BREGMAN OPERATOR SPLITTING WITH VARIABLE STEPSIZE FOR TOTAL GENERALIZED VARIATION BASED MULTI-CHANNEL MRI RECONSTRUCTION

by

BENJAMIN COWEN

Submitted in partial fulfillment of the requirements for the degree of Master of Science

Department of Mathematics, Applied Mathematics, and Statistics
CASE WESTERN RESERVE UNIVERSITY

August, 2015
CASE WESTERN RESERVE UNIVERSITY
SCHOOL OF GRADUATE STUDIES

We hereby approve the thesis of
Benjamin Cowen,
candidate for the degree of
Master of Science*.

Committee Chair
Weihong Guo

Committee Member
Steven Izen

Committee Member
Julia Dobrosotskaya

Date of Defense
April 29th, 2015

*We also certify that written approval has been obtained for any proprietary material contained therein.
## Contents

1 Introduction .......................................................... 7

2 Preliminaries .......................................................... 7
   2.1 Partial Parallel Imaging ........................................... 7
   2.2 TGV ................................................................. 8

3 Proposed Model and Algorithm ..................................... 10
   3.1 Model ............................................................... 10
   3.2 Algorithm ......................................................... 12

4 Numerical Results and Comparisons ............................. 16

5 Conclusions ............................................................ 30
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Image smoothing results of TV and TGV. The TGV result has sharper edges and less oil painting artifacts. Figure is extracted from [6].</td>
</tr>
<tr>
<td>2</td>
<td>Reconstructions of a synthetic test image using a 30 coil sensitivity map and 7.87% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.</td>
</tr>
<tr>
<td>3</td>
<td>Reconstructions of a synthetic test image using a 30 coil sensitivity map and 14.48% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.</td>
</tr>
<tr>
<td>4</td>
<td>Reconstructions of a foot MRI using a 30 coil sensitivity map and 7.87% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.</td>
</tr>
<tr>
<td>5</td>
<td>Reconstructions of a pancreas MRI using a 30 coil sensitivity map and 7.87% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.</td>
</tr>
<tr>
<td>6</td>
<td>Reconstructions of a pancreas MRI using a 30 coil sensitivity map and 14.48% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.</td>
</tr>
<tr>
<td>7</td>
<td>Reconstructions of a shoulder MRI using a 30 coil sensitivity map and 7.87% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.</td>
</tr>
<tr>
<td>8</td>
<td>Reconstructions of a shoulder MRI using a 30 coil sensitivity map and 14.48% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.</td>
</tr>
<tr>
<td>9</td>
<td>Reconstructions of a photograph of a train using a 30 coil sensitivity map and 14.48% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.</td>
</tr>
</tbody>
</table>
List of Tables

1. Undersampling Ratios for 128x128 target image, and an example of a $P$ corresponding to 40 radial projections (after reshaping by stacking its columns). .................................................. 17
2. Parameter Selections ........................................................................................................ 18
3. Relative errors achieved by TGV and TV, respectively, under various sampling conditions when reconstructing the smoothly shaded test image. .................................................. 19
4. Relative errors achieved by TGV and TV, respectively, under various sampling conditions when reconstructing a foot MR image. ................................................................. 23
5. Relative errors achieved by TGV and TV, respectively, under various sampling conditions when reconstructing the pancreas MR image. ............................................................... 23
6. Relative errors achieved by TGV and TV, respectively, under various sampling conditions when reconstructing the shoulder MR image. ................................................................. 28
7. Relative errors achieved by TGV and TV, respectively, under various sampling conditions when reconstructing the photograph of a train. ............................................................... 28
Acknowledgement

I place on record my thanks to Nicole Seiberlich and Jesse Hamilton (Biomedical Engineering Department, Case Western Reserve University) for providing me with the MRI data necessary for the simulations described below. I am also grateful for MR-Technology Information Portal (www.mrtip.com) for providing free magnetic resonance images and information for academic use.

A special thanks is extended to the coordinators of the Department of Mathematics, Applied Mathematics, and Statistics for providing me with facilities and high-performance machines which I relied on to conduct my research. Particularly, Catherine Pizzano and Juwanda Rowell.

I wish to express my sincere thanks to the many Department faculty members who have contributed to my understanding of mathematics, imaging, and my own abilities over the past 5 years.

I take this opportunity to express the deepest gratitude to my advisor, professor, and friend, Dr. Weihong Guo. I am extremely thankful and indebted to her for her expertise and sincere, valuable guidance and encouragement extended to me.
Bregman Operator Splitting with Variable Stepsize for Total Generalized Variation Based Multi-Channel MRI Reconstruction

Abstract by

BENJAMIN ETHAN COWEN

This paper presents a fast algorithm for total generalized variation (TGV) based image reconstruction of magnetic resonance images collected by a technique known as partial parallel imaging (PPI). TGV is a generalization of the commonly employed total variation (TV) regularizer. TV reconstructs piecewise constant images and is known to produce oil-painting artifacts, while TGV reconstructs images with piecewise polynomial intensities and largely avoids this issue. The proposed algorithm combines the Bregman Operator Splitting with Variable Stepsize (BOSVS) approach derived by Chen, Hager, et al. [8] with the closed-form expressions for the TGV subproblem that arises in the alternating directional method of multipliers, derived by Guo, Qin and Yin [13]. The ill-conditioned inversion matrix that comes from PPI is approximated according to a stepsize rule similar to that in BOSVS. The stepsize rule starts with a Barzilai-Borwein step, then uses a line search to ensure convergence and efficiency. The proposed regularizer is shown to achieve better results than TV, especially for reconstructing smooth details, in sampling conditions as low as 7.87%.
1 Introduction

Inverse problems related to image reconstruction have triggered the development of many novel optimization algorithms to compensate for noisy, blurry, and/or incomplete imaging measurements. The general form of this problem is given by

\[ u^* = \arg \min_u H(u) + \phi(u), \]  

where \( H(u) \) is a fidelity term related to the data and \( \phi(u) \) is a regularization term constructed to impose prior knowledge on the solution to avoid non-uniqueness, e.g. by excluding non-smooth solutions. When \( H(u) \) represents an underdetermined system, an appropriate choice of regularizer is crucial to recovering images with high accuracy. Total variation (TV) is a simple, commonly used regularization term that considers first order derivatives of solutions and reconstructs functions with piecewise constant values. In image processing it effectively preserves sharp edges but also causes oil painting artifacts. Total generalized variation (TGV) largely avoids these artifacts by demanding smoothness from partial derivatives of various orders of the solution. TGV generalizes TV and leads to piecewise polynomial intensities, allowing more precise reconstruction of smooth regions while retaining the edge preserving benefits of TV.

The focus of this paper is to present an algorithm that solves (1.1) with highly ill-conditioned \( H(u) \) such as those which arise from a technique in magnetic resonance (MR) imaging called partial parallel imaging (PPI). MR imaging is a versatile technology that provides better contrast between soft tissues than most other imaging modalities, but is very expensive and thus has much to benefit from increases in efficiency and speed. MR images are obtained by placing a subject in a strong magnetic field and then repeatedly transmitting radio frequency electromagnetic fields to stimulate hydrogen nuclei within the body. When these nuclei relax, a detectable frequency signal is produced. These signals are detected by a receiver coil and then inverse-transformed to image space to obtain an image of the scanned object. PPI is an emerging MR technique that surrounds the target with receiver coils that collect in parallel portions of the Fourier components. In PPI, the scan time is reduced by taking incomplete measurements while the image quality is increased. However, taking very sparse measurements without using any regularization results in aliased images. Many image processing algorithms have been developed to address the reduction of such artifacts [3,7,19,22].

