

Stochastic optimization of three-dimensional non-Cartesian sampling trajectory

Guanhua Wang¹   | Jon-Fredrik Nielsen^{1,2}  | Jeffrey A. Fessler^{1,2}  | Douglas C. Noll¹ 

¹Biomedical Engineering, University of Michigan, Ann Arbor, Michigan, United States

²EECS, University of Michigan, Ann Arbor, Michigan, United States

Correspondence

Guanhua Wang, Biomedical Engineering, University of Michigan, Ann Arbor, MI, USA.

Email: guanhuaw@umich.edu

Funding information

National Institutes of Health, Grant/Award Numbers: R01 EB023618, U01 EB026977, S10OD026738-01; National Science Foundation, Grant/Award Number: IIS 1838179

Purpose: Optimizing three-dimensional (3D) k-space sampling trajectories is important for efficient MRI yet presents a challenging computational problem. This work proposes a generalized framework for optimizing 3D non-Cartesian sampling patterns via data-driven optimization.

Methods: We built a differentiable simulation model to enable gradient-based methods for sampling trajectory optimization. The algorithm can simultaneously optimize multiple properties of sampling patterns, including image quality, hardware constraints (maximum slew rate and gradient strength), reduced peripheral nerve stimulation (PNS), and parameter-weighted contrast. The proposed method can either optimize the gradient waveform (spline-based freeform optimization) or optimize properties of given sampling trajectories (such as the rotation angle of radial trajectories). Notably, the method can optimize sampling trajectories synergistically with either model-based or learning-based reconstruction methods. We proposed several strategies to alleviate the severe nonconvexity and huge computation demand posed by the large scale. The corresponding code is available as an open-source toolbox.

Results: We applied the optimized trajectory to multiple applications including structural and functional imaging. In the simulation studies, the image quality of a 3D kooshball trajectory was improved from 0.29 to 0.22 (NRMSE) with Stochastic optimization framework for 3D Non-Cartesian samPling trajectorY (SNOPY) optimization. In the prospective studies, by optimizing the rotation angles of a stack-of-stars (SOS) trajectory, SNOPY reduced the NRMSE of reconstructed images from 1.19 to 0.97 compared to the best empirical method (RSOS-GR). Optimizing the gradient waveform of a rotational EPI trajectory improved participants' rating of the PNS from "strong" to "mild."

Conclusion: SNOPY provides an efficient data-driven and optimization-based method to tailor non-Cartesian sampling trajectories.

KEYWORDS

data-driven optimization, deep learning, MRI, non-Cartesian sampling, image acquisition

1 | INTRODUCTION

Most MRI systems sample data in the frequency domain (k-space) following prescribed sampling trajectories. Efficient sampling strategies can accelerate acquisition and improve image quality. Many well-designed sampling strategies and their variants, such as spiral, radial, CAIPIRINHA, and PROPELLER,¹⁻⁴ have enabled MRI's application to many areas.⁵⁻⁸ Sampling patterns in k-space are either located on the Cartesian raster or arbitrary locations (non-Cartesian sampling). This paper focuses on optimizing three-dimensional (3D) non-Cartesian trajectories and introduces a generalized gradient-based optimization method for automatic trajectory design or tailoring.

The design of sampling patterns usually considers certain properties of k-space signals. For instance, the variable density spiral trajectory⁹ samples more densely in the central k-space where more energy is located. For higher spatial frequency regions, the variable density spiral trajectory uses larger gradient strengths and slew rates to cover k-space as quickly as possible. Compared to two-dimensional (2D) sampling, designing 3D sampling analytically is more challenging for several reasons. The number of parameters increases in 3D, and the parameter selection is more difficult due to the larger search space. For example, a 3D radial trajectory with 10 000 spokes has 20 000 degrees of freedom, while its 2D multislice counterpart with 200 spokes per slice has only 200 degrees of freedom. Additionally, analytical designs usually are based on the Shannon–Nyquist relationship¹⁰⁻¹² that might not fully consider properties of sensitivity maps and nonlinear reconstruction methods. For 3D sampling patterns with high undersampling (acceleration) ratios, there are limited analytical tools for designing sampling patterns with an anisotropic field of view (FOV) and resolution. The peripheral nerve stimulation (PNS) effect¹³ is also more severe in 3D imaging because of the additional spatial encoding gradient, further complicating manual designs. For these reasons, automatic designs of 3D sampling trajectories are crucial for efficient acquisition.

Many 3D sampling approaches exist. The “stack-of-2D” strategy stacks 2D sampling patterns in the slice direction.^{6,12} This approach is easier to implement and enables slice-by-slice 2D reconstruction. Another design applies Cartesian sampling in the frequency-encoding direction and non-Cartesian sampling in the phase-encoding direction.^{14,15} However, these approaches do not fully explore the design space in three dimensions and may not perform as well as true 3D sampling trajectories.¹⁶

Recently, 3D SPARKLING¹⁶ proposes to optimize 3D sampling trajectories based on the goal of conforming to a given density while distributing samples as uniformly

as possible.¹⁷ That method demonstrated improved image quality compared to the “stack-of-2D-SPARKLING” approach. In both 2D and 3D, the SPARKLING approach uses a prespecified sampling density in k-space that is typically an isotropic radial function. This density function cannot readily capture distinct energy distributions of different imaging protocols, therefore adaptive density functions were recently proposed.¹⁸ In SPARKLING, the PNS effects are not controlled explicitly, and the user may need to lower the slew rate to reduce PNS. SPARKLING optimizes the location of every sampling point, or the gradient waveform (freeform optimization), and cannot optimize parameters of existing sampling patterns.

In addition to analytical methods, learning-based methods have been investigated for designing trajectories. Since different anatomies have distinct energy distributions in the frequency domain, an algorithm may learn to optimize sampling trajectories from training datasets. Several studies have shown that different anatomies produce distinct optimized sampling patterns, and these optimized sampling trajectories can improve image quality.¹⁹⁻²⁷ Some methods can optimize sampling trajectories with respect to specific reconstruction algorithms to further enhance reconstruction image quality.^{14,28} Several recent studies also applied learning-based approaches to 3D non-Cartesian trajectory design. J-MoDL¹⁴ proposes to learn sampling patterns and model-based deep learning (MoDL) reconstruction algorithms jointly. J-MoDL optimizes the sampling locations in the phase-encoding direction, to avoid the computation cost of nonuniform Fourier transform. PILOT/3D-FLAT^{22,29} jointly optimizes freeform 3D non-Cartesian trajectories and a reconstruction neural network with gradient-based methods. These studies use the standard auto-differentiation approach to calculate the gradient used in optimization, which can be inaccurate and lead to suboptimal optimization results.²⁸

This work extends our previous methods^{20,28} and introduces a generalized Stochastic optimization framework for 3D Non-Cartesian samPling trajectorY (SNOPY). The proposed method can automatically tailor given trajectories and learn k-space features from training datasets. We present several optimization objectives, including image quality, hardware constraints, PNS effect suppression and image contrast. Users can simultaneously optimize one or multiple characteristics of a given sampling trajectory. Similar to previous learning-based methods,^{14,20-22} the sampling trajectory can be jointly optimized with trainable reconstruction algorithms to improve image quality. The joint optimization approach can thus exploit the synergy between acquisition and reconstruction, and learn optimized trajectories specific for different anatomies and reconstruction methods.^{14,20,28,30,31} The

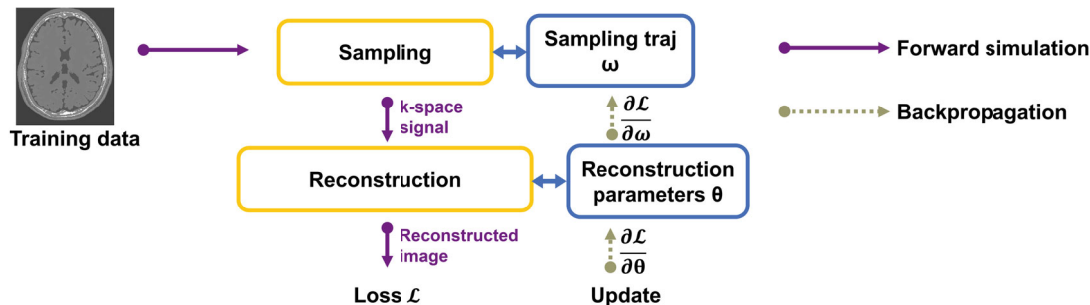


FIGURE 1 Diagram of SNOPI. The sampling trajectory (and possibly reconstruction parameters) are updated using gradient methods. The training/optimization process uses differentiable programming to obtain the gradient necessary for the update.

algorithm can optimize various properties of a sampling trajectory, such as readout waveforms, or rotation angles of readout shots, making it more practical and applicable. We also introduced several techniques to improve efficiency, enabling large-scale 3D trajectory optimization. We tested the proposed methods with multiple imaging applications, including structural and functional imaging. These applications benefited from the SNOPI-optimized sampling trajectories in both simulation and prospective studies.

