SIAM J. IMAGING SCIENCES
Vol. 11, No. 2, pp. 1179–1204

PET-MRI Joint Reconstruction by Joint Sparsity Based Tight Frame Regularization*

Jae Kyu Choi†, Chenglong Bao‡, and Xiaoqun Zhang§

Abstract. Recent technical advances lead to the coupling of PET and MRI scanners, enabling one to acquire functional and anatomical data simultaneously. In this paper, we propose a tight frame based PET-MRI joint reconstruction model via the joint sparsity of tight frame coefficients. In addition, a nonconvex balanced approach is adopted to take the different regularities of PET and MRI images into account. To solve the nonconvex and nonsmooth model, a proximal alternating minimization algorithm is proposed, and the global convergence is present based on the Kurdyka–Łojasiewicz property. Finally, the numerical experiments show that our proposed models achieve better performance over the existing PET-MRI joint reconstruction models.

Key words. joint reconstruction, positron emission tomography, magnetic resonance imaging, (tight) wavelet frames, data driven tight frames, joint sparsity, proximal alternating schemes

AMS subject classifications. 65K15, 68U10, 90C90, 92C55

DOI. 10.1137/17M1131453

1. Introduction.

Medical imaging is the technique and process of visualizing the anatomy of a body for clinical analysis and medical intervention, as well as the functions of some organs and tissues. For many decades, single medical imaging modality has been available and widely used to image either function or anatomy. For example, positron emission tomography (PET) and single positron emission computed tomography (SPECT) provide information on the distribution of radioisotopes and tracers in living tissue. These distributions allow clinicians to determine certain functions of tissue [33, 35]. On the other hand, magnetic resonance imaging (MRI) and X-ray computed tomography (CT) can provide the information on anatomical structure.

Recent technical advances have allowed for the coupling of PET and MRI, leading to significant advantages over traditional PET-CT and stand-alone MRI [16]. The resulting PET-MRI modalities are able to acquire functional data and anatomical data simultaneously, while PET-CT modalities acquire data sequentially [16, 18, 26, 52]. Since both PET and MRI data stem from the same underlying object, it would be more desirable to explore the

*Received by the editors May 22, 2017; accepted for publication (in revised form) January 2, 2018; published electronically May 8, 2018.

http://www.siam.org/journals/siims/11-2/M113145.html

Funding: The research of the first author was partially supported by General Financial Grant from the China Postdoctoral Science Foundation (2017M611539). The research of the third author was partially supported by NSFC (11771288), the Young Top-notch Talent program of China, and 973 program (2015CB856004)

†Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai 200240, China (jaycjk@sjtu.edu.cn).

‡Corresponding author. Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, China (clbao@math.tsinghua.edu.cn).

§Corresponding author. Institute of Natural Sciences, School of Mathematical Sciences, and MOE-LSC, Shanghai Jiao Tong University, Shanghai 200240, China (xqzhang@sjtu.edu.cn).

relation between two modality images and develop a joint reconstruction model, rather than restore images independently. Typically, the major drawback of PET is low signal-to-noise ratio and low spatial resolution in the visualized image [35]. On the other hand, MRI provides an anatomical image with superior spatial resolution, whereas the so-called k-space data acquisition process is so time consuming that there has been increasing demand for methods reducing the amount of acquired data without degrading the image quality [43]. However, when the k-space data is undersampled, the Nyquist sampling criterion is violated, and this inevitably leads to the aliasing in the reconstructed image [43]. Nevertheless, through the joint reconstruction process, the complementary information in MRI and PET images can be borrowed to improve restoration quality, especially in the case of degraded and/or incomplete data. In this paper, we only consider the two dimensional PET-MRI joint reconstruction, as the same modeling can be easily applied to the three dimensional cases, and even to other joint reconstruction problems such as SPECT-CT.

Let u_1 and u_2 denote the two dimensional PET and MRI images, and let A and $\mathcal{F}_p$ denote the acquisition process operator of PET and MRI, respectively. The acquired PET data are assumed to be an instance of the Poisson distributed vector with expectation $Au_1 + c$, where c denotes the mean number of background counts [9, 26]. Meanwhile, the noise in MRI can be modeled as a complex white Gaussian noise [26]. Then, by the maximum likelihood principle, the PET-MRI joint reconstruction problem is formulated as the following minimization problem:

$$\min_{u_1\in\mathcal{C}_1, u_2\in\mathcal{C}_2} \langle 1, Au_1 + c\rangle - \langle f, \ln(Au_1 + c)\rangle + \frac{\kappa}{2}\|\mathcal{F}_p u_2 - g\|_2^2 + \mathcal{R}(u_1, u_2), \tag{1}$$

where $\mathcal{C}_1$, $\mathcal{C}_2$ are feasible sets, κ is a positive constant related to the noise level of MRI data, and $\mathcal{R}(u_1, u_2)$ is a nonseparable regularization term for u_1 and u_2.

In relation to the joint reconstruction, the above model (1) naturally raises the question of what is the proper choice of $\mathcal{R}(u_1, u_2)$. As the function cannot be independent from structure [26, 45], it is reasonable to assume that the images to be reconstructed have *structural correlation* [26]. In the literature, the structure of an image is mathematically modeled as a singularity such as edge and ridge, and the idea of structural correlation based joint reconstruction was first explored in the color image processing as the color channels in general share the singularities [8, 25, 29, 36, 47, 50, 53]. The successful examples include the vectorial total variation [8, 29] and the parallel level set (PLS) [25]. Motivated by the works on the color image processing, the authors in [26] recently applied joint total variation (JTV) [26, 31] and PLS priors to the PET-MRI joint reconstruction. Their analysis shows that the Gâteaux derivatives of both JTV and PLS lead to the nonlinear, inhomogeneous diffusion, which promotes the edge-enhancing effect. However, even though both JTV and PLS methods showed significant improvements over the individual reconstruction methods, the regularization term $\mathcal{R}(u_1, u_2)$ has to be further improved as the artifacts appear from the other modality image when the images have complex structures.