The remainder of the paper is organized as follows. We start with a brief mathematical overview of TGV and PPI in Section 2. Section 3 contains the derivation of our proposed model and algorithm. Section 4 contains numerical results. Finally, conclusions and discussion are provided in Section 5.

2 Preliminaries

2.1 Partial Parallel Imaging

PPI is often based on the following equation:

\[ PFS_j u = f_j, \]  

where \( S_j \in \mathbb{C}^{N \times N} \) is a diagonal matrix corresponding to the sensitivity map of the \( j \)th receiver coil, \( F \) is the Fourier transform, \( f_j \in \mathbb{C}^N \) is the partial k-space data collected by the \( j \)th receiver coil, and \( u \in \mathbb{C}^N \) is the \( m \times n = N \) target image reshaped by vertically stacking its \( n \) columns. \( P \) is the selection matrix: a \( q \times N \) binary mask obtained by extracting from the identity matrix rows which correspond to the measured Fourier coefficients. In compressed sensing and PPI we assume \( q << N \), corresponding to significant undersampling. Sensitivity maps are diagonal matrices used to estimate the impact of a pixel of the target image on the measured Fourier coefficients. Each diagonal element of \( S_j \) corresponds to a pixel of the target image.
(2.1) suggests the following fidelity term:

$$H(u) = \frac{1}{2} \|Au - f\|_2^2,$$

where the operator $A$ and data $f$ are defined as:

$$Au = \begin{pmatrix} PFS_1 u \\ PFS_2 u \\ \vdots \\ PFS_R u \end{pmatrix}, \quad f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_R \end{pmatrix}$$

where $R$ is the number of coils. Minimization of (2.2) corresponds to finding the best fit solution between the model $A$ and data $f$. In many applications including PPI and compressive sensing, we assume (2.3) is highly underdetermined, i.e. $q << N$. In PPI, $q$ corresponds to the extent of undersampling in an MR scan.

### 2.2 TGV

Total variation (TV) is a widely used regularizer in mathematical image processing. For $\ell_1$ minimization it is formulated as

$$\|u\|_{TV} \equiv \|D_1 u\|_1 + \|D_2 u\|_1,$$

where $D_j$ is the forward finite-difference matrix along the $j$th dimension of $u$. By minimizing a sum of this quantity and an appropriate fidelity term such as (2.2), sharp edges and piecewise constant intensity patterns are efficiently recovered from the target image. However, TV is notorious for producing oil-painting (or stair-casing) artifacts, especially if the underlying signal has smooth details. Many efforts have been made to improve the performance of TV [16–18, 20]. In particular, TGV, a generalization of TV, has been proposed [6]. With order greater than or equal to two, it takes into account higher order partial derivatives of the image and efficiently avoids oil painting artifacts. Figure 1, extracted from [6], displays a comparison between TV and TGV on a denoising problem. The TGV algorithm results in a much smoother recovery.

TGV of order $k$ with positive weights $\alpha = (\alpha_0, \ldots, \alpha_{k-1})$ is defined as

$$\text{TGV}^{k}_{\alpha}(u) = \sup \left\{ \int_{\Omega} u \text{div}^k v \, dv \, \left| \, v \in C_c^k(\Omega, \text{Sym}^k(\mathbb{R}^d)), \|\text{div}^j v\|_\infty \leq \alpha_j, \, j = 0, \ldots, k - 1 \right\},$$

Figure 1: Image smoothing results of TV and TGV. The TGV result has sharper edges and less oil painting artifacts. Figure is extracted from [6].
where $C^k_\ell(\Omega, \text{Sym}^k(\mathbb{R}^d))$ is the space of compactly supported symmetric tensor fields and $\text{Sym}^k(\mathbb{R}^d)$ is the space of symmetric tensors on $\mathbb{R}^d$. When $k, \alpha = 1$, and $\text{Sym}^1(\mathbb{R}^d) = \mathbb{R}^d$, $\text{TGV}_1$ is identical to TV. The improvement in image reconstruction quality using $k$ minimization, extracted from [6,13]. For notational convenience, let

$$TGV_\alpha^2(u) = \sup \left\{ \int_\Omega u \text{div}^2 w \, dx \mid w \in C^2_c(\Omega, S^{d \times d}), \|w\|_\infty \leq \alpha_0, \|\text{div} w\|_\infty \leq \alpha_1 \right\}, \quad (2.6)$$

where the divergences are defined as $(\text{div} w)_h = \sum_{j=1}^d \frac{\partial w_{hj}}{\partial x_j}, 1 \leq h \leq d, \text{div}^2 w = \sum_{h,j=1}^d \frac{\partial^2 w_{hj}}{\partial x_h \partial x_j}$, and the infinity norms are given by:

$$\|w\|_\infty = \sup_{l \in \Omega} \left( \sum_{h,j=1}^d |w_{hj}(l)|^2 \right)^{1/2}, \quad \|\text{div} w\|_\infty = \sup_{l \in \Omega} \left( \sum_{j=1}^d |(\text{div} w)_j(l)|^2 \right)^{1/2}.$$ We define the space of bounded generalized variation to illustrate that image reconstruction with TGV leads to piecewise polynomial intensities:

$$\text{BGV}^k(\Omega) = \left\{ u \in L^1(\Omega) \mid TGV_\alpha^k(u) < \infty \right\}, \quad \|u\|_{\text{BGV}^k} = \|u\|_1 + TGV_\alpha^k(u)$$

$\text{BGV}^k(\Omega)$ is a Banach space independent of the weight vector $\alpha$. Note that $TGV_\alpha^k$ is a semi-norm that is zero for all polynomials of degree up to $k - 1$. TGV is convex, making its addition to the minimization objective function computationally feasible. We refer to [6] for further theoretical details and comparisons.

To be efficiently implemented in our algorithm, we present a reformulation of TGV in terms of $\ell_1$ minimization, extracted from [6,13]. For notational convenience, let

$$U = C^2_c(\Omega, \mathbb{R}), \quad V = C^2_c(\Omega, \mathbb{R}^2), \quad W = C^2_c(\Omega, S^{2 \times 2}),$$

and $v = \text{div} w$. Then the discretized $TGV_\alpha^2$ of $u \in U$ can be written as

$$TGV_\alpha^2(u) = \max_{v \in V, w \in W} \{ (u, \text{div} v) | \text{div} w = v, \|w\|_\infty \leq \alpha_0, \|v\|_\infty \leq \alpha_1 \}$$

with

$$\text{div} w = \begin{bmatrix} \partial_x w_{11} + \partial_y w_{12} \\ \partial_x w_{21} + \partial_y w_{22} \end{bmatrix}.$$ We introduce the indicator functional of a closed set $B$:

$$\mathcal{I}_B = \begin{cases} 0, & x \in B \\ \infty, & \text{else} \end{cases}$$

and recall the fact $\mathcal{I}(\{0\})(\cdot) = -\min_y \langle y, \cdot \rangle$. For further notational convenience let $O$ denote the maximization conditions $\|w\|_\infty \leq \alpha_0, w \in W, \|v\|_\infty \leq \alpha_1, v \in V$. Then the discrete $TGV^2$ can be repre-
The posed model is given by:

\[
TGV_\alpha^2(u) = \min_{p \in V} \max_{\|w\|_\infty \leq \alpha_0, w \in W} \langle u, \text{div} v \rangle + \langle p, v - \text{div} w \rangle
\]