2 | THEORY

This section describes the proposed gradient-based methods for trajectory optimization. We use the concept of differentiable programming to compute the descent gradient with respect to sampling trajectories required in the gradient-based methods. The sampling trajectory and reconstruction parameters are differentiable parameters, whose gradients can be computed by auto-differentiation. To learn or update these parameters, one may apply (stochastic) gradient descent algorithms. Figure 1 illustrates the basic idea. The sampling trajectories can be optimized in conjunction with the parameters of learnable reconstruction algorithms so that the learned sampling trajectories and reconstruction methods are in synergy and produce high-quality images. The SNOPI algorithm combines several optimization objectives to ensure that the optimized sampling trajectories have desired properties. Section 2.1 delineates these objective functions. Section 2.3 shows that the proposed method is applicable to multiple scenarios with different parameterization strategies. For non-Cartesian sampling, the system model usually involves non-uniform fast Fourier transforms (NUFFT). Section 2.4 briefly describes an efficient and accurate way to calculate the gradient involving NUFFTs. Section 2.5 suggests several engineering approaches to make this large-scale optimization problem solvable and efficient.

2.1 | Optimization objectives

This section outlines the optimization objectives in SNOPI. As SNOPI is a stochastic gradient descent-like algorithm, the objective function, or loss function, is by default defined on a mini-batch of data. The final loss function can be a linear combination of following loss terms to ensure the optimized trajectory possesses multiple required properties.

2.1.1 | Image quality

For many MRI applications, efficient acquisition and reconstruction aim to produce high-quality images. Consequently, the learning objective should encourage images reconstructed from sampled k-space signals to match the reference images. We formulate this similarity objective as the following image quality training loss:

$$\mathcal{L}_{\text{recon}} = \ell(f_{\theta, \mathbf{c}}(\mathbf{A}(\boldsymbol{\omega}(\mathbf{c}))\mathbf{x} + \boldsymbol{\varepsilon}) - \mathbf{x}). \quad (1)$$

Here, $\boldsymbol{\omega}(\mathbf{c}) \in \mathbb{R}^{N_{\text{fe}} \times N_{\text{s}} \times N_{\text{d}}}$ denotes the trajectory to be optimized, with N_{s} shots, N_{fe} sampling points in each shot, and N_{d} image dimensions. For 3D MRI, $N_{\text{d}} = 3$. $\boldsymbol{\varepsilon}$ is simulated complex Gaussian noise. $\mathbf{A}(\boldsymbol{\omega})$ is the forward system matrix for sampling trajectory $\boldsymbol{\omega}(\mathbf{c})$.³² \mathbf{c} denotes the parameterization coefficients of sampling trajectories $\boldsymbol{\omega}$, which is introduced in Section 2.3. In this study, \mathbf{A} also incorporated multicoil sensitivity information.³³ \mathbf{x} denotes the reference image from the training set \mathcal{X} , which is typically reconstructed from fully sampled signals. In addition to contrast-weighted imaging, if the training dataset \mathcal{X} includes *quantitative parameter maps*, one may also simulate \mathbf{x} using the Bloch equation, and \mathbf{A} can subsequently consider imaging physics such as relaxation. $f_{\theta, \boldsymbol{\omega}}(\cdot)$ is the reconstruction algorithm to be delineated in Section 2.2. $\boldsymbol{\theta}$ denotes the reconstruction algorithm's parameters. It can be kernel weights in a convolutional neural network, or the regularizer coefficient in a model-based reconstruction

method. The similarity term $\ell(\cdot)$ can be ℓ_1 norm, ℓ_2 norm, or a combination of both. There are also other ways to measure the distance between \mathbf{x} and $f_{\theta,\omega}(\mathbf{A}(\omega)\mathbf{x} + \varepsilon)$, such as the structural similarity index.³⁴ For simplicity, this work used a linear combination of ℓ_1 norm and square-of- ℓ_2 norm, which is a common practice in deep learning-based image reconstruction.³⁵

2.1.2 | Hardware limits

The gradient system of MR scanners has physical constraints, namely maximum gradient strength and slew rate. Ideally, we would like to enforce a set of constraints of the form

$$\|\mathbf{g}_i[j, :]\|_2 \leq g_{\max}, \quad \mathbf{g}_i = \mathbf{D}_1\omega[:, i, :]/(\gamma\Delta t) \in \mathbb{R}^{(N_{\text{fe}}-1) \times N_d},$$

for every shot $i = 1, \dots, N_s$ and time sample $j = 1, \dots, N_{\text{fe}}$, where \mathbf{g}_i denotes the gradient strength of the i shot and g_{\max} denotes the desired gradient upper bound. One may use a Euclidean norm along the spatial axis so that any rotation of the sampling trajectory still obeys the constraint. Applying the penalty to each individual gradient axis is also feasible. A similar constraint is enforced on the Euclidean norm of the slew rate $\mathbf{s}_i = \mathbf{D}_2\omega[:, i, :]/(\gamma\Delta t^2)$, where \mathbf{D}_1 and \mathbf{D}_2 denote first-order and second-order finite difference operators applied along the readout dimension. Δt denotes the raster time interval and γ denotes the gyromagnetic ratio.

To make the optimization more practical, we follow previous studies,^{20,22} and formulate the hardware constraint as a soft penalty term:

$$\mathcal{L}_g = \sum_{i=1}^{N_s} \sum_{j=1}^{N_{\text{fe}}-1} \phi_{g_{\max}}(\|\mathbf{g}_i[j, :]\|_2). \quad (2)$$

$$\mathcal{L}_s = \sum_{i=1}^{N_s} \sum_{j=1}^{N_{\text{fe}}-2} \phi_{s_{\max}}(\|\mathbf{s}_i[j, :]\|_2). \quad (3)$$

Here ϕ is a penalty function, and we use a simple soft-thresholding function $\phi_\lambda(x) = \max(|x| - \lambda, 0)$, because it is subdifferentiable and easy to implement. It is possible to use more sophisticated functions. Since ϕ here is a soft penalty and the optimization results may exceed the threshold, s_{\max} and g_{\max} can be slightly lower than the scanner's actual physical limits to ensure that the optimization results are feasible on the scanner. Applying a sanity check before sequence programming is also useful. In addition to the soft-penalty approach, recent studies³⁶ also used projection-based methods.

2.1.3 | Suppression of PNS effect

The additional gradient axis in 3D imaging can result in stronger PNS effects compared to 2D imaging. To quantify possible PNS effects of candidate gradient waveforms, SNOPI uses a convolution model described in Reference 37:

$$R_{id}(t) = \frac{1}{s_{\min}} \int_0^t \frac{\mathbf{s}_{id}(\theta)c}{(c+t-\theta)^2} d\theta, \quad (4)$$

where R_{id} denotes the PNS effect of the gradient waveform from the i th shot and the d th dimension. \mathbf{s}_{id} is the slew rate of the i th shot in the d th dimension. c (chronaxie) and s_{\min} (minimum stimulation slew rate) are scanner-specific parameters.

Likewise, we discretize the convolution model and use a soft penalty term as the following loss function:

$$\mathcal{L}_{\text{pns}} = \sum_{i=1}^{N_s} \sum_{j=1}^{N_{\text{fe}}} \phi_{p_{\max}} \left(\left(\sum_{d=1}^{N_d} \mathbf{p}_{id}[j]^2 \right)^{\frac{1}{2}} \right). \quad (5)$$

Again, ϕ denotes the soft-thresholding function, with PNS threshold p_{\max} (usually $\leq 80\%$ ³⁷). This model combines the three spatial axes via the sum-of-squares manner and does not consider anisotropic characteristics of PNS.³⁸ The implementation may use an FFT (with zero padding) for efficient convolution.

2.1.4 | Image contrast

In many applications, the optimized sampling trajectory should maintain certain parameter-weighted contrasts. For example, ideally the (gradient) echo time (TE) should be identical for each shot. Again, we replace this hard constraint with an TE penalty. Other parameters, like repetition time (TR) and inversion time, can be predetermined in the RF pulse design. Specifically, the corresponding loss function encourages the sampling trajectory to cross the k-space center at certain time points:

$$\mathcal{L}_c = \sum_{\{i,j,d\} \in C} \phi_0(|\omega[i, j, d]|), \quad (6)$$

where C is a collection of gradient time points that are constrained to cross the k-space zero point. ϕ is still a soft-thresholding function, with threshold 0.

The total loss function is a linear combination of the above terms

$$\mathcal{L} = \lambda_{\text{recon}} \mathcal{L}_{\text{recon}} + \lambda_g \mathcal{L}_g + \lambda_s \mathcal{L}_s + \lambda_{\text{pns}} \mathcal{L}_{\text{pns}} + \lambda_c \mathcal{L}_c.$$

Note that not every term is required. For example, experiment Section 3.2.2 only used the $\mathcal{L}_{\text{recon}}$. Section 5 further discusses how to choose linear weights λ s.

