CROSS-VALIDATION AND PREDICTED RISK ESTIMATION FOR NONLINEAR ITERATIVE REWEIGHTED LEAST-SQUARES MRI RECONSTRUCTION

Sathish Ramani[*], *Jon-Fredrik Nielsen*[†] *and Jeffrey A. Fessler*[*]

[*]EECS Department, University of Michigan, Ann Arbor, MI, USA.
[†]fMRI Laboratory, University of Michigan, Ann Arbor, MI, USA.

ABSTRACT

Regularization is an effective means of reducing noise and artifacts in MR image reconstruction from undersampled k-space data. Proper application of regularization demands appropriate selection of the associated regularization parameter. Generalized cross-validation (GCV) is a popular parameter tuning technique especially for linear reconstruction methods, but its application to nonlinear iterative MRI reconstruction is more involved as it demands the evaluation of the Jacobian matrix of the reconstruction algorithm with respect to complex-valued data. We derive analytical expressions for recursively updating this Jacobian matrix for an iterative reweighted least-squares reconstruction algorithm. Our method can also be used to calculate a predicted risk estimate (PSURE) for MRI based on Stein's principle. We demonstrate with simulations and experiments with real data that regularization parameter selection based on GCV and PSURE provides near-MSE-optimal results for nonlinear MRI reconstruction from undersampled k-space data using ℓ_1-regularization.

Index Terms— MRI reconstruction, regularization, cross-validation, Stein's unbiased risk estimate, Jacobian matrix.

1. INTRODUCTION

Magnetic resonance image (MRI) reconstruction from single-coil undersampled k-space data is an ill-posed problem and requires regularization to provide meaningful reconstruction results. Non-quadratic regularizers that promote sparsity (e.g., ℓ_1-regularization) or that preserve edges (e.g., total variation) are attractive as they can effectively reduce noise and artifacts in the reconstruction [1]. However, a successful application of such regularization criteria demands proper selection of associated regularization parameters that control the balance between noise-amplification and image-smoothing. This task is generic to many regularized reconstruction problems including that of MRI and is usually performed manually.

In this paper, we focus on two quantitative techniques: *generalized cross-validation* (GCV) [2, 3] and the *predicted* form [4] of *Stein's unbiased risk estimate* [5] (PSURE), for automated selection of the regularization parameter for iterative nonlinear MRI reconstruction from undersampled k-space data. Both GCV and SURE-type approaches have been used in image restoration problems, especially involving linear reconstruction algorithms [2]. Their computation for nonlinear reconstruction problems is however complicated by the nonlinearity of the associated algorithms [3, 6–9]. Specifically, they require the evaluation of the Jacobian matrix [3, 4, 6, 8, 9] of the estimator with respect to complex-valued MR data [4] for MRI reconstruction. We propose to evaluate this Jacobian matrix for a fast variant of iterative reweighted least-squares (IRLS) algorithm [1] that is gradient-based and is capable of accommodating a variety of smooth and nonsmooth regularization criteria. To the best of our knowledge, the derivation of this complex-valued Jacobian matrix has not been documented for nonlinear iterative MRI reconstruction with IRLS-type algorithms. We demonstrate using experiments with realistically simulated nonCartesian MR data [10] and with Cartesian MR data acquired with a 3T GE scanner that the use of GCV and PSURE lead to near-optimal selections for the regularization parameter for MRI reconstruction from undersampled data.

This work was supported by the National Institutes of Health under grant P01 CA87634.

2. REGULARIZED MRI RECONSTRUCTION

2.1. Problem Formulation

We formulate regularized MRI reconstruction as a discretized optimization problem,

$$\mathbf{u}_\lambda(\mathbf{y}) = \arg\min_{\mathbf{u}} \left\{ \Psi(\mathbf{u}) = \frac{1}{2}\|\mathbf{y} - \mathbf{F}\,\mathbf{u}\|_2^2 + \lambda\,\Psi(\mathbf{R}\mathbf{u}) \right\}, \quad (1)$$

where the $M \times 1$ vector $\mathbf{y}$ represents (possibly undersampled) Cartesian / nonCartesian data from a single-channel receiver coil, $\mathbf{F}$ is the $M \times N$ Fourier encoding matrix associated with $\mathbf{y}$ such that $M \leq N$, $\mathbf{u}_\lambda(\mathbf{y})$ denotes the $N \times 1$ reconstruction (that is an implicit function of $\mathbf{y}$), Ψ is a suitable regularization, $\mathbf{R} \in \mathbb{R}^{P\times 1}$ is a matrix denoting the regularization operator, e.g., frames, finite differences, etc., and $\lambda > 0$ denotes the regularization parameter. For simplicity of exposition, we consider $\Psi(\mathbf{R}\mathbf{u}) = \sum_{l=1}^{P} \Phi(|[\mathbf{R}\mathbf{u}]_l|)$, where Φ is a suitable convex potential function, e.g., $\Phi(x) = x$ leads to an ℓ_1-regularization. The methods developed in this paper can be also extended to handle other regularizers, e.g., total variation [8].