Combining the fidelity (2.2) and regularization (2.7) terms justified in the previous section, our proposed model is thus

\[
\text{3.1 Model}
\]

Thus, the continuous version of our regularization function is defined as

\[
\phi(u) = TGV_\alpha^2(u) = \alpha_1 \|\nabla u - p\|_1 + \alpha_0 \|\bar{\mathcal{E}}(p)\|_1
\]

Here we must assume that \(u, p\) are absolutely continuous. The interchangeability of the max / min in the first equation can be proved by applying the sufficient conditions for the max-min equality in [4]. In the third equation, we exploit the symmetry property of \(\|v\|_\infty \leq \alpha_1\) about zero by replacing \(v\) with \(-v\). The operators \(\nabla\) and \(\bar{\mathcal{E}}(p)\) are given by:

\[
\nabla : U \to V, \quad \nabla u = \begin{bmatrix} \partial_x u \\ \partial_y u \end{bmatrix}
\]

\[
\bar{\mathcal{E}} : V \to W, \quad \bar{\mathcal{E}}(p) = \begin{bmatrix} \partial_x p_1 \\ \frac{1}{2} (\partial_y p_1 + \partial_z p_2) \\ \frac{1}{2} (\partial_y p_1 + \partial_z p_2) \\ \partial_y p_2 \end{bmatrix}.
\]

Thus, the continuous version of our regularization function is defined as

\[
\phi(u) = TGV_\alpha^2(u) = \alpha_1 \|\nabla u - p\|_1 + \alpha_0 \|\bar{\mathcal{E}}(p)\|_1
\] (2.7)

\[
\text{3 Proposed Model and Algorithm}
\]

\[
\text{3.1 Model}
\]

Combining the fidelity (2.2) and regularization (2.7) terms justified in the previous section, our proposed model is given by:

\[
(u^*, p^*) = \arg \min_{u, p} \|Au - f\|_2^2 + \alpha_1 \|\nabla u - p\|_1 + \alpha_0 \|\bar{\mathcal{E}}(p)\|_1
\] (3.1)

After discretization, the problem can be efficiently solved by Alternating Direction Method of Multipliers (ADMM). ADMM is a common approach that iteratively minimizes a problem’s augmented Lagrangian one variable at a time [10,11]. In our case, \(u\) and \(p\) are solved together in a \((u, p)\) subproblem.

To discretize the problem, we approximate \(\nabla u\) by \(Du\), where \(Du = D_1u + D_2u\) and \(D_j\) is the forward difference operator in the \(j\)th dimension. Then \(\bar{\mathcal{E}}(p)\) can be approximated by

\[
\bar{\mathcal{E}}(p) = \begin{bmatrix} D_1 p_1 \\ \frac{1}{2} (D_2 p_1 + D_1 p_2) \\ \frac{1}{2} (D_2 p_1 + D_1 p_2) \\ D_2 p_2 \end{bmatrix}.
\]

The discretized version of our proposed model is thus

\[
(u^*, p^*) = \arg \min_{u, p} \|Au - f\|_2^2 + \alpha_1 \|Du - p\|_1 + \alpha_0 \|\bar{\mathcal{E}}(p)\|_1
\] (3.2)
and we now focus on deriving its solution.

Starting with Bregman operator splitting [12], we introduce two auxiliary variables and change the problem into a constrained minimization:

$$\min_{u,p,y,z} \|Au - f\|_2^2 + \alpha_1\|y\|_1 + \alpha_0\|z\|_1$$
subject to $y = Du - p$, $z = \mathcal{E}(p)$, \hspace{1cm} (3.3)

where $\|y\|_1$ is the sum of $\ell_2$-norms of $2 \times 1$ vectors and $\|z\|_1$ is the sum of the Frobenius norms of $2 \times 2$ matrices. The augmented Lagrangian associated with (3.3) is then:

$$\mathcal{L}(u,p,y,z,b_y,b_z) = \|Au - f\|_2^2 + \alpha_1\|y\|_1 + \alpha_0\|z\|_1 + \frac{\alpha_1\mu_1}{2}\|y - (Du - p) - b_y\|_2^2 + \frac{\alpha_0\mu_0}{2}\|z - \mathcal{E}(p) - b_z\|_2^2$$ \hspace{1cm} (3.4)

where $b_y, b_z$ are Lagrange multipliers. If $(b_y^{(k)}, b_z^{(k)})$ is an approximation to $(b_y, b_z)$ at the $k^{th}$ iteration, then the gradient version of the method of multipliers iteratively minimizes $\mathcal{L}(u,p,y,z,b_y,b_z)$ with respect to $(u,p,y,z)$ and updates each multiplier as follows:

$$\begin{align*}
(u^{k+1}, p^{k+1}, y^{k+1}, z^{k+1}) &= \arg\min_{u,p,y,z} \mathcal{L}(u,p,y,z,b_y^{(k)}, b_z^{(k)}) \\
b_y^{k+1} &= b_y^{(k)} + \gamma((Du^{k+1} - p^{k+1} - y^{k+1}) \\
b_z^{k+1} &= b_z^{(k)} + \gamma(\mathcal{E}(y^{k+1}) - z^{k+1})
\end{align*}$$

We now introduce two modifications to the $(u,p,y,z)$ minimization to ensure uniqueness and increase convergence speed as suggested by [8]. First we insert a proximal term for each auxiliary variable:

$$\begin{align*}
(u^{k+1}, p^{k+1}, y^{k+1}, z^{k+1}) &= \arg\min_{u,p,y,z} \|Au - f\|_2^2 \\
&\hspace{1cm} + \alpha_1\|y\|_1 + \frac{\alpha_1\mu_1}{2}\|y - (Du - p) - b_y\|_2^2 \\
&\hspace{1cm} + \frac{\alpha_0\mu_0}{2}\|z - \mathcal{E}(p) - b_z\|_2^2
\end{align*}$$ \hspace{1cm} (3.5)

where $\beta_1, \beta_0 > 0$. The second modification is in anticipation of a highly ill-conditioned inversion matrix coming from the term

$$H(u) = \frac{1}{2}\|Au - f\|_2^2.$$ \hspace{1cm} (3.6)

The expansion of $H$ in a Taylor series around $u^k$ can be expressed as

$$H(u) = H(u^k) + (u - u^k)^T \nabla H(u^k) + \frac{1}{2}(u - u^k)^T \nabla^2 H(u^k)(u - u^k),$$ \hspace{1cm} (3.7)

where $\nabla H(u^k) = A^T(Au^k - f)$ and $\nabla^2 H(u^k) = A^TA$. In PPI and other applications $A^TA$ may be large and dense. We make the approximation $A^TA \approx \delta_k\|u\|_2^2$:

$$H(u) \approx \frac{1}{2}\|Au - f\|_2^2 + (u - u^k)^T A^T(Au^k - f) + \frac{\delta_k}{2}\|u - u^k\|_2^2$$

