brain test set; mean ± s.e.m. over 16 volumes



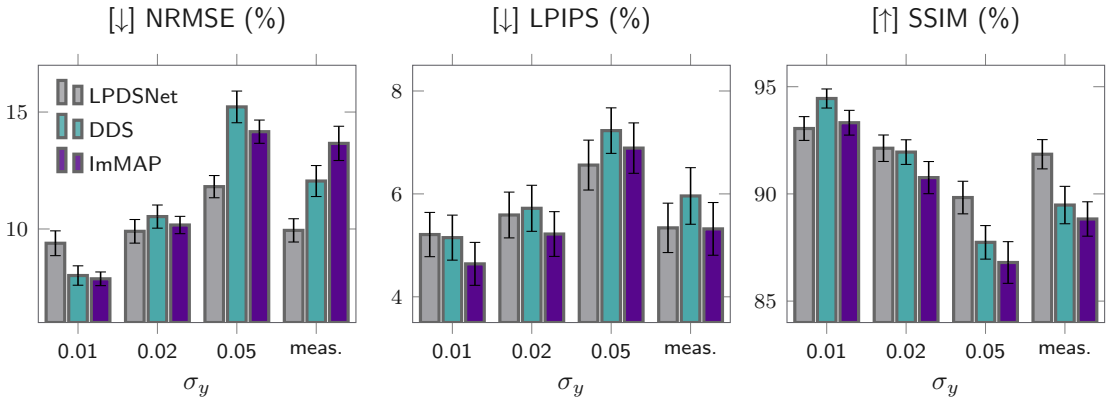
Quantitative results

8×, T1w brain test set; mean ± s.e.m. over 16 volumes



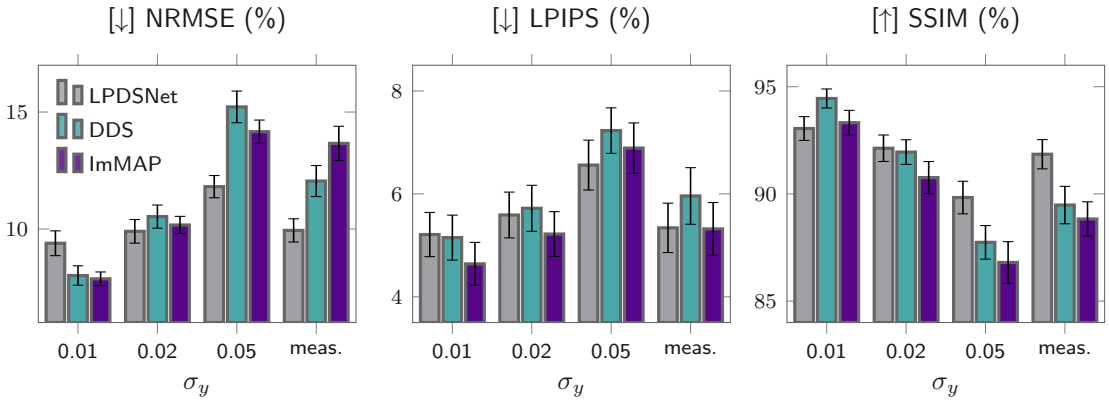
Quantitative results

8×, T1w brain test set; mean ± s.e.m. over 16 volumes



Quantitative results

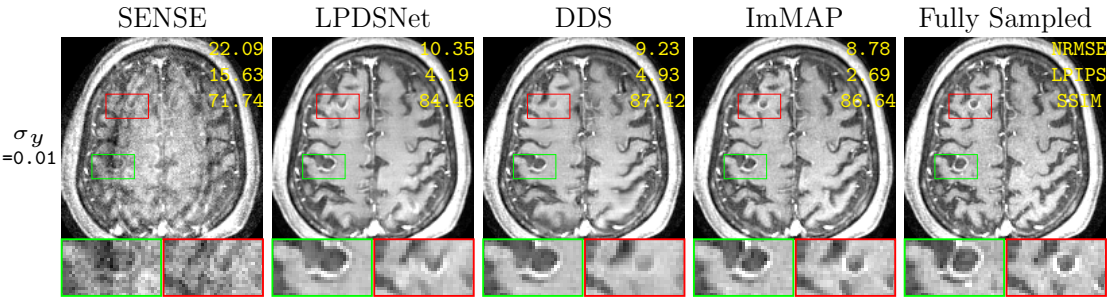
8×, T1w brain test set; mean ± s.e.m. over 16 volumes



