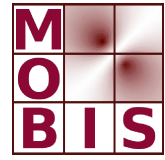




SpezialForschungsBereich F 32



Karl-Franzens Universität Graz  
Technische Universität Graz  
Medizinische Universität Graz



# A Total Variation Based Approach to Correcting Surface Coil Magnetic Resonance Images

S.L. Keeling      M. Hintermueller      F. Knoll  
D. Kraft      A. Laurain

SFB-Report No. 2010-016

April 2010

A-8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the  
Austrian Science Fund (FWF)



SFB sponsors:

- **Austrian Science Fund (FWF)**
- **University of Graz**
- **Graz University of Technology**
- **Medical University of Graz**
- **Government of Styria**
- **City of Graz**



# A Total Variation Based Approach to Correcting Surface Coil Magnetic Resonance Images

Stephen L. Keeling<sup>1</sup>, Michael Hintermüller<sup>2</sup>, Florian Knoll<sup>3</sup>,  
Daniel Kraft<sup>1</sup>, Antoine Laurain<sup>1</sup>

**Abstract.** Magnetic resonance images which are corrupted by noise and by smooth modulations are corrected using a variational formulation incorporating a total variation like penalty for the image and a high order penalty for the modulation. The optimality system is derived and numerically discretized. The cost functional used is non-convex, but it possesses a bilinear structure which allows the ambiguity among solutions to be resolved technically by regularization and practically by normalizing the maximum value of the modulation. Since the cost is convex in each single argument, convex analysis is used to formulate the optimality condition for the image in terms of a primal-dual system. To solve the optimality system, a nonlinear Gauss-Seidel outer iteration is used in which the cost is minimized with respect to one variable after the other using an inner generalized Newton iteration. Favorable computational results are shown for artificial phantoms as well as for realistic magnetic resonance images. Reported computational times demonstrate the feasibility of the approach in practice.

**Keywords:** Magnetic resonance imaging, denoising, modulation correction, bilinear residual, primal dual, total variation regularization

## 1 Introduction

Magnetic Resonance Imaging (MRI) is a medical imaging method in which radio frequency coils, such as those shown in Figure 1, are used for both nuclear excitation and for signal detection in order to measure the hydrogen atom density distribution in the human body. Displaying this distribution permits visualization of tissues with varying density. For a full discussion of the principles of MRI, see, e.g., [5]. On the one hand, a large coil such as the body coil shown in Figure 1a, may be used for the uniform illumination of a volume as shown in Figure 2a. On the other hand, a smaller surface coil such as those shown in Figure 1b can be used to resolve local details with greater sensitivity near the coil center but with an illumination falling off with the distance from the coil center as seen in Figure 2b.

By using an undersampling strategy, surface coils can also be implemented in parallel to achieve the high temporal resolution necessary for dynamic examinations. However, in this case the surface coil images are corrupted by aliasing in addition to the modulation effect seen in Figure 2. The parallel imaging problem is considered in the companion paper [18], while the present work is focused purely on the correction of surface coil images such as seen in Figure 2. The accepted model for the corruptions of such images is that the surface coil measurement  $\tilde{u}$  may be represented in terms of the underlying uncorrupted image  $u^*$  according to  $\tilde{u} = \sigma u^* + n$  where  $\sigma$  is a smooth modulation and  $n$  is noise [3].

The estimation of the smooth modulation  $\sigma$ , or bias field inhomogeneity, has been researched extensively, as can be seen in the review found in [4]. In some approaches it is assumed that, in addition to the surface image, a body coil image has also been measured, and the modulation is determined roughly from their quotient; see, e.g., the authors' previous works [17] and [19] and

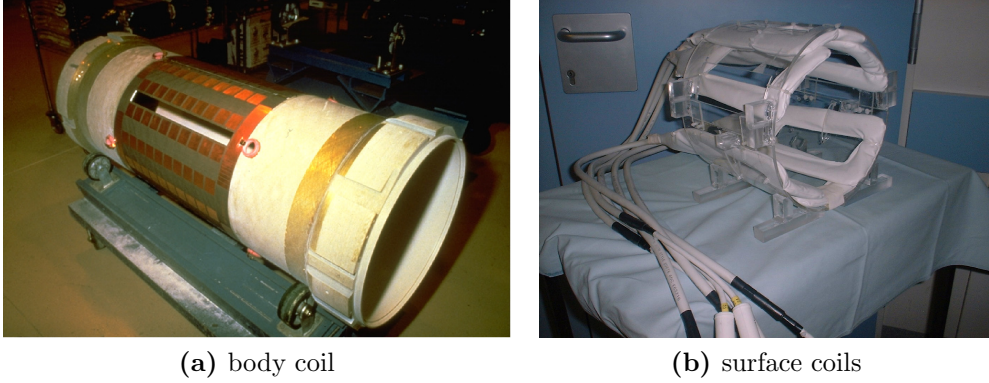
---

<sup>1</sup>Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Heinrichstraße 36, 8010 Graz, Austria; email: [firstname.lastname@uni-graz.at](mailto:firstname.lastname@uni-graz.at).

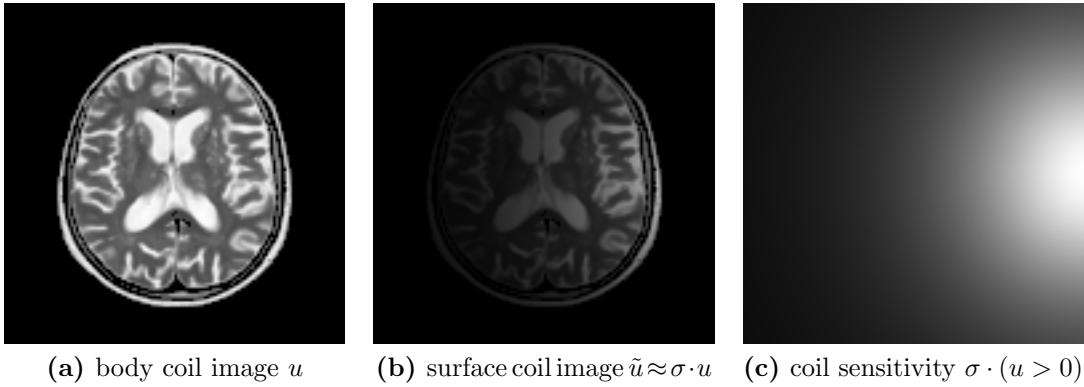
<sup>2</sup>Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany; email: [hint@mathematik.hu-berlin.de](mailto:hint@mathematik.hu-berlin.de).

<sup>3</sup>Institut für Medizintechnik, Technische Universität Graz, Kronesgasse 5/II, 8010 Graz, Austria; email: [florian.knoll@tugraz.at](mailto:florian.knoll@tugraz.at).

All authors are supported by the Austrian Science Fund *Fond zur Förderung der Wissenschaftlichen Forschung* (FWF) under grant SFB F032 ("Mathematical Optimization and Applications in Biomedical Sciences" <http://math.uni-graz.at/mobis>).



**Figure 1:** (a) Shown on the left is a magnetic resonance body coil in which a reclining patient may be situated. (Used with permission of GE Medical Systems.) (b) Shown on the right are smaller surface coils mounted on a head rack. (See [17].)



**Figure 2:** (a) Shown on the left is a body coil image  $u$  corresponding to the apparatus seen in Figure 1a. (b) Shown in the middle is a surface coil image  $\tilde{u}$  corresponding to the apparatus seen in Figure 1b. Here, the brightest point of the image is nearest to the coil center. (c) Shown on the right is the coil sensitivity  $\sigma$  associated with the middle image. Here,  $\sigma$  is displayed only on  $(u > 0)$ , the support of the first image.

the references cited therein. Other methods do not require a reference body coil image. These involve to represent the modulation in terms of smooth basis functions such as polynomials [25] or B-splines [20] and to apply a regularization approach based upon low-pass filtering [27], on statistical measures such as entropy [23] or on spatial derivatives [26]. There are also multiresolution schemes based upon wavelets [21] or pyramidal schemes [22]. Yet another class of methods involves to segment the surface coil image while estimating the modulation [24]. The present work is related to that of [24] and [26] in the sense that the regularization used here involves a segmentation of the estimated image  $u$  as well as a high-order penalty for the modulation  $\sigma$ , but simultaneously the corrected image  $u$  is estimated here using total variation regularization.

The approach proposed here is formulated variationally so that the residual  $\sigma u - \tilde{u}$  is minimized together with regularization which allows the two unknowns  $\sigma$  and  $u$  to be determined simultaneously. A second order derivative penalty is used for the modulation and a total variation like penalty is used for the reconstructed image. Specifically, the formulation for the modulation is based upon the authors' work in [17] and [19], and the total variation formulation is based upon the authors' primal-dual formulation in [13], as opposed to others such as [6] or [14]. As explained below, the computation of the modulation  $\sigma$  is improved by using a segmentation of the estimated image  $u$  to account for the edges in the data  $\tilde{u}$ . Also, the segmentation

approach used here is based on the use of topological derivatives as in the work of the authors [15].

With respect to the required minimization, note that the derivative of the residual  $|\sigma u - \tilde{u}|^2$  is bilinear in the unknowns  $\sigma$  and  $u$ . The difference in imposed regularity allows the two unknowns to be separated but leads to poor scaling in the cost functional gradient components corresponding to the modulation as opposed to the image [7]. Thus, a pure gradient descent scheme can be inefficient for the minimization. On the other hand, the non-convexity of the cost functional makes its Hessian in general indefinite, and thus a pure Newton's method is not suitable [7]. In this work a nonlinear Gauss-Seidel outer iteration is used in which the formulated cost is minimized with respect to one variable after the other using an inner generalized Newton iteration.

The paper is organized as follows. In Section 2 the variational formulation is presented and its optimality system is derived in Section 3. Then in Section 4 the numerical discretization and solution of the optimality system is presented. In Section 5, computational results are shown for both phantom images and for measured magnetic resonance images. By simulating the corruptions for which the proposed approach is designed, a measure of accuracy in the reconstruction is available. Using this measure as well as a visual demonstration, it is shown that the proposed approach performs very well to reconstruct images corrupted by noise and smooth modulations.