(complete the square:) \hspace{1cm} $= \frac{1}{2}\|Au - f\|_2^2 - \frac{1}{2\delta_k}\|A^T(Au^k - f)\|_2^2 + \frac{\delta_k}{2}\|u - u^k + \delta_k^{-1}A^T(Au^k - f)\|_2^2$. \hspace{1cm} (3.8)

The last term of (3.8) is the only one relevant to minimization since $u_k$ is constant at each iteration. Plugging this into (3.5), we obtain

$$\begin{align*}
(u^{k+1}, p^{k+1}, y^{k+1}, z^{k+1}) &= \arg\min_{u,p,y,z} \delta_k \frac{1}{2}\|u - u^k + \delta_k^{-1}A^T(Au^k - f)\|_2^2 \\
&\hspace{1cm} + \alpha_1\|y\|_1 + \frac{\alpha_1\mu_1}{2}\|y - (Du - p) - b_y\|_2^2 \\
&\hspace{1cm} + \frac{\alpha_0\mu_0}{2}\|z - \mathcal{E}(p) - b_z\|_2^2
\end{align*}$$ \hspace{1cm} (3.9)
In this paper we rely on a $\delta_k$ that is chosen through a line search. The details will be reviewed after further simplification of the minimization.

### 3.2 Algorithm

Applying ADMM to the problem yields:

$$
\begin{aligned}
\begin{cases}
    y^{k+1} &= \arg \min_y \|y\|_1 + \frac{\alpha}{2} \|y - (Du^k - p^k) - b^k_y\|^2_2 + \frac{\beta_1}{2\alpha_1} \|y - y^k\|^2_2 \\
    z^{k+1} &= \arg \min_z \|z\|_1 + \frac{\alpha}{2} \|z - \mathcal{E}(p^k) - b^k_z\|^2_2 + \frac{\beta_0}{2\alpha_0} \|z - z^k\|^2_2 \\
    (u^{k+1}, p^{k+1}) &= \arg \min_{u, p} \frac{\alpha}{2} \|u - u^k + \delta_k^{-1} A^T (Au^k - f)\|^2_2 + \frac{\alpha \mu_k}{2} \|y^{k+1} - (Du - p) - b^k_y\|^2_2 \\
    b_y^{k+1} &= b_y^k + \gamma(Du^{k+1} - p^{k+1} - y^{k+1}) \\
    b_z^{k+1} &= b_z^k + \gamma(\mathcal{E}(p^{k+1}) - z^{k+1})
\end{cases}
\end{aligned}
$$

(3.10)

Some convergence proofs for the ADMM and its variants can be found in [2,5,9,15]. Since $y$ and $z$ are component-wise separable, the first two subproblems have explicit solutions given by soft thresholding (also known as shrinkage). For completeness, we present the derivation of a shrinkage operator that includes proximal terms added in (3.5).

Let $w$ be a dummy variable to represent a component of either $y$ or $z$ with $\rho, \beta > 0$. Then the goal is to determine $w^{k+1} = \arg \min_w \Psi_k(w)$, where

$$
\Psi_k(w) = \|w\|_1 + \frac{\rho}{2} \|w - B - b_w\|_2 + \frac{\beta}{2} \|w - w^k\|^2_2.
$$

The first order necessary condition is $\frac{d}{dw} \Psi_k(w) = 0$. There are two cases to consider:

**Case 1: $w \neq 0$**

$$
\frac{d}{dw} \Psi_k(w) = 0 = sgn(w) + \rho (w - B - b_w) + \beta (w - w^k)
$$

$$
\Rightarrow w = -\frac{sgn(w)}{\rho + \beta} + \frac{\rho}{\rho + \beta} (B + b_w) + \frac{\beta}{\rho + \beta} w^k
$$

$$
\Rightarrow \begin{cases}
    w < 0 : \frac{\rho}{\rho + \beta} (B + b_w) + \frac{\beta}{\rho + \beta} w^k < -\frac{1}{\rho + \beta} \\
    w > 0 : \frac{\rho}{\rho + \beta} (B + b_w) + \frac{\beta}{\rho + \beta} w^k > \frac{1}{\rho + \beta}
\end{cases}
$$

Let $S = \frac{\rho}{\rho + \beta} (B + b_w) + \frac{\beta}{\rho + \beta} w^k$. Then

$$
\Rightarrow w^{k+1} = \begin{cases}
    S + \frac{1}{\rho + \beta} , \text{ if } S < -\frac{1}{\rho + \beta} S + \frac{1}{\rho + \beta} , \text{ if } S > \frac{1}{\rho + \beta}
\end{cases}
$$

**Case 2: $w = 0$.** We use the subdifferential $\frac{d}{dw} \|w\|_1 \equiv [-1, 1]$, and hence require that

$$
0 \in [-1, 1] - \rho (B + b_w) - \beta w^k
$$

$$
\Rightarrow \rho (B + b_w) + \beta w^k \in [-1, 1]
$$

$$
\Rightarrow w^{k+1} = 0 \text{ if } |S| < \frac{1}{\rho + \beta}
$$

Together, we have

$$
w^{k+1} = sgn(S) \max\{|S| - \frac{1}{\rho + \beta}, 0\}.
$$
Now we can define for the $y$-subproblem:

$$y^{k+1}(l) = \text{shrink}_2 \left( \left[ \begin{array}{c} \mu_1 \\ \mu_1 + \frac{\mu_1}{\alpha_1} \end{array} \right] (D_{u}^{k}(l) - p^{k}(l) + b_{y}^{k}(l)) + \left[ \begin{array}{c} \beta_1 \\ \mu_1 + \frac{\beta_1}{\alpha_1} \end{array} \right] y^{k}, \frac{1}{\mu_1 + \frac{\beta_1}{\alpha_1}} \right),$$

(3.11)

where $y^{k+1}(l)$ corresponds to pixel $l$ and

$$\text{shrink}_2(a, b) = \begin{cases} 0, & a = 0 \\ \frac{\alpha}{\|a\|_2} - \frac{a}{\|a\|_2} - b, & a \neq 0 \end{cases}.$$

Likewise, for the $z$-subproblem we have:

$$z^{k+1}(l) = \text{shrink}_F \left( \left[ \begin{array}{c} \mu_0 \\ \mu_0 + \frac{\beta_0}{\alpha_0} \end{array} \right] (E(p)(l) + b_{z}^{k}(l)) + \left[ \begin{array}{c} \beta_0 \\ \mu_0 + \frac{\beta_0}{\alpha_0} \end{array} \right] z^{k}, \frac{1}{\mu_0 + \frac{\beta_0}{\alpha_0}} \right),$$

(3.13)

where $z^{k+1}(l)$ corresponds to pixel $l$ and

$$\text{shrink}_F(a, b) = \begin{cases} 0, & a = 0 \\ \frac{\alpha}{\|a\|_F} - \frac{a}{\|a\|_F} - b, & a \neq 0 \end{cases}.$$

Note that 0 here is a $2 \times 2$ matrix. In our implementation, because of the complex relationship between coefficients, we multiply each threshold (the second inputs to shrink2 and shrink_F) by an experimentally determined scalar $w$ to easily control the size of our solutions.