ImMAP is competitive with DDS on all metrics and favored by LPIPS at low noise; on measurement noise the noisy reference favors the smoother end-to-end LPDSNet

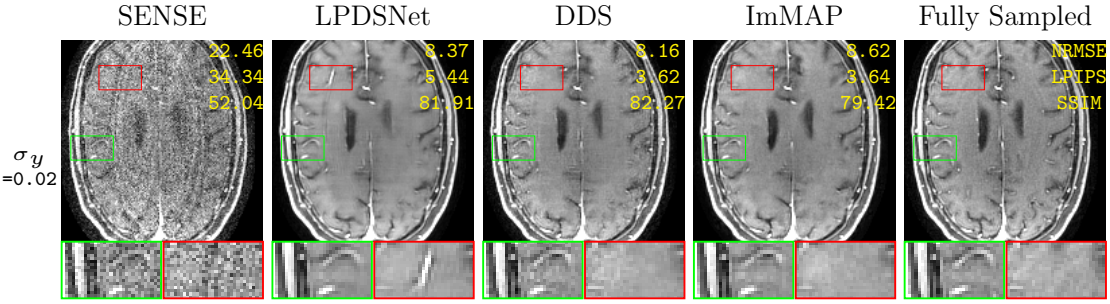
Visual comparison

$8\times$, $\sigma_y = 0.01$; NRMSE/LPIPS/SSIM (%) in yellow



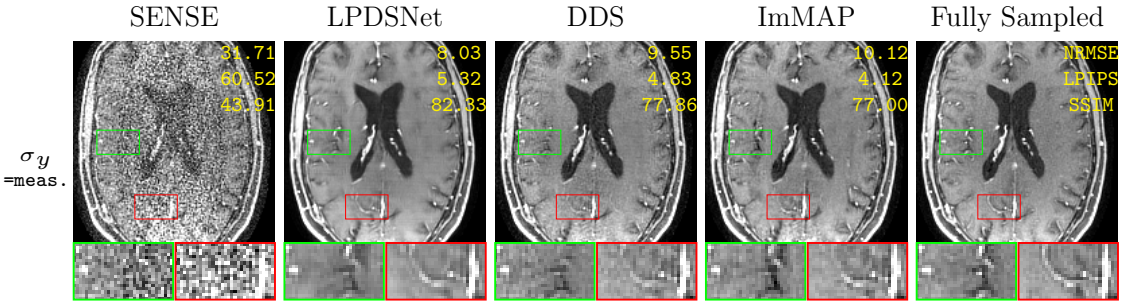
Visual comparison

$8\times$, $\sigma_y = 0.02$; NRMSE/LPIPS/SSIM (%) in yellow



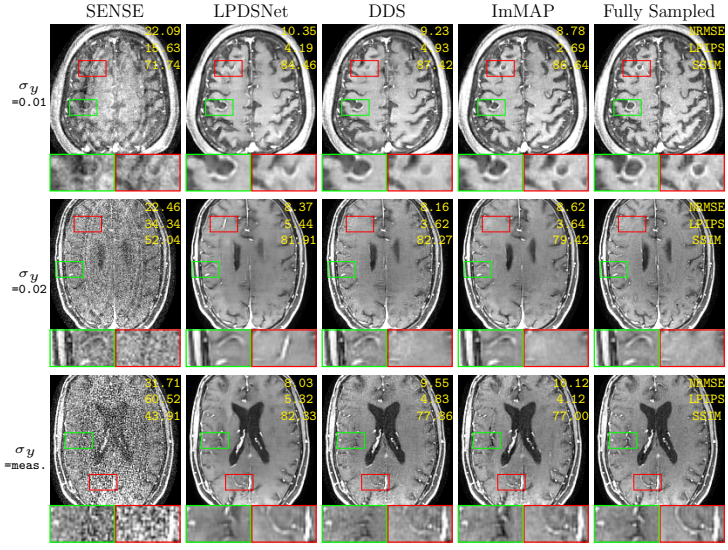
Visual comparison

$8\times$, σ_y = meas.; NRMSE/LPIPS/SSIM (%) in yellow