2.2 | Reconstruction

In (1), the reconstruction algorithm $f_{\theta, \omega}(\cdot)$ can be various algorithms. Consider a typical cost function for regularized MR image reconstruction

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{A}(\omega)\mathbf{x} - \mathbf{y}\|_2^2 + \mathcal{R}(\mathbf{x}). \quad (7)$$

$\mathcal{R}(\mathbf{x})$ here can be a Tikhonov regularization $\nu\|\mathbf{x}\|_2^2$ (CG-SENSE³⁹), a sparsity penalty $\nu\|\mathbf{T}\mathbf{x}\|_1$ (compressed sensing,⁴⁰ \mathbf{T} is a finite-difference operator), a roughness penalty $\nu\|\mathbf{T}\mathbf{x}\|_2^2$ (penalized least squares, PLS), or a neural network (MoDL⁴¹). Section 4 shows that different reconstruction algorithms lead to distinct optimized sampling trajectories. In training, \mathbf{y} is retrospectively simulated as $\bar{\mathbf{y}} = \mathbf{A}(\omega)\mathbf{x} + \varepsilon$ (following (1)). In prospective studies, \mathbf{y} is the acquired k-space data.

To get a reconstruction estimation $\hat{\mathbf{x}}$, one may use iterative reconstruction algorithms. Specifically, the algorithm should be step-wise differentiable (or subdifferentiable) to enable differentiable programming. The backpropagation uses the chain rule to traverse every step of the iterative algorithm to calculate gradients with respect to variables such as ω .

2.3 | Parameterization

As is shown in Reference 20, directly optimizing every k-space sampling location (or equivalently every gradient waveform time point) may lead to suboptimal results. Additionally, in many applications, one may need to optimize certain properties of existing sampling patterns, such as the rotation angles of a multishot spiral trajectory, so that the optimized trajectory can be easily integrated into existing workflows. For these needs, we propose two parameterization strategies.

The first approach, spline-based freeform optimization, represents the sampling pattern using a linear basis, that is, $\omega = \mathbf{B}\mathbf{c}$, where \mathbf{B} is a matrix of samples of a basis such as quadratic B-spline kernels and \mathbf{c} denotes the coefficients to be optimized.^{20,22} This approach fully exploits the generality of a gradient system. Using a linear parameterization like B-splines reduces degrees of freedom and facilitates applying hardware constraints.^{20,42} Additionally, the parameterization can be combined with multiscale optimization to avoid suboptimal local minima and further improve optimization results.^{17,20,22} However, freely

optimized trajectories could introduce implementation challenges. For example, some MRI systems can not store hundreds of different gradient waveforms.

The second approach is to optimize attributes \mathbf{c} of existing trajectories, where $\omega(\mathbf{c})$ is a differentiable function of the attributes \mathbf{c} . For example, many applications use radial trajectories, where the rotation angles can be optimized. Suppose $\mathbf{s} \in \mathbb{R}^{3 \times N}$ is one radial sampling spoke, and consider an M -shot 3D radial trajectory,

$$\omega = \left[\mathbf{R}_1 \cdots \mathbf{R}_M \right] \mathbf{I}_M \otimes \mathbf{s}, \quad (8)$$

where $\mathbf{R}_i \in \mathbb{R}^{3 \times 3}$ denotes a rotation matrix, \mathbf{I}_M denotes an identity matrix of size M , and \otimes denotes the Kronecker product. In this case, the list of \mathbf{R}_i is the coefficient to be optimized. This approach is easier to implement on scanners, and can work with existing workflows.

2.4 | Efficient and accurate Jacobian calculation

In the similarity loss (1), the sampling trajectory is embedded in the forward system matrix \mathbf{A} . The system matrix for non-Cartesian sampling usually includes NUFFT operators.³² Updating the sampling trajectory in each optimization step requires the Jacobian, or the gradient with respect to the sampling trajectory. The NUFFT operator typically involves interpolation in the frequency domain, which is nondifferentiable due to rounding operations. Several previous works used auto-differentiation (with subgradients) to calculate an approximate numerical gradient,^{22,29} but that approach is inaccurate and slow.²⁸ We derived an efficient and accurate Jacobian approximation method.²⁸ For example, the efficient Jacobian of a forward system model \mathbf{A} is:

$$\frac{\partial \mathbf{A}\mathbf{x}}{\partial \omega^{[d]}} = -\iota \text{diag} \{ \mathbf{A}(\mathbf{x} \odot \mathbf{r}^{[d]}) \}, \quad (9)$$

where $d \in \{1, 2, 3\}$ denotes the spatial dimensions, $\mathbf{r}^{[d]}$ denotes the Euclidean spatial grid, \odot denotes the Hadamard product, and ι is the imaginary unit. Calculating this Jacobian simply uses another NUFFT, which is more efficient than the auto-differentiation approach. See Reference 28 for more cases, such as $\frac{\partial \mathbf{A}'\mathbf{A}\mathbf{x}}{\partial \omega^{[d]}}$ and the detailed derivation.

2.5 | Efficient optimization

2.5.1 | Optimizer

Generally, to optimize the sampling trajectory ω and other parameters (such as reconstruction parameters θ)

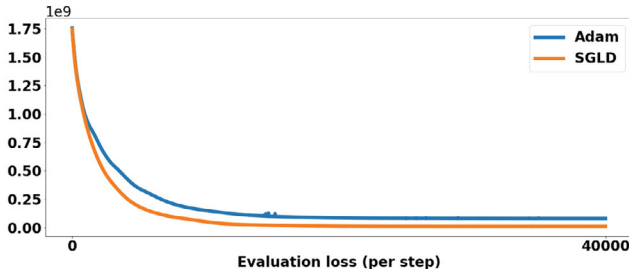


FIGURE 2 The evaluation loss curve for SGLD and Adam. The training process costs ~ 1 h.

via stochastic gradient descent-like methods, each update takes a step (in the simplest form)

$$\begin{aligned}\theta^K &= \theta^{K-1} - \eta_\theta \frac{\partial \mathcal{L}}{\partial \theta}(\omega^{K-1}, \theta^{K-1}) \\ \omega^K &= \omega^{K-1} - \eta_\omega \frac{\partial \mathcal{L}}{\partial \omega}(\omega^{K-1}, \theta^{K-1}),\end{aligned}$$

where \mathcal{L} is the loss function described in Section 2.1 and where η_θ and η_ω denote step-size parameters.

The optimization is highly nonconvex and may suffer from suboptimal local minima. We used stochastic gradient Langevin dynamics (SGLD)⁴³ as the optimizer to improve results and accelerate training. Each update of SGLD injects Gaussian noise into the gradient to introduce randomness

$$\begin{aligned}\theta^K &= \theta^{K-1} - \eta_\theta \frac{\partial \mathcal{L}}{\partial \theta^{K-1}} + \sqrt{2\eta_\theta} \mathcal{N}(0, 1) \\ \omega^K &= \omega^{K-1} - \eta_\omega \frac{\partial \mathcal{L}}{\partial \omega^{K-1}} + \sqrt{2\eta_\omega} \mathcal{N}(0, 1).\end{aligned}\quad (10)$$

Across most experiments, we observed that SGLD led to improved results and faster convergence compared with SGD or Adam.⁴⁴ Figure 2 shows a loss curve of SGLD and Adam of experiment Section 3.2.3.

2.5.2 | Memory saving techniques

Due to the large dimension, the memory cost for naive 3D trajectory optimization would be prohibitively intensive. We developed several techniques to reduce memory use and accelerate training.

As discussed above, the efficient Jacobian approximation uses only 10% of the memory required by the standard auto-differentiation approach.²⁸ We also used in-place operations in certain reconstruction steps, such as the conjugate gradient (CG) method, because with careful design it will not interrupt auto-differentiation. (See our open-source code* for details.) The primary memory

bottleneck relates to 3D NUFFT operators. One can pre-calculate the Toeplitz embedding kernel to save memory and accelerate computation.^{45,46} In the training phase, we used NUFFTs with lower accuracy, for instance, with a smaller oversampling ratio for gridding.²⁸ Table 1 shows the incrementally improved efficiency achieved with these techniques. Without the proposed techniques, optimizing 3D trajectories would require hundreds of gigabytes of memory, which would be impractical for a single node. SNOPI enables solving this otherwise prohibitively large problem on a single GPU.

3 | METHODS

3.1 | Datasets

We used two publicly available datasets; both of them contain 3D multicoil raw k-space data. SKM-TEA⁴⁷ is a 3D quantitative double-echo steady-state⁴⁸ knee dataset. It was acquired by 3T GE MR750 scanners and 15/16-channel receiver coils. SKM-TEA includes 155 subjects. We used 132 for training, 10 for validation, and 13 for the test. Calgary brain dataset⁴⁹ is a 3D brain T1w MP-RAGE⁵⁰ k-space dataset. It includes 67 available subjects, acquired by an MR750 scanner and 12-channel head coils. We used 50 volumes for training, six for validation, and seven for testing. All sensitivity maps were calculated by ESPIRiT.⁵¹

3.2 | Simulation experiments

We experimented with multiple scenarios to show the broad applicability of the proposed method. All the experiments used a node equipped with an Nvidia Tesla A40 GPU for training.

3.2.1 | Optimizing 3D gradient waveform

We optimized the sampling trajectory with a 3D radial (“kooshball”) initialization.^{52,53} As is described in Section 2.3, the experiment optimized the readout waveform of each shot with B-spline parameterization, to reduce the number of degrees of freedom and enable multiscale optimization. The initial 3D radial trajectory had a 5.12 ms long readout (raster time = 4 μ s) and 1024 shots (8 \times acceleration), using the rotation angle described in Reference 16. The training used the SKM-TEA dataset. The retrospectively cropped FOV was 158 \times 158 \times 51 mm³ with 0.76 \times 0.62 \times 1.6 mm³ simulated resolution. The receiver bandwidth was ± 125 kHz (dwell time = 4 μ s).