Wavelet frames are known to be effective in capturing image singularities from degraded images. This is due to the multiscale structure of the wavelet frame systems, short supports of the framelet functions with varied vanishing moments, and the presence of both low pass and

high pass filters in the wavelet frame filter banks, which are desirable in sparsely approximating images [24]. In addition, due to the redundancy of systems, wavelet frames are more robust to errors than the (bi)orthogonal wavelet bases in [44]. Finally, the special structure of filter banks enables the construction of a so-called *data driven tight frame* to provide a better sparse approximation of a given input image adaptively [14, 54, 56, 57], with the computational efficiency over the K-SVD method in [1]. The successful applications in the medical image reconstruction can be found in the sparse angle X-ray CT reconstruction (wavelet frames in [21] and data driven tight frames in [56]), the limited view CT reconstruction [19], the sparse MRI reconstruction [40], and the data driven tight frame based CT reconstruction in [57], etc.

In this paper, we propose a new tight frame (wavelet frame and data driven tight frame) based PET-MRI joint reconstruction model. Our proposed model is also based on the observation that two modality images have a strong correlation in the image singularities. As the singularities are sparsely captured by the (data driven) wavelet tight frame, our model enforces the *joint sparsity* [28] of frame coefficients to exploit such a structural correlation. Related examples can be found in the previous work in [54], where the authors use the *vector* ℓ_0 *norm* of frame coefficients to construct a multichannel data driven tight frame enforcing the structural correlation between image channels. Indeed, their applications to color image denoising and joint color-depth image reconstruction show the improvements over the independent reconstruction. Motivated by their success in color image processing, we further explore the application in the PET-MRI joint reconstruction. More precisely, using the vector ℓ_0 norm of frame coefficients, we propose a balanced approach [10, 12, 17] for the joint reconstruction from the data with Poisson noise (PET) and the undersampled k-space data (MRI). Finally, as it can be observed that the two modality images have different regularity, we expect the use of a balanced approach will help suppress the artifacts which the existing models in [26] may introduce.

The rest of this paper is organized as follows. In section 2, we review the basic concepts of wavelet frame and data driven tight frame. In section 3, we introduce our tight frame based PET-MRI joint reconstruction model, followed by the proximal alternating minimization algorithm, and the convergence analysis is given at the end of this section. In section 4, we present some experimental results, and the concluding remarks are given in section 5.

2. Preliminaries and related works.

2.1. Wavelet tight frames.

In this subsection, we briefly introduce the concept of tight frames and wavelet tight frames. Interested readers may consult [11, 22, 23, 46, 48] for details. Let $\mathcal{H}$ be a Hilbert space equipped with the inner product $\langle\cdot,\cdot\rangle$ and the norm $\|\cdot\|$. A sequence $\{\varphi_n : n \in \mathbb{Z}\} \subseteq \mathcal{H}$ is called a tight frame on $\mathcal{H}$ if

$$\|u\|^2 = \sum_{n\in\mathbb{Z}} |\langle u, \varphi_n\rangle|^2 \qquad \text{for all} \quad u \in \mathcal{H}. \tag{2}$$

Given $\{\varphi_n : n \in \mathbb{Z}\} \subseteq \mathcal{H}$, the analysis operator $W : \mathcal{H} \to \ell_2(\mathbb{Z})$ is defined as

$$u \in \mathcal{H} \mapsto Wu = \{\langle u, \varphi_n\rangle : n \in \mathbb{Z}\} \in \ell_2(\mathbb{Z}),$$

and the synthesis operator $W^T : \ell_2(\mathbb{Z}) \to \mathcal{H}$ is defined as

$$v \in \ell_2(\mathbb{Z}) \mapsto W^T v = \sum_{n\in\mathbb{Z}} v[n]\varphi_n \in \mathcal{H}.$$

Thus, $\{\varphi_n : n \in \mathbb{Z}\}$ is a tight frame on $\mathcal{H}$ if and only if $W^T W = I$ with $I : \mathcal{H} \to \mathcal{H}$ being the identity. This means that for a given tight frame $\{\varphi_n : n \in \mathbb{Z}\}$, we have the following canonical expression

$$u = \sum_{n\in\mathbb{Z}} \langle u, \varphi_n\rangle \varphi_n,$$

where $Wu = \{\langle u, \varphi_n\rangle : n \in \mathbb{Z}\}$ is called the canonical tight frame coefficients. Hence, the tight frames are generalizations of orthonormal bases to the redundant systems. In fact, a tight frame is an orthonormal basis if and only if $\|\varphi_n\| = 1$ for all $n \in \mathbb{Z}$.