2.2. An IRLS-MIL Algorithm

We use an iterative reweighted least-squares (IRLS) type algorithm [1] for solving (1). IRLS-type algorithms are gradient-based and provide a general framework for solving reconstruction problems like (1) for a variety of regularization criteria. In the basic form of IRLS, at iteration i, one needs to solve a linear system of equations of the form [1]

$$\mathbf{A}_{(i)}\mathbf{u}_{(i+1)} = \mathbf{F}^{\mathrm{H}}\mathbf{y}, \quad (2)$$

where $\mathbf{A}_{(i)} = \mathbf{F}^{\mathrm{H}}\mathbf{F} + \mathbf{R}^\top\mathbf{\Gamma}_{(i)}^{-1}\mathbf{R}$, $(\cdot)^\top$ and $(\cdot)^{\mathrm{H}}$ represent standard and Hermitian-transposes, respectively, $\mathbf{\Gamma}_{(i)} = \mathrm{diag}\{\boldsymbol{\gamma}_{(i)}\}$ and the l-th element of $\boldsymbol{\gamma}_{(i)} \in \mathbb{R}_+^P$ is $\gamma_{(i)l} = \left(\frac{\lambda}{t}\frac{\mathrm{d}\Phi(t)}{\mathrm{d}t}|_{t=|[\mathbf{R}\mathbf{u}_{(i)}]_l|}\right)^{-1}$ [1, 8]. For ℓ_1-regularization (i.e., $\Phi(x) = x$), this leads to [8]

$$\gamma_{(i)l} = \lambda^{-1}|[\mathbf{R}\mathbf{u}_{(i)}]_l|. \quad (3)$$

To manage nonsmooth regularization (such as the considered ℓ_1-regularization) "corner-rounding" is often administered to (2), i.e., a small positive constant, $\epsilon > 0$, is added to $\{\gamma_{(i)l}\}_l$ to ensure that

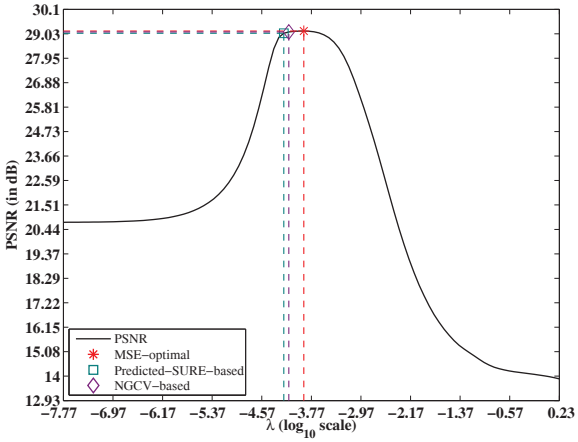

Fig. 1. Experiment with the analytical Shepp-Logan phantom [10] using a nonCartesian spiral trajectory. Plot of PSNR(λ) versus λ indicating λ-values that minimize NGCV(λ), PSURE(λ) and true MSE(λ).

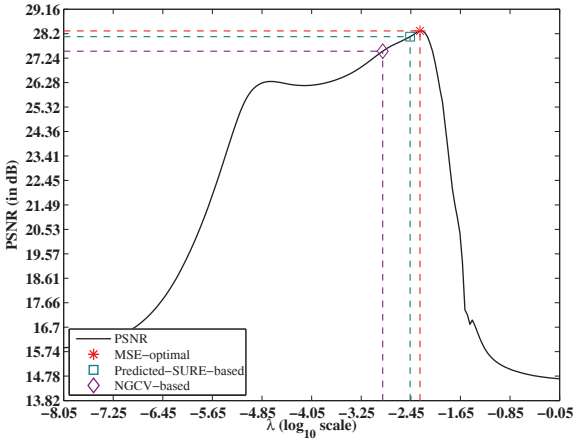

Fig. 2. Experiment with real Cartesian data. Plot of PSNR(λ) versus λ indicating λ-values that minimize NGCV(λ), PSUREλ) and true MSE(λ).

$\gamma_{(i)l} > 0 \; \forall \; l$, and that $\mathbf{A}_{(i)}$ is well-conditioned [1, 8]. As useful as it can be, the inclusion of ϵ sets a tradeoff [11, Sec. VI-A]: small ϵ-values lead to slow convergence while large values result in fast convergence to a solution different from that of (1).

