Fast Approximators for Least-Norm Reconstructions of Undersampled Non-Cartesian MRI Data

Joshua D. Trzasko1, Yunhong Shu2, Armando Manduca3, and Matt A Bernstein2

1Physiology and Biomedical Engineering, Mayo Clinic, Rochester, MN, United States, 2Department of Radiology, Mayo Clinic, Rochester, MN, United States

Introduction: Least-norm reconstruction of undersampled Cartesian MRI data, commonly referred to as a “zero filling” [1], remains a popular strategy due to its simple and efficient implementation, and readily-characterized behavior. However, despite its mathematical simplicity, least-norm reconstruction of undersampled non-Cartesian MRI data is computationally intensive, requiring use of iterative methods that may be impractical for clinical use [2]. Here, we describe a novel and efficient numerical framework for generating accurate, non-iterative (i.e., direct) approximators of least-norm reconstructions of non-Cartesian MRI data.

Theory: Suppose we model our signal acquisition as \( g = Hf \), where \( H \) is the \( N \rightarrow K \) discrete-time Fourier transform (DTFT) and \( f \) is the (discrete approximation of the) image of interest. To mitigate rank-deficiency challenges, we can consider the equivalent system, \( Rg = RHf \), where \( R \) “prunes” redundant rows of \( H \) and \( g \) (via averaging) such that \( RH \) is full row rank. When \( K \ll N \), the problem of recovering \( f \) from \( g \) is underdetermined. A standard approach for such scenarios is least-norm estimation,

\[
\hat{u} = \arg \min_{d \in \mathbb{R}^N} \| H^* R H \|_2 R g = \arg \min_{d \in \mathbb{R}^N} J(d) \tag{1}
\]

For Cartesian imaging, the crosstalk matrix is the identity scaled by a constant and factors out of (1), yielding “zero filling.” For non-Cartesian, however, this is not the case, and inversion of the crosstalk matrix is computationally challenging. Thus, we approximate it by a diagonal matrix, \( D = \text{diag}(d) \), so as to permit a direct reconstruction, namely \( \hat{u} = H^* R \hat{H} \), \( \hat{H} \) is the identity scaled by a constant on the diagonal. Since \( \hat{H} H^* \approx I \) is positive definite, \( D \) should be as well; and so \( d \) should be real and non-negative. Adopting a least squares loss function, the (constrained) least norm approximator design problem is then

\[
d = \arg \min_{d \in \mathbb{R}^N} \| H^* R \hat{H} \|_2 R g - H^* R \hat{H} \|_2 \hat{u} = \arg \min_{d \in \mathbb{R}^N} J(d) \tag{2}
\]

Ideally, \( g \) would be simultaneously optimized to maximize \( J \), akin to the construction in [6] for gridding kernel optimization; however, this would resort (2) to spectral optimization which is computationally impractical. Instead, we employ \( g = e \), the unit vector and expected spectral response from sampling of a delta function. The gradient of (2) under \( g = e \) is \( \nabla J(d) = 2(\hat{R} H^* \hat{R} d - e) \). As (2) is convex, in this work we adopt Nesterov’s proximal gradient method [7] for solving this problem, which is defined as follows:

\[
\text{INIT : } d_0 = 0, t_0 = 1, y_0 = d_0 \quad \text{REPEAT : } \quad 1) d_{k+1} = P\left(y_k - \frac{1}{t_k} \nabla J(y_k)\right) \quad 2) t_{k+1} = \frac{1}{2} \left(1+\sqrt{1+4t_k}\right) \quad 3) y_{k+1} = \frac{1}{t_{k+1}} \left(y_k - \frac{t_k}{t_{k+1}}(d_k - d_{k+1}) \right)
\]

where \( P \) is the non-negative real projector and the Lipschitz constant, \( L \), can be determined via power iteration. Note that the positive-definiteness of \( \hat{R} H^* \hat{R} \) only asserts \( d^T e = \text{Re}(d^T e) > 0 \), and so the solution to \( \nabla J(d) = 0 \) and (2) are not necessarily equivalent. Fortunately, Nesterov’s scheme provides a very efficient mechanism for determining a solution to the constrained optimization problem. Interestingly, despite possessing a different mathematical objective, the proposed least norm approximator is semantically similar to several existing methods for direct non-Cartesian reconstruction. For example, several authors have considered solving \( \nabla J(d) = 0 \) (and closely-related forms) based on point-spread function (PSF) optimization arguments [8-10]. Also based on PSF arguments, Samsonov et al. [11] suggested a non-negative projected gradient descent that is related to our approach for solving (2), and Bydder et al. [12] later incorporated an added fidelity term (to Jackson et al.’s [8] estimate) that encourages, but does not explicitly enforce, non-negativity.

Example: Fig. 1 shows example 256x256 reconstructions of an 16-shot Archimede spiral acquisition (single-channel, 4096 samples/shot, repeat sampling at the origin) of a resolution phantom obtained by directly solving (1), using the proposed approximation in (2), and, for reference, the iterative method of Ref. [9]. A 1.125x oversampled non-uniform FFT (NUFFT) [6] employing a \( W \times 6 \) Kaiser-Bessel kernel [13] was used to realize \( H \). Equation (1) was solved via 50 conjugate gradient iterations, which required 0.7176s by a multithreaded C++ implementation (OpenMP, FFTW) running on a dual 6-core 3.0GHz machine. Conversely, direct non-Cartesian reconstruction via adjoint NUFFT requires ~9.088 s on the same machine, which is almost 90x faster. Estimation of \( d \) according to (2), via 50 iterations of Nesterov’s algorithm, required only 0.8163 s of computation, and yet can be reused for later reconstructions. Similarly, 50 iterations of the method in Ref. 9 required only 0.4024s. A 1.125x oversampled non-uniform FFT (NUFFT) [6] employing a \( W \times 6 \) Kaiser-Bessel kernel [13] was used to realize \( H \). Equation (1) was solved via 50 conjugate gradient iterations, which required 0.7176s by a multithreaded C++ implementation (OpenMP, FFTW) running on a dual 6-core 3.0GHz machine. Conversely, direct non-Cartesian reconstruction via adjoint NUFFT requires ~9.088 s on the same machine, which is almost 90x faster. Estimation of \( d \) according to (2), via 50 iterations of Nesterov’s algorithm, required only 0.8163 s of computation, and yet can be reused for later reconstructions. Similarly, 50 iterations of the method in Ref. 9 required only 0.4024s.

Summary: We have proposed a novel and efficient numerical strategy for generating fast and accurate approximators of least norm reconstructions of undersampled non-Cartesian MRI data. Beyond standalone application, the proposed method can also be used to reduce the computational complexity of iterative non-Cartesian reconstruction methods requiring repeated calculation of least-norm estimates, such as equality-constrained Compressive Sensing strategies [14].