One of the most widely used class of tight frames is the discrete wavelet frame generated by a set of finitely supported filters $\{q_1, \ldots, q_m\}$. In this paper, we only discuss the undecimated wavelet frame, which is also known as the translation invariant wavelet frame transform. Given $q \in \ell_1(\mathbb{Z})$, define a convolution operator $\mathcal{S}_q : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$ by

$$(\mathcal{S}_q u)[n] = (q * u)[n] = \sum_{k\in\mathbb{Z}} q[n-k]u[k] \qquad \text{for} \quad u \in \ell_2(\mathbb{Z}).$$

Given a set of finitely supported filters $\{q_1, \ldots, q_m\}$, define the analysis operator W and the synthesis operator W^T, respectively, by

$$W = \left[\mathcal{S}^T_{q_1[-\cdot]}, \mathcal{S}^T_{q_2[-\cdot]}, \ldots, \mathcal{S}^T_{q_m[-\cdot]}\right]^T, \tag{3}$$

$$W^T = \left[\mathcal{S}_{q_1}, \mathcal{S}_{q_2}, \ldots, \mathcal{S}_{q_m}\right]. \tag{4}$$

Then, the rows of W form a tight frame on $\ell_2(\mathbb{Z})$ if and only if $W^T W = I$, i.e., the filters $\{q_1, \ldots, q_m\}$ satisfy one of the so-called *unitary extension principle* condition [32]:

$$\sum_{l=1}^{m}\sum_{k\in\mathbb{Z}} q_l[n+k]q_l[k] = \delta_n = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{if } n \neq 0. \end{cases} \tag{5}$$

Once the one dimensional filters generate a wavelet tight frame on $\ell_2(\mathbb{Z})$, the higher dimensional wavelet tight frame could be obtained via the tensor product of the one dimensional filters.

2.2. Data driven tight frames. Even though the tensor product wavelet frame based approach is simple to implement and able to achieve sparse representation of a piecewise smooth image, the major disadvantage is that these framelets mostly focus on singularities

along the horizontal and vertical directions [14, 39]. For example, when an image has complex geometries such as directional textures, the tensor product wavelet frame coefficients may not be sparse enough [14].

In order to explore the adaptive sparse approximation of images, various data driven methods have been proposed (e.g., [1, 14]). Among these methods, one promising transformation is the so-called data driven tight frame [14]. It aims to construct a tight frame which sparsely approximates a given image u adaptively. The tight frame $W = W(q_1, \dots, q_{r^2})$ is generated by finitely supported real valued filters $\{q_1, \dots, q_{r^2}\}$ satisfying (5). More concretely, given an image u, the data driven tight frame is obtained via solving the following minimization,

$$\min_{v,W} \ \|v - Wu\|_2^2 + \lambda^2 \|v\|_0 \qquad \text{subject to} \qquad W^T W = I, \tag{6}$$

where the ℓ_0 norm $\|v\|_0$ encodes the number of nonzero entries in the coefficient vector v. To solve (6), we first reformulate it in the following way. Reshape all $r \times r$ patches of u into $G \in \mathbb{R}^{r^2 \times p}$, where p denotes the number of total patches. Let $D \in \mathbb{R}^{r^2 \times r^2}$ be the matrix generated by concatenating filters $\{q_1, \dots, q_{r^2}\}$ into column vectors $\{\vec{q}_1, \dots, \vec{q}_{r^2}\}$. Denote $V \in \mathbb{R}^{r^2 \times p}$ as the frame coefficients. Hence, we have

$$\begin{aligned} u \Leftrightarrow G &= (\vec{g}_1, \dots \vec{g}_p) \in \mathbb{R}^{r^2 \times p}, \\ W \Leftrightarrow D &= (\vec{q}_1, \dots, \vec{q}_{r^2}) \in \mathbb{R}^{r^2 \times r^2}, \\ v \Leftrightarrow V &= (\vec{v}_1, \dots, \vec{v}_p) \in \mathbb{R}^{r^2 \times p}. \end{aligned}$$

Under this setting, (6) is equivalent to

$$\min_{V,D} \ \left\|V - D^T G\right\|_F^2 + \lambda^2 \|V\|_0 \qquad \text{subject to} \qquad DD^T = I, \tag{7}$$

where $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. For solving (7), the alternating minimization with closed form solutions is presented in [14], and the proximal alternating minimization (PAM) scheme with global convergence property is proposed in [4].

Recently, such a patch-based adaptive construction of tight frame filters has been extended to the following data driven tight frame for multichannel imaging in [54]:

$$\begin{aligned} &\min_{v,\{W_i\}_{i=1}^c} \ \sum_{i=1}^{c} w_i \|W_i u_i - v_i\|_2^2 + \lambda \|v\|_{2,0} \\ &\text{subject to} \ \ W_i^T W_i = I, \quad i = 1, \dots, c. \end{aligned} \tag{8}$$

Here, W_i, u_i, and v_i for $i = 1, 2, \dots, c$, respectively, denote the tight frame filters, ith channel image, and frame coefficients, and $\|v\|_{2,0} = \left|\{j : |v_1[j]|^2 + \cdots + |v_c[j]|^2 \neq 0\}\right|$ with $v = (v_1, \dots, v_c)$ is the vector ℓ_0 norm to promote the joint sparsity. This vector ℓ_0 norm $\|\cdot\|_{2,0}$ encodes the structural correlation among the channel images, which demonstrates the advantages in multichannel image restoration [54].

3. Model and algorithm.

3.1. PET-MRI joint reconstruction model. Let u_1 and u_2, respectively, denote the PET and MRI images to be reconstructed. In the literature, the observed PET data $f \in \mathbb{R}_+^M$ are corrupted by the Poisson noise while the MRI data (or k-space data) $g \in \mathbb{C}^L$ are corrupted by white complex Gaussian noise. Assuming that the noise is independent for each pixel, we have

$$P(f|u_1) = \prod_{j=1}^{M} \frac{(Au_1 + c)[j]^{f[j]} e^{-(Au_1+c)[j]}}{f[j]!},$$

where the PET forward operator $A : \mathbb{R}^N \to \mathbb{R}^M$ is, in general, modeled as the discrete attenuated Radon transform [5, 9], and c denotes the mean number of background counts. Meanwhile, for MRI, we have

$$P(g|u_2) \propto \prod_{j=1}^{L} \exp\left\{ -\frac{|(\mathcal{F}_p u_2)[j] - g[j]|^2}{2\sigma^2} \right\},$$

where the MRI forward operator $\mathcal{F}_p : \mathbb{R}^N \to \mathbb{C}^L$ is modeled as the unitary discrete Fourier transform followed by a projection onto the measured frequencies, and σ is the standard deviation of noise in k-space data. Given that f and g are conditionally independent on u_1 and u_2, we have

$$P(f, g|u_1, u_2) = P(f|u_1, u_2)P(g|u_1, u_2) = P(f|u_1)P(g|u_2). \tag{9}$$

Then, together with $P(u_1, u_2) \propto \exp(-\mathcal{R}(u_1, u_2))$ and the Bayes' rule, the maximum a posteriori is equivalent to the minimization problem (1).