We will now derive a solution for the $(u, p)$ subproblem, assuming an appropriate $\delta_k$ has been determined. The line-search procedure used to choose $\delta_k$ is described after the closed-form expression for $(u^{k+1}, p^{k+1})$ is presented. The derivation is largely extracted from [13]. The first-order necessary conditions for optimality are found by setting the gradient of the $(u, p)$ subproblem in (3.10) to zero:

$$\begin{cases} 0 = \alpha_1 \mu_1 \sum_{j=1}^{2} D_j^T(D_j u - p_j - y_j^{k+1} + b_{y}^{k+1}) + \delta_k (u - u^k + \delta_k A^T(Au^k - f)) \\ 0 = \alpha_1 \mu_1 (p_1 - D_1 u + y_1^{k+1} - b_{y}^{k+1}) + \alpha_0 \mu_0 (D_1^T(D_1 p_1 - z_1^{k+1} + b_{z}^{k+1}) + \frac{1}{2} D_2^T (D_2 p_1 + D_1 p_2 - 2 z_3^{k+1} + 2 b_{z}^{k+1}) \\ 0 = \alpha_1 \mu_1 (p_2 - D_2 u + y_2^{k+1} - b_{y}^{k+1}) + \alpha_0 \mu_0 (D_2^T(D_2 p_2 - z_2^{k+1} + b_{z}^{k+1}) + \frac{1}{2} D_1^T (D_1 p_2 + D_2 p_1 - 2 z_3^{k+1} + 2 b_{z}^{k+1}) \\ 0 = \alpha_1 \mu_1 \sum_{j=1}^{2} D_j^T(D_j y_j - b_{y}^{k}) + \delta_k u - A^T(Au - f) \\ 0 = \alpha_1 \mu_1 (b_{y}^{k-1} - y_1^{k}) + \alpha_0 \mu_0 (D_1^T(D_1 z_1 - b_{z}^{k}) + \frac{1}{2} D_2^T (2 z_3^{k} - 2 b_{z}^{k})) \\ 0 = \alpha_1 \mu_1 (b_{y}^{k-1} - y_2^{k}) + \alpha_0 \mu_0 (D_2^T(D_2 z_2 - b_{z}^{k}) + \frac{1}{2} D_1^T (2 z_3^{k} - 2 b_{z}^{k})) \\ \end{cases}$$

(3.15)

We can diagonalize the coefficient matrix arising from this linear system block-wise under the Fourier transform, by the fact that all circulant matrices can be diagonalized under the Fourier transform. Thus, $F D_j F^*$ and $F D_j^T D_k F^*$ are diagonal for $j, k = 1, 2$, where $F^*$ is the inverse Fourier transform.

We group $u$, $p_1$, and $p_2$ terms on the left-hand sides of each equation in (3.15) to obtain the following linear system:

$$\begin{bmatrix} d_1 & d_4 & d_7 \\ d_2 & d_5 & d_8 \\ d_3 & d_6 & d_9 \end{bmatrix} \begin{bmatrix} u \\ p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix}.$$

(3.16)

Each element $d_i$, $B_i$ is a matrix as defined below.

$$\begin{cases} d_1 = \alpha_1 \mu_1 \sum_{j=1}^{2} D_j^T D_j + \delta_k I \\ d_2 = \alpha_1 \mu_1 \mu_0 D_1^T D_1 + \frac{1}{2} D_2^T D_2 \\ d_3 = \alpha_1 \mu_1 \mu_0 D_2^T D_2 + \frac{1}{2} D_1^T D_1 \\ d_4 = -\alpha_1 \mu_1 D_1 \\ d_5 = -\alpha_1 \mu_1 D_2 \\ d_6 = \frac{1}{2} D_1^T D_2 \\ B_1 = \alpha_1 \mu_1 \sum_{j=1}^{2} D_j^T (y_j^{k+1} - b_{y}^{k}) + \delta_k u - A^T(Au - f) \\ B_2 = \alpha_1 \mu_1 (b_{y}^{k-1} - y_1^{k}) + \alpha_0 \mu_0 D_1^T (z_1^{k+1} - b_{z}^{k}) + \frac{1}{2} D_2^T (2 z_3^{k+1} - 2 b_{z}^{k}) \\ B_3 = \alpha_1 \mu_1 (b_{y}^{k-1} - y_2^{k}) + \alpha_0 \mu_0 D_2^T (z_2^{k+1} - b_{z}^{k}) + \frac{1}{2} D_1^T (2 z_3^{k+1} - 2 b_{z}^{k}) \end{cases}$$
We now apply a preconditioner matrix to (3.16) to achieve diagonalization:

\[
\begin{bmatrix}
F & 0 & 0 \\
0 & F & 0 \\
0 & 0 & F
\end{bmatrix}
\begin{bmatrix}
d_1 & d_4^* & d_5^* \\
d_4 & d_2 & d_6 \\
d_5 & d_6 & d_3
\end{bmatrix}
\begin{bmatrix}
F^* & 0 & 0 \\
0 & F^* \\
0 & 0 & F^*
\end{bmatrix}
\begin{bmatrix}
F \eta_u \\
F \eta_{p_1} \\
F \eta_{p_2}
\end{bmatrix}
= 
\begin{bmatrix}
F & 0 & 0 \\
0 & F & 0 \\
0 & 0 & F
\end{bmatrix}
\begin{bmatrix}
B_1 \\
B_2 \\
B_3
\end{bmatrix}
\]  
(3.17)

Denote \( \tilde{a}_j = \text{diag}(Fd_jF^*) \), \( \tilde{d}_j^* = \text{diag}(Fd_j^*F^*) = \text{conj}(\text{diag}(Fd_jF^*)) \), and let .\( ^* \) denote component-wise multiplication. Then we can write (3.17) as:

\[
\begin{bmatrix}
\tilde{d}_1^*(Fu) + \tilde{d}_4^*(Fp_1) + \tilde{d}_5^*(Fp_2) \\
\tilde{d}_2^*(Fu) + \tilde{d}_4^*(Fp_1) + \tilde{d}_5^*(Fp_2) \\
\tilde{d}_3^*(Fu) + \tilde{d}_4^*(Fp_1) + \tilde{d}_5^*(Fp_2)
\end{bmatrix}
= 
\begin{bmatrix}
FB_1 \\
FB_2 \\
FB_3
\end{bmatrix}
\]

Now \( Fu, Fp_1, \) and \( Fp_2 \) can be obtained via Cramer’s Rule, giving us the following closed forms:

\[
\begin{align*}
\{ u & = F^* \left( \begin{array}{ccc}
FB_1 & \tilde{d}_1^* & \tilde{d}_5^* \\
FB_2 & \tilde{d}_2^* & \tilde{d}_5^* \\
FB_3 & \tilde{d}_6 & \tilde{d}_5^*
\end{array} \right) / \text{denom} \\
p_1 & = F^* \left( \begin{array}{ccc}
\tilde{d}_1 & FB_1 & \tilde{d}_5^* \\
\tilde{d}_4 & FB_2 & \tilde{d}_5^* \\
\tilde{d}_5 & FB_3 & \tilde{d}_5^*
\end{array} \right) / \text{denom} \\
p_2 & = F^* \left( \begin{array}{ccc}
\tilde{d}_1 & \tilde{d}_2 & FB_1 \\
\tilde{d}_4 & \tilde{d}_2 & FB_2 \\
\tilde{d}_5 & \tilde{d}_6 & FB_3
\end{array} \right) / \text{denom}
\end{align*}
\]  
(3.18)

where division is component wise and

\[
\text{denom} = \begin{vmatrix}
\tilde{d}_1 & \tilde{d}_4^* & \tilde{d}_5^* \\
\tilde{d}_4 & \tilde{d}_2 & \tilde{d}_5^* \\
\tilde{d}_5 & \tilde{d}_6 & \tilde{d}_5^*
\end{vmatrix}.
\]

Here, \(| \cdot |_\ast \) is defined to be

\[
\begin{vmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33}
\end{vmatrix}_\ast
= a_{11} \cdot a_{22} \cdot a_{33} + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32} - a_{12} \cdot a_{22} \cdot a_{33} - a_{13} \cdot a_{21} \cdot a_{32} - a_{23} \cdot a_{31} \cdot a_{13}
\]

where \( a_{ij} \in \mathbb{R}^n \).