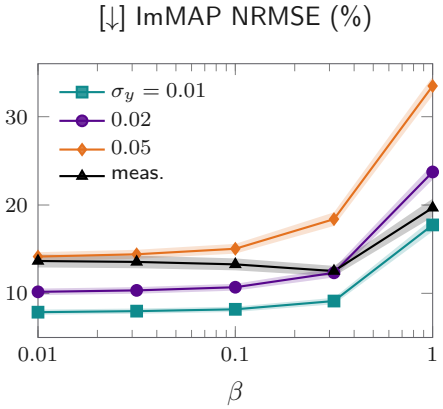
Visual comparison

8×; NRMSE/LPIPS/SSIM (%) in yellow



Ablation: ImMAP noise injection β

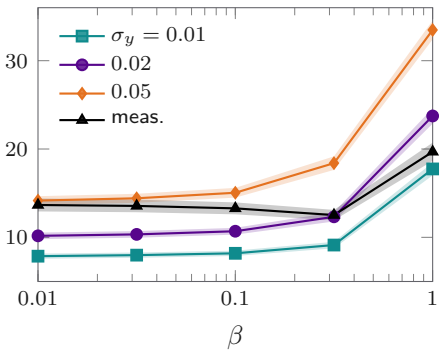
8×; bands: s.e.m. over 16 volumes



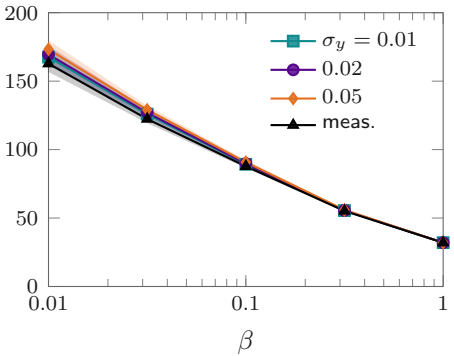
Ablation: ImMAP noise injection β

8×; bands: s.e.m. over 16 volumes

[↓] ImMAP NRMSE (%)



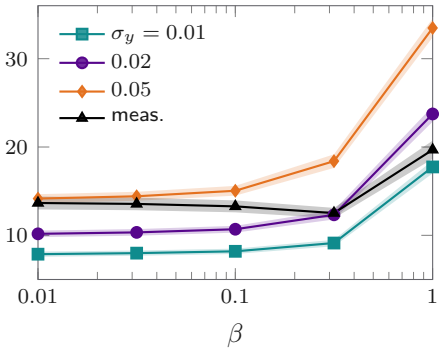
[↓] ImMAP iterations



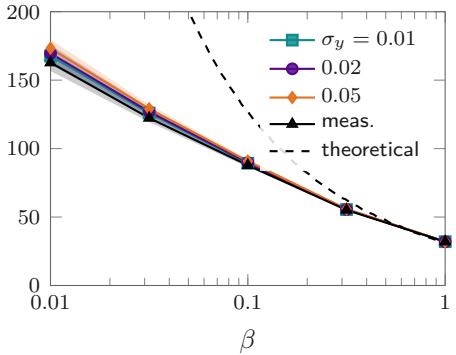
Ablation: ImMAP noise injection β

8×; bands: s.e.m. over 16 volumes

[↓] ImMAP NRMSE (%)