Let v_1 and v_2 denote the wavelet frame coefficients of u_1 and u_2 under a given wavelet frame transform W, respectively. As the singularities of u_1 and u_2 are mostly correlated, we assume that the sparsity of v_1 and v_2 is correlated. Exploiting the idea of joint sparsity in [54], we propose our joint sparsity tight frame (JSTF) PET-MRI joint reconstruction model as follows:

$$\min_{u_1 \in \mathcal{C}_1, u_2 \in \mathcal{C}_2, v} \Phi_1(u_1) + \Phi_2(u_2) + \frac{\mu_1}{2}\|Wu_1 - v_1\|_2^2 + \frac{\mu_2}{2}\|Wu_2 - v_2\|_2^2 + \lambda\|v\|_{2,0}, \tag{10}$$

where

$$\Phi_1(u_1) = \langle 1, Au_1 + c\rangle - \langle f, \ln(Au_1 + c)\rangle,$$
$$\Phi_2(u_2) = \frac{\kappa}{2}\|\mathcal{F}_p u_2 - g\|_2^2,$$

and $\kappa = 1/\sigma^2$. Here, $\mathcal{C}_1 = [0, a_1]^N$ and $\mathcal{C}_2 = [0, a_2]^N$, respectively, denote feasible sets which reflect the physical properties of PET and MRI images.

As each modality image may contain information on the different features despite the structural correlation, we may have to actively learn the joint sparse approximation of u_1 and

u_2 [54]. Hence, we also propose the following joint sparsity based data driven tight frame (JSDDTF) joint reconstruction model

$$\begin{aligned} &\min_{\substack{u_1\in\mathcal{C}_1,u_2\in\mathcal{C}_2,\\ W_1,W_2,v}} \Phi_1(u_1)+\Phi_2(u_2)+\frac{\mu_1}{2}\|W_1u_1-v_1\|_2^2+\frac{\mu_2}{2}\|W_2u_2-v_2\|_2^2+\lambda\|v\|_{2,0} \\ &\text{subject to}\quad W_i^TW_i=I,\qquad i=1,2, \end{aligned} \tag{11}$$

where the tight frames W_1 and W_2 are learned from the two modality images u_1 and u_2, respectively.

In the literature, the joint sparsity of tight frame coefficients can also be achieved via the following joint analysis (JAnal) based approach

$$\min_{u_1\in\mathcal{C}_1,u_2\in\mathcal{C}_2} \Phi_1(u_1)+\Phi_2(u_2)+\lambda\|Wu\|_{2,1} \tag{12}$$

with

$$\|Wu\|_{2,1}=\left\|\left(\|Wu_1\|_2^2+\|Wu_2\|_2^2\right)^{1/2}\right\|_1,\qquad u=(u_1,u_2).$$

Indeed, under properly chosen parameters, the above JAnal model (12) can be viewed as a discretization of a variational model including the following JTV model,

$$\min_{u_1\in\mathcal{C}_1,u_2\in\mathcal{C}_2} \Phi_1(u_1)+\Phi_2(u_2)+\lambda\left\|\left(\|\nabla u_1\|_2^2+\|\nabla u_2\|_2^2\right)^{1/2}\right\|_1, \tag{13}$$

while achieving the restoration results superior to the standard finite difference discretization of JTV (see, e.g., [13] for the related details). Meanwhile, our model takes the balanced approach (e.g., [10, 12, 17]) as a special case; by fixing $u_1=W^Tv_1$ and $u_2=W^Tv_2$, the JSTF model (10) reduces to

$$\begin{aligned} \min_{v,W^Tv_1\in\mathcal{C}_1,W^Tv_2\in\mathcal{C}_2}\ &\Phi_1\left(W^Tv_1\right)+\Phi_2\left(W^Tv_2\right)\\ &+\frac{\mu_1}{2}\left\|\left(I-WW^T\right)v_1\right\|_2^2+\frac{\mu_2}{2}\left\|\left(I-WW^T\right)v_2\right\|_2^2+\lambda\|v\|_{2,0}, \end{aligned}$$

and similar reasoning can be applied to the JSDDTF model (11) as well, by fixing $u_i=W_i^Tv_i$ for $i=1,2$.