*<https://github.com/guanhuaw/Bjork>

TABLE 1 The memory/time use reduction brought by proposed techniques.

Plain	+Efficient Jacobian	+In-place ops	+Toeplitz embedding	+Low-res NUFFT
5.7GB / 10.4s	272 MB / 1.9 s	253 MB / 1.6 s	268 MB / 0.4 s	136 MB / 0.2 s

Notes: Here we used a 2D 400 × 400 test case, and CG-SENSE reconstruction (20 iterations). “+” means adding the technique to previous columns.

The training loss was

$$\mathcal{L} = \mathcal{L}_{\text{recon}} + 10^3 \mathcal{L}_g + 10^3 \mathcal{L}_s + \mathcal{L}_{\text{pns}}.$$

The gradient strength (g_{max}) and slew rate (s_{max}) were 50 mT/m and 150 mT/m/ms (for individual axis). The PNS threshold (p_{max}) was 80%. The simulated noise ϵ was 0. The batch size was 3. The learning rate η_{ω} decayed from 10^{-4} to 0 linearly. For multilevel optimization, we used 3 levels (with B-spline kernel widths = 32, 16, and 8), and each level used 200 epochs. The total training time was approximately 240 h. The trajectory was optimized with respect to several image reconstruction algorithms. We used a regularizer weight $\nu = 10^{-3}$ and 30 CG iterations for CG-SENSE and PLS. For learning-based reconstruction, we used the MoDL⁴¹ network that alternates between a neural network-based denoiser and data consistency updates. We used a 3D version of the denoising network,⁵⁴ 20 CG iterations for the data consistency update, and six outer iterations. Similar to previous investigations,^{14,20} SNOPI jointly optimized the neural network’s parameters and the sampling trajectory using (10).

3.2.2 | Optimizing rotation angles of stack-of-stars trajectory

This experiment optimized the rotation angles of a stack-of-stars trajectory, which is a widely used volumetric imaging sequence. The training used the Calgary brain dataset. We used PLS as the reconstruction method for simplicity, with $\nu = 10^{-3}$ and 30 iterations. The simulated noise ϵ was 0 and the batch size was 12. We used 200 epochs and a learning rate linearly decaying from 10^{-4} to 0. The FOV was retrospectively cropped to $256 \times 218 \times 32 \text{ mm}^3$ with 1 mm^3 resolution. We used 40 spokes per kz location (6× acceleration), and 1280 spokes in total. The readout length was 3.5 ms. The receiver bandwidth was $\pm 125 \text{ kHz}$ (dwell time = $4 \mu\text{s}$). The trajectory was a stack of 32 kz planes, hence SNOPI optimized 1280 rotation angles in this case.

Since optimizing rotation angles does not impact the gradient strength, slew rate, PNS, and image contrast, we only used the reconstruction loss $\mathcal{L} = \mathcal{L}_{\text{recon}}$. We regarded the method (RSOS-GR) proposed in previous works¹² as the best empirical scheme. We applied 200 epochs with a

linearly decaying learning rate from 10^{-3} to 0. The training cost approximately 20 h.

3.2.3 | PNS suppression of 3D rotational EPI trajectory for functional imaging

The third application optimizes the rotation EPI (REPI) trajectory,⁵⁵ which provides an efficient sampling strategy for fMRI. For high resolution (i.e., $\leq 1 \text{ mm}$), we found that subjects may experience strong PNS effects introduced by REPI. This experiment aimed to reduce the PNS effect of REPI while preserving the original image contrast. We optimized one shot of REPI, being parameterized by B-spline kernels (width = 16). The optimized readout shot was rotated using the angle scheme similar to Reference 55 for multishot acquisition.

We designed the REPI readout for an oscillating steady steady imaging sequence, a novel fMRI signal model that can improve the SNR.^{56,57} The FOV was $200 \times 200 \times 12 \text{ mm}^3$, with 1 mm^3 isotropic resolution, TR = 16 ms, and TE = 7.4 ms. The readout length was 10.6 ms. The receiver bandwidth was $\pm 250 \text{ kHz}$ (dwell time = $2 \mu\text{s}$). The gradient strength (g_{max}), and slew rate (s_{max}) constraints were 58 mT/m and 200 mT/m/ms (three axes combined).

To accelerate training, the loss term here excluded the reconstruction loss $\mathcal{L}_{\text{recon}}$:

$$\mathcal{L} = 10^{-2} \mathcal{L}_g + 10^{-2} \mathcal{L}_s + \mathcal{L}_{\text{pns}} + 10^2 \mathcal{L}_c.$$

The training used 40 000 steps, with the learning rate decaying linearly from 10^{-4} to 0. The training cost approximately 1 h.

3.3 | In vivo experiments

We implemented the optimized trajectory prospectively on a GE UHP 3.0T scanner equipped with a Nova Medical 32-channel head coil. Participants gave informed consent under local IRB approval. Since the cache space in this MR system cannot load hundreds of distinct gradient waveforms, the experiment 3.2.1 was not implemented prospectively. Readers may refer to the corresponding 2D prospective studies²⁰ for image quality improvement and correction of eddy current effects. For experiment 3.2.2,

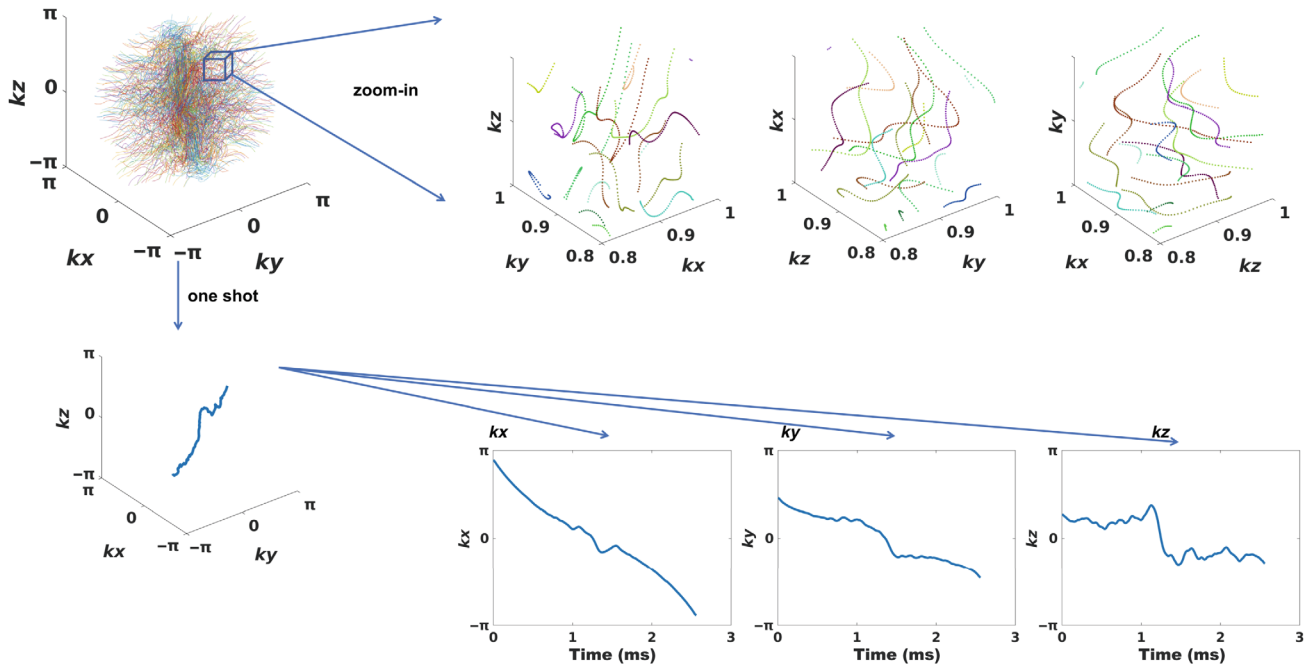


FIGURE 3 The optimized sampling trajectory of experiment Section 3.2.1. The training process involves the SKM-TEA dataset and CG-SENSE reconstruction. The upper row shows a zoomed-in region from different viewing perspectives. The lower row displays one shot from different perspectives.

we programmed the sampling trajectory with a 3D T1w fat-saturated GRE sequence,⁵⁸ with TR/TE = 14/3.2 ms and FA = 20°. The experiment included 4 healthy subjects. For experiment 3.2.3, to rate the PNS effect, we asked three participants to score the nerve stimulation with a 5-point Likert scale from “mild tingling” to “strong muscular twitch.”

3.4 | Reproducible research

The code is publicly available[†]. As an accompanying project, MIRTorch[‡] facilitates applying differentiable programming to MRI sampling and reconstruction.