Recently, we proposed a matrix-splitting approach to circumvent the use of ϵ: the basic idea is to add and subtract a matrix $\mathbf{C}_{(i)}$ on both sides of (2) to obtain

$$\mathbf{B}_{(i)}\mathbf{u}_{(i+1,j+1)} = \mathbf{F}^{\mathrm{H}}\mathbf{y} + [\mathbf{C}_{(i)} - \mathbf{F}^{\mathrm{H}}\mathbf{F}]\mathbf{u}_{(i+1,j)}, \tag{4}$$

where $\mathbf{B}_{(i)} \triangleq \mathbf{C}_{(i)} + \mathbf{R}^{\top}\mathbf{\Gamma}_{(i)}^{-1}\mathbf{R}$. Thus,

$$\mathbf{u}_{(i+1,j+1)} = \mathbf{B}_{(i)}^{-1}(\mathbf{F}^{\mathrm{H}}\mathbf{y} + [\mathbf{C}_{(i)} - \mathbf{F}^{\mathrm{H}}\mathbf{F}]\mathbf{u}_{(i+1,j)}). \tag{5}$$

The sequence of iterations indexed by j in (5) is guaranteed to converge to a solution of (2) when $\mathbf{C}_{(i)} - \mathbf{F}^{\mathrm{H}}\mathbf{F} \succ \mathbf{0}$ [1]. Applying Sherman-Morrison-Woodbury matrix inversion lemma (MIL) to $\mathbf{B}_{(i)}^{-1}$ leads to $\mathbf{B}_{(i)}^{-1} = \mathbf{C}_{(i)}^{-1} - \mathbf{C}_{(i)}^{-1}\mathbf{R}^{\top}\mathbf{G}_{(i)}^{-1}\mathbf{R}\mathbf{C}_{(i)}^{-1}$ and the following update-rules for (5):

$$\mathbf{u}_{(i+1,j+1)} = \mathbf{b}_{(i+1,j)} - \mathbf{C}_{(i)}^{-1}\mathbf{R}^{\top}\mathbf{v}_{(i+1,j)}, \tag{6}$$

$$\text{solve } \{\mathbf{G}_{(i)}\mathbf{v}_{(i+1,j)} = \mathbf{R}\mathbf{b}_{(i+1,j)}\} \text{ for } \mathbf{v}_{(i+1,j)}, \tag{7}$$

where $\mathbf{b}_{(i+1,j)} \triangleq \mathbf{C}_{(i)}^{-1}\mathbf{F}^{\mathrm{H}}\mathbf{y} + [\mathbf{I}_N - \mathbf{C}_{(i)}^{-1}\mathbf{F}^{\mathrm{H}}\mathbf{F}]\mathbf{u}_{(i+1,j)}$ and $\mathbf{G}_{(i)} \triangleq \mathbf{\Gamma}_{(i)} + \mathbf{R}\mathbf{C}_{(i)}^{-1}\mathbf{R}^{\top}$. Unlike $\mathbf{A}_{(i)}$ that depends on $\mathbf{\Gamma}_{(i)}^{-1}$ (2), $\mathbf{G}_{(i)}$ depends on $\mathbf{\Gamma}_{(i)}$, so we do not need to use "corner-rounding" in our scheme. We apply the following iterative solver for (7) that is linear in both $\mathbf{b}_{(\cdot)}$ and $\mathbf{v}_{(\cdot)}$:

$$\mathbf{v}_{(i+1,j,k+1)} = \mathbf{D}_{\mathbf{\Gamma}_{(i)}}^{-1}[\mathbf{R}\mathbf{b}_{(i+1,j)} + \mathbf{H}_{\rho}\mathbf{v}_{(i+1,j,k)}], \tag{8}$$

where $\mathbf{H}_{\rho} \triangleq \rho\mathbf{I}_P - \mathbf{R}\mathbf{C}_{(i)}^{-1}\mathbf{R}^{\top}$, $\mathbf{I}_P$ is the identity matrix of size P, and $\mathbf{D}_{\mathbf{\Gamma}_{(i)}} \triangleq \mathbf{\Gamma}_{(i)} + \rho\mathbf{I}_P$ is a diagonal matrix. In practice, we perform K iterations of (8) and apply the final update $\mathbf{v}_{(i+1,j,K)}$ in place of $\mathbf{v}_{(i+1,j)}$ in (6). Thus, (6) and (8) together form the IRLS-MIL algorithm that we propose to employ for performing MRI reconstruction based on (1). The update-rule (8) guarantees convergence to a solution of (7) when $\rho > \max \text{eigval}\{\mathbf{R}\mathbf{C}_{(i)}^{-1}\mathbf{R}^{\top}\}$ [1] and also simplifies the computation of GCV and PSURE as elucidated in Section 4.

3. CROSS-VALIDATION & RISK ESTIMATION

The quality of reconstruction $\mathbf{u}_{\lambda}$ greatly depends on the regularization parameter λ in (1); careful selection of λ is crucial as too small or too large a λ-value can result in either noisy reconstructions or loss of details, respectively. In the following, we focus on two quantitative techniques for automated adjustment of λ for MRI reconstruction using the IRLS-MIL algorithm described in Section 2.2.