When $\mu_1,\ \mu_2=\infty$, the JSTF model (10) can be rewritten as the following nonconvex variant of the JAnal model:

$$\min_{u_1\in\mathcal{C}_1,u_2\in\mathcal{C}_2} \Phi_1(u_1)+\Phi_2(u_2)+\lambda\|Wu\|_{2,0} \tag{14}$$

as $v_i=Wu_i$ for $i=1,2$. Hence, the main difference between (10) and the JAnal model

(12) lies in the contribution of $\|Wu_i - v_i\|_2^2$, which measures the distance between v_i (frame coefficients) and Wu_i (canonical coefficients of u_i). Since the canonical coefficients, in general, may not be sparse enough due to the complex geometry of the human body, we use the term $\|Wu_i - v_i\|_2^2$ to provide the flexibility in the sparse approximation of two modality images. Most importantly, under some mild conditions on the tight frame W, the canonical coefficients not only provide the information on the image singularities, but their magnitude also reflects the regularity of an image (see [7] for details). This means that, when the JAnal model (12) is used, we implicitly assume that the two modality images should have the same regularity, as well as the exact coincidence of singularities. However, even though the singularities of two modality images have a strong correlation, they may not exactly coincide with each other as different modality images reflect different physical properties of a human body. In addition, when two images have different regularities, the JAnal model can cause the singularities in one modality image to affect the smooth region in the other even if their singularities coincide. As a consequence, the JAnal model may introduce the artifacts in the reconstructed images. In contrast, μ_1 and μ_2 in our proposed model play a role of balancing the structural correlation of two modality images and their different regularity. As we can observe that MRI images are less noisy compared to PET images, in our proposed model, μ_2 is chosen to be larger than μ_1, and we expect that this different choice of μ_i will help to suppress such artifacts. Finally, the same arguments are applied to the JSDDTF model (11) as it has the same modeling philosophy as the JSTF model (10) except that we learn two tight frames adaptively from u_1 and u_2.

3.2. Alternating minimization algorithm. We propose the PAM algorithms [2] to solve both (10) and (11). As the two schemes are similar except that W_1 and W_2 are additionally updated in (11), we only consider the algorithm for solving (11). Given initializations u_1^0 and u_2^0, the initializations of v, W_1, and W_2 are obtained via

$$\begin{aligned} &\min_{v,W_1,W_2} \ \mu_1 \left\|W_1 u_1^0 - v_1\right\|_2^2 + \mu_2 \left\|W_2 u_2^0 - v_2\right\|_2^2 + \widetilde{\lambda}\|v\|_{2,0} \\ &\text{subject to} \ \ W_i^T W_i = I, \quad i = 1, 2, \end{aligned} \tag{15}$$

using the algorithm in [14]. After the initializations, we optimize $\{u_1, u_2, v, W_1, W_2\}$ by solving the model (11) alternatively. The full details are described in Algorithm 1. Note that we consider the subproblems separately whenever they are separable.

It is easy to see that each problem in (16) is strongly convex and smooth, and there are numerous algorithms to solve it. In our algorithm, u_1^{k+1} is updated using the projected scaled gradient method [34] which is a special case of the preconditioned alternating projection algorithm [37]; Let $u_1^{k,0} = u_1^k$. For $j = 0, 1, 2, \ldots,$

$$\begin{aligned} M^j &= \operatorname{diag}\left(u^j / A^T 1\right), \\ u_1^{j+1/2} &= u_1^j - \rho_1^j M^j \left[A^T\left(1 - \frac{f}{Au_1^j + c}\right) + \mu_1\left(u_1^j - \left(W_1^k\right)^T v_1^k\right) + \alpha_1^k\left(u_1^j - u_1^k\right)\right], \\ u_1^{j+1} &= \min\left\{\max\left(u_1^{j+1/2}, 0\right), a_1\right\}. \end{aligned}$$

Algorithm 1 PAM Algorithm for (11).

Initialization: u_1^0, u_2^0, v^0, W_1^0, W_2^0

for $k = 0, 1, 2, \cdots$ **do**

(1) Optimize u_1 and u_2:

$$
\begin{aligned}
u_1^{k+1} &= \underset{u_1 \in \mathcal{C}_1}{\operatorname{argmin}} \ \Phi_1(u_1) + \frac{\mu_1}{2} \left\| W_1^k u_1 - v_1^k \right\|_2^2 + \frac{\alpha_1^k}{2} \left\| u_1 - u_1^k \right\|_2^2, \\
u_2^{k+1} &= \underset{u_2 \in \mathcal{C}_2}{\operatorname{argmin}} \ \Phi_2(u_2) + \frac{\mu_2}{2} \left\| W_2^k u_2 - v_2^k \right\|_2^2 + \frac{\alpha_2^k}{2} \left\| u_2 - u_2^k \right\|_2^2.
\end{aligned} \tag{16}
$$

(2) Optimize W_1 and W_2:

$$
\begin{aligned}
W_1^{k+1} &= \underset{W_1^T W_1 = I}{\operatorname{argmin}} \ \frac{\mu_1}{2} \left\| W_1 u_1^{k+1} - v_1^k \right\|_2^2 + \frac{\beta_1^k}{2} \left\| W_1 - W_1^k \right\|_2^2, \\
W_2^{k+1} &= \underset{W_2^T W_2 = I}{\operatorname{argmin}} \ \frac{\mu_2}{2} \left\| W_2 u_2^{k+1} - v_2^k \right\|_2^2 + \frac{\beta_2^k}{2} \left\| W_2 - W_2^k \right\|_2^2.
\end{aligned} \tag{17}
$$

(3) Optimize $v = (v_1, v_2)$:

$$
\begin{aligned}
v^{k+1} = \underset{v}{\operatorname{argmin}} \ & \lambda \|v\|_{2,0} + \frac{\mu_1}{2} \left\| v_1 - W_1^{k+1} u_1^{k+1} \right\|_2^2 \\
& + \frac{\mu_2}{2} \left\| v_2 - W_2^{k+1} u_2^{k+1} \right\|_2^2 + \frac{\gamma^k}{2} \left\| v - v^k \right\|_2^2.
\end{aligned} \tag{18}
$$

end for

Similarly, u_2^{k+1} is updated using the projected gradient method [42]. Let $u_2^{k,0} = u_2^k$. For $j = 0, 1, 2, \ldots$

$$
u_2^{j+1/2} = u_2^j - \rho_2^j \left[\kappa \mathcal{F}_p^* \left(\mathcal{F}_p u_2^j - g \right) + \mu_2 \left(u_2^j - \left(W_2^k \right)^T v_2^k \right) + \alpha_2^k \left(u_2^j - u_2^k \right) \right],
$$

$$
u_2^{j+1} = \min \left\{ \max \left(u_2^{j+1/2}, 0 \right), a_2 \right\},
$$

where $\mathcal{F}_p^*$ is the conjugate transpose of $\mathcal{F}_p$. In any case, we omit the outer iteration k on M, u_1, and u_2. Since the constraints $\mathcal{C}_1$ and $\mathcal{C}_2$ are convex, the above iterations converge to the global minimizers for appropriately chosen ρ_1^j and ρ_2^j [34, 42].