4 | RESULTS

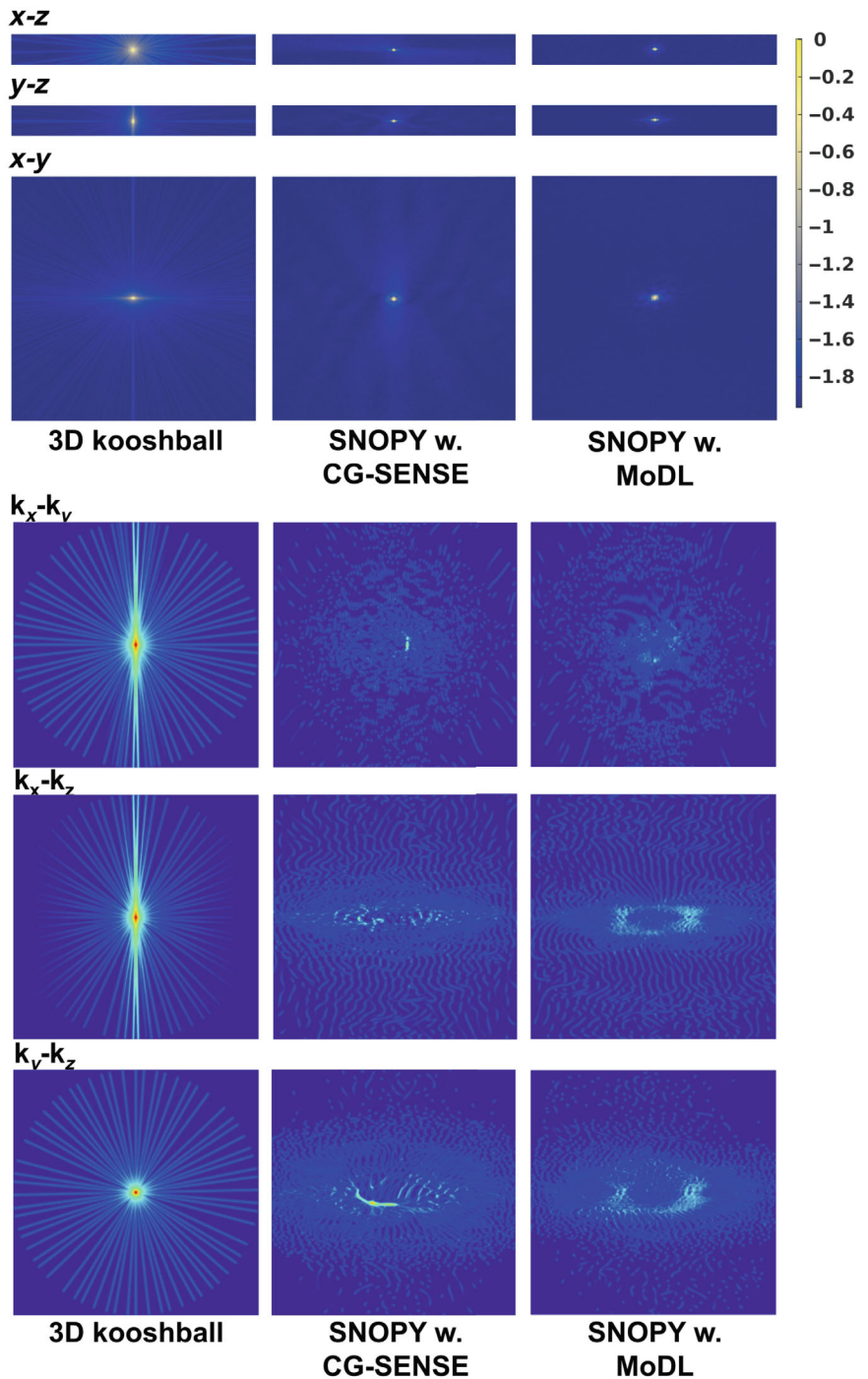
For the spline-based freeform optimization experiment delineated in Section 3.2.1, Figure 3 presents an example of the optimized trajectory, along with zoomed-in regions and plots of a single shot. Similar to the 2D case²⁰ and SPARKLING,^{16,17} the multilevel B-spline optimization generates a swirling trajectory that can cover more k-space in the fixed readout time, to reduce large gaps between sampling locations and, consequently, aliasing

artifacts. Notably, the zoomed-in region highlights that different shots were automatically learned not to overlap with each other, which implicitly improved the sampling efficiency.¹⁷ Figure 4 displays point spread functions (PSFs) of trajectories jointly optimized with different reconstruction algorithms. To visualize the sampling density in different regions of k-space, we convolved the trajectory with a Gaussian kernel, and Figure 4 shows the density of central profiles from different views. Compared with 3D kooshball, the SNOPI optimization led to fewer radial patterns in PSFs, corresponding to fewer streak artifacts in Figure 5. Trajectories optimized with different reconstruction algorithms generated different PSFs and densities, which agrees with previous studies.^{28,30,31} Table 2 lists the quantitative reconstruction quality of different trajectories. The image quality metric is the average PSNR of the test set. SNOPI led to approximately 4 dB higher PSNR than the kooshball initialization. Figure 5 includes examples of reconstructed images. Compared to kooshball, SNOPI’s reconstructed images have reduced artifacts and blurring. Though MoDL (and its variants) are well-performing NN-based reconstruction algorithms according to the open fastMRI reconstruction challenge,⁵⁹ many important structures are distorted using the kooshball trajectory. Using the SNOPI-optimized trajectory, a simple model-based reconstruction (CG-SENSE) can reconstruct such structures. The gradient strength and the slew rate of optimized sampling trajectories are exhibited in Appendix S1. SNOPI solves a nonconvex

[†]<https://github.com/guanhuaw/SNOPI>

[‡]<https://github.com/guanhuaw/MIRTorch>

FIGURE 4 Visualization of the optimized trajectory in experiment Section 3.2.1. The upper subfigure displays PSFs (log-scaled, single-coil) of trajectories optimized with different reconstruction methods. The lower subfigure shows the density of sampling trajectories, obtained by convolving the sampling points with a Gaussian kernel. Three rows are central profiles from three perspectives.



problem; therefore, its results depend on the initialization. Appendix S1 compare optimization results with different initializations.

For experiment 3.2.2, Figure 6 shows the PSF of the optimized and RSOS-GR schemes.¹² For the in-plane ($x-y$) PSF, the SNOPI rotation shows noticeably reduced streak-like patterns. In the $y-z$ direction, SNOPI optimization leads to a narrower central lobe and suppressed aliasing artifacts. The prospective in vivo experiments also support this theoretical finding. In Figure 6, the example slices (reconstructed by PLS) from prospective studies show that

SNOPI reduced streaking artifacts. The average PSNR of SNOPI and RSOS-GR for the four participants were 39.23 and 37.84 dB, respectively. Appendix S1 shows the rotation angles before and after SNOPI optimization.

In experiment 3.2.3, we tested three settings: unoptimized REPI, optimized with PNS threshold (p_{\max} in (5)) = 80%, and optimized with $p_{\max} = 70\%$. Figure 7 shows one slice of reconstructed images by the CS-SENSE algorithm, as well as the subjective ratings of PNS. Though SNOPI suppressed the PNS effect, the image contrast was well preserved by the image contrast regularizer (6). Figure 8

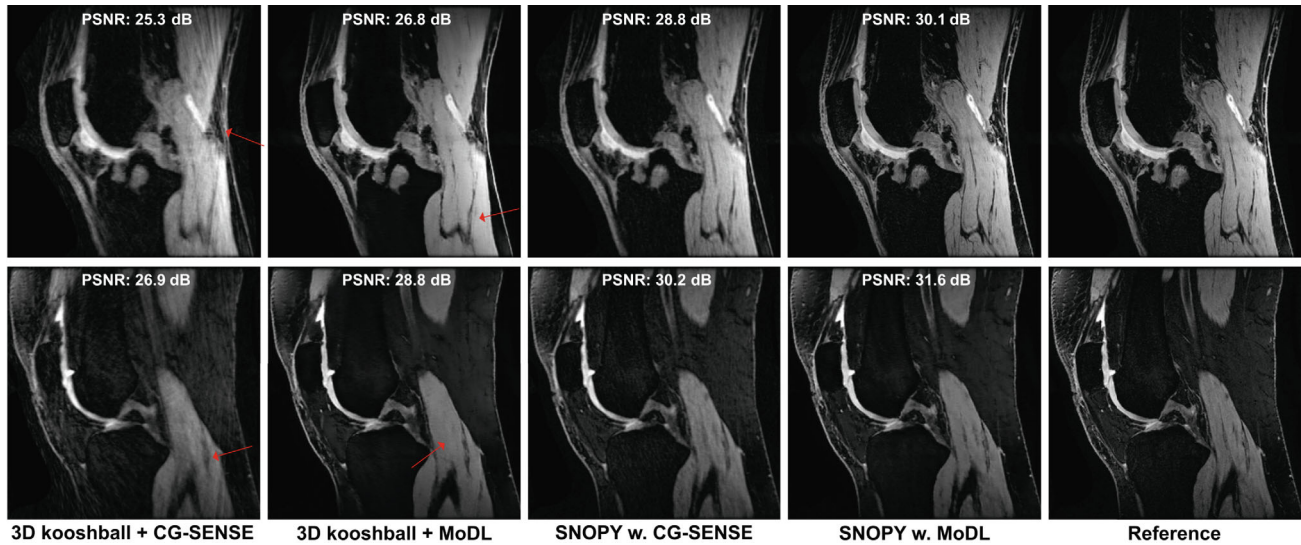


FIGURE 5 Examples of the reconstructed images for two knee slices in Experiment 3.2.1.