## 2 Overview of the Variational Approach

Given simply the image  $\tilde{u}$ , as in Figure 2b, which satisfies  $\sigma u \approx \tilde{u}$  for an image  $u$ , as in Figure 2a, and a modulation  $\sigma$ , as in Figure 2c, the desired unknowns  $u$  and  $\sigma$  are determined here by minimizing the sum of a residual term plus regularization terms:

$$J(u, \sigma) = \frac{1}{2} \int_{\Omega} |\sigma u - \tilde{u}|^2 d\mathbf{x} + \frac{\nu}{2} \int_{\Omega} |\nabla^2 \sigma|^2 d\mathbf{x} + \frac{\kappa}{2} \int_{\Omega} u^2 d\mathbf{x} + \mu \int_{\Omega} \phi_{\epsilon}(|Du|) d\mathbf{x} \quad (2.1)$$

Here  $\Omega = (0, 1)^d$  is the image domain, with  $d = 2$  in the examples of this work, but there is no fundamental restriction on the dimension. For simplicity, the arguments of  $J$  in (2.1) are assumed to be real-valued<sup>5</sup>. Furthermore, they are considered to be mappings  $u, \sigma : \Omega \rightarrow [0, 1]$ , although the restriction of range is not explicitly enforced in (2.1), as would be the case, e.g., if barrier functions were added to (2.1). For (2.1), the  $\ell_2$  norm of the  $n$ th order derivative is given by:

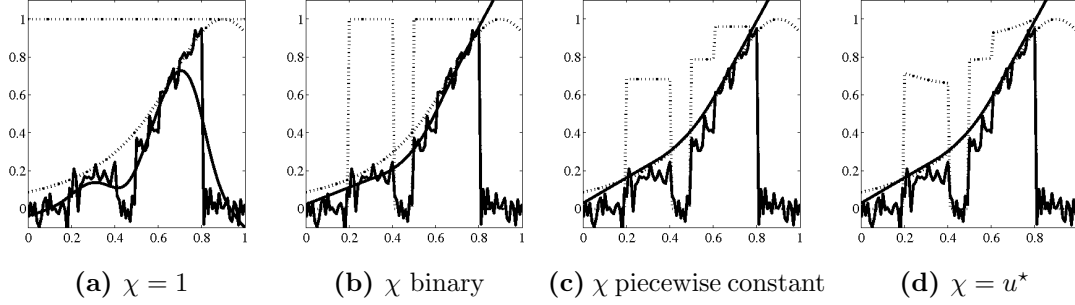
$$|\nabla^n \sigma|^2 = \nabla^n \sigma \cdot \nabla^n \sigma, \quad \nabla^n \sigma_1 \cdot \nabla^n \sigma_2 = \sum_{|\alpha|=n} \binom{n}{\alpha!} \partial^{\alpha} \sigma_1 \partial^{\alpha} \sigma_2 \quad (2.2)$$

The penalty on the second derivative of the modulation  $\sigma$  seen in (2.1) is based upon work in [17] and [19]. Among the key points is first the fact that the modulation  $\sigma$  is much smoother than the image  $u$ . Also, to avoid that  $\sigma$  have values outside  $[0, 1]$ , barrier functions have been considered as seen in [12] and [15], but it is found here in practice that the modulation is non-negative on the image support, where it can also be scaled to be less than one. The values of  $\sigma$  outside the support of  $u$  are not important except in the way that they influence values inside the support indirectly through the smoothness of the modulation. In particular, the high order natural boundary conditions on  $\sigma$  reduce disturbances of the modulation at the domain boundary and thus also at the boundary of the image support. For example, when a penalty such as  $\int_{\Omega} |\Delta \sigma|^2 d\mathbf{x}$  is used, then harmonic functions are in the kernel of the penalty and boundary disturbances appear as seen in [17]. Also, when a spectral penalty is used such as

<sup>5</sup>While the raw measured data are complex-valued [5], it is assumed here for simplicity that their inverse Fourier Transform is real-valued and therefore agrees with the magnitude images such as shown in Figure 2. When this is not the case, real and imaginary parts of the residual  $\sigma u - \tilde{u}$  may be processed separately as seen in [17].

$\sum_{\mathbf{k}} (1 + |\mathbf{k}|^2)^2 |\omega(\mathbf{k})|^2$ , where  $\{\omega(\mathbf{k})\}$  are Fourier or trigonometric series coefficients of  $\sigma$ , then  $\sigma$  is implicitly continued by periodicity outside of  $\Omega$ , and finite dimensional approximations lead to boundary disturbances.

The computation of the modulation  $\sigma$  is also improved by using a segmentation of the estimated image  $u$ , as can be seen from the 1D example shown in Figure 3. Here,  $\tilde{u}$  is shown



**Figure 3:**  $J$  in (2.1) is minimized with respect to  $\sigma$  for fixed  $u = \chi$ , where 4 different choices for  $\chi$  are shown. The solid noisy curve is  $\tilde{u}$ . The dotted smooth bell-shaped curve is the exact  $\sigma^*$ . The dotted piecewise smooth curve in (d) is the exact  $u^*$ . The curve  $\chi$  also appears as the dotted piecewise smooth curve in each graph. In each case the minimizer  $\sigma$  is the solid noiseless curve.

in each plot as the solid noisy curve. Using these data, the modulation is to be estimated by minimizing  $J$  in (2.1) with respect to  $\sigma$  for fixed  $u = \chi$ , where 4 different choices for  $\chi$  are tested. Each  $\chi$  is shown in the respective plot as a dotted piecewise smooth curve. The exact modulation  $\sigma^*$  appears in each plot as the dotted smooth bell-shaped curve. The exact image  $u^*$  is shown in Figure 3d where it agrees with  $\chi$ . By neglecting all information about the exact image  $u^*$  (or a practical estimation  $u$ ) Figure 3a shows that using  $\chi = 1$  leads to an excessive smoothing of the estimation  $\sigma$  over the edges of the data. One sees an immediate improvement in Figure 3b by using only a binary segmentation  $\chi = S_{\text{bin}}(u^*)$ . Yet another improvement is seen in Figure 3c by using a piecewise constant segmentation  $\chi = S_{\text{pc}}(u^*)$ . Finally, note that by using  $\chi = u^*$  for Figure 3d, the improvement obtained in relation to Figure 3c is quite marginal. Also note, as discussed in the previous paragraph, that the estimation errors in the modulation seen outside the support of  $\chi$  will not affect the modulation correction once the data are masked using the segmentation according to  $\tilde{u} \leftarrow (\chi > \chi_{\min}) \cdot \tilde{u}$ .

To regularize the image  $u$ , the function  $\phi_\epsilon$  in (2.1) is the Gauss-TV penalty used by the authors in [13] and [16],

$$\phi_\epsilon(s) = \begin{cases} s^2/(2\epsilon), & |s| \leq \epsilon \\ |s| - \epsilon/2, & |s| \geq \epsilon \end{cases} \quad (2.3)$$

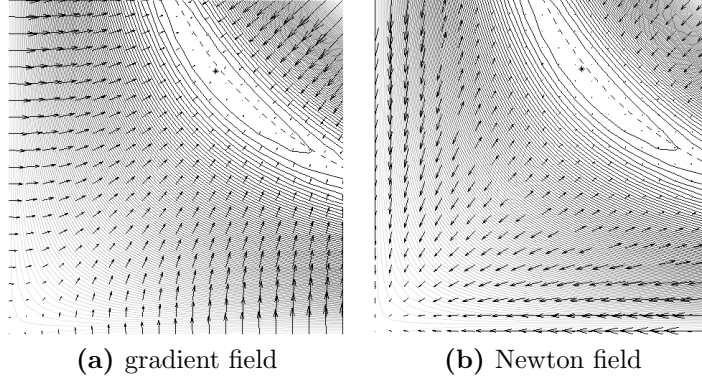
which emerges naturally from the duality formulation as shown in [13] and as seen below in (A.2). For  $u \in \text{BV}(\Omega)$ ,  $Du$  in  $\phi_\epsilon(|Du|)$  is understood as a measure [1], and otherwise  $J(u, \sigma)$  is unbounded. The  $L^2$  regularization in (2.1) is included so that  $(\kappa + \sigma^2)^{-1}$  exists for (3.11) below, even when  $\sigma$  becomes very small.

Note that an additive decomposition,  $\tilde{u} \rightarrow u + v$ , has been accomplished previously through different regularizations on  $u$  and  $v$ ; see, e.g., [2]. In the present work, as well as in [7] and [18], the multiplicative decomposition,  $\tilde{u} \rightarrow \sigma u$ , is also accomplished through different regularizations on  $\sigma$  and  $u$ . Note that this structure makes the derivative of the residual in (2.1) bilinear in the unknown image and modulation. To examine the landscape of the functional in (2.1), consider the minimization of the following model function:

$$f(x, y) = \frac{1}{2}(yx - z)^2 + \frac{1}{2}\nu y^2 + \mu|x| \quad (2.4)$$

A contour plot of  $f$  is shown in Figure 4 along with the vector fields  $-\nabla f$  and  $-\nabla^2 f^{-1} \nabla f$ .





**Figure 4:** Contour plots of  $f$  in (2.4) (with  $z = 0.5$ ,  $\mu = \nu = 0.01$ ) are shown with the vector fields (a)  $-\nabla f$  on the left and (b)  $-\nabla^2 f^{-1} \nabla f$  on the right. The curve  $yx = z$  is shown dashed in both plots, and the minimizer for  $f$  is shown as the asterisk near the dashed curve.

Without regularization from  $\mu$  and  $\nu$ , the whole curve  $yx = z$  would minimize  $f$ ; however, with positive regularization a unique minimizer exists, although it lies in a flat and elongated region of the landscape. The location of the unique minimizer depends of course entirely on the regularization. Note that the gradient direction field  $-\nabla f$  points strongly toward  $yx = z$ , but the field is rather weak in a near neighborhood of the curve. On the other hand, the Newton direction field  $-\nabla^2 f^{-1} \nabla f$  actually points away from  $yx = z$  unless sufficiently near to the curve, where the Newton direction field is actually weaker than the gradient direction field. Thus, it is not surprising that computational experiments using Newton's method to minimize functionals such as (2.1) have had limited success; see, e.g., [7] for related investigations. Since the cost is convex in each single argument, it can be minimized rapidly with respect to one argument while the other is held fixed; furthermore, employing such a procedure alternately has performed better than carrying out line searches along gradient directions for the joint functional. Thus, a nonlinear Gauss-Seidel outer iteration is used in which the formulated cost is minimized with respect to one variable after the other using an inner generalized Newton iteration.

This iterative method arrives quickly at the flat elongated region of the landscape. However, any of the above iterations can stall in such a region, leading numerically to an effective non-uniqueness in minimizers. To distinguish among such numerical minimizers, iterations may be guided by additional information. For instance, since the arguments of  $J$  are expected to have range in  $[0, 1]$ , the modulation is normalized here in each iteration to achieve a value of 1 at the brightest point of the data. Note that a pointwise normalization of the modulation is more stable than that of the image since the modulation is smoother. Also, on the basis of the discussion of Figure 3, the modulation is regularized in early iterations by using a segmentation  $\chi$  of the image  $u$  instead of the image itself. Such projection and regularization techniques have been found here to perform better, for instance, than treating the nonconvexity of  $J$  by starting iterations with larger and ending with smaller regularization parameters in (2.1).

### 3 Optimality Conditions

In this work  $H^k(\Omega)$  denotes the Sobolev space of functions with distributional derivatives up to order  $k$  in  $L^2(\Omega)$ ; see [10] for further information about these function spaces. As explained in [17], the optimality condition for (2.1) with respect to  $\sigma$  for fixed  $u$  is:

$$B(u)\sigma = u\tilde{u}, \quad \sigma \in H^2(\Omega) \quad (3.1)$$

which in weak form is defined by:

$$\int_{\Omega} [\nu \nabla^2 \sigma \cdot \nabla^2 \bar{\sigma} + u^2 \sigma \bar{\sigma}] d\mathbf{x} = \int_{\Omega} \bar{\sigma} u \tilde{u} d\mathbf{x}, \quad \forall \bar{\sigma} \in H^2(\Omega). \quad (3.2)$$

According to [17] there is a unique weak solution  $\sigma \in H^2(\Omega)$  when  $\tilde{u} \in L^2(\Omega)$  holds and  $u \in L^\infty(\Omega)$  has a support with positive measure. The additional regularity  $\sigma \in H^4(\Omega)$  is shown in [19]. In strong form,  $B(u) = \nu \Delta^2 + u^2$ , and a smooth solution  $\sigma$  satisfies the natural boundary conditions  $\partial_n^3 \sigma = \partial_n^2 \sigma = \partial_n \partial_\tau \sigma = 0$ ,  $\partial \Omega$ , where  $\partial_n$  and  $\partial_\tau$  are the normal and tangential derivatives respectively [17].