To solve (17) and (18), we introduce

$$
\{u_1, W_1, v_1\} \Leftrightarrow \{G_1, D_1, V_1\},
$$

$$
\{u_2, W_2, v_2\} \Leftrightarrow \{G_2, D_2, V_2\},
$$

using $r \times r$ patches of u_1 and u_2. Under the above reformulation, (17) and (18) become

$$
\begin{aligned}
D_1^{k+1} &= \underset{D_1 D_1^T = I}{\operatorname{argmin}} \ \frac{\mu_1}{2} \left\| D_1^T G_1^{k+1} - V_1^k \right\|_F^2 + \frac{\beta_1^k}{2} \left\| D_1 - D_1^k \right\|_F^2, \\
D_2^{k+1} &= \underset{D_2 D_2^T = I}{\operatorname{argmin}} \ \frac{\mu_2}{2} \left\| D_2^T G_2^{k+1} - V_2^k \right\|_F^2 + \frac{\beta_2^k}{2} \left\| D_2 - D_2^k \right\|_F^2,
\end{aligned} \tag{19}
$$

and

$$
\begin{aligned}
V^{k+1} = \underset{V}{\operatorname{argmin}} \ & \lambda \|V\|_{2,0} + \frac{\mu_1}{2} \left\| V_1 - \left(D_1^{k+1} \right)^T G_1^{k+1} \right\|_F^2 \\
& + \frac{\mu_2}{2} \left\| V_2 - \left(D_2^{k+1} \right)^T G_2^{k+1} \right\|_F^2 + \frac{\gamma^k}{2} \left\| V - V^k \right\|_F^2,
\end{aligned} \tag{20}
$$

where $V = (V_1, V_2)$. Hence, to solve (19), we use the following closed form formulas:

$$
\begin{aligned}
D_1^{k+1} &= X_1 Y_1^T, \quad \text{where} \ \ X_1, \Sigma_1, Y_1 \ \ \text{is the SVD of} \ \ G_1^{k+1} \left(V_1^k \right)^T + \frac{\beta_1^k}{\mu_1} D_1^k, \\
D_2^{k+1} &= X_2 Y_2^T, \quad \text{where} \ \ X_2, \Sigma_2, Y_2 \ \ \text{is the SVD of} \ \ G_2^{k+1} \left(V_2^k \right)^T + \frac{\beta_2^k}{\mu_2} D_2^k.
\end{aligned} \tag{21}
$$

The closed form solution to (20) is given by

$$
V^{k+1} = \mathcal{T}_{2\lambda, \mu+\gamma^k} \left[\left(\frac{\mu_1 (D_1^{k+1})^T G_1^{k+1} + \gamma^k V_1^k}{\mu_1 + \gamma^k}, \frac{\mu_2 (D_2^{k+1})^T G_2^{k+1} + \gamma^k V_2^k}{\mu_2 + \gamma^k} \right) \right], \tag{22}
$$

where $\mathcal{T}_{\lambda,\mu}$ with $\mu = (\mu_1, \mu_2)$ is the generalized hard thresholding formula [54] defined as

$$
\left(\mathcal{T}_{\lambda, w}[(U_1, U_2)] \right)_j = \begin{cases} (U_{1,j}, U_{2,j}) & \text{if} \ \ \sum_{i=1}^2 w_i |U_{i,j}|^2 \geq \lambda, \\ 0 & \text{otherwise.} \end{cases} \tag{23}
$$

Here, $(\mathcal{T}_{\lambda,w}[(U_1, U_2)])_j$ and $U_{i,j}$ denote the jth row vector of $\mathcal{T}_{\lambda,w}[(U_1, U_2)]$ and U_i, respectively.

3.3. Convergence analysis. This section is devoted to the convergence analysis of our alternating minimization algorithm. Due to the same reason as the PAM algorithm, we focus on the convergence analysis of sequence $\{(u_1^k, u_2^k, W_1^k, W_2^k, v^k) : k \in \mathbb{N}\}$ in (11) generated by Algorithm 1. To do this, we assume the following:

A1. $f > 0$ and A satisfies

$$
A[i,j] \geq 0, \quad A1 > 0, \quad \text{and} \quad A^T 1 > 0,
$$

where $A[i,j]$ denotes the (i,j)th entry of A.

A2. $c > 0$ is a constant vector.

A3. There exists $0 < L < U$ such that $L \leq \alpha_1^k, \alpha_2^k, \beta_1^k, \beta_2^k, \gamma^k \leq U$ for all $k \in \mathbb{N}$.

We first introduce some basic notation and definitions.