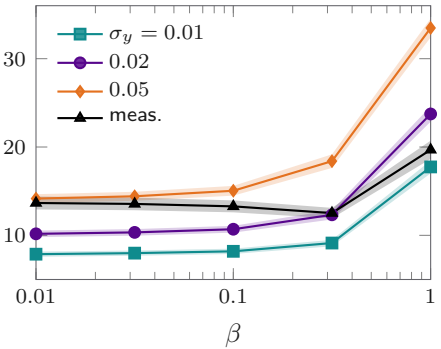


[↓] ImMAP iterations

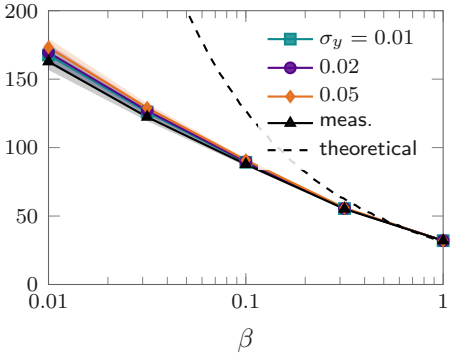


Ablation: ImMAP noise injection β

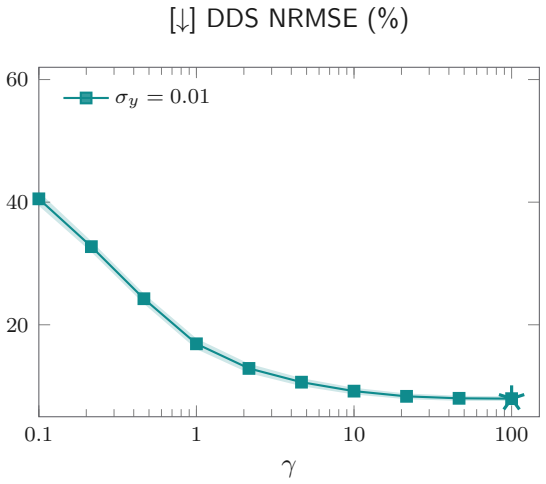
[↓] ImMAP NRMSE (%)



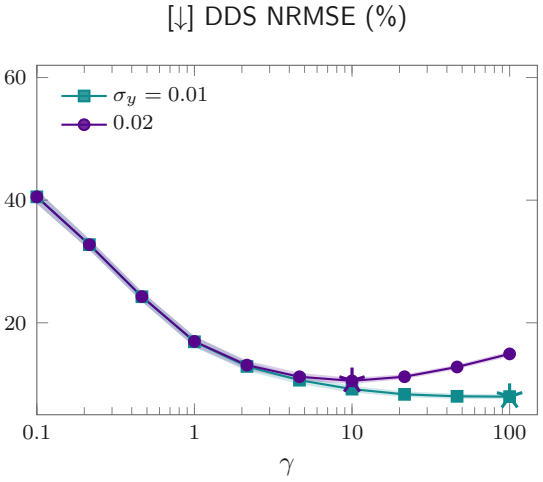
[↓] ImMAP iterations



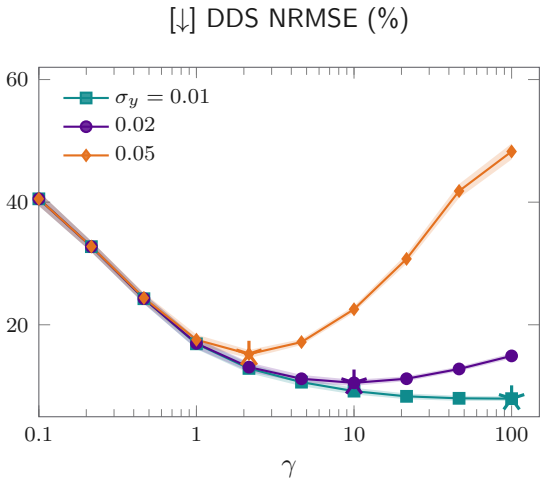
$\beta \to 0$ is best and stable up to $\beta \approx 0.3$; convergence takes far fewer iterations than the schedule predicts. Set β as low as the compute budget allows (0.01 suffices).