In initial iterations of the nonlinear Gauss-Seidel method (3.1) is solved by replacing the image  $u$  with a segmentation  $S(u)$  based upon [15]. Specifically, the image is approximated by:

$$S(u) = \sum_{i=1}^M c_i \chi_i \quad (3.3)$$

where  $\chi_i$  is the characteristic function for a subdomain  $\Omega_i \subseteq \Omega$  in which the segmentation possesses the grey level  $c_i$ . These subdomains are disjoint and are determined so that the cost  $J(\Omega_1, \dots, \Omega_M) = \int_{\Omega} [S(u) - u]^2 d\mathbf{x}$  is minimized and so cannot be reduced by transferring part of one subdomain to another. The topological derivative  $T_{ij}(\mathbf{x})$  measures the rate of change of  $J(\Omega_1, \dots, \Omega_M)$  when a ball  $B(\mathbf{x}, r)$  of vanishingly small radius  $r$  is transferred from  $\Omega_i$  to  $\Omega_j$ :

$$T_{ij}(\mathbf{x}) = \lim_{|B(\mathbf{x}, r)| \rightarrow 0} \frac{J(\Omega_1, \dots, \Omega_i \setminus B(\mathbf{x}, r), \dots, \Omega_j \cup B(\mathbf{x}, r), \dots, \Omega_M) - J(\Omega_1, \dots, \Omega_M)}{|B(\mathbf{x}, r)|} \quad (3.4)$$

Here and below,  $|B|$  denotes the measure of the set  $B$ . The segmentation (3.3) is determined so that all derivatives  $T_{ij}$  are non-negative. As shown in [15] they are given explicitly as:

$$T_{i,j}(\mathbf{x}) = \begin{cases} [c_j - u(\mathbf{x})]^2 - [c_i - u(\mathbf{x})]^2, & |\Omega_j| \neq 0 \\ -[c_i - u(\mathbf{x})]^2, & |\Omega_j| = 0 \end{cases} \quad (3.5)$$

for  $\mathbf{x} \in \Omega_i$ , and  $T_{i,j}(\mathbf{x}) = 0$  for  $\mathbf{x} \notin \Omega_i$ . In order that the topological derivative approach operate in a more global fashion, a point  $\mathbf{x}$  is transferred from  $\Omega_i$  to  $\Omega_j$  only when

$$T_{i,j}(\mathbf{x}) < \gamma \min_{\mathbf{y} \in \Omega_j} T_{i,j}(\mathbf{y}) \quad (3.6)$$

where  $\gamma \in (0, 1)$ . Given the subdomains, the grey levels  $\{c_i\}$  are given by the following:

$$c_i = \frac{1}{|\Omega_i|} \int_{\Omega_i} u d\mathbf{x}; \quad |\Omega_i| \neq 0, \quad c_i = 0, \quad |\Omega_i| = 0. \quad (3.7)$$

See Algorithm 2 for the details of determining  $S(u)$ .

To establish an optimality condition for (2.1) with respect to  $u$  for fixed  $\sigma$ , define the functionals  $\mathcal{F} : L^2(\Omega) \rightarrow \mathbf{R} \cup \{\infty\}$ ,

$$\mathcal{F}(u) = \frac{1}{2} \int_{\Omega} |\sigma u - \tilde{u}|^2 d\mathbf{x} + \frac{\kappa}{2} \int_{\Omega} u^2 d\mathbf{x} \quad (3.8)$$

and  $\mathcal{G} : L^2(\Omega) \rightarrow \mathbf{R} \cup \{\infty\}$ ,

$$\mathcal{G}(u) = \frac{\mu}{2} \int_{\Omega} \phi_{\epsilon}(|Du|) d\mathbf{x} \quad (3.9)$$

so that the dependence on  $u$  in  $J$  is given by  $\mathcal{F}(u) + \mathcal{G}(u)$ .<sup>6</sup> Note that these operators satisfy the conditions of the Fenchel Duality Theorem, and the desired optimality condition is thus

---

<sup>6</sup>The authors wish to thank Otmar Scherzer for his suggestion that these functionals be defined on  $L^2$ ; see also the recent book [11].



given by [9]:

$$\begin{cases} \mathcal{F}(u) + \mathcal{F}^*(v) &= \int_{\Omega} uv d\mathbf{x} \\ \mathcal{G}(u) + \mathcal{G}^*(-v) &= - \int_{\Omega} uv d\mathbf{x} \end{cases} \quad (3.10)$$

where the convex conjugates  $\mathcal{F}^*$  and  $\mathcal{G}^*$  in (3.10) are given as follows; see Appendix A for the details. First,  $\mathcal{F}^* : L^2(\Omega) \rightarrow \mathbf{R} \cup \{\infty\}$  is given by:

$$\mathcal{F}^*(v) = \frac{1}{2} \int_{\Omega} [(\kappa + \sigma^2)^{-1}(v + \sigma \tilde{u})^2 - \tilde{u}^2] d\mathbf{x} \quad (3.11)$$

Secondly,  $\mathcal{G}^* : L^2(\Omega) \rightarrow \mathbf{R} \cup \{\infty\}$  is given by:

$$\mathcal{G}^*(v) = \begin{cases} \frac{\epsilon}{2\mu} \int_{\Omega} |\mathbf{p}|^2 d\mathbf{x}, & \text{if } v = \nabla \cdot \mathbf{p}, \quad \mathbf{p} \in S_{\mu} \\ +\infty, & \text{else} \end{cases} \quad (3.12)$$

where

$$S_{\mu} = \{\mathbf{p} \in H_0(\text{div}) : |\mathbf{p}| \leq \mu, \text{ a.e. in } \Omega\} \quad (3.13)$$

$$H_0(\text{div}) = \{\mathbf{p} \in L^2(\Omega, \mathbf{R}^d) : \nabla \cdot \mathbf{p} \in L^2(\Omega), \mathbf{n} \cdot \mathbf{p} = 0, \partial\Omega\} \quad (3.14)$$

and  $\mathbf{n}$  is the outwardly directed unit normal vector at  $\partial\Omega$ . As seen in Appendix A, the functionals (3.8), (3.9), (3.11) and (3.12) lead to the following formulation of the optimality system (3.10):

$$\begin{cases} (\kappa + \sigma^2)u - \nabla \cdot \mathbf{p} &= \sigma \tilde{u} \\ -\mu Du + [Du]_{\epsilon} \mathbf{p} &= 0 \end{cases} \quad u \in \text{BV}(\Omega), \quad \mathbf{p} \in S_{\mu} \quad (3.15)$$

where  $[Du]_{\epsilon} = \max\{\epsilon, |Du|\}$ .

## 4 Numerical Methods

The discretization of the optimality conditions in Subsection 3 begins with a division of  $\Omega$  into  $N^d = 2^{pd}$  (dimension  $d = 2$ ) cells, each with unit aspect ratio and width  $h = 2^{-p}$ . Specifically, with the integer component multi-indices  $\mathbf{j} = (j_1, j_2, \dots)$ ,  $\mathbf{0} = (0, 0, \dots)$ , and  $\mathbf{1} = (1, 1, \dots)$ , the cell centroids are  $\mathbf{x}_{\mathbf{j}} = (\mathbf{j} - \frac{1}{2})h$ ,  $\mathbf{1} \leq \mathbf{j} \leq N \cdot \mathbf{1}$ . Then,  $U_{\mathbf{j}} \approx u(\mathbf{x}_{\mathbf{j}})$  and  $\mathbf{U}$  denotes the vector of values  $\{U_{\mathbf{j}}\}$  according to the lexicographic ordering in which  $j_1$  increments first from 1 to  $N$ , then  $j_2$ , and so on. Also, let  $D(\mathbf{U})$  denote the diagonal matrix with the values  $\{U_{\mathbf{j}}\}$  situated along the diagonal according to the lexicographic ordering.

Following [17], (3.1) is discretized by  $B(u) \approx \nu B_h + D(\mathbf{U})^2$ , where  $B_h$  is a finite difference approximation to the biharmonic operator with natural boundary conditions. Specifically, the stencil values (weights for neighboring cells) for  $B_h$  are given explicitly (for  $d = 2$ ) as follows for the cells with centroids  $\{\mathbf{x}_{\mathbf{j}} : \mathbf{1} \leq \mathbf{j} \leq 3 \cdot \mathbf{1}\}$ , where stencil weights are obtained by dividing the following by  $2800h^4$ :

0	0	-152	424	208	0	424	920	848	208	208	848	768	848	208
0	0	-592	-2176	848	0	-2176	-3920	-4352	848	848	-4352	-4512	-4352	848
0	0	4368	-2256	768	0	-2256	20400	-4512	768	768	-4512	24768	-4512	768
0	0	-592	-2176	848	0	-2176	-3920	-4352	848	848	-4352	-4512	-4352	848
0	0	-152	424	208	0	424	920	848	208	208	848	768	848	208
0	0	-152	424	208	0	424	920	848	208	208	848	768	848	208
0	0	-592	-2176	848	0	-2176	-3920	-4352	848	848	-4352	-4512	-4352	848
0	0	3440	-1960	920	0	-1960	16960	-3920	920	920	-3920	20400	-3920	920
0	0	-296	-1088	424	0	-1088	-1960	-2176	424	424	-2176	-2256	-2176	424
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	-152	424	208	0	424	920	848	208	208	848	768	848	208
0	0	-296	-1088	424	0	-1088	-1960	-2176	424	424	-2176	-2256	-2176	424
0	0	928	-296	-152	0	-296	3440	-592	-152	-152	-592	4368	-592	-152
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

For given  $u$  and  $\tilde{u}$  set  $\mathbf{U} = \{u(\mathbf{x}_j)\}$  and  $\tilde{\mathbf{U}} = \{\tilde{u}(\mathbf{x}_j)\}$ . Then the numerical solution to (3.1) is given as  $\mathbf{S} = \{S_j\}$ ,  $S_j \approx \sigma(\mathbf{x}_j)$  which solves:

$$[\nu B_h + D(\mathbf{U})^2] \mathbf{S} = D(\mathbf{U}) \tilde{\mathbf{U}} \quad (4.1)$$

The matrix on the left side can be stored in sparse format, and the equation is solved in GNU Octave<sup>\*</sup> in the present work using backslash. As shown in the authors' work [19], a multigrid approach can be used advantageously for large images.

The segmentation computations in (3.3) - (3.7) are carried out by evaluating each function at the cell centroids  $\mathbf{x} = \mathbf{x}_j$ . See Algorithm 2 for the details of determining  $S(\mathbf{U})$ . Note that this algorithm was implemented in C++ using the associated GNU Octave interface.