Definition 3.1. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be a proper and lower semicontinuous (lsc) function.*

1. *The domain of f, denoted as* $\mathrm{dom}(f)$*, is defined as*
$$\mathrm{dom}(f) = \{x \in \mathbb{R}^n : f(x) < \infty\}.$$
2. *For each $x \in \mathrm{dom}(f)$, the Fréchet subdifferential of f at x is defined as*
$$\partial_F f(x) = \left\{ s \in \mathbb{R}^n : \liminf_{y \to x} \frac{f(y) - f(x) - \langle s, y - x\rangle}{\|y - x\|} \geq 0 \right\}.$$
If $x \notin \mathrm{dom}(f)$, then we set $\partial_F f(x) = \emptyset$.
3. *The (limiting) subdifferential of f at x is defined as*
$$\partial f(x) = \{s \in \mathbb{R}^n : \exists x_n \ s.t.\ f(x_n) \to f(x) \ \&\ s_n \in \partial_F f(x_n) \to s\}.$$
The domain of ∂f is defined as $\mathrm{dom}(\partial f) = \{x \in \mathbb{R}^n : \partial f(x) \neq \emptyset\}$.
4. *$x \in \mathrm{dom}(f)$ is a critical point of f if $0 \in \partial f(x)$.*

In the literature, the global convergence of the alternating schemes is established on the framework of the Kurdyka–Lojasiewicz (KL) property in [38, 41]. Even though it is challenging to verify whether a given function f satisfies the KL property, there are some functions of special cases which have proven to satisfy the KL property, such as analytic functions and semialgebraic functions (e.g., [2, 3, 6, 55]).

To apply the framework of the KL property to our model, we let
$$x = (x_1, x_2, x_3, x_4, x_5) = (u_1, u_2, W_1, W_2, v)$$
for notational simplicity. Recall $\mathcal{C}_1 = [0, a_1]^N$ and $\mathcal{C}_2 = [0, a_2]^N$, and let
$$\mathcal{D} = \{W \in \mathbb{R}^{r^2 N \times N} : W^T W = I_{r^2 \times r^2}\}.$$
Then we define
$$\begin{aligned} P(x) &= \Phi_1(u_1) + \Phi_2(u_2) + \frac{\mu_1}{2}\|W_1 u_1 - v_1\|_2^2 + \frac{\mu_2}{2}\|W_2 u_2 - v_2\|_2^2, \\ r_i(x_i) &= \imath_{\mathcal{C}_i}(x_i) \qquad i = 1, 2, \\ r_i(x_i) &= \imath_{\mathcal{D}}(x_i) \qquad i = 3, 4, \\ r_5(x_5) &= \lambda \|x_5\|_{2,0}, \end{aligned}$$
where $\imath_{\mathcal{A}}$ is the indicator function of a set $\mathcal{A}$: $\imath_{\mathcal{A}}(x) = 0$ if $x \in \mathcal{A}$, $\imath_{\mathcal{A}}(x) = \infty$ otherwise.

Using the above notation, we reformulate the model (11) as
$$\min_x \ H(x) := P(x) + \sum_{i=1}^{5} r_i(x_i). \tag{24}$$

Note that it is not hard to verify that $\Phi_1(u_1)$ and the indicator functions are analytic, and

$r_5(x_5)$ is a semialgebraic function. Together with that the remaining terms are all polynomials, we can see that the functional H defined as in (24) satisfies the KL property.

Under this reformulation, we can present the following result on the global convergence.

Theorem 3.2. *Assume that* A1–A3 *hold. Let $H(x)$ be the objective function defined as in (24). Then the sequence $\{x^k : k \in \mathbb{N}\}$ generated by Algorithm 1 converges to a critical point of H. Moreover, $\{x^k\}_{k\in\mathbb{N}}$ has the following finite length property:*

$$\sum_{k=0}^{\infty} \left\| x^{k+1} - x^k \right\|_2 < \infty. \tag{25}$$

The proof of Theorem 3.2 is similar to the framework given in [3, Theorem 3.7]. In fact, $H(x)$ defined as in (24) satisfies the KL property, and noting that $\Phi_1(x_1)$ is analytic, proper, lsc, strictly convex, and coercive [27], it is not hard to see that $H(x)$ satisfies the first condition in [3, Theorem 3.7]. (See [3] for the details.) Hence, the proof is completed provided that we show that the sequence $\{x^k : k \in \mathbb{N}\}$ generated by Algorithm 1 is bounded. In fact, since the constraints $\mathcal{C}_1$, $\mathcal{C}_2$, and $\mathcal{D}$ are compact sets, it suffices to show that v^k is bounded.

Lemma 3.3. *Assume that* A1–A3 *hold. Let $H(x)$ be the objective function defined as in (24). For $\{x^k : k \in \mathbb{N}\}$ generated by Algorithm 1, there exist R_1, $R_2 > 0$ such that $\|v_1^k\|_2 \le R_1$ and $\|v_2^k\|_2 \le R_2$ for all $k \ge 0$.*

Proof. Note that the constraint on u_1 and u_2 leads to

$$\left\|u_1^k\right\|_2 \le a_1\sqrt{N} \quad \text{and} \quad \left\|u_2^k\right\|_2 \le a_2\sqrt{N}$$

for all $k \ge 0$. If we choose $R_1 = a_1\sqrt{N}$ and $R_2 = a_2\sqrt{N}$, then the proof is completed by mathematical induction. For $k = 0$, (23) implies

$$\left\|v_1^0\right\|_2 \le \left\|W_1^0 u_1^0\right\|_2 = \left\|u_1^0\right\|_2 \le R_1,$$

$$\left\|v_2^0\right\|_2 \le \left\|W_2^0 u_2^0\right\|_2 = \left\|u_2^0\right\|_2 \le R_2,$$

where the equality comes from the fact that W_1^0 and W_2^0 are tight frames.