Cross-validation is a popular method for parameter selection in inverse problems especially involving linear algorithms [2]. Since IRLS-MIL is nonlinear, we specifically consider the GCV measure in [3] that is applicable to nonlinear algorithms. We denote this by NGCV (i.e., GCV for nonlinear methods):

$$\text{NGCV}(\lambda) \triangleq \frac{M^{-1}\|\mathbf{y} - \mathbf{F}\mathbf{u}_{\lambda}(\mathbf{y})\|_2^2}{(1 - M^{-1}\Re\{\text{tr}\{\mathbf{F}\mathbf{J}(\mathbf{u}_{\lambda};\mathbf{y})\}\})^2}, \tag{9}$$

where $\text{tr}\{\cdot\}$ denotes the trace of a matrix. For MRI reconstruction, $\mathbf{J}(\mathbf{u}_{\lambda};\mathbf{y}) \in \mathbb{C}^{N\times M}$ denotes the complex-valued Jacobian matrix of $\mathbf{u}_{\lambda}$ with respect to $\mathbf{y}$ that is defined in terms of its elements as [4, 12]

$$[\mathbf{J}(\mathbf{u}_{\lambda};\mathbf{y})]_{nm} \triangleq \frac{1}{2}\left(\frac{\partial u_{\lambda,n}(\mathbf{z})}{\partial z_{\Re m}} - \iota\frac{\partial u_{\lambda,n}(\mathbf{z})}{\partial z_{\mathcal{I}m}}\right)\Bigg|_{\mathbf{z}=\mathbf{y}}, \tag{10}$$

where $z_{\Re m}$ and $z_{\mathcal{I}m}$ denote the m-th component of the real and imaginary parts of a vector $\mathbf{z} \in \mathbb{C}^M$. When $\mathbf{u}_{\lambda}$ is specified in terms of $\mathbf{y}$ and $\mathbf{y}^{\star}$ (the complex conjugate of $\mathbf{y}$), $\mathbf{J}(\mathbf{u}_{\lambda};\mathbf{y})$ can be evaluated treating $\mathbf{y}$ as a variable and $\mathbf{y}^{\star}$ as a constant [4, 12]. Similarly, $\mathbf{J}(\mathbf{u}_{\lambda};\mathbf{y}^{\star})$ can be evaluated treating $\mathbf{y}$ as constant [4, 12]. We take the real part, $\Re\{\cdot\}$, in the denominator of (9) to avoid spurious complex entries while computing NGCV(λ) numerically.

As an alternative to NGCV(λ), we also consider the *predicted mean squared-error*, PMSE(λ) $\triangleq M^{-1}\|\mathbf{F}(\mathbf{u}_{\text{true}} - \mathbf{u}_{\lambda}(\mathbf{y})\|_2^2$, for quantifying image-quality where $\mathbf{u}_{\text{true}}$ is the unknown (deterministic) true image. Since PMSE(λ) depends on $\mathbf{u}_{\text{true}}$, it cannot be directly used and must be estimated in practice. Assuming that noise in $\mathbf{y}$ is zero-mean white complex-Gaussian with variance σ^2, Stein's principle [5] can be used to obtain[1] a *predicted Stein's unbiased risk*

[1]Derivation of (11) requires the hypotheses that the components $\{u_{\lambda,n}(\mathbf{y})\}_{n=1}^{N}$ of $\mathbf{u}_{\lambda}(\mathbf{y})$ do not grow faster than an appropriate exponential function [7, Theorem 1] and be (weakly) differentiable with respect to the real and imaginary parts of the components $\{y_m\}_{m=1}^{M}$ of $\mathbf{y}$. Evaluation of NGCV(λ) also requires the differentiability hypothesis.

estimate [4, 5, 7]: PSURE(λ) $\triangleq$

$$M^{-1}\|\mathbf{y} - \mathbf{F}\mathbf{u}_\lambda(\mathbf{y})\|_2^2 - \sigma^2 + 2\sigma^2 M^{-1}\Re\{\mathrm{tr}\{\mathbf{F}\mathbf{J}(\mathbf{u}_\lambda;\mathbf{y})\}\}, \quad (11)$$

that is unbiased in the sense that $\mathcal{E}_\mathbf{y}\{\text{PMSE}(\lambda)\} = \mathcal{E}_\mathbf{y}\{\text{PSURE}(\lambda)\}$ [4, 5, 7], where $\mathcal{E}_\mathbf{y}\{\cdot\}$ represents expectation with respect to $\mathbf{y}$.