Using this closed form expression, we can efficiently solve the system multiple times in an inner-loop for the determination of an appropriate \( \delta_k \) using a line search, which we will now describe in detail. At each outer-loop iteration, we set \( \delta_k = \eta^j \delta_{0,k} \), where \( j \geq 0 \) is the smallest integer such that \( Q_{k+1} \geq \frac{C}{\varepsilon^2} \), with the following definitions:

\[
\begin{align*}
\delta_{0,k} & = \max \left\{ \delta_{\min}, \frac{\| A(u^{k}-u^{k-1}) \|^2}{\| u^{k}-u^{k-1} \|^2} \right\} \\
Q_{k+1} & = \eta_k Q_k + \Delta_k \\
\eta_k & \in [0, (1 - k^{-2})^2], \text{ for } k > 1 \\
(u^{k+1}, p^{k+1}) & \text{ are given by (3.18)} \\
\Delta_k & = \sigma(\delta_k \| u^{k+1} - u^k \|^2) + \alpha_1 \mu_1 \| Du^{k+1} - p^{k+1} - y^{k+1} \|^2 + \alpha_0 \mu_0 \| \mathcal{E}(p^{k+1}) - z^{k+1} \|^2 - \| A(u^{k+1} - u^k) \|^2
\end{align*}
\]  
(3.19)

Notice that our initial guess for \( \delta_k (\delta_{0,k}) \) is the safeguarded Barzilai-Borwein (BB) approximation to the Hessian [1]. The stepsize rule above is related to but different from the rule found in the monotone
SpaRSA algorithm (Sparse Reconstruction by Separable Approximation), introduced in [14]. This work was expanded in [21], where it is shown that there exists a sufficiently large $\delta_k$ to decrease the objective function at each iteration $k$. The stepsize rule was adjusted for a problem similar to ours in [8], by Chen et al. They proved that $\delta_k$ is sufficient when

$$
\|A(u^{k+1} - u^k)\|^2 \leq \sigma \left( \delta_k \|u^{k+1} - u^k\|^2 + \rho \|Bu^{k+1} - u^k\|^2 \right) + \eta_k Q_k + \frac{C}{k^2} \tag{3.20}
$$

where $B$ is a linear operator (such as $D$) and $w$ is an auxiliary variable (such as $y$ or $z$). TV is exploited in [8]. Chen et al. refer to this algorithm as Bregman Operator Splitting with Variable Stepsize (BOSVS). To combine the BOSVS approach with TGV, we propose the following natural extension of (3.20):

$$
\|A(u^{k+1} - u^k)\|^2 \leq \sigma \left( \delta_k \|u^{k+1} - u^k\|^2 + \eta_1 \|Du^{k+1} - p^{k+1} - y^k\|^2 + \mu_0 \|E(p^{k+1}) - z^k\|^2 \right) + \eta_k Q_k + \frac{C}{k^2}. \tag{3.21}
$$

Although we have not independently proved convergence of this scheme, we found that BOSVS with TGV experimentally produces comparable or better results than BOSVS with TV in PPI reconstruction, particularly in very undersampled conditions. Convergence of BOSVS is contingent on the asymptotic monotonicity of $\delta_k$. Thus we have a penalty step in the algorithm, which increases $\delta_{\min}$ by a factor $\tau > 1$ whenever $\delta_k$ is not monotone decreasing. Hence, if the monotonicity of $\delta_k$ continues to be violated, then we eventually have

$$
\delta_{\min} > \frac{\|A\|^2}{\sigma} \geq \frac{\|A(u^k - u^{k-1})\|^2}{\sigma \|u^k - u^{k-1}\|^2} \geq \frac{\frac{1}{2} \|A(u^k - u^{k-1})\|^2 - \alpha_1 \mu_1 \|Du^{k+1} - p^{k+1} - y^k\|^2 - \alpha_0 \mu_0 \|E(p^{k+1}) - z^k\|^2}{\|u^k - u^{k-1}\|^2},
$$

and so $\delta_{0,k} = \delta_{\min}$ whenever $\delta_{\min} > \|A\|^2$, in which case the line search stops with $j = 0$. This can be seen by observing that (3.22) implies that $\Delta_k \geq 0$, and thus

$$
Q_{k+1} \geq \eta_k Q_k + \Delta_k \geq \eta_k Q_k \geq \frac{-\eta_k C}{(k-1)^2} \geq \frac{-C}{k^2}
$$

since $\eta_k \leq \frac{(k-1)^2}{C}$, and thus the stepsize rule is satisfied. In this case, the convergence analysis from [23] applies. The entirety of BOSVS with TGV is summarized in Algorithm 1.
Algorithm 1 BOSVS with TGV

Given \( \tau, C, \eta, \gamma > 1 \), \( \alpha_j, \mu_j, \beta_j > 0 \) for \( j = \{0, 1\} \), \( \delta_{\min} > 0 \), \( \sigma \in (0, 1) \), as well as initial guesses for \( \delta_0, u^1, p^1, y^1, x^1, b^1_x, b^1_y \). Set \( k = 1 \), \( Q_1 = 0 \).

**STEP 1:** \( y^{k+1} = \arg\min_y \|y\|_1 + \frac{\mu_1}{2} \| y - (Du^k - p^k) - b^k_y \|_2^2 + \frac{\beta_1}{2\sigma_0} \| y - y^k \|_2^2 \)

**STEP 2:** \( z^{k+1} = \arg\min_z \|z\|_1 + \frac{\mu_0}{2} \| z - E(p^k) - b^k_z \|_2^2 + \frac{\beta_0}{2\sigma_0} \| z - z^k \|_2^2 \)

**STEP 3.1:** Set \( \delta_k = \eta^j \delta_{0,k} \) according to (3.19)

**STEP 3.2:** \( (u^{k+1}, p^{k+1}) = \arg\min_{u,p} \frac{\mu}{2} \| u - u^k + \delta_k^{-1} A^T (Au^k - f) \|_2^2 + \frac{\alpha_1 \mu_1}{2} \| y^{k+1} - (Du - p) - b^k_y \|_2^2 \)

\( + \frac{\alpha_0 \mu_0}{2} \| z^{k+1} - E(p^k) - b^k_z \|_2^2 \) (given by (3.18))