The optimality system (3.15) is solved using a generalized Newton method. The system,

$$\begin{bmatrix} (\kappa + \sigma^2) & -\nabla \cdot \\ [-\mu I + \frac{(|\nabla u| > \epsilon)}{[\nabla u]_\epsilon} \nabla u \mathbf{p}^T] \nabla & [\nabla u]_\epsilon \end{bmatrix} \begin{bmatrix} \delta u \\ \delta \mathbf{p} \end{bmatrix} = - \begin{bmatrix} (\kappa + \sigma^2)u - \nabla \cdot \mathbf{p} - \sigma \tilde{u} \\ -\mu \nabla u + [\nabla u]_\epsilon \mathbf{p} \end{bmatrix} \quad (4.2)$$

is simplified by first eliminating the second equation to obtain:

$$\begin{aligned} \left\{ (\kappa + \sigma^2) - \nabla \cdot \left[ \frac{1}{[\nabla u]_\epsilon} \left( \mu I - \frac{(|\nabla u| > \epsilon)}{2[\nabla u]_\epsilon} [\nabla u \mathbf{p}^T + \mathbf{p} \nabla u^T] \right) \right] \nabla \right\} \delta u = \\ -(\kappa + \sigma^2)u + \sigma \tilde{u} + \mu \nabla \cdot \left( \frac{\nabla u}{[\nabla u]_\epsilon} \right) \end{aligned} \quad (4.3)$$

and the eliminated equation becomes:

$$\delta \mathbf{p} = \frac{1}{[\nabla u]_\epsilon} \left( \mu I - \frac{(|\nabla u| > \epsilon)}{2[\nabla u]_\epsilon} [\nabla u \mathbf{p}^T + \mathbf{p} \nabla u^T] \right) \nabla \delta u + \mu \frac{\nabla u}{[\nabla u]_\epsilon} - \mathbf{p} \quad (4.4)$$

Note that the term  $\nabla u \mathbf{p}^T$  has been symmetrized in (4.3) and (4.4) with  $[\nabla u \mathbf{p}^T + \mathbf{p} \nabla u^T]/2$ .

To discretize (4.3) and (4.4) the discrete derivative matrices  $\nabla_h^{(x_i)}$  are formed using forward differences, where Neumann boundary conditions are implemented implicitly by setting the stencil weights of  $\nabla_h^{(x_i)}$  to zero at cells adjacent to the boundary where  $x_i = 1$ . In  $\mathbf{R}^2$  the gradient and divergence are then approximated by:

$$\nabla \approx \nabla_h = \begin{pmatrix} \nabla_h^{(x_1)} \\ \nabla_h^{(x_2)} \end{pmatrix}, \quad \nabla \cdot \approx -\nabla_h^T \quad (4.5)$$

Also, functions of  $\nabla_h \mathbf{U}$  are defined according to:

$$\begin{aligned} |\nabla_h \mathbf{U}| &= \left\{ [(\nabla_h^{(1)} \mathbf{U})_j^2 + (\nabla_h^{(2)} \mathbf{U})_j^2]^{\frac{1}{2}} \right\} & (|\nabla_h \mathbf{U}| > \epsilon) &= \{(|\nabla_h \mathbf{U}|)_j > \epsilon\} \\ \frac{1}{[\nabla_h \mathbf{U}]_\epsilon} &= \left\{ \frac{1}{\max\{(|\nabla_h \mathbf{U}|)_j, \epsilon\}} \right\} & \frac{(|\nabla_h \mathbf{U}| > \epsilon)}{[\nabla_h \mathbf{U}]_\epsilon} &= D(|\nabla_h \mathbf{U}| > \epsilon) \frac{1}{[\nabla_h \mathbf{U}]_\epsilon} \end{aligned} \quad (4.6)$$

and  $\mathbf{P} = (\mathbf{P}_1; \mathbf{P}_2)$  is understood below as a column vector of column vectors  $\mathbf{P}_1$  and  $\mathbf{P}_2$ . Thus, (4.3) is discretized as:

$$\begin{aligned} \left\{ \kappa I + D(\mathbf{S})^2 + \right. \\ \left. \nabla_h^T \left[ D \left( \frac{1}{[\nabla_h \mathbf{U}]_\epsilon} \right) \left( \mu I - D \left( \frac{(|\nabla_h \mathbf{U}| > \epsilon)}{2[\nabla_h \mathbf{U}]_\epsilon} \right) [\nabla_h \mathbf{U} \mathbf{P}_\mu^T + \mathbf{P}_\mu \nabla_h \mathbf{U}^T] \right) \right] \nabla_h \right\} \delta \mathbf{U} = \\ - [\kappa I + D(\mathbf{S})^2] \mathbf{U} + D(\mathbf{S}) \tilde{\mathbf{U}} - \mu \nabla_h^T \left[ D \left( \frac{1}{[\nabla_h \mathbf{U}]_\epsilon} \right) \nabla_h \mathbf{U} \right] \end{aligned} \quad (4.7)$$

<sup>\*</sup> See <http://www.gnu.org/software/octave/>

---

**Algorithm 1** Surface Coil Image Correction
 

---

```

1: Input:  $\tilde{U}$ ,  $\epsilon$ ,  $\kappa$ ,  $\mu$ ,  $\nu$ ,  $M$ ,  $\gamma$ ,  $\delta$ 
2: Output:  $U$ ,  $S$ 

3: Initialization:
4:   segment  $\tilde{U}$  to obtain  $S(\tilde{U})$  from Algorithm 2, using  $3M$  instead of  $M$ 
5:   mask the data according to  $\tilde{U} \leftarrow \tilde{U} \cdot (S(\tilde{U}) > \min(S(\tilde{U})))$ 
6:   starting image:  $U = \tilde{U}$ 
7:   starting modulation:  $S = 1$ 
8:   for method=1,2 do ▷ (method=1  $\Rightarrow$  purely  $L_2$ , method=2  $\Rightarrow$  with TV)
9:     Outer Iteration: start with  $t_0 = \|U\|$ ,  $t = 2\delta \cdot t_0$ 
10:    while ( $t > \delta \cdot t_0$ ) do
11:      save  $\hat{U} = U$ 
12:      segment  $U$  to obtain  $S(U)$  from Algorithm 2,
13:        in the first iteration with 2 grey levels, increasing incrementally to  $M$ 
14:      solve (4.1) for  $S$  where  $U$  is replaced by the segmentation  $S(U)$ 
15:      normalize  $S = S/S_i$  where  $\tilde{U}_i = \max_j\{\tilde{U}_j\}$ 
16:      if method=1 then
17:        set  $U = S\tilde{U}/(\kappa + S^2)$ 
18:      else if method=2 then
19:        Inner Iteration: start with  $s_0 = \|U\|$ ,  $s = 2\delta \cdot s_0$ ,  $P = 0$ 
20:        while ( $s > \delta \cdot s_0$ ) do
21:          solve (4.7) for  $\delta U$  and set  $U = U + \delta U$ 
22:          set  $\delta P$  with (4.9), set  $P = P + \delta P$  and set  $P_\mu$  according to (4.8)
23:          update  $s = \|\delta U\|$ 
24:        end while
25:      end if
26:      update  $t = \|U - \hat{U}\|$ 
27:    end while
28:  end for

```

---

where, following [13],  $P_\mu$  is a projection of  $P$  truncated to have a cellwise magnitude not more than  $\mu$ :

$$\begin{aligned}
P_\mu &= (D(|P| < \mu)P_1; D(|P| < \mu)P_2) \\
&+ \mu(D(|P| \geq \mu)D(|P|)^{-1}P_1; D(|P| \geq \mu)D(|P|)^{-1}P_2),
\end{aligned} \tag{4.8}$$

$$|P| = \{\sqrt{(P_1)_j^2 + (P_2)_j^2}\}$$

As shown in [13], using the truncation  $P_\mu$  in (4.7) guarantees that  $\delta U$  provides a descent direction for the cost functional with fixed modulation. The matrix on the left in (4.7) can be stored in sparse format and the system is solved in GNU Octave in the present work using backslash. The update for  $P$  itself is given by a discretization of (4.4):

$$\begin{aligned}
\delta P = D\left(\frac{1}{[\nabla U]_\epsilon}\right) &\left(\mu I - D\left(\frac{(|\nabla U| > \epsilon)}{2[\nabla U]_\epsilon}\right) [\nabla_h U P_\mu^T + P_\mu \nabla_h U^T]\right) \nabla_h \delta U \\
&+ \mu D\left(\frac{1}{[\nabla U]_\epsilon}\right) \nabla_h U - P
\end{aligned} \tag{4.9}$$

The above numerical formulations are applied algorithmically as shown in Algorithm 1 which calls Algorithm 2 for segmentation. Computational results using these algorithms are shown in the next subsection.

---

**Algorithm 2** Image Segmentation

---

```
1: Input:  $U, M, \gamma, \delta$ 
2: Output:  $\Omega_i, c_i, i = 1, \dots, M$ 

3: Initialization:
4:   starting subdomains:  $\chi_1 = 1, \chi_j = 0$  for  $j \neq 1$ 
5:   starting grey levels:  $c_j$  determined from (3.7)
6:   starting topological derivatives:
7:   for  $j = 2, \dots, M$  do
8:     for  $i = 1, \dots, (j-1)$  do
9:       set  $T_{i,j}$  according to (3.5) with  $|\Omega_j| = 0$ 
10:      transfer  $\{x \in \Omega_i : (3.6) \text{ holds}\}$  from  $\Omega_i$  to  $\Omega_j$ 
11:      update  $c_i$  and  $c_j$  with (3.7)
12:    end for
13:  end for

14: Iteration: start with  $t_0 = 1 + \sum_{i,j}^M \|T_{i,j}\|, t = 2\delta \cdot t_0$ 
15: while  $(t > \delta \cdot t_0)$  do
16:   for  $i = 1, \dots, M$  do
17:     set  $T_{i,j}$  for  $i \neq j = 1, \dots, M$  with (3.5) if  $|\Omega_i| \neq 0$  and otherwise  $T_{i,j} = 0$ 
18:     project  $T_{i,j}(\mathbf{x}) = \min\{T_{i,j}(\mathbf{x}), 0\}$ 
19:     define  $T_i(\mathbf{x}) = \min_{1 \leq j \leq M} T_{i,j}(\mathbf{x})$ 
20:     transfer  $\{x \in \Omega_i : T_{i,j}(\mathbf{x}) = T_i(\mathbf{x}) \text{ and } T_i(\mathbf{x}) < \gamma \cdot \min_{\mathbf{y} \in \Omega_i} T_i(\mathbf{y})\}$ 
21:       from  $\Omega_i$  to  $\Omega_j, i \neq j = 1, \dots, M$ 
22:     update  $c_j, j = 1, \dots, M$ , with (3.7)
23:   end for
24:   update  $t = \sum_{i,j}^M \|T_{i,j}\|$ 
25: end while
```

---

## 5 Computational Results

In this section, given images are artificially corrupted with noise and with a smooth modulation, and Algorithm 1 is used to correct for these corruptions. A reconstruction can thereby be compared to a known desired result.