TABLE 2 The quantitative reconstruction quality (PSNR) of the test set.

	CG-SENSE	PLS	MoDL
3D kooshball	28.2 dB	28.2 dB	30.1 dB
SNOPY	32.3 dB	32.4 dB	33.6 dB

presents one shot before and after the optimization, and one plot of simulated PNS effects. The SNOPY optimization effectively reduced subjective PNS effects of given REPI readout gradients in both simulation and in-vivo experiments. Intuitively, SNOPY smoothed the trajectory to avoid a constantly high slew rate, preventing a strong PNS effect.

5 | DISCUSSION

SNOPY presents a novel and intuitive approach to optimizing non-Cartesian sampling trajectories. Via differentiable programming, SNOPY enables the application of gradient-based and data-driven methods to trajectory design. Various applications and in-vivo experiments demonstrated the applicability and robustness of SNOPY and its 2D predecessor.²⁰

Experiments 3.2.1 and 3.2.2 used SNOPY to tailor sampling trajectories according to specific training datasets and reconstruction algorithms, by formulating reconstruction image quality as a training loss. One concern was whether the learned trajectories would overfit the training dataset. In experiment 3.2.2, the training set used an MP-RAGE sequence, while the prospective sequence was an RF-spoiled GRE. Similarly, 2D

prospective and retrospective experiments²⁰ showed that trajectories learned with particular pulse sequences and hardware still improved the image quality of other sequences and hardware, and the NN-based reconstruction did not require fine-tuning with respect to prospective experiments. These empirical studies suggest that trajectory optimization is robust to moderate distribution shifts between training and inference. An intuitive explanation is that SNOPY can improve the PSF by reducing aliasing, and such improvements are universally beneficial. Future investigations will explore the robustness of SNOPY in more diverse settings, such as optimizing trajectories with healthy controls and prospectively testing them with pathological participants to examine image quality for pathologies. It will also be desirable to test SNOPY with different FOVs, resolutions, and B_0 strengths.

Our experiments demonstrated that iterative reconstruction with simple analytical regularizers, such as CG-SENSE, can benefit from the SNOPY-optimized sampling trajectories. As depicted in Figure 3, CG-SENSE with SNOPY optimization can successfully reconstruct many anatomical structures that were blurred in the MoDL reconstruction without SNOPY trajectory. This result is consistent with previous studies,²⁸ where compressed sensing algorithms with trajectory optimization also outperformed NN-based reconstruction. These findings indicate untapped potentials of model-based reconstruction by optimizing sampling trajectories.

MRI systems are prone to imperfections such as field inhomogeneity⁶⁰ and eddy currents.⁶¹ Many correction approaches exist, such as B_0 -informed reconstruction⁴⁵ and trajectory mapping.^{62,63} SNOPY-optimized trajectories are compatible with existing correction methods. For instance, we demonstrated the feasibility of implementing

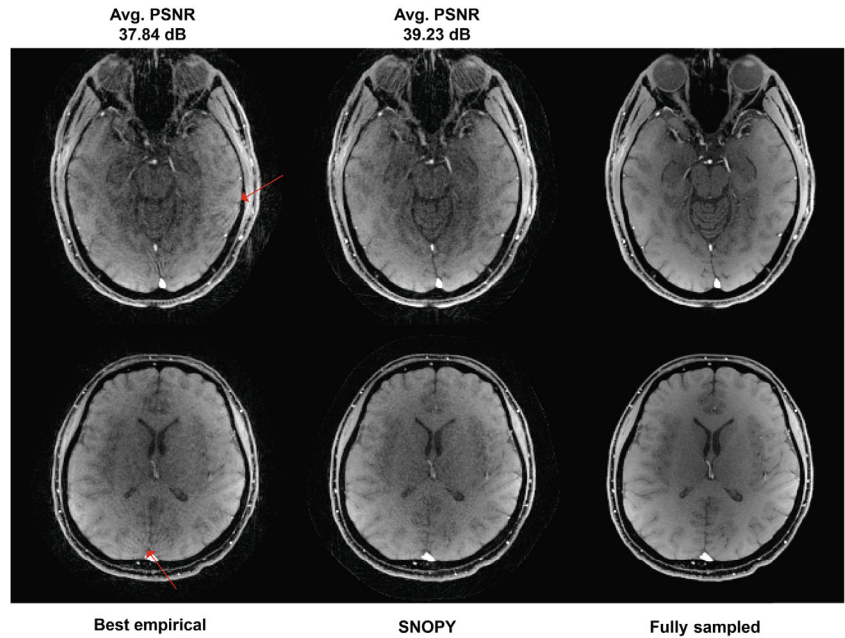


FIGURE 6 Prospective results of Section 3.2.2, optimizing the rotation angles of the stack-of-stars (6× acceleration). “Best empirical” uses the design from a previous study.¹² The upper subfigure shows two slices from prospective in-vivo experiments. The reconstruction algorithm was penalized least squares (PLS). **Avg. PSNR** is the average peak signal-to-noise ratio (PSNR) of the four subjects compared to the fully sampled reference. The lower subfigure shows the log-scaled PSF (single-coil) of two trajectories.

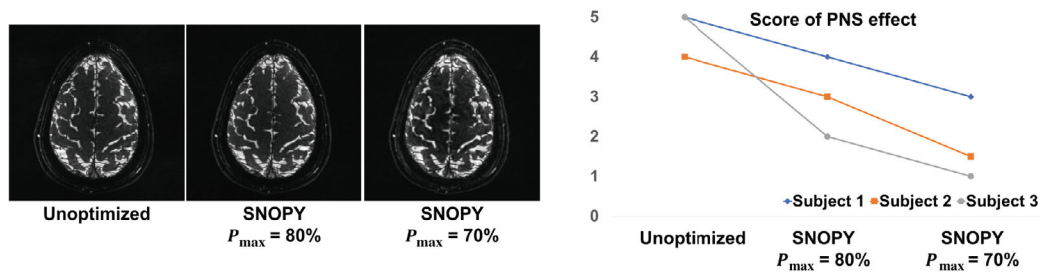


FIGURE 7 Prospective results of Section 3.2.3. We showed three different trajectories: the unoptimized rotation EPI, as well as SNOPY-optimized with peripheral nerve stimulation (PNS) thresholds of 80% and 70%. The left subfigure shows one slice of reconstructed images. The reconstruction used penalized least squares (PLS) and 120 shots (volume TR = 2s). The right subfigure shows subjective scores of the PNS effect.

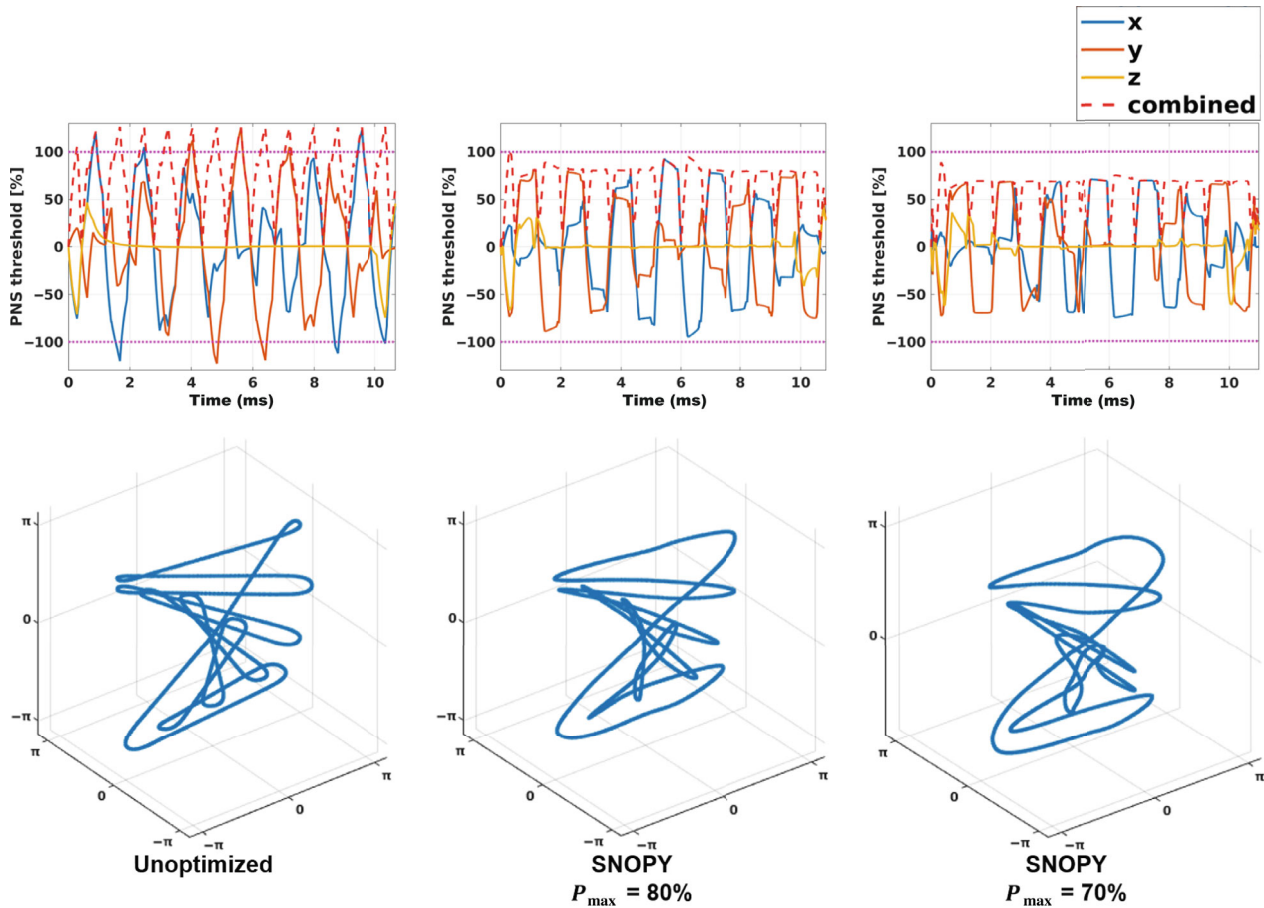


FIGURE 8 The first row of plots displays the peripheral nerve stimulation (PNS) effect calculated by the convolution model (5) used in Experiment 3.2.3. The second row shows the corresponding readout trajectories before and after SNOPI optimization.