**STEP 3.3:** If \( \delta_k > \delta_{k-1} \), replace \( \delta_{\min} = \delta_{\min} \eta \)

**STEP 4:** \( b^{k+1}_y = b^k_y + \gamma (Du^{k+1} - p^{k+1} - y^{k+1}) \)

**STEP 5:** \( b^{k+1}_z = b^k_z + \gamma (E(p^{k+1}) - z^{k+1}) \)

**STEP 6:** If a stopping criterion is satisfied, STOP. Else, set \( k = k + 1 \) and go back to **STEP 1**

4 Numerical Results and Comparisons

We tested Algorithm 1 against its equivalent formulation that uses the TV regularizer [8]. We obtained a set of radial PPI sensitivity maps from Nicole Seiberlich (Biomedical Engineering Department, Case Western Reserve University), corresponding to a 30-coil PPI array. Raw PPI data was simulated for each phantom image \( I \in \mathbb{R}^{128^2} \) as follows:

\[
f_j = P F S_j I, \tag{4.1}
\]

where \( S_j \) is again the diagonal sensitivity matrix for the \( j^{\text{th}} \) coil. In our dataset, \( S_j \in \mathbb{C}^{128^2 \times 128^2} \), Radial selection matrices \( P \) were generated to simulate various levels of undersampling. Since \( P \) and \( S_j \) are diagonal matrices, and because each subproblem is either separable or diagonalizable, our implementation avoids the creation of \( 128^2 \times 128^2 \) matrices by reordering their diagonal entries into \( 128 \times 128 \) matrices. Table 1 displays a visualization of \( P \) that corresponds to 40 radial projections. We can then rewrite (4.1) as

\[
f_j = P \otimes F \otimes S_j \otimes I, \tag{4.2}
\]

where \( \otimes \) denotes the Hadamard product (i.e. componentwise multiplication) and \( f_j, S_j, P, F \) and \( I \) are each \( 128 \times 128 \) matrices. The rows of the identity matrix that were not included in \( P \) are now represented as zeros.

We tested the algorithm on series of images including synthetically shaded shapes, real MR images, and a photograph. Our primary measure of accuracy was relative error with the ground truth, given by

\[
\text{relative error} = \frac{\| u_{\text{reconstructed}} - u_{\text{ground truth}} \|_2}{\| u_{\text{ground truth}} \|_2} \tag{4.3}
\]

Note that the shape of our target image must match the dimension of \( S_j \), which is \( 128 \times 128 \).

Implementation of our algorithm includes the selection of many parameters, whose optimal values depend on the problem of interest. Our chosen parameter values were determined through brute-force testing on just one image (the foot MRI) and are displayed in Table 2. A brief overview of
<table>
<thead>
<tr>
<th>Number of Radial Projections</th>
<th>Percent of Frequencies Sampled</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>7.87</td>
</tr>
<tr>
<td>16</td>
<td>14.48</td>
</tr>
<tr>
<td>32</td>
<td>26.82</td>
</tr>
<tr>
<td>40</td>
<td>32.62</td>
</tr>
<tr>
<td>70</td>
<td>51.65</td>
</tr>
<tr>
<td>90</td>
<td>63.06</td>
</tr>
</tbody>
</table>

Table 1: Undersampling Ratios for 128x128 target image, and an example of a $P$ corresponding to 40 radial projections (after reshaping by stacking its columns).
the parameters’ interpretations is now provided, and their roles can also be reviewed in (3.10) and (3.19). \((\alpha_0, \alpha_1) = (10^{-3}, 10^{-3})\) comes from the definition of TGV\(_2\) (see Equation 2.6). \((\mu_0, \mu_1) = (10^{-4}, 10^{-1})\) is used to balance the weights of the Bregman penalty terms in the Lagrangian (3.4). \((\beta_1, \beta_0) = (10^{-4}, 10^{-4})\) weights the relative importance of the proximal terms inserted to the auxiliary variable subproblems. \(\delta_0 = 60\) is traditionally set to 1, but we found that the higher value results in slightly better performance. \(\eta = 1.01\) is used to scale \(\delta_{0,k}\) during the line search and \(\sigma = 0.9999\) is a proportional term in the line search criterion given in (3.21). \(w = 10\) is multiplied to the second input of our shrinkage operators, the thresholds, given in (3.11) and (3.13) in order to easily control the range of their outputs. \(\gamma = 10^{-3}\) is a parameter used to control the size of the Bregman penalty constants \(b_y\) and \(b_z\). \(\delta_{\text{min}} = 10^{-3}\) comes from the Barzilai-Borwein Hessian approximation

\[
\delta_{0,k} = \max \left\{ \delta_{\text{min}}, \frac{||A(u^k - u^{k-1})||^2}{||u^k - u^{k-1}||^2} \right\},
\]

and is the minimum value we consider in the line search at each iteration. \(\delta_{\text{min}}\) is increased by a factor of \(\tau = 1.01\) for each \(k: \delta_k > \delta_{k-1}\). This is to say that if \(\delta_k\) repeatedly fails to decrease monotonically then eventually \(\delta_k\) becomes constant and equal to \(\delta_{\text{min}}\) in all future iterations. This penalty step is proposed by [8] to ensure convergence of BOSVS using TV, using the convergence analysis of [23]. We terminated the algorithm if either of the following stopping criteria evaluated as true:

1. Subsequent iterations had a relative error of less than \(10^{-7}\).

2. 1000 iterations have been computed.

Figure 2 displays TGV vs. TV reconstructions of a smoothly shaded testing image using the 30-coil sensitivity map under 7.87% frequency sampling conditions, and Figure 3 displays a similar comparison under 14.48% sampling conditions. TGV produces significantly better reconstructions at these low sampling rates. Numerical comparisons are provided in Table 3, where we display the relative error of the reconstructed image with respect to the ground truth given by (4.3). One can observe substantially more artifacts in reconstructions from TV than in those from TGV, particularly in low-sampling conditions and with respect to smooth details. Even when TV performed reasonably well under 14.48% (30-coil) sampling, oil-painting artifacts are apparent that do not appear in the TGV reconstructions. Figure 4 displays TGV vs. TV reconstructions of a foot MRI (from www.mr-tip.com) using the 30-coil sensitivity map under 7.87% frequency sampling conditions. Artifacts from TV are particularly apparent in the smooth heel and dark tissue at the bottom of the foot, which are much less noticeable in the TGV reconstructions. TGV provides high quality solutions at very low sampling rates. Table 4 provides numerical comparisons in terms of relative error.

Next, we test the algorithms on a pancreas MRI (from www.mr-tip.com). Figure 5 displays a comparison under 7.87% sampling conditions, in which you can see that the artifacts from TV obscure the

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Accepted Value</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>((\alpha_0, \alpha_1))</td>
<td>((10^{-3}, 10^{-3}))</td>
<td>TGV Parameters</td>
</tr>
<tr>
<td>((\mu_0, \mu_1))</td>
<td>((10^{-4}, 10^{-1}))</td>
<td>Regularizer weights</td>
</tr>
<tr>
<td>((\beta_1, \beta_0))</td>
<td>((10^{-4}, 10^{-4}))</td>
<td>Proximal Term Weights</td>
</tr>
<tr>
<td>(\delta_{\text{min}})</td>
<td>(10^{-3})</td>
<td>Minimum Hessian Approx.</td>
</tr>
<tr>
<td>(\delta_0)</td>
<td>60</td>
<td>Initial Hessian Approx.</td>
</tr>
<tr>
<td>(\eta)</td>
<td>1.01</td>
<td>(\delta_k = \eta \delta_{0,k})</td>
</tr>
<tr>
<td>(\sigma)</td>
<td>0.9999</td>
<td>Line Search Factor</td>
</tr>
<tr>
<td>(\tau)</td>
<td>1.01</td>
<td>Penalty for (\delta_{\text{min}})</td>
</tr>
<tr>
<td>(w)</td>
<td>10</td>
<td>Threshold Factor</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>(10^{-3})</td>
<td>Bregman Update</td>
</tr>
</tbody>
</table>

Table 2: Parameter Selections
details on the left portion of the pancreas MRI beyond recognition, while TGV successfully recovers the smooth portions in-between the small structures. Under 14.48% sampling conditions in Figure 6, both TV and TGV do a good job at reconstructing the smooth area on the right side of the pancreas MRI. However, if you carefully examine the bottom of the close-up frame, you will see that the smooth region contains some oil-painting artifacts. While it does not ruin the image, it is similar in terms of intensity to the details on the left-bottom part of the close-up region. Especially without a side-by-side comparison with the ground-truth, it could be misinterpreted. Table 5 provides numerical comparisons in terms of relative error.