For the example of Figure 5, the exact image of Figure 2a was used. To obtain this image, a fully sampled T2 weighted Turbo Spin Echo Scan of the brain of a multiple sclerosis patient was acquired on a clinical 1.5T scanner. Written informed consent was obtained prior to the examination. Sequence parameters were repetition time  $TR = 3845\text{ms}$ , echo time  $TE = 80\text{ms}$ , turbo factor 13, 2 signal averages, matrix size  $(x, y, z) = 512 \times 512, 24$  slices with a slice thickness of 5mm and an in-plane resolution of  $0.44\text{mm} \times 0.44\text{mm}$ . The data set was exported and resized to  $128 \times 128$  for ease of computational experimentation.

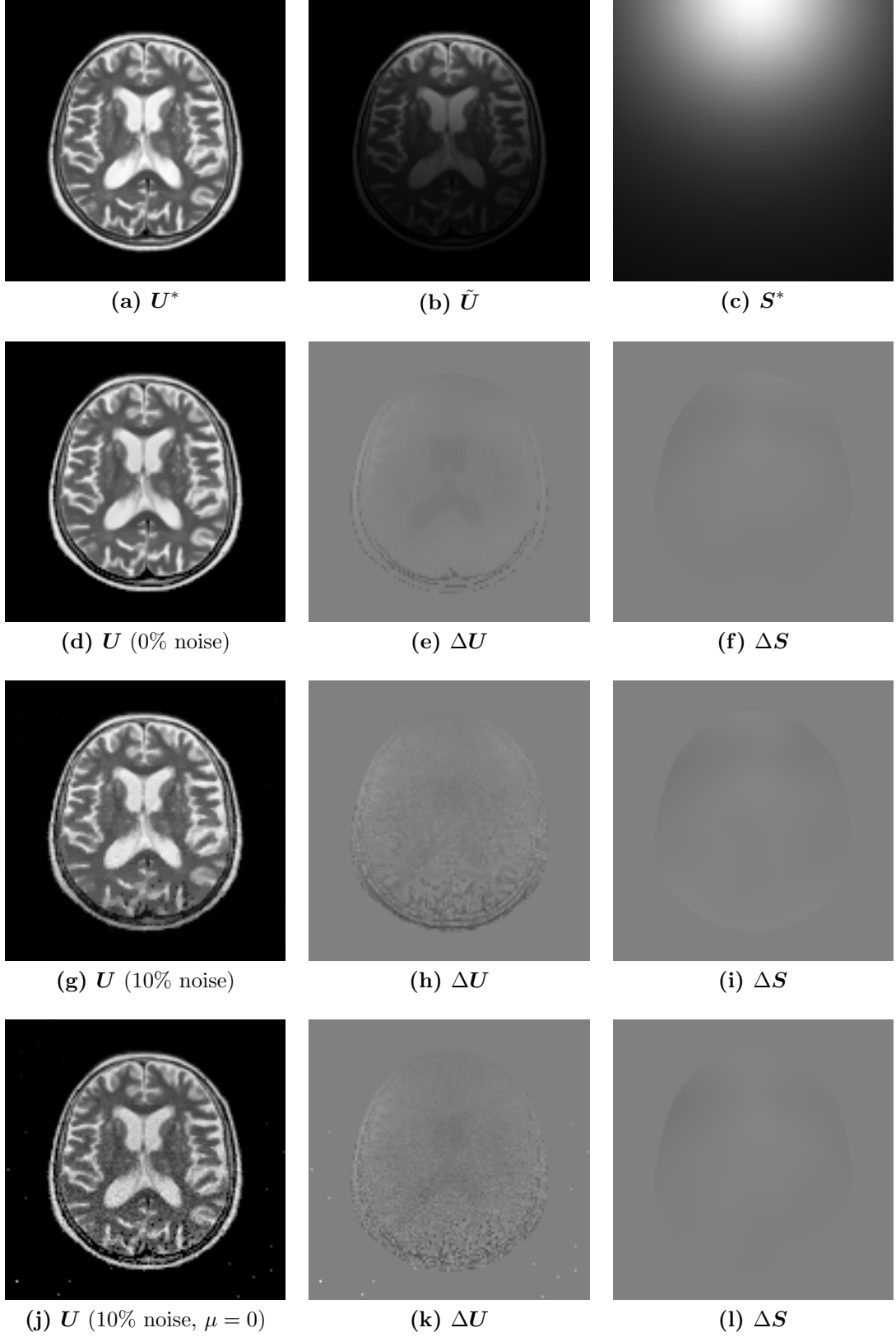
Based upon the authors' considerations in [12] of appropriate sensitivity parameterizations founded on the Biot-Savart Law, functions of the following form are used for artificial modulations:

$$\sigma(\mathbf{x}) = \frac{1}{[1 + \alpha \|\mathbf{x} - \mathbf{x}_0\|^2]^{\frac{3}{2}}} \quad (5.1)$$

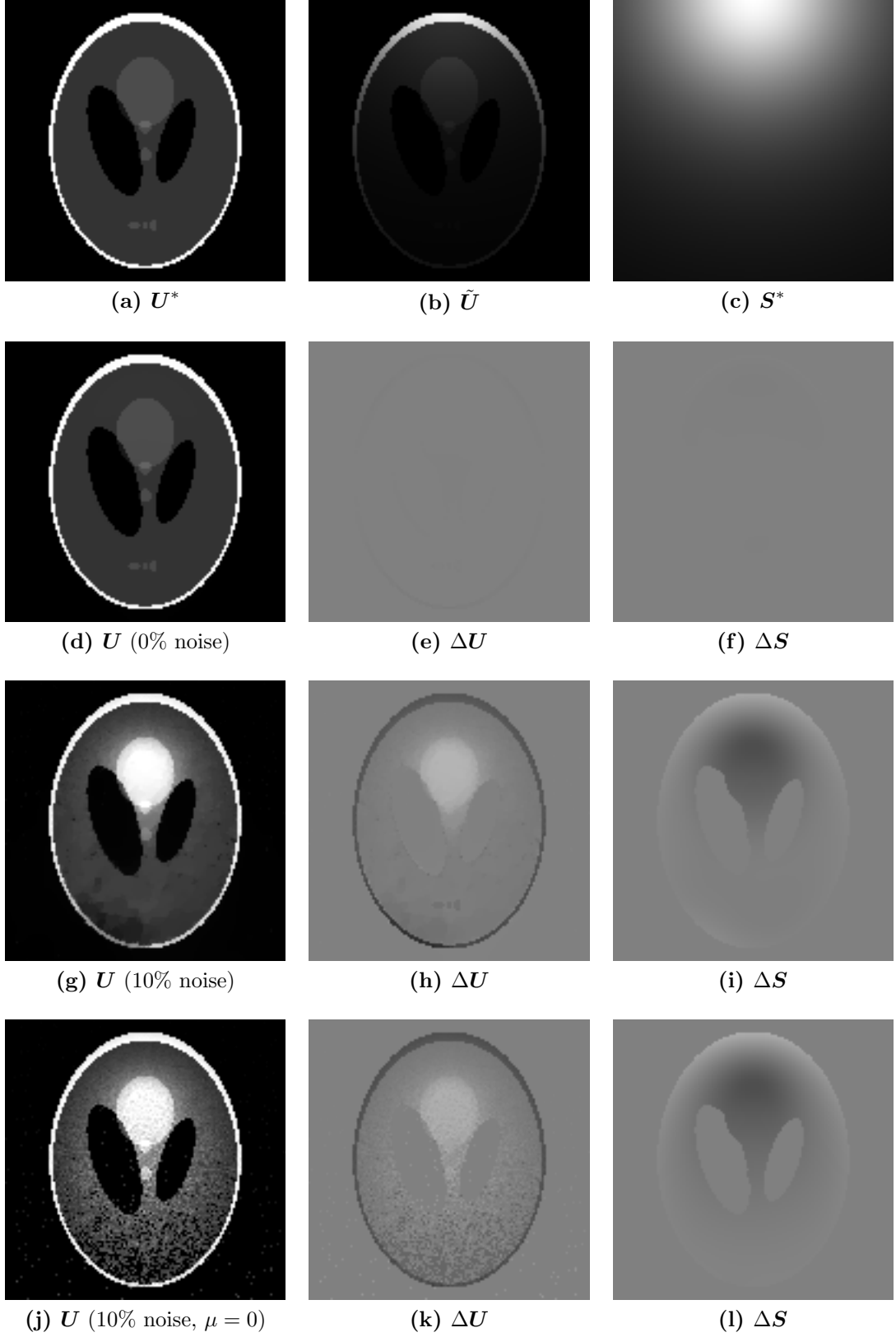
For the example of Figure 5, a modulation (5.1) was used with parameters,

$$\mathbf{x}_0 = (\tfrac{1}{2}, \tfrac{1}{2}) + r(\cos(\theta), \sin(\theta)), \quad r = \tfrac{3}{4}\sqrt{\tfrac{2}{2}}, \quad \theta = \tfrac{\pi}{2}, \quad \alpha = 5 \quad (5.2)$$

This modulation  $\mathbf{S}^*$  is shown in Figure 5c. The exact image  $\mathbf{U}^*$  appears in Figure 5a for comparison purposes. The product of  $\mathbf{U}^*$  and  $\mathbf{S}^*$ , denoted by  $\tilde{\mathbf{U}}$ , is shown in Figure 5b.



**Figure 5:** The top row shows the exact image  $U^*$  and the exact modulation  $S^*$  whose product forms the data  $\tilde{U}$ . In the subsequent rows, the reconstruction  $U$  is shown in the first column for the indicated conditions, and the corresponding error images  $\Delta U$  and  $\Delta S$  are shown respectively in the second and third columns. Reconstructed images are shown on the scale  $[0, 1]$  while error images are shown on the scale  $[-1, 1]$ .



**Figure 6:** The top row shows the exact image  $U^*$  and the exact modulation  $S^*$  whose product forms the data  $\tilde{U}$ . In the subsequent rows, the reconstruction  $U$  is shown in the first column for the indicated conditions, and the corresponding error images  $\Delta U$  and  $\Delta S$  are shown respectively in the second and third columns. Reconstructed images are shown on the scale  $[0, 1]$  while error images are shown on the scale  $[-1, 1]$ .