eddy currents correction for a 2D freeform optimized trajectory in Reference 20. Additionally, incorporating system imperfections into the forward learning/optimization phase, such as off-resonance maps in the system model \mathbf{A} (as defined in (1)), may enhance the intrinsic robustness of the optimized trajectory. However, this approach requires the distribution of system imperfections, which is typically scanner-specific. To address this limitation, we plan to investigate prospective simulation approaches in future studies. The model mismatch may also happen at the digitization level: the training set typically consists of concrete discrete-space images, whereas real objects are continuous. This inverse crime is common in learning-based methods and may lead to suboptimal results. Future research should investigate strategies for mitigating this issue.

SNOPI uses a relatively simplified model of PNS. More precise models, such as Reference 38, may lead to improved PNS suppression results. SNOPI can also incorporate other optimization objectives to encourage properties such as robustness to field inhomogeneity and reduction of acoustic noise.

The training process incorporates several loss terms, including image quality, PNS suppression, hardware limits, and image contrast. By combining these terms, the optimization can lead to trajectories that have multiple desired characteristics. One may alter the optimization results by controlling the coefficients. For example, with a larger coefficient of the hardware constraint loss, the trajectory will better conform to s_{\max} and g_{\max} . The Appendix S1 contains an example of optimization results using different combinations of weights. Setting the weights of several terms can be complicated. Empirically, the weight of soft constraints, including hardware (\mathcal{L}_g and \mathcal{L}_s), PNS suppression (\mathcal{L}_{pns}), and contrast (\mathcal{L}_c) can be tuned to a higher value if the optimized trajectory significantly violates these constraints. Additionally, the training losses may sometimes contradict each other, and the optimization process would get stuck in a local minimum. To address this, several empirical tricks have been employed. Similar to SPARKLING,¹⁷ the constraint on maximum gradient strength can be relaxed using a higher receiver bandwidth. Bayesian optimization is another option for finding optimal loss weights, but may increase training time. Using

SGLD can introduce randomness that helps escape local minima. In spline-based optimization, one can use a larger B-spline kernel width in the early stages of a coarse-to-fine search.

Trajectory optimization is a nonconvex problem. SNOPI uses several methods, including effective Jacobian approximation, parameterization, multilevel optimization, and SGLD, to alleviate the nonconvexity and achieve better optimization results. These methods were also found to be effective in previous studies.^{20,28} Initialization is also important for non-convex problems, as demonstrated in Appendix S1. SNOPI can leverage existing knowledge of MR sampling as a benign initialization. For instance, our experiments used the widely accepted golden-angle stack-of-stars as optimization bases. The SNOPI algorithm can sequentially improve these skillfully designed trajectories to combine the best of both stochastic optimization and researchers' insights.

SNOPI has a wide range of potential applications, including dynamic and quantitative imaging, particularly if large-scale quantitative datasets are available. These new applications may require task-specific optimization objectives in addition to the ones described in Section 2.1. In particular, if the reconstruction method is not easily differentiable, such as the MR fingerprinting reconstruction based on dictionary matching,⁶⁴ one needs to design a surrogate objective for image quality.

ACKNOWLEDGMENTS

The authors thank Dr. Melissa Haskell, Dr. Shouchang Guo, Dinank Gupta, and Yuran Zhu for the helpful discussion.

DATA AVAILABILITY STATEMENT

Following Magnetic Resonance in Medicine's reproducibility goal, the open-source code is available at <https://github.com/guanhuaw/SNOPI>. The training used publicly available datasets, as described in the main text.

ORCID

Guanhua Wang  <https://orcid.org/0000-0002-1622-5664>

Jon-Fredrik Nielsen  <https://orcid.org/0000-0002-2058-3579>

Jeffrey A. Fessler  <https://orcid.org/0000-0001-9998-3315>

Douglas C. Noll  <https://orcid.org/0000-0002-0983-3805>

TWITTER

Guanhua Wang  @GuanhuaW97

REFERENCES

- Ahn CB, Kim JH, Cho ZH. High-speed spiral-scan Echo planar NMR imaging-I. *IEEE Trans Med Imaging*. 1986;5:2-7.
- Lauterbur PC. Image formation by induced local interactions: examples employing nuclear magnetic resonance. *Nature*. 1973;242:190-191.
- Breuer FA, Blaimer M, Mueller MF, et al. Controlled aliasing in volumetric parallel imaging (2D CAIPIRINHA). *Magn Reson Med*. 2006;55:549-556.
- Pipe JG. Motion correction with PROPELLER MRI: application to head motion and free-breathing cardiac imaging. *Magn Reson Med*. 1999;42:963-969.
- Glover GH, Law CS. Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility Artifacts. *Magn Reson Med*. 2001;46:515-522.
- Johansson A, Balter J, Cao Y. Rigid-body motion correction of the liver in image reconstruction for Golden-angle stack-of-stars DCE MRI. *Magn Reson Med*. 2018;79:1345-1353.
- Yu MH, Lee JM, Yoon JH, Kiefer B, Han JK, Choi BI. Clinical application of controlled aliasing in parallel imaging results in a higher acceleration (CAIPIRINHA)-volumetric interpolated Breathhold (VIBE) sequence for Gadoteric acid-enhanced liver MR imaging. *J Magn Reson Imaging*. 2013;38:1020-1026.
- Forbes KPN, Pipe JG, Bird CR, Heiserman JE. PROPELLER MRI: clinical testing of a novel technique for quantification and compensation of head motion. *J Magn Reson Imaging*. 2001;14:215-222.
- Lee JH, Hargreaves BA, Hu BS, Nishimura DG. Fast 3D imaging using variable-density spiral trajectories with applications to limb perfusion. *Magn Reson Med*. 2003;50:1276-1285.
- Gurney PT, Hargreaves BA, Nishimura DG. Design and analysis of a practical 3D cones trajectory. *Magn Reson Med*. 2006;55:575-582.
- Johnson KM. Hybrid radial-cones trajectory for accelerated MRI. *Magn Reson Med*. 2017;77:1068-1081.
- Zhou Z, Han F, Yan L, Wang DJJ, Hu P. Golden-ratio rotated stack-of-stars acquisition for improved volumetric MRI. *Magn Reson Med*. 2017;78:2290-2298.
- Ham CLG, Engels JML, van de Wiel GT, Machielsen A. Peripheral nerve stimulation during MRI: effects of high gradient amplitudes and switching rates. *J Magn Reson Imaging*. 1997;7:933-937.
- Aggarwal HK, Jacob M. Joint optimization of sampling patterns and deep priors for improved parallel MRI. Paper presented at: 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2020; Barcelona, Spain:8901-8905.
- Bilgic B, Gagoski B, Cauley SF, Fan AP. Wave-CAIPI for highly accelerated 3D imaging. *Magn Reson Med*. 2015;73:2152-2162.
- Chaithya GR, Weiss P, Daval-Fr erot G, Massire A, Vignaud A, Ciuciu P. Optimizing full 3D SPARKLING trajectories for high-resolution magnetic resonance imaging. *IEEE Trans Med Imaging*. 2022;41:2105-2117.
- Lazarus C, Weiss P, Chauffert N, et al. SPARKLING: variable-density K-space filling curves for accelerated T2*-weighted MRI. *Mag Res Med*. 2019;81:3643-3661.
- Chaithya G, Ramzi Z, Ciuciu P. Learning the sampling density in 2D SPARKLING MRI acquisition for optimized image reconstruction. Paper presented at: 2021 29th European Signal Processing Conference (EUSIPCO). IEEE; 2021; Dublin, Ireland:960-964.
- Huijben IAM, Veeling BS, van Sloun RJG. Learning sampling and model-based signal recovery for compressed