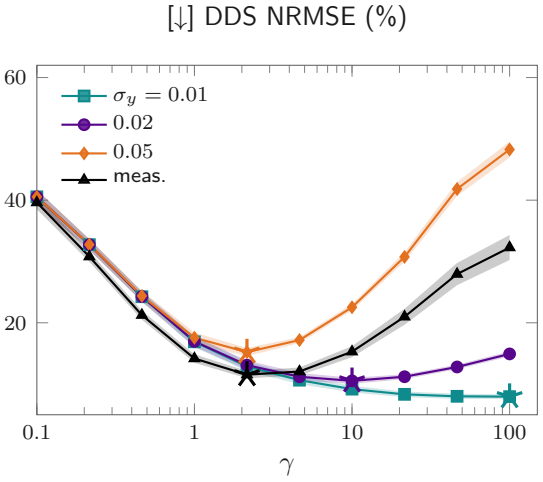
- γ: Lagrange multiplier in DDS's few-step CG
  (data consistency vs. denoiser)



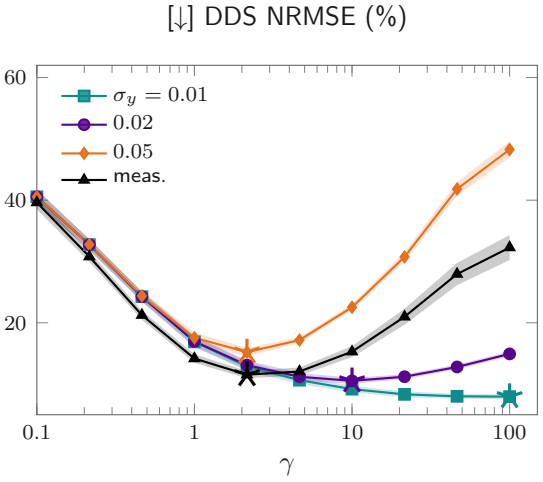
- γ: Lagrange multiplier in DDS's few-step CG
  (data consistency vs. denoiser)
- the optimum moves with the noise level:
  $\gamma^\star = 100 \to 10 \to 2.2 \to 2.2$



- γ: Lagrange multiplier in DDS's few-step CG
  (data consistency vs. denoiser)
- the optimum moves with the noise level:
  $\gamma^\star = 100 \to 10 \to 2.2 \to 2.2$



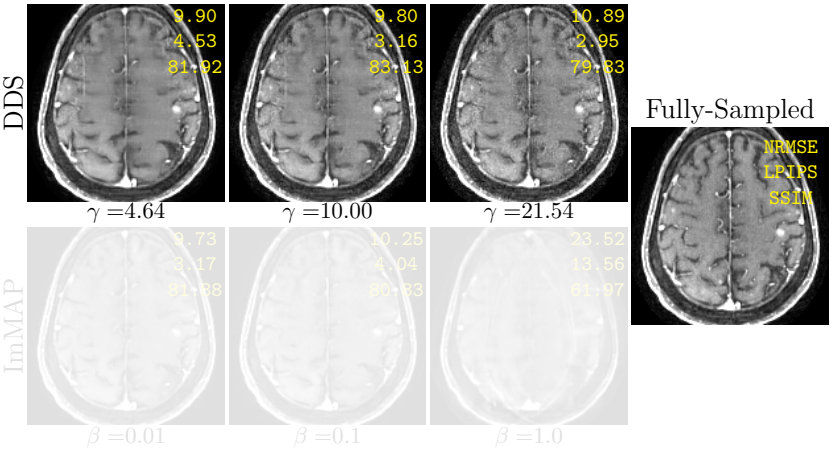
- γ: Lagrange multiplier in DDS's few-step CG
  (data consistency vs. denoiser)
- the optimum moves with the noise level:
  $\gamma^\star = 100 \to 10 \to 2.2 \to 2.2$
- tuning needs *reference data at the inference noise level*



- γ: Lagrange multiplier in DDS's few-step CG
  (data consistency vs. denoiser)
- the optimum moves with the noise level:
  $\gamma^\star = 100 \to 10 \to 2.2 \to 2.2$
- tuning needs *reference data at the inference noise level*
- ImMAP: the choice $\beta \to 0$ is unambiguous