The data image  $\tilde{U}$  was used as input for Algorithm 1 along with the empirically optimal parameters  $\epsilon = 10^{-3}$ ,  $\kappa = 10^{-5}$ ,  $\mu = 10^{-4}$ ,  $\nu = 10^2$ ,  $M = 5$ ,  $\gamma = 0.5$  and  $\delta = 10^{-3}$ . Also, the dimension of the images is  $N = 128$ , but the domain is normalized to  $(1, N)^2$  so that  $h = 1$ . The reconstruction  $U$  is shown in Figure 5d with intensity scale  $[0, 1]$ , where lower intensities are shown darker while higher intensities are shown brighter. To avoid the effect of an error in intensity scale, the reconstruction error is measured with the metrics,

$$\begin{aligned} d_2(U, U^*) &= \frac{1}{N} \min_{s \in \mathbf{R}} \|sU - U^*\|_{\ell_2} = \frac{1}{N} \|s^*U - U^*\|_{\ell_2}, \quad s^* = U \cdot U^* / \|U\|_{\ell_2}^2 \\ d_\infty(U, U^*) &= \|\Delta U\|_{\ell_\infty}, \quad \Delta U = s^*U - U^* \end{aligned} \quad (5.3)$$

The error images  $\Delta U = s^*U - U^*$  and  $\Delta S = s^*S - S^*$  are shown respectively in Figures 5e and 5f, each with the intensity scale  $[-1, 1]$ . The quantitative errors,  $d_p(U, U^*)$  and  $d_p(\hat{S}, \hat{S}^*)$ , are relatively small in this case, and they are listed in line 1 of Table 1. Also shown in Table 1 is the elapsed computational time, the number of purely  $L_2$  iterations and the number of additional TV iterations. Note that the computations were performed on an Intel Core 2 Q9400 (four cores) with 4GB RAM, and the code was implemented in Octave as noted in Section 4. The simulation was then repeated with 10% noise added as follows with  $n = 0.10$ :

$$\text{FFT}(\tilde{U}) \rightarrow \hat{U}, \quad \hat{U} + \frac{n}{N} \|\hat{U}\|_2 \mathbf{X} \rightarrow \hat{U}, \quad \mathfrak{RFFT}^{-1}(\hat{U}) \rightarrow \tilde{U} \quad (5.4)$$

where  $\mathbf{X}$  is  $N \times N$  with elements normally distributed around mean 0 with variance 1. With this noise added,  $\mu = 10^{-3}$  was used on input. The resulting reconstruction is shown in Figure 5g, and the error images  $\Delta U$  and  $\Delta S$  are shown respectively in Figures 5h and 5i. The quantitative errors in this case are listed in line 2 of Table 1 along with the computational costs. Note that the image reconstruction error is higher than in the previous case although the quality of the reconstructed image is quite acceptable. Notice however that the modulation reconstruction error is comparable to that in the previous case, owing necessarily to the smoothness of the modulation. The computational costs shown in Table 1 follow the expected trends, and these are all acceptable. The simulation was finally repeated with 10% noise added but without using TV regularization. The resulting reconstruction is shown in Figure 5j, and the error images  $\Delta U$  and  $\Delta S$  are shown respectively in Figures 5k and 5l. The quantitative errors in this case are listed in line 3 of Table 1 along with the computational costs. Notice again that the modulation reconstruction error is comparable to that in the previous case, owing necessarily to the smoothness of the modulation. The reconstructed image shown in Figure 5j is accurately demodulated but evidently noisier than the result shown in Figure 5g, particularly in the region where the modulation intensity is low. Thus, it can be argued that the  $L_2$  type metrics of (5.3) are not adequate for measuring image reconstruction quality.

Simulations with phantoms are next considered to demonstrate how the method performs on images with a very different histogram than that of Figure 5. In the following example, the exact image  $U^*$  is that shown in Figure 6a. The exact modulation is shown in Figure 6c. The product of these two images, denoted by  $\tilde{U}$ , is shown in Figure 6b. The data image  $\tilde{U}$  was used as input for Algorithm 1, and other input parameters were the same as with Figure 5. The reconstruction  $U$  is shown in Figure 6d, and the error images  $\Delta U$  and  $\Delta S$  are shown respectively in Figures 6e and 6f. The quantitative errors are relatively small in this case, and they are listed along with computational costs in line 1 of Table 2. Note that the computational costs for this simple noise-free phantom example are considerably higher than in any other cases, owing indeed perhaps to the simplicity itself. The simulation was then repeated with 10% noise added to the image in Figure 6a. The resulting reconstruction is shown in Figure 6g and the error images  $\Delta U$  and  $\Delta S$  are shown respectively in Figures 6h and 6i. The quantitative errors in this case are listed in line 2 of Table 2 along with the computational costs. Note that the accuracy of modulation reconstruction for this simple phantom example is conspicuously less



Figure 5 row	$d_2(\mathbf{U}, \mathbf{U}^*)$	$d_\infty(\mathbf{U}, \mathbf{U}^*)$	$d_2(\hat{\mathbf{S}}, \hat{\mathbf{S}}^*)$	$d_\infty(\hat{\mathbf{S}}, \hat{\mathbf{S}}^*)$	time	$L_2$	TV
2	0.029	0.27	0.011	0.069	6.64	11	4
3	0.050	0.49	0.013	0.064	17.8	19	15
4	0.052	0.96	0.010	0.064	8.30	19	—

**Table 1:** Shown for each case of Figure 5 are the quantitative errors, measured in the metrics of (5.3), as well as the computational time in seconds, the number of purely  $L_2$  iterations and the number of additional TV iterations. Here,  $\hat{\mathbf{S}} = \mathbf{S} \cdot (\mathbf{U}^* > 0)$  and  $\hat{\mathbf{S}}^* = \mathbf{S}^* \cdot (\mathbf{U}^* > 0)$ .

Figure 6 row	$d_2(\mathbf{U}, \mathbf{U}^*)$	$d_\infty(\mathbf{U}, \mathbf{U}^*)$	$d_2(\hat{\mathbf{S}}, \hat{\mathbf{S}}^*)$	$d_\infty(\hat{\mathbf{S}}, \hat{\mathbf{S}}^*)$	time	$L_2$	TV
2	0.0027	0.034	0.0023	0.018	59.2	135	11
3	0.11	0.72	0.083	0.39	44.7	61	15
4	0.11	0.55	0.084	0.39	33.6	61	—

**Table 2:** Shown for each case of Figure 6 are the quantitative errors, measured in the metrics of (5.3), as well as the computational time in seconds, the number of purely  $L_2$  iterations and the number of additional TV iterations. Here,  $\hat{\mathbf{S}} = \mathbf{S} \cdot (\mathbf{U}^* > 0)$  and  $\hat{\mathbf{S}}^* = \mathbf{S}^* \cdot (\mathbf{U}^* > 0)$ .

than that seen for the realistic example of Figure 5, owing indeed perhaps to the simplicity of the phantom image. The simulation was finally repeated with 10% noise added but without using TV regularization. The resulting reconstruction is shown in Figure 6j and the error images  $\Delta\mathbf{U}$  and  $\Delta\mathbf{S}$  are shown respectively in Figures 6k and 6l. The quantitative errors in this case are listed in line 3 of Table 2 along with the computational costs. Again, the reconstructed image shown in Figure 6j is accurately demodulated but evidently noisier than the result shown in Figure 6g, particularly in the region where the modulation intensity is low. Here again, it can be argued that the  $L_2$  type metrics of (5.3) are not adequate for measuring image reconstruction quality.

These results can be summarized as follows. Reconstructions are generally more accurate for realistic images than for artificial images, except when an artificial image is piecewise constant and noise-free. The quality of image reconstruction is naturally higher where the intensity of the modulation is higher. Image quality is clearly improved with use of TV regularization, although this improvement is not dramatically quantified with the use of  $L_2$  type metrics. The quality of the modulation reconstruction for realistic images seems nearly independent of the variance of normally distributed noise, owing evidently to the smoothness of the modulation. The computational costs of the methods developed are not excessive, and these permit the approach to be used in a practical environment.

## References

- [1] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing*, Partial Differential Equations and the Calculus of Variations, Springer, Applied Mathematical Sciences, Vol. 147, second edition, 2006.
- [2] J.-F. Aujol and A. Chambolle, *Dual norms and image decomposition models*, Int. J. Comp. Vision, Vol. 63, No. 1, pp. 85 – 104, 2005.
- [3] L. Axel, J. Constantini and J. Listerud, *Intensity Correction in Surface Coil MR Imaging*, Am. J. Roentgenol., Vol. 148, pp. 418 – 420, 1987.
- [4] B. Belaroussi, J. Milles, S. Carme, Y. Min Zhu and H. Benoit, *Intensity non-uniformity correction in MRI: existing methods and their validation*, Medical Image Analysis, Vol. 10, pp. 234 – 246, 2006.

- [5] M.A. Bernstein, K.F. King and X.J. Zhou, *Handbook of MRI Pulse Sequences*, Academic Press, 2004.
- [6] T. Chan, G. Golub and P. Mulet, *A Nonlinear Primal-Dual Method for Total Variation-Based Image Restoration*, SIAM J. Sci. Comput., Vol. 20, No. 6, pp. 1964 – 1977, 1999.
- [7] C. Clason and G. Von Winckel, *On a bilinear optimization problem in parallel magnetic resonance imaging*, Applied Mathematics and Computation, DOI: 10.1016/j.amc.2010.02.047, 2010.
- [8] F. Demengel and R. Témam, *Convex Functions of a Measure and Applications*, Indiana University Mathematics Journal, Vol. 33, pp. 673 – 709, 1984.
- [9] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, Vol. 28 of Classics in Applied Mathematics, SIAM, Philadelphia, PA, English ed., 1999. (Translated from French.)
- [10] L.C. Evans, *Partial Differential Equations*, American Mathematical Society, Providence, 1998.
- [11] M. Grasmair, H. Grossauer, M. Haltmeier, F. Lenzen and O. Scherzer, *Variational Methods in Imaging*, Applied Mathematical Sciences, Vol. 167, Springer Verlag, New York, 2008.
- [12] M. Hintermüller, S.L. Keeling and A. Laurain, *Modulation recovery and image reconstruction in MRI: a structural study by parameterization*, SFB-Report No. 2008-023, Graz, Austria, 2008, submitted.
- [13] M. Hintermüller and G. Stadler, *A primal-dual algorithm for TV-based inf-convolution-type image restoration*, SIAM J. Sci. Comput., Vol. 28, No. 1, pp. 1 – 23, 2006.
- [14] M. Hintermüller and K. Kunisch, *Total bounded variation regularization as a bilaterally constrained optimization problem*, SIAM J. Appl. Math., Vol. 64, pp. 1311 – 1333, 2004.
- [15] M. Hintermüller and A. Laurain, *Multiphase image segmentation and modulation recovery based on shape and topological sensitivity*, J. Math. Imaging and Vision, Vol. 35, No. 1, pp. 1 – 22, 2009.
- [16] S.L. Keeling, *Total Variation Based Convex Filters for Medical Imaging*, Applied Mathematics and Computation, Vol. 139, pp. 101 – 119, 2003.
- [17] S.L. Keeling and R. Bammer, *A Variational Approach to Magnetic Resonance Coil Sensitivity Estimation*, Applied Mathematics and Computation, Vol. 158, No. 2, pp. 53 – 82, 2004.
- [18] S.L. Keeling, C. Clason, M. Hintermüller, F. Knoll, A. Laurain, G. Von Winckel, *An Image Space Approach to Cartesian Based Parallel MR Imaging with Total Variation Regularization*, SFB-Report 2008-022, Graz, Austria, 2008, submitted.
- [19] S.L. Keeling and G. Haase, *Geometric Multigrid for High-Order Regularizations of Early Vision Problems*, Applied Mathematics and Computation, Vol. 184, pp. 536 – 556, 2007.
- [20] S.-H. Lai and M. Fang, *A new variational shape-from-orientation approach to correcting intensity inhomogeneities in magnetic resonance images*, Med. Image Anal., Vol. 3, No. 4, pp. 409 – 424, 1999.
- [21] F.-H. Lin, Y.-J. Chen, J. Belliveau and L. Wald, *A wavelet-based approximation of surface coil sensitivity profiles for correction of image intensity inhomogeneity and parallel imaging reconstruction*, Hum. Brain Mapp., Vol. 19, No. 2, pp. 96 – 111, 2003.