- sensing MRI. Paper presented at: 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2020:8906-8910.
20. Wang G, Luo T, Nielsen JF, Noll DC, Fessler JA. B-spline parameterized joint optimization of reconstruction and K-space trajectories (BJORK) for accelerated 2D MRI. *IEEE Trans Med Imaging*. 2022;41:2318-2330.
 21. Bahadir CD, Wang AQ, Dalca AV, Sabuncu MR. Deep-learning-based optimization of the under-sampling pattern in MRI. *IEEE Trans Comput Imaging*. 2020;6:1139-1152.
 22. Weiss T, Senouf O, Vedula S, Michailovich O, Zibulevsky M, Bronstein A. PILOT: Physics-Informed Learned Optimized Trajectories for Accelerated MRI. *J Mach Learn Biomed Imag*. 2021;6:1-23.
 23. Sanchez T, Gozcu B, van Heeswijk RB, et al. Scalable learning-based sampling optimization for compressive dynamic MRI. Paper presented at: ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2020; Barcelona, Spain:8584-8588.
 24. Jin KH, Unser M, Yi KM. Self-supervised deep active accelerated MRI. 2020 <http://arxiv.org/abs/1901.04547>
 25. Zeng D, Sandino C, Nishimura D, Vasanawala S, Cheng J. Reinforcement learning for online undersampling pattern optimization. Paper presented at: Proceedings of the International Society of Magnetic Resonance in Medicine (ISMRM); 2019; Montreal, Canada:1092.
 26. Sherry F, Benning M, Reyes JCD, et al. Learning the sampling pattern for MRI. *IEEE Trans Med Imaging*. 2020;39:4310-4321.
 27. Gözcü B, Mahabadi RK, Li YH, et al. Learning-based compressive MRI. *IEEE Trans Med Imaging*. 2018;37:1394-1406.
 28. Wang G, Fessler JA. Efficient approximation of Jacobian matrices involving a non-uniform fast Fourier transform (NUFFT). *IEEE Trans Comput Imaging*. 2023;9:43-54.
 29. Alush-Aben J, Ackerman-Schraier L, Weiss T, Vedula S, Senouf O, Bronstein A. 3D FLAT: feasible learned acquisition trajectories for accelerated MRI. In: Deeba F, Johnson P, Würfl T, Ye JC, eds. *Machine Learning for Medical Image Reconstruction*. Lecture Notes in Computer Science. Springer International Publishing; 2020:3-16.
 30. Gözcü B, Sanchez T, Cevher V. Rethinking sampling in parallel MRI: a data-driven approach. Paper presented at: 2019 27th European Signal Processing Conference (EUSIPCO); 2019;A Coruna, Spain:1-5.
 31. Zibetti MVW, Herman GT, Regatte RR. Fast data-driven learning of parallel MRI sampling patterns for large scale problems. *Sci Rep*. 2021;11:19312.
 32. Fessler JA, Sutton BP. Nonuniform fast Fourier transforms using min-max interpolation. *IEEE Trans Signal Process*. 2003;51:560-574.
 33. Pruessmann KP, Weiger M, Börner P, Boesiger P. Advances in sensitivity encoding with arbitrary k-space trajectories. *Magn Reson Med*. 2001;46:638-651.
 34. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process*. 2004;13:600-612.
 35. Zhao H, Gallo O, Frosio I, Kautz J. Loss functions for image restoration with neural networks. *IEEE Trans Comput Imaging*. 2017;3:47-57.
 36. Radhakrishna CG, Ciuciu P. Jointly learning non-Cartesian k-space trajectories and reconstruction networks for 2D and 3D MR imaging through projection. *Bioengineering*. 2023;10:158.
 37. Schulte RF, Noeske R. Peripheral nerve stimulation-optimal gradient waveform design. *Magn Reson Med*. 2015;74:518-522.
 38. Davids M, Guérin B, Vom Endt A, Schad LR, Wald LL. Prediction of peripheral nerve stimulation thresholds of MRI gradient coils using coupled electromagnetic and neurodynamic simulations. *Magn Reson Med*. 2019;81:686-701.
 39. Maier O, Baete SH, Fyrdahl A, et al. CG-SENSE revisited: results from the first ISMRM reproducibility challenge. *Magn Reson Med*. 2021;85:1821-1839.
 40. Lustig M, Donoho DL, Santos JM, Pauly JM. Compressed sensing MRI. *IEEE Signal Process Mag*. 2008;25:72-82.
 41. Aggarwal HK, Mani MP, Jacob M. MoDL: model-based deep learning architecture for inverse problems. *IEEE Trans Med Imaging*. 2019;38:394-405.
 42. Hao S, Fessler JA, Noll DC, Nielsen JF. Joint design of excitation k-space trajectory and RF pulse for small-tip 3D tailored excitation in MRI. *IEEE Trans Med Imaging*. 2016;35:468-479.
 43. Welling M, Teh YW. Bayesian learning via stochastic gradient Langevin dynamics. *Proceedings of 28th International Conference on Machine Learning(ICML)*. Omnipress; 2011:681-688.
 44. Kingma DP, Ba J. Adam: a method for stochastic optimization. 2017 <http://arxiv.org/abs/1412.6980>
 45. Fessler JA, Lee S, Olafsson VT, Shi HR, Noll DC. Toeplitz-based iterative image reconstruction for MRI with correction for magnetic field inhomogeneity. *IEEE Trans Signal Process*. 2005;53:3393-3402.
 46. Muckley MJ, Stern R, Murrell T, Knoll F. TorchKbNufft: a high-level, hardware-agnostic non-uniform fast fourier transform. Paper presented at: ISMRM Workshop on Data Sampling & Image Reconstruction; 2020; Sedona, AZ, USA.
 47. Desai AD, Schmidt AM, Rubin EB, et al. SKM-TEA: a dataset for accelerated MRI reconstruction with dense image labels for quantitative clinical evaluation. 2022 <http://arxiv.org/abs/2203.06823>
 48. Welsch GH, Scheffler K, Mamisch TC, et al. Rapid estimation of cartilage T2 based on double Echo at steady state (DESS) with 3 tesla. *Magn Reson Med*. 2009;62:544-549.
 49. Souza R, Lucena O, Garrafa J, et al. An open, multi-vendor, multi-field-strength brain MR dataset and analysis of publicly available skull stripping methods agreement. *Neuroimage*. 2018;170:482-494.
 50. Brant-Zawadzki M, Gillan GD, Nitz WR. MP RAGE: a three-dimensional, T1-weighted, gradient-Echo sequence-initial experience in the brain. *Radiology*. 1992;182:769-775.
 51. Uecker M, Lai P, Murphy MJ, Virtue P, Elad M, Pauly JM. ESPIRiT - an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA. *Magn Reson Med*. 2014;71:990-1001.
 52. Barger AV, Block WF, Toropov Y, Grist TM, Mistretta CA. Time-resolved contrast-enhanced imaging with isotropic resolution and broad coverage using an Undersampled 3D projection trajectory. *Magn Reson Med*. 2002;48:297-305.
 53. Herrmann KH, Krämer M, Reichenbach JR. Time efficient 3D radial UTE sampling with fully automatic delay compensation on a clinical 3T MR scanner. *PLoS One*. 2016;11:e0150371.

54. Yu S, Park B, Jeong J. Deep iterative down-up CNN for image denoising. Paper presented at: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 2019; Long Beach, CA.
55. Rettenmeier CA, Maziero D, Stenger VA. Three dimensional radial echo planar imaging for functional MRI. *Magn Reson Med.* 2022;87:193-206.
56. Guo S, Noll DC. Oscillating steady-state imaging (OSSI): a novel method for functional MRI. *Magn Reson Med.* 2020;84:698-712.
57. Guo S, Fessler JA, Noll DC. High-resolution oscillating steady-state fMRI using patch-tensor low-rank reconstruction. *IEEE Trans Med Imaging.* 2020;39:4357-4368.
58. Nielsen JF, Noll DC. TOPPE: a framework for rapid prototyping of MR pulse sequences. *Magn Reson Med.* 2018;79:3128-3134.
59. Muckley MJ, Riemenschneider B, Radmanesh A, et al. Results of the 2020 fastMRI challenge for machine learning MR image reconstruction. *IEEE Trans Med Imaging.* 2021;40:2306-2317.
60. Sutton BP, Noll DC, Fessler JA. Fast, iterative image reconstruction for MRI in the presence of field Inhomogeneities. *IEEE Trans Med Imaging.* 2003;22:178-188.
61. Ahn CB, Cho ZH. Analysis of the eddy-current induced artifacts and the temporal compensation in nuclear magnetic resonance imaging. *IEEE Trans Med Imaging.* 1991;10:47-52.
62. Duyn JH, Yang Y, Frank JA, van der Veen JW. Simple correction method for k-space trajectory deviations in MRI. *J Magn Reson.* 1998;132:150-153.
63. Robison RK, Li Z, Wang D, Ooi MB, Pipe JG. Correction of B0 Eddy current effects in spiral MRI. *Magn Reson Med.* 2019;81:2501-2513.
64. Ma D, Gulani V, Seiberlich N, et al. Magnetic resonance fingerprinting. *Nature.* 2013;495:187.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

Appendix S1. Supporting information

How to cite this article: Wang G, Nielsen J-F, Fessler JA, Noll DC. Stochastic optimization of three-dimensional non-Cartesian sampling trajectory. *Magn Reson Med.* 2023;90:417-431. doi: 10.1002/mrm.29645