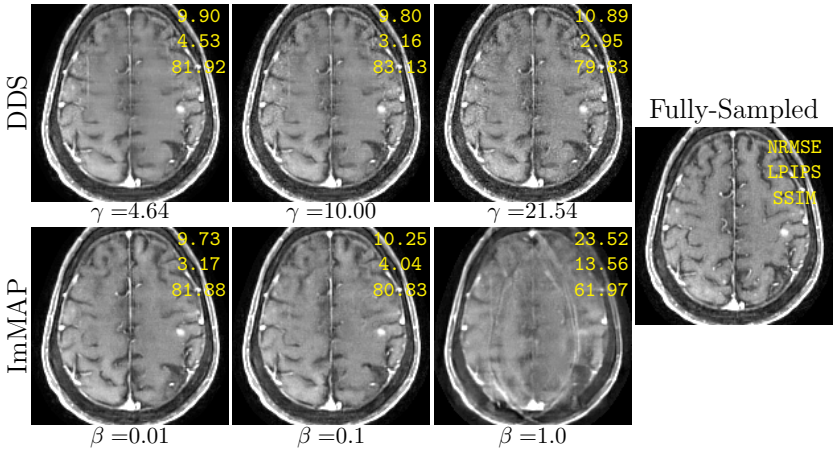
Ablation: visual

$8\times$, $\sigma_y = 0.02$; NRMSE/LPIPS/SSIM (%) on each image



Ablation: visual

$8\times$, $\sigma_y = 0.02$; NRMSE/LPIPS/SSIM (%) on each image



DDS: smoothing and artifacts as γ decreases, and the metrics disagree on the best γ.
ImMAP: same visual quality at $\beta = 0.01$ and 0.1, best metrics at the lower β.

Part 4

Conclusion

NYU Langone Health

Conclusion

- **ImMAP**: diffusion-style MRI reconstruction that integrates the acquisition physics $(\boldsymbol{A}, \boldsymbol{\Sigma}_y)$ directly into MAP estimation, derived from Tweedie's formula alone

Conclusion

- **ImMAP**: diffusion-style MRI reconstruction that integrates the acquisition physics $(\boldsymbol{A}, \boldsymbol{\Sigma}_y)$ directly into MAP estimation, derived from Tweedie's formula alone
- Competitive with end-to-end (LPDSNet) and diffusion (DDS) methods on simulated and real scanner noise; preserves detail without noise amplification or oversmoothing

Conclusion

- **ImMAP**: diffusion-style MRI reconstruction that integrates the acquisition physics $(\boldsymbol{A}, \boldsymbol{\Sigma}_y)$ directly into MAP estimation, derived from Tweedie's formula alone
- Competitive with end-to-end (LPDSNet) and diffusion (DDS) methods on simulated and real scanner noise; preserves detail without noise amplification or oversmoothing
- Robust to its single hyperparameter β; DDS requires per-noise-level tuning of γ

Conclusion

- **ImMAP**: diffusion-style MRI reconstruction that integrates the acquisition physics $(\boldsymbol{A}, \boldsymbol{\Sigma}_y)$ directly into MAP estimation, derived from Tweedie's formula alone
- Competitive with end-to-end (LPDSNet) and diffusion (DDS) methods on simulated and real scanner noise; preserves detail without noise amplification or oversmoothing
- Robust to its single hyperparameter β; DDS requires per-noise-level tuning of γ
- A reliable, interpretable baseline that clarifies the practical utility and limits of diffusion models for accelerated MRI

Thank you

Noisy MRI Reconstruction via MAP Estimation with an Implicit Deep-Denoiser Prior

Nikola Janjušević, Amirhossein Khalilian-Gourtani, Yao Wang, Li Feng

nikola.janjusevic@nyulangone.org

Supported in part by NIH R01EB030549, EB031083, R21EB032917, and P41EB017183, under the rubric of the Center for Advanced Imaging Innovation and Research (CAI^{2}R), an NIBIB National Center for Biomedical Imaging and Bioengineering.

ImMAP paper

arXiv:2511.11963

all publications

nikopj.github.io/publications

NYU Langone Health

CAI^{2}R Center for Advanced Imaging Innovation and Research

NYU TANDON SCHOOL OF ENGINEERING