- [22] J.V. Manjóna, J.J. Lulla, J. Carbonell-Caballeroa, G. García-Martía, L. Martí-Bonmatíb and M. Robles, *A nonparametric MRI inhomogeneity correction method*, Medical Image Analysis, Vol. 11, pp. 336 – 345, 2007.
- [23] J. Milles, Y. Zhu, N. Chen, L. Panych, G. Gimenez and C. Guttman, *MRI intensity nonuniformity correction using simultaneously spatial and gray-level histogram information*, SPIE Medical Imaging, pp. 734 – 742, 2004.
- [24] D. Pham and J. Prince, *An adaptative fuzzy C-means algorithm for image segmentation in the presence of intensity inhomogeneities*, Pattern Recogn. Lett., Vol. 20, No. 1, pp. 57 – 68, 1999.
- [25] M. Styner, C. Brechbuhler, G. Szekely and G. Gerig, *Parametric estimate of intensity inhomogeneities applied to MRI*, IEEE Trans. Med. Imaging, Vol. 19, No. 3, pp. 153 – 165, 2000.
- [26] U. Vovk, F. Pernus and B. Likar, *MRI intensity inhomogeneity correction by combining intensity and spatial information*, Phys. Med. Biol., Vol. 49, No. 17, pp. 4119 – 4133, 2004.
- [27] L. Zhou, Y. Zhu, C. Bergot, A.-M. Laval-Jeantet, V. Bousson, J.-D. Laredo and M. Laval-Jeantet, *A method of radio-frequency inhomogeneity correction for brain tissue segmentation*, MRI. Comput. Med. Imag. Grap., Vol. 25, No. 5, pp. 379 – 389, 2001.

## A Derivation of Primal-Dual Optimality System

For completeness, the Fenchel Duality Theorem is stated as follows; see [9] for the details. Let  $J(u) = \mathcal{F}(\Lambda u) + \mathcal{G}(u)$  be a decomposition of the cost  $J$ , where the operators  $\mathcal{F} : Y \rightarrow \mathbf{R} \cup \{\infty\}$ ,  $\mathcal{G} : V \rightarrow \mathbf{R} \cup \{\infty\}$  and  $\Lambda : V \rightarrow Y$  satisfy the following. First  $\Lambda$  is a bounded linear operator from  $V$  into  $Y$ . Also  $\mathcal{F}$  and  $\mathcal{G}$  are convex, lower semicontinuous functionals on  $Y$  and  $V$ , and there exists a  $u_0 \in V$  such that  $J(u_0) < \infty$ , and  $\mathcal{F}$  is continuous in  $Y$  at  $\Lambda u_0$ . Then the minimization of  $J$  is characterized by  $\inf_{u \in V} [\mathcal{F}(\Lambda u) + \mathcal{G}(u)] = \sup_{p \in Y^*} [-\mathcal{F}^*(-p) - \mathcal{G}^*(\Lambda^* p)]$  where  $\mathcal{F}^*$  and  $\mathcal{G}^*$  are the convex conjugates defined by  $\mathcal{F}^*(q) = \sup_{p \in Y} [\langle p, q \rangle_{Y, Y^*} - \mathcal{F}(p)]$  and  $\mathcal{G}^*(v) = \sup_{u \in V} [\langle u, v \rangle_{V, V^*} - \mathcal{G}(u)]$ . Also,  $\Lambda^*$  is the adjoint of  $\Lambda$ . Furthermore, the optimality system for this problem is given by  $\mathcal{F}(\Lambda u) + \mathcal{F}^*(-p) = -\langle \Lambda u, p \rangle_{Y, Y^*}$  and  $\mathcal{G}(u) + \mathcal{G}^*(\Lambda^* p) = \langle u, \Lambda^* p \rangle_{V, V^*}$ .

As opposed to the structure used in [13] and [14], for instance, in the present work  $V = Y = L^2(\Omega)$  and  $\Lambda = I$  are used. So the pairings  $\langle \cdot, \cdot \rangle_{Y, Y^*}$  and  $\langle \cdot, \cdot \rangle_{V, V^*}$  are both the inner product on  $L^2(\Omega)$ . It is shown below that  $\mathcal{G}$  in (3.9) and  $\mathcal{F}$  (3.8) as well as their conjugates are convex and lower semicontinuous on  $L^2(\Omega)$ . For  $u_0 = 0$  the cost satisfies  $J(u_0) < \infty$ , and the residual  $\mathcal{F}$  is continuous in  $L^2(\Omega)$  at  $u_0$ . For these claims it is noted that if  $\mathcal{H}$  a functional defined on a Banach space where it is convex, lower semicontinuous and proper ( $\mathcal{H}(h) > -\infty$ ,  $\forall h$ , and  $\exists h$  where  $\mathcal{H}(h) < +\infty$ ), then  $\mathcal{H}^*$  is also convex, lower semicontinuous and proper and  $\mathcal{H}^{**} = \mathcal{H}$  [9].

The convex conjugate of  $\mathcal{F}$  in (3.8) is given as follows:

$$\mathcal{F}^*(v) = \sup_{u \in L^2(\Omega)} \int_{\Omega} \left[ uv - \frac{1}{2} |\sigma u - \tilde{u}|^2 - \frac{\kappa}{2} u^2 \right] d\mathbf{x}$$

The directional derivative of the integral with respect to  $u$  for a perturbation  $\tilde{u}$  is  $\int_{\Omega} [v - \sigma(\sigma u - \tilde{u}) - \kappa u] d\mathbf{x}$ . Since  $\kappa > 0$ , the critical function is given by  $u^* = (v + \sigma \tilde{u}) / (\kappa + \sigma^2)$ . Inserting this  $u^*$  above gives the maximum as seen in (3.11). Note that  $\mathcal{F}$  is quadratic and continuous on  $L^2(\Omega)$ , and so  $\mathcal{F}$  is convex and lower semicontinuous. Also,  $\mathcal{F}$  is clearly proper and thus  $\mathcal{F}^*$  is convex, lower semicontinuous and proper, and  $\mathcal{F}^{**} = \mathcal{F}$  holds.

Next, define the functional on  $L^2(\Omega)$ ,

$$\tilde{\mathcal{G}}(v) = \begin{cases} \frac{\epsilon}{2\mu} \int_{\Omega} |\mathbf{p}|^2 d\mathbf{x}, & \text{if } v = \nabla \cdot \mathbf{p}, \quad \mathbf{p} \in S_{\mu} \\ +\infty, & \text{else} \end{cases} \quad (\text{A.1})$$

which agrees with the formula in (3.12) where  $S_{\mu}$  is defined by (3.13). It will be shown that  $\tilde{\mathcal{G}}$  is convex, lower semicontinuous and proper on  $L^2(\Omega)$ , so that  $\tilde{\mathcal{G}}^{**} = \tilde{\mathcal{G}}$  holds. Once it is shown that  $\tilde{\mathcal{G}}^* = \mathcal{G}$  in (3.9), it follows that  $\mathcal{G}^* = \tilde{\mathcal{G}}^{**} = \tilde{\mathcal{G}}$  in (A.1) and (3.12). Clearly,  $\tilde{\mathcal{G}}$  is convex and proper. To show the lower semicontinuity of  $\tilde{\mathcal{G}}$ , let  $\{v_k\}$  be a sequence converging in  $L^2(\Omega)$  to  $v$ . For the nontrivial case, suppose  $v_k = \nabla \cdot \mathbf{p}_k$  where  $\mathbf{p}_k \in S_{\mu}$ ,  $\forall k$ . Since  $\{\mathbf{p}_k\}$  is bounded in  $L^{\infty}(\Omega, \mathbf{R}^d)$  according to  $|\mathbf{p}_k| \leq \mu$ , there is a subsequence converging in a weak- $\star$  sense to  $\mathbf{p} \in L^{\infty}(\Omega, \mathbf{R}^d)$  which necessarily satisfies  $|\mathbf{p}| \leq \mu$ . Let the subsequence be denoted again by  $\{\mathbf{p}_k\}$ . As a consequence of the  $L^2(\Omega)$  strong convergence of  $\{v_k\}$  and the  $L^{\infty}(\Omega, \mathbf{R}^d)$  weak- $\star$  convergence of  $\{\mathbf{p}_k\}$ ,

$$\int_{\Omega} v \psi d\mathbf{x} = \lim_{k \rightarrow \infty} \int_{\Omega} v_k \psi d\mathbf{x} = \lim_{k \rightarrow \infty} - \int_{\Omega} \mathbf{p}_k \cdot \nabla \psi d\mathbf{x} = - \int_{\Omega} \mathbf{p} \cdot \nabla \psi d\mathbf{x}, \quad \forall \psi \in C^{\infty}(\bar{\Omega})$$

where the second equality follows with  $v_k = \nabla \cdot \mathbf{p}_k$  and no boundary term emerges since  $\mathbf{p}_k \in H_0(\text{div})$  as defined in (3.14). Thus,  $\mathbf{p} \in L^{\infty}(\Omega, \mathbf{R}^d) \subset L^2(\Omega, \mathbf{R}^d)$  and  $\nabla \cdot \mathbf{p} = v \in L^2(\Omega)$ . Combining the left and right sides of the last equation gives:

$$\int_{\partial\Omega} \psi \mathbf{p} \cdot \mathbf{n} d\mathbf{x} = \int_{\Omega} \mathbf{p} \cdot \nabla \psi d\mathbf{x} + \int_{\Omega} \psi v d\mathbf{x} = 0, \quad \forall \psi \in C^{\infty}(\bar{\Omega})$$

Thus,  $\mathbf{p} \in H_0(\text{div})$ . It follows that  $v = \nabla \cdot \mathbf{p}$  with  $\mathbf{p} \in S_{\mu}$ . As a consequence of the  $L^{\infty}(\Omega, \mathbf{R}^d)$  weak- $\star$  convergence of  $\{\mathbf{p}_k\}$ ,

$$\int_{\Omega} |\mathbf{p}|^2 d\mathbf{x} = \lim_{k \rightarrow \infty} \int_{\Omega} \mathbf{p} \cdot \mathbf{p}_k d\mathbf{x} \leq \|\mathbf{p}\|_{L^2(\Omega, \mathbf{R}^d)} \lim_{k \rightarrow \infty} \|\mathbf{p}_k\|_{L^2(\Omega, \mathbf{R}^d)}$$

where  $\mathbf{p}, \{\mathbf{p}_k\} \subset L^{\infty}(\Omega, \mathbf{R}^d) \subset L^1(\Omega, \mathbf{R}^d)$ . Dividing both sides of the above inequality by  $\|\mathbf{p}\|_{L^2(\Omega, \mathbf{R}^d)}$  shows finally that  $\tilde{\mathcal{G}}(v) \leq \tilde{\mathcal{G}}(v_k)$  and thus  $\tilde{\mathcal{G}}$  is lower semicontinuous.