NGCV(λ), PSURE(λ) (and PMSE(λ)) are computed in the measurement-domain: for MRI, this corresponds to evaluating these performance-measures at sample locations in k-space. While these are not equivalent to the usual image-domain error-measure MSE(λ) $\triangleq N^{-1}\|\mathbf{u}_{\mathrm{true}} - \mathbf{u}_\lambda(\mathbf{y})\|_2^2$, it must be noted that it is impossible to estimate MSE(λ) for ill-posed inverse problems in general [7, 9]. This is apparent for MRI reconstruction from undersampled k-space data, where $\mathbf{y}$ contains only partial information about $\mathbf{u}_{\mathrm{true}}$ due to undersampling. We demonstrate through numerical experiments that NGCV(λ) and PSURE(λ) can still provide near-MSE-optimal λ for MRI reconstruction (1). NGCV(λ) does not require require σ^2 unlike PSURE(λ). However, both NGCV(λ) and PSURE(λ) necessitate the evaluation of $\mathbf{J}(\mathbf{u}_\lambda;\mathbf{y})$. We propose to evaluate $\mathbf{J}(\mathbf{u}_\lambda;\mathbf{y})$ analytically for the IRLS-MIL scheme (6), (8) for MRI reconstruction (1).

4. EVALUATION OF THE JACOBIAN MATRIX $\mathbf{J}(\mathbf{u}_\lambda;\mathbf{y})$

Due to the iterative nature of IRLS-MIL, we compute $\mathbf{J}(\mathbf{u}_\lambda;\mathbf{y})$ recursively. We also need $\mathbf{J}(\mathbf{u}_\lambda;\mathbf{y}^\star)$ as it plays an equal role in the derivation of $\mathbf{J}(\mathbf{u}_\lambda;\mathbf{y})$. In the sequel, we drop the subscript λ in $\mathbf{u}_\lambda$ and use $\mathbf{z}$ to represent either $\mathbf{y}$ or $\mathbf{y}^\star$ where applicable, for ease of notation. We begin with (6) and use linearity of (10) to obtain, at the end of K iterations of (8), that

$$\mathbf{J}(\mathbf{u}^{(i+1,j+1)};\mathbf{z}) = \mathbf{J}(\mathbf{b}^{(i+1,j)};\mathbf{z}) - \mathbf{C}_{(i)}^{-1}\mathbf{R}^\top\mathbf{J}(\mathbf{v}^{(i+1,j,K)};\mathbf{z}), \quad (12)$$

where $\mathbf{J}((\cdot)^{(i,j)};\cdot)$ indicates a Jacobian matrix update at a (nested) iteration indexed by i and j. From the definition of complex-valued Jacobian matrices (Section 3), we have that

$$\begin{aligned}\mathbf{J}(\mathbf{b}^{(i+1,j)};\mathbf{y}) &= \mathbf{C}_{(i)}^{-1}\mathbf{F}^{\mathrm{H}} + [\mathbf{I}_N - \mathbf{C}_{(i)}^{-1}\mathbf{F}^{\mathrm{H}}\mathbf{F}]\,\mathbf{J}(\mathbf{u}^{(i+1,j)};\mathbf{y}),\\ \mathbf{J}(\mathbf{b}^{(i+1,j)};\mathbf{y}^\star) &= [\mathbf{I}_N - \mathbf{C}_{(i)}^{-1}\mathbf{F}^{\mathrm{H}}\mathbf{F}]\,\mathbf{J}(\mathbf{u}^{(i+1,j)};\mathbf{y}^\star). \quad (13)\end{aligned}$$

Using product rule for Jacobian matrices [4, 12] and the fact that (8) involves only a diagonal matrix $\mathbf{D}_{\mathbf{\Gamma}_{(i)}}^{-1}$, we get from (8) that

$$\begin{aligned}\mathbf{J}(\mathbf{v}^{(i+1,j,k+1)};\mathbf{z}) = \mathbf{D}_{\mathbf{\Gamma}_{(i)}}^{-1}[\mathbf{R}\mathbf{J}(\mathbf{b}^{(i+1,j)};\mathbf{z}) + \mathbf{H}_\rho\mathbf{J}(\mathbf{v}^{(i+1,j,k)};\mathbf{z})]\\ -\mathbf{D}_\mathbf{v}\mathbf{D}_{\mathbf{\Gamma}_{(i)}}^{-2}\mathbf{J}(\mathbf{\gamma}^{(i)};\mathbf{z}), \quad (14)\end{aligned}$$

where $\mathbf{D}_\mathbf{v} \triangleq \mathrm{diag}\{\mathbf{R}\mathbf{b}_{(i+1,j)} + \mathbf{H}_\rho\mathbf{v}_{(i+1,j,k)}\}$. Since $\mathbf{\gamma}_{(i)}$ is a function of $\mathbf{y}$ (and $\mathbf{y}^\star$) via $\mathbf{u}^{(i)}$, we apply chain rule [4, 12] for Jacobian matrices to get that

$$\mathbf{J}(\mathbf{\gamma}^{(i)};\mathbf{z}) = \mathbf{J}(\mathbf{\gamma}^{(i)};\mathbf{u}_{(i)})\mathbf{J}(\mathbf{u}^{(i)};\mathbf{z}) + (\mathbf{J}(\mathbf{\gamma}^{(i)};\mathbf{u}_{(i)})\,\mathbf{J}(\mathbf{u}^{(i)};\mathbf{z}^\star))^\star.$$