Next, we test the algorithms on an MRI of a shoulder (from www.mr-tip.com). Figures 7 and 8 show the reconstruction comparisons under 7.87% and 14.48% sampling conditions, respectively. In both cases, the TV reconstruction loses details in-between larger structures to artifacts. TGV successfully recovers these details, as well as the smoothness characterizing most of the shoulder image. Table 6 provides numerical comparisons in terms of relative error.

Finally, we test the algorithms on a photograph of a train, extracted from [6]. A comparison is displayed in Figure 9. The image has smooth regularities similar to those in the synthetic image tested above, which TGV successfully recovers. TV produces a grainy image that fails to capture the smooth gradients in this real-life image. Table 7 provides numerical comparisons in terms of relative error.

<table>
<thead>
<tr>
<th>Sampling Rate</th>
<th>TGV error</th>
<th>TV error</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.87</td>
<td>4.84 × 10^{-3}</td>
<td>3.61 × 10^{-2}</td>
</tr>
<tr>
<td>14.48</td>
<td>4.8 × 10^{-4}</td>
<td>1.61 × 10^{-2}</td>
</tr>
<tr>
<td>26.82</td>
<td>2.36 × 10^{-5}</td>
<td>5.01 × 10^{-3}</td>
</tr>
</tbody>
</table>

Table 3: Relative errors achieved by TGV and TV, respectively, under various sampling conditions when reconstructing the smoothly shaded test image.
Ground Truth:

Figure 2: Reconstructions of a synthetic test image using a 30 coil sensitivity map and 7.87% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.
30 coils, 14.48% sampling

Ground Truth:

TV:

TGV:

Figure 3: Reconstructions of a synthetic test image using a 30 coil sensitivity map and 14.48% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.
Figure 4: Reconstructions of a foot MRI using a 30 coil sensitivity map and 7.87% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.
Table 4: Relative errors achieved by TGV and TV, respectively, under various sampling conditions when reconstructing a foot MR image.

<table>
<thead>
<tr>
<th>Sampling Rate</th>
<th>TGV error</th>
<th>TV error</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.87</td>
<td>3.11 × 10^{-2}</td>
<td>9.32 × 10^{-2}</td>
</tr>
<tr>
<td>14.48</td>
<td>3.17 × 10^{-3}</td>
<td>3.99 × 10^{-2}</td>
</tr>
<tr>
<td>26.82</td>
<td>4.44 × 10^{-5}</td>
<td>8.98 × 10^{-3}</td>
</tr>
</tbody>
</table>

Table 5: Relative errors achieved by TGV and TV, respectively, under various sampling conditions when reconstructing the pancreas MR image.

<table>
<thead>
<tr>
<th>Sampling Rate</th>
<th>TGV error</th>
<th>TV error</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.87</td>
<td>4.55 × 10^{-2}</td>
<td>1.26 × 10^{-1}</td>
</tr>
<tr>
<td>14.48</td>
<td>3.69 × 10^{-3}</td>
<td>4.97 × 10^{-2}</td>
</tr>
<tr>
<td>26.82</td>
<td>4.58 × 10^{-5}</td>
<td>1.20 × 10^{-2}</td>
</tr>
</tbody>
</table>
30 coils, 7.87% sampling

Ground Truth:

TV:

TGV:

Figure 5: Reconstructions of a pancreas MRI using a 30 coil sensitivity map and 7.87% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.
30 coils, 14.48% sampling

Ground Truth:

TV:

TGV:

Figure 6: Reconstructions of a pancreas MRI using a 30 coil sensitivity map and 14.48% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.
Figure 7: Reconstructions of a shoulder MRI using a 30 coil sensitivity map and 7.87% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.
30 coils, 14.48% sampling

Ground Truth:

Figure 8: Reconstructions of a shoulder MRI using a 30 coil sensitivity map and 14.48% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.
Table 6: Relative errors achieved by TGV and TV, respectively, under various sampling conditions when reconstructing the shoulder MR image.

<table>
<thead>
<tr>
<th>Sampling Rate</th>
<th>TGV error</th>
<th>TV error</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.87</td>
<td>$1.98 \times 10^{-2}$</td>
<td>$8.50 \times 10^{-2}$</td>
</tr>
<tr>
<td>14.48</td>
<td>$1.26 \times 10^{-3}$</td>
<td>$3.30 \times 10^{-2}$</td>
</tr>
<tr>
<td>26.82</td>
<td>$3.89 \times 10^{-5}$</td>
<td>$7.69 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

Table 7: Relative errors achieved by TGV and TV, respectively, under various sampling conditions when reconstructing the photograph of a train.

<table>
<thead>
<tr>
<th>Sampling Rate</th>
<th>TGV error</th>
<th>TV error</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.87</td>
<td>$1.66 \times 10^{-2}$</td>
<td>$5.95 \times 10^{-2}$</td>
</tr>
<tr>
<td>14.48</td>
<td>$1.28 \times 10^{-3}$</td>
<td>$2.13 \times 10^{-2}$</td>
</tr>
<tr>
<td>26.82</td>
<td>$3.89 \times 10^{-5}$</td>
<td>$5.09 \times 10^{-3}$</td>
</tr>
</tbody>
</table>
30 coils, 14.48% sampling

Ground Truth:

TV:

TGV:

Figure 9: Reconstructions of a photograph of a train using a 30 coil sensitivity map and 14.48% sampling. The ground truth image is at the top, reconstructions from TV are in the middle, and those from TGV are at the bottom. Images on the right are close-up views of the regions enclosed by squares in the left images.
5 Conclusions

We proposed an algorithm based on Bregman operator splitting with variable stepsize that uses TGV as a regularizer. Since TGV represents regularities of solutions with higher order partial derivatives, it is better equipped to reconstruct images with smooth gradient intensities than TV. We used ADMM to deal with the non-differentiable terms in our model, and a line search to compute a trustworthy approximation to the ill-conditioned Hessian expected from PPI reconstruction. The proposed algorithm achieved higher quality PPI reconstructions than its TV counterpart on a synthetically shaded image, real MR images, and a real photograph. In comparison to TV, TGV performs particularly well in highly undersampled conditions (as low as 7.87%). Numerical results indicate that TGV consistently reconstructs images closer to the target solution than TV, in terms of relative error. In [13] it is shown that TGV performs better when combined with Shearlets. We focused on showing that TGV by itself is sufficient because there are additional complications when using Shearlets on non-square images.
References