Now the convex conjugate of  $\tilde{\mathcal{G}}$  is given by:

$$\tilde{\mathcal{G}}^*(u) = \sup_{\substack{v = \nabla \cdot \mathbf{p} \\ \mathbf{p} \in S_{\mu}}} - \int_{\Omega} \left[ u \nabla \cdot \mathbf{p} + \frac{\epsilon}{2\mu} |\mathbf{p}|^2 \right] d\mathbf{x} = \sup_{\substack{v = \nabla \cdot \mathbf{p} \\ \mathbf{p} \in S_{\mu}}} \int_{\Omega} \left[ \mathbf{p} \cdot Du - \frac{\epsilon}{2\mu} |\mathbf{p}|^2 \right] d\mathbf{x}$$

Now define  $\varphi(\boldsymbol{\xi}) = \mu \phi_{\epsilon}(|\boldsymbol{\xi}|)$ . A straightforward calculation shows that  $\varphi^*(\boldsymbol{\eta}) = \epsilon |\boldsymbol{\eta}|^2 / (2\mu)$  for  $|\boldsymbol{\eta}| \leq \mu$  and otherwise  $\varphi^*(\boldsymbol{\eta}) = +\infty$ . Then the convex function  $\varphi$  of the measure  $Du$  is characterized by [8]:

$$\int_{\Omega} \mu \phi_{\epsilon}(|Du|) = \int_{\Omega} \varphi(Du) = \sup_{\mathbf{p} \in \mathcal{D}_{\varphi^*}(X)} \int_{\Omega} \mathbf{p} \cdot Du - \int_{\Omega} \varphi^*(\mathbf{p}) d\mathbf{x}$$

where  $\mathcal{D}_{\varphi^*}(X) = \{\mathbf{p} \in X : \varphi^* \circ \mathbf{p} \in L^1(\Omega)\}$  and  $X$  may be taken indifferently as  $X = C_0^{\infty}(\Omega, \mathbf{R}^d)$  or  $L^1(\Omega, \mathbf{R}^d)$ . Since  $C_0^{\infty}(\Omega, \mathbf{R}^d) \subset H_0(\text{div}) \subset L^1(\Omega, \mathbf{R}^d)$ , the supremum may be taken over  $X = H_0(\text{div})$ . Then the last two equations show that

$$\tilde{\mathcal{G}}^*(u) = \int_{\Omega} \mu \phi_{\epsilon}(|Du|) = \mathcal{G}(u)$$

Thus,  $\mathcal{G}^*(v) = \tilde{\mathcal{G}}^{**}(v) = \tilde{\mathcal{G}}(v)$  is given according to (A.1) and (3.12).

With these convex conjugates the primal problem is related to the dual problem as follows:

$$\begin{aligned}
& \inf_{u \in L^2(\Omega)} \left\{ \frac{1}{2} \int_{\Omega} |\sigma u - \tilde{u}|^2 d\mathbf{x} + \frac{\kappa}{2} \int_{\Omega} u^2 d\mathbf{x} + \frac{\mu}{2} \int_{\Omega} \phi_{\epsilon}(|\nabla u|) d\mathbf{x} \right\} \\
&= \inf_{u \in L^2(\Omega)} [\mathcal{F}(u) + \mathcal{G}(u)] = \sup_{v \in L^2(\Omega)} [-\mathcal{F}^*(-v) - \mathcal{G}^*(v)] = \\
& \sup_{\substack{v = \nabla \cdot \mathbf{p} \\ \mathbf{p} \in S_{\mu}}} \left\{ \frac{1}{2} \int_{\Omega} \{ (v + \sigma \tilde{u})^2 / (\kappa + \sigma^2) - \tilde{u}^2 \} d\mathbf{x} + \frac{\epsilon}{2\mu} \int_{\Omega} |\mathbf{p}|^2 d\mathbf{x} \right\}
\end{aligned} \tag{A.2}$$

Here it can be seen from the last term that  $\epsilon > 0$  in (2.3) provides a regularization in the dual problem, which motivates its use in [13].

The optimality system (3.10) is obtained with the constructions above. For  $\mathcal{F}$  in (3.8) and  $\mathcal{F}^*$  in (3.11) the equation  $\mathcal{F}(u) + \mathcal{F}^*(v) = \int_{\Omega} uv d\mathbf{x}$  can be written as follows:

$$\begin{aligned}
0 &= \frac{1}{2} \int_{\Omega} \{ [\sigma u - \tilde{u}]^2 + \kappa u^2 + (v + \sigma \tilde{u})^2 / (\kappa + \sigma^2) - \tilde{u}^2 - 2uv \} d\mathbf{x} \\
&= \frac{1}{2} \int_{\Omega} [(\kappa + \sigma^2)u - v - \sigma \tilde{u}]^2 / (\kappa + \sigma^2) d\mathbf{x}
\end{aligned}$$

which is equivalent to the first equation in (3.15) with  $v = \nabla \cdot \mathbf{p}$ .

For the next optimality condition let  $Du = \nabla u d\mathbf{x} + Du_s$  denote the Lebesgue decomposition of the measure  $Du$ , where  $\nabla u d\mathbf{x}$  and  $Du_s$ , respectively, are absolutely continuous and singular with respect to Lebesgue measure. For  $\mathcal{G}$  in (3.9) and  $\mathcal{G}^*$  in (3.12) the equation  $\mathcal{G}(u) + \mathcal{G}^*(-v) = -\int_{\Omega} uv d\mathbf{x}$  can be written as follows in terms of  $v = \nabla \cdot \mathbf{p} \in S_{\mu}$ :

$$\begin{aligned}
0 &= \int_{\Omega} \mu \phi_{\epsilon}(|Du|) + \int_{\Omega} \left[ \frac{\epsilon}{2\mu} |\mathbf{p}|^2 + u \nabla \cdot \mathbf{p} \right] d\mathbf{x} \\
&= \int_{|\nabla u| < \epsilon} \left[ \frac{\mu}{2\epsilon} |\nabla u|^2 + \frac{\epsilon}{2\mu} |\mathbf{p}|^2 - \nabla u \cdot \mathbf{p} \right] d\mathbf{x} - \int_{\Omega} Du_s \cdot \mathbf{p} \\
&+ \int_{|\nabla u| \geq \epsilon} \left[ \mu |\nabla u| - \frac{\mu\epsilon}{2} + \frac{\epsilon}{2\mu} |\mathbf{p}|^2 - \nabla u \cdot \mathbf{p} \right] d\mathbf{x} + \int_{\Omega} \mu |Du_s| \\
&= \int_{|\nabla u| < \epsilon} \frac{\epsilon}{2\mu} \left| \mathbf{p} - \frac{\mu}{\epsilon} \nabla u \right|^2 d\mathbf{x} \quad (\geq 0)
\end{aligned} \tag{A.3}$$

$$+ \int_{|\nabla u| \geq \epsilon} \left[ \frac{\epsilon}{2\mu} \left| \mathbf{p} - \frac{\mu}{\epsilon} \nabla u \right|^2 - \frac{\mu}{2\epsilon} (|\nabla u| - \epsilon)^2 \right] d\mathbf{x} \quad (\geq 0) \tag{A.4}$$

$$+ \int_{\Omega} [\mu |Du_s| - Du_s \cdot \mathbf{p}] \quad (\geq 0) \tag{A.5}$$

The non-negativity of (A.3) is evident. The non-negativity of (A.5) follows since  $Du_s \cdot \mathbf{p} \leq |Du_s| |\mathbf{p}| \leq |Du_s| \mu$ . The non-negativity of (A.4) can be seen from the following estimate:

$$\int_{|\nabla u| \geq \epsilon} \left[ \frac{\epsilon}{2\mu} \left| \mathbf{p} - \frac{\mu}{\epsilon} \nabla u \right|^2 - \frac{\mu}{2\epsilon} (|\nabla u| - \epsilon)^2 \right] d\mathbf{x} \tag{A.6}$$

$$\geq \int_{|\nabla u| \geq \epsilon} \frac{\mu}{2\epsilon} \left[ \left( \frac{\epsilon}{\mu} |\mathbf{p}| - |\nabla u| \right)^2 - (|\nabla u| - \epsilon)^2 \right] d\mathbf{x} \tag{A.7}$$

$$= \int_{|\nabla u| \geq \epsilon} (\mu - |\mathbf{p}|) \left[ \frac{\epsilon}{2\mu} (\mu - |\mathbf{p}|) + (|\nabla u| - \epsilon) \right] d\mathbf{x} \geq 0 \tag{A.8}$$

Since each integrand in (A.3) and (A.4) is non-negative and the sum is zero, each integrand must be pointwise zero. That the integrand in (A.8) must vanish means that  $|\mathbf{p}| = \mu$  must hold when  $|\nabla u| \geq \epsilon$  holds. That the two integrals in (A.6) and (A.7) are zero implies, after removing identical terms in each, that

$$0 = \int_{|\nabla u| \geq \epsilon} [|\nabla u||\mathbf{p}| - \nabla u \cdot \mathbf{p}] d\mathbf{x}$$

where the integrand is non-negative. Thus,  $\nabla u \cdot \mathbf{p} = |\nabla u||\mathbf{p}|$  holds, i.e.,  $\nabla u/|\nabla u|$  and  $\mathbf{p}/|\mathbf{p}|$  are parallel unit vectors, so  $|\nabla u|\mathbf{p} = |\mathbf{p}|\nabla u = \mu\nabla u$  holds when  $|\nabla u| \geq \epsilon$  holds. That the integrand in (A.3) is non-negative means that  $\epsilon\mathbf{p} = \mu\nabla u$  holds when  $|\nabla u| < \epsilon$  holds, and it follows that  $\epsilon|\mathbf{p}| = \mu|\nabla u| < \epsilon\mu$  or  $|\mathbf{p}| < \mu$  holds. These conditions can be summarized as follows:

$$\begin{aligned} (|\nabla u| \geq \epsilon \Rightarrow) \quad & |\mathbf{p}| = \mu, \quad |\nabla u|\mathbf{p} = \mu\nabla u \\ (|\nabla u| < \epsilon \Rightarrow) \quad & |\mathbf{p}| < \mu, \quad \epsilon\mathbf{p} = \mu\nabla u \end{aligned}$$

or:

$$\max\{\epsilon, |\nabla u|\}\mathbf{p} = \mu\nabla u \tag{A.9}$$

Since (A.5) is zero,  $\mu|Du_s| = Du_s \cdot \mathbf{p} \leq |Du_s||\mathbf{p}|$  implies that  $|\mathbf{p}| = \mu$  on the support of  $Du_s$ . Thus, on this support,  $Du_s/|Du_s|$  and  $\mathbf{p}/\mu = \mathbf{p}/|\mathbf{p}|$  are parallel unit vectors where

$$|Du_s|\mathbf{p} = |\mathbf{p}|Du_s = \mu Du_s \tag{A.10}$$

Finally define the function  $\psi(\boldsymbol{\xi}) = \max\{\epsilon, |\boldsymbol{\xi}|\}$ , and according to [8], the convex function  $\psi$  of the measure  $Du$  is defined by  $\psi(Du) = \psi(\nabla u)d\mathbf{x} + \psi_\infty(Du_s) = \max\{\epsilon, |\nabla u|\}d\mathbf{x} + |Du_s|$  since  $\psi_\infty(\boldsymbol{\xi}) = \lim_{t \rightarrow \infty} \psi(t\boldsymbol{\xi})/t = |\boldsymbol{\xi}|$ . Thus, combining (A.9) and (A.10) gives  $\max\{\epsilon, |Du|\}\mathbf{p} = \mu Du$  as seen in the second equation of (3.15).