Thus, evaluation of $\mathbf{J}(\mathbf{u}^{(\cdot)};\mathbf{y})$ requires $\mathbf{J}(\mathbf{u}^{(\cdot)};\mathbf{y}^\star)$ as mentioned earlier. It only remains to evaluate $\mathbf{J}(\mathbf{\gamma}^{(i)};\mathbf{u}_{(i)})$: for the considered ℓ_1-regularization, we have from (3) that $\mathbf{\gamma}_{(i)} = \{\lambda^{-1}|[\mathbf{R}\mathbf{u}_{(i)}]_l|\}_l = \{\lambda^{-1}\sqrt{[\mathbf{R}\mathbf{u}_{(i)}]_l[\mathbf{R}\mathbf{u}_{(i)}^\star]_l}\}_l$, which (treating $\mathbf{u}_{(i)}^\star$ as a constant) leads to [4, 12]

$$\mathbf{J}(\mathbf{\gamma}^{(i)};\mathbf{u}_{(i)}) = \mathrm{diag}\{(2\lambda^2\mathbf{\gamma}_{(i)})^{-1}\}\,\mathrm{diag}\{\mathbf{R}\mathbf{u}_{(i)}^\star\}\,\mathbf{R}. \quad (15)$$

For typical reconstruction sizes, the Jacobian matrices will be enormous and cannot be stored and manipulated directly. So we use a Monte-Carlo procedure [6] to estimate the traces in (9) and (11): specifically, we store and update $N \times 1$ matrix-vector products of the form $\mathbf{J}(\mathbf{u}^{(\cdot)};\cdot)\mathbf{n}$ and $\mathbf{J}(\mathbf{v}^{(\cdot)};\cdot)\mathbf{n}$ where $\mathbf{n} \in \mathbb{R}^M$ is an i.i.d. zero-mean random vector with unit variance. Then, we estimate the traces in (9) and (11) at every iteration as

$$\mathrm{tr}\{\mathbf{F}\mathbf{J}(\mathbf{u}^{(i,j)};\mathbf{y})\} \approx \hat{t} \triangleq \mathbf{n}^\top\mathbf{F}\mathbf{J}(\mathbf{u}^{(i,j)};\mathbf{y})\mathbf{n}. \quad (16)$$

It can be shown that $\hat{t}$ is unbiased [4], i.e.,

$$\mathcal{E}_\mathbf{n}\{\hat{t}\} = \mathcal{E}_\mathbf{n}\{\mathrm{tr}\{\mathbf{F}\mathbf{J}(\mathbf{u}^{(i,j)};\mathbf{y})\}\},$$

and its variance can be minimized by using a binary random vector $\mathbf{n} = \mathbf{n}_{\pm 1}$ whose entries are either $+1$ or -1 with probability 0.5 [4] compared to using a Gaussian $\mathbf{n}$ [6]. To summarize our scheme, we run the sequence of iterations (6), (8) to obtain the reconstruction $\mathbf{u}_{(\cdot)}$ and simultaneously use (12)-(16) with $\mathbf{n}_{\pm 1}$ to numerically compute NGCV(λ) and PSURE(λ) at every iteration. Since NGCV(λ) and PSURE(λ) both require $\mathrm{tr}\{\mathbf{F}\mathbf{J}(\mathbf{u}^{(i,j)};\mathbf{y})\}$, their computational complexity is very similar.

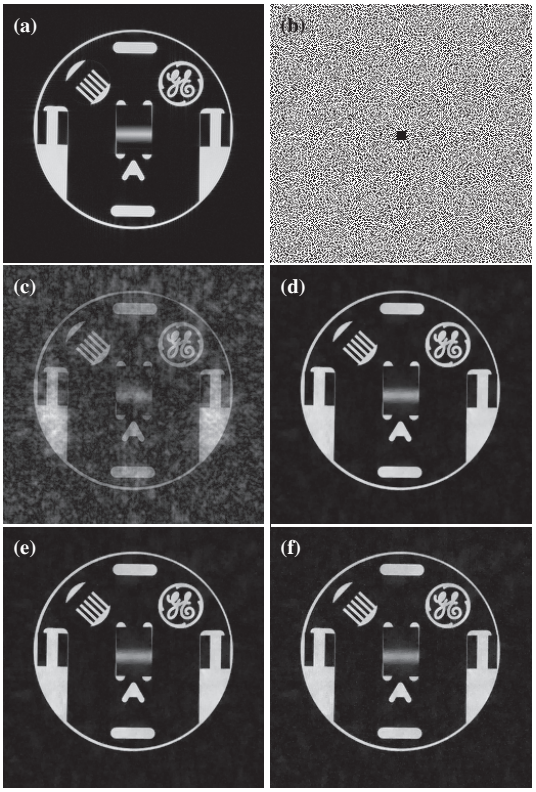

Fig. 3. Experiment with real Cartesian data. (a) Magnitude of reference (i.e., $\mathbf{u}_{\mathrm{true}}$); (b) *Retrospective* random sampling (60% undersampling); Magnitude of (c) zero-filled iFFT reconstruction (15.82 dB) and ℓ_1-regularized reconstructions with λ selected to minimize (d) true MSE(λ) (28.31 dB); (e) PSURE(λ) (28.08 dB); (f) NGCV(λ) (27.52 dB).

5. EXPERIMENTAL SETUP & RESULTS

We implemented products with $\mathbf{F}$ using NUFFT [13] for nonCartesian data and FFT for Cartesian data. We chose $\mathbf{C}_{(i)}\ \forall\ i$ to be a circulant matrix and implemented $\mathbf{C}_{(i)}^{-1}$ using FFTs in IRLS-MIL. For nonCartesian data, we used $\mathbf{C}_{(i)} = \mathbf{C}_{\mathrm{Frob}} + \nu\mathbf{I}_N\ \forall\ i$, where

$\mathbf{C}_{\mathrm{Frob}}$ is a circulant matrix that is "closest" to $\mathbf{F}^{\mathrm{H}}\mathbf{F}$ in Frobenius norm [14] and ν was selected such that $\nu\mathbf{I} \succ \mathbf{F}^{\mathrm{H}}\mathbf{F} - \mathbf{C}_{\mathrm{Frob}}$. For Cartesian data, we simply set $\mathbf{C}_{(i)} = \mathbf{F}^{\mathrm{H}}\mathbf{F} + \nu\mathbf{I}_N \ \forall \ i$ for some $\nu > 0$, since $\mathbf{F}^{\mathrm{H}}\mathbf{F}$ is already circulant in this case. We used finite differences for $\mathbf{R}$ in (1) and performed 10 j-, 10 k- and 7 i-iterations of IRLS-MIL (6), (8) using $\mathbf{F}^{\mathrm{H}}\mathbf{y}$ as the initialization. In all experiments, the overall compute time for obtaining the reconstruction and simultaneously evaluating NGCV(λ) and PSURE(λ) for a given λ under these settings was 50 seconds on a 8-core PC with 2.80-GHz Intel Xeon processors. For the purpose of illustration, we evaluated NGCV(λ) and PSURE(λ) for a wide range of λ (see Figs. 1 and 2), while in practice, golden-section search can be used to optimize λ with relatively fewer evaluations of NGCV(λ) and PSURE(λ).

In the first experiment, we used a spiral trajectory (with 64 leaves and 512 samples per leaf corresponding to 50% undersampling) and simulated realistic nonCartesian MR data of 40 dB SNR using the analytical Shepp-Logan phantom of Guerquin-Kern *et al.* [10]. We reconstructed 256×256 images of the Shepp-Logan phantom by running the IRLS-MIL algorithm for varying λ. We assumed σ^2 was available for computing PSURE(λ). Fig. 1 plots PSNR(λ) $\triangleq$ $10\log_{10}(\max\{\mathbf{u}_{\mathrm{true}}\}^2/\mathrm{MSE}(\lambda))$ as a function λ and indicates λ-values that minimize NGCV(λ), PSURE(λ) and the true MSE(λ). We see that both NGCV- and PSURE-based λ-selections lead to reconstructions with PSNRs almost close to that of the minimum-MSE reconstruction.

Next, we acquired 10 independent sets of fully-sampled 2-D Cartesian data (256×256) of a GE-phantom using a 3T GE scanner (GRE sequence with flip angle = 35°, T_R = 200 ms, T_E = 7 ms, FOV = 15 cm). These fully-sampled Cartesian datasets were used to reconstruct (using iFFT) 2-D images that were then averaged to obtain a reference image that served as the "unknown" $\mathbf{u}_{\mathrm{true}}$ (Fig. 3a) for computing MSE(λ) and PSNR(λ). We estimated σ^2 for use in PSURE(λ) from separately acquired dummy-data (with the same scan setting) when no RF field was applied. We *retrospectively* undersampled data from one of the 10 sets (Fig. 3b) and ran the IRLS-MIL algorithm minimizing NGCV(λ) and PSURE(λ) individually. We plot PSNR(λ) versus λ in Fig. 2 for this experiment. Again, PSURE-based selection is close to the minimum of MSE and the corresponding PSURE-based reconstruction (Fig. 3e) is visually similar to the MSE-optimal one (Fig. 3d). Although NGCV-selection is sub-optimal, it leads to only a marginal decrease in PSNR and yields a fairly accurate result (Fig. 3f) that is visually comparable to the PSURE-based reconstruction (Fig. 3e).

6. SUMMARY, CONCLUSION & DISCUSSION

In this paper, we presented a method for computing the generalized cross-validation measure (NGCV) [3] and a predicted risk estimate (PSURE) [4] for nonlinear MRI reconstruction using an iterative reweighted least-squares-type (IRLS-MIL [1]) algorithm. Both NGCV and PSURE require the evaluation of a complex-valued Jacobian matrix [3, 7] that we carried out analytically for the IRLS-MIL algorithm. We presented numerical results using simulated (nonCartesian) and real (Cartesian) MR data and illustrated that both NGCV and PSURE are able to provide near-optimal selections of the regularization parameter for regularized nonlinear MRI reconstruction.

Our methods can be directly applied to 3-D Cartesian and non-Cartesian MRI reconstruction (albeit with increased computation). The principles underlying our work can also be extended to other algorithms, e.g., the split-Bregman algorithm [4].

Evaluating NGCV(λ) and PSURE(λ) requires additional memory and computation of the same order as the IRLS-MIL algorithm used for reconstruction. Although these performance measures can be optimized using golden-section search, multiple evaluations of NGCV(λ) and PSURE(λ) may be necessary. Alternatively, MRI reconstruction can be formulated as a constrained minimization of a suitable (nonquadratic) regularization subject to a data-consistency constraint: this avoids the need for searching an appropriate λ, however such constrained problems are harder to tackle. We are currently investigating a comparison of the methods proposed in this paper with the constrained formulation both in terms of reconstruction quality and computation time.

7. REFERENCES

[1] S. Ramani and J. A. Fessler, "An accelerated iterative reweighted least squares algorithm for compressed sensing MRI," in *Proc. IEEE Intl. Symp. Biomed. Imag.*, 2010, pp. 257–60.

[2] S. J. Reeves, "Optimal space-varying regularization in iterative image restoration," *IEEE Trans. Im. Proc.*, vol. 3, no. 3, pp. 319–23, May 1994.

[3] D. A. Girard, "The fast Monte-Carlo Cross-Validation and C_L procedures: Comments, new results and application to image recovery problems - Rejoinder," *Computation. Stat.*, vol. 10, pp. 251–258, 1995.

[4] S. Ramani, Z. Liu, J. Rosen, J. F. Nielsen, and J. A. Fessler, "Regularization parameter selection for nonlinear iterative image restoration and MRI reconstruction using GCV and SURE-based methods," *IEEE Trans. Im. Proc.*, DOI: 10.1109/TIP.2012.2195015, 2012, to appear.

[5] C. Stein, "Estimation of the mean of a multivariate normal distribution," *Ann. Stat.*, vol. 9, no. 6, pp. 1135–51, Nov. 1981.

[6] C. Vonesch, S. Ramani, and M. Unser, "Recursive risk estimation for non-linear image deconvolution with a wavelet-domain sparsity constraint," *Proc. IEEE Intl. Conf. Img. Proc.*, pp. 665–8, 2008.

[7] Y. C. Eldar, "Generalized SURE for exponential families: applications to regularization," *IEEE Trans. Sig. Proc.*, vol. 57, no. 2, pp. 471–81, Feb. 2009.

[8] S. Ramani, J. Rosen, Z. Liu, and J. A. Fessler, "Iterative weighted risk estimation for nonlinear image restoration with analysis priors," *Proc. SPIE Elec. Img.*, vol. 8296, pp. 82960N–1–12, 2012.

[9] R. Giryes, M. Elad, and Y. C. Eldar, "The projected GSURE for automatic parameter tuning in iterative shrinkage methods," *Appl. Comp. Harm. Anal.*, vol. 30, no. 3, pp. 407–22, 2011.

[10] M. Guerquin-Kern, L. Lejeune, K. P. Pruessmann, and M. Unser, "Realistic analytical phantoms for parallel magnetic resonance imaging," *IEEE Trans. Med. Img.*, vol. 31, no. 3, pp. 626–36, 2012.

[11] S. Ramani and J. A. Fessler, "Parallel MR image reconstruction using augmented Lagrangian methods," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 694–706, 2011.

[12] A. Hjorungnes and D. Gesbert, "Complex-valued matrix differentiation: techniques and key results," *IEEE Trans. Sig. Proc.*, vol. 55, no. 6, pp. 2740–6, 2007.

[13] J. A. Fessler and B. P. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Trans. Sig. Proc.*, vol. 51, no. 2, pp. 560–74, Feb. 2003.

[14] T. F. Chan, "An optimal circulant preconditioner for Toeplitz systems," *SIAM J. Sci. Stat. Comp.*, vol. 9, no. 4, pp. 766–71, July 1988.