For the mathematical induction, we assume that $\|v_1^k\|_2 \le R$ and $\|v_2^k\|_2 \le R$ for $k \ge 0$. Again, by (23), we have

$$\left\|v_1^{k+1}\right\|_2 \le \left\|\frac{\mu_1 W_1^{k+1}u_1^{k+1} + \gamma^k v_1^k}{\mu_1 + \gamma^k}\right\|_2 \le \frac{\mu_1}{\mu_1+\gamma^k}\left\|W_1^{k+1}u_1^{k+1}\right\|_2 + \frac{\gamma^k}{\mu_1+\gamma^k}\left\|v_1^k\right\|_2 \le R_1,$$

$$\left\|v_2^{k+1}\right\|_2 \le \left\|\frac{\mu_2 W_2^{k+1}u_2^{k+1} + \gamma^k v_2^k}{\mu_2 + \gamma^k}\right\|_2 \le \frac{\mu_2}{\mu_2+\gamma^k}\left\|W_2^{k+1}u_2^{k+1}\right\|_2 + \frac{\gamma^k}{\mu_2+\gamma^k}\left\|v_2^k\right\|_2 \le R_2$$

from the fact that W_1^k, $W_2^k \in \mathcal{D}$ for all $k \in \mathbb{N}$. This completes the proof. ∎

4. Numerical results. In this section, we present some experimental results to compare our JSTF model (10) and JSDDFT model (11) with the several existing methods. We choose to compare it with the the following analysis based individual reconstruction models:

$$\min_{u_1\in\mathcal{C}_1} \Phi_1(u_1)+\lambda_1\|Wu_1\|_1, \tag{26}$$

$$\min_{u_2\in\mathcal{C}_2} \Phi_2(u_2)+\lambda_2\|Wu_2\|_1, \tag{27}$$

both of which are solved by the split Bregman method [15], the data driven tight frame (DDTF) individual reconstruction models:

$$\min_{u_1\in\mathcal{C}_1,v_1,W_1} \Phi_1(u_1)+\frac{\mu_1}{2}\|W_1u_1-v_1\|_2^2+\lambda_1\|v_1\|_0 \ \text{ subject to } \ W_1^TW_1=I, \tag{28}$$

$$\min_{u_2\in\mathcal{C}_2,v_2,W_2} \Phi_2(u_2)+\frac{\mu_2}{2}\|W_2u_2-v_2\|_2^2+\lambda_2\|v_2\|_0 \ \text{ subject to } \ W_2^TW_2=I, \tag{29}$$

solved by the PAM algorithm, and the quadratic PLS (QPLS) method in [26]. We also compare it with the JAnal model (12) solved by the split Bregman method, as a replacement for the JTV model (13).

The experiments are conducted with 256×256 PET and MRI images taking values in $[0,1]$, which are available on the Harvard Whole Brain Atlas webpage.[1] Throughout this paper, we set both $\mathcal{C}_1$ and $\mathcal{C}_2$ at $[0,1]^N$. Note that MRI images are classified into the positron density (PD) weighted image, the T1 weighted image, and the T2 weighted image according to the pulse sequence design [30] (see Figure 1). In the following, we perform experiments using different MRI images for different u_2's, which we shall refer to as PET-PD, PET-T1, and PET-T2, respectively. To measure the quality of restored images, we compute the relative error, the peak signal-to-noise ratio (PSNR), and the correlation between the reconstructed image $\widetilde{u}_i$ and the true image u_i, respectively, defined as

$$\begin{aligned}
\mathrm{RelErr}(u_i,\widetilde{u}_i)&=\frac{\|u_i-\widetilde{u}_i\|_2}{\|u_i\|_2},\\
\mathrm{PSNR}(u_i,\widetilde{u}_i)&=-10\log_{10}\frac{\|u_i-\widetilde{u}_i\|_2^2}{N},\\
\mathrm{Corr}(u_i,\widetilde{u}_i)&=\frac{\langle u_i-\overline{u}_i,\widetilde{u}_i-\overline{\widetilde{u}}_i\rangle}{\|u_i-\overline{u}_i\|_2\|\widetilde{u}_i-\overline{\widetilde{u}}_i\|_2},
\end{aligned}$$

where $\overline{u}_i$ and $\overline{\widetilde{u}}_i$ denote the mean of u_i and $\widetilde{u}_i$, respectively.

To synthesize the data, we generate the forward PET operator A as described in [5], and we set $c=1$ in all cases. To generate 256×256 PET data f with Poisson noise, we use

[1] http://www.med.harvard.edu/AANLIB/home.html.

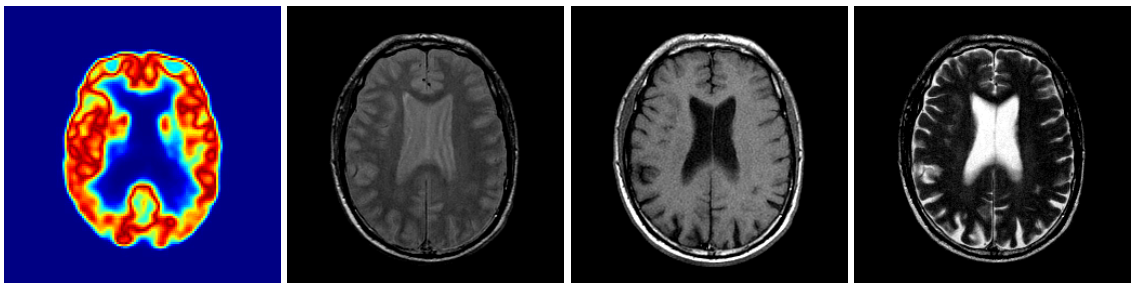

Figure 1. *True PET image (first column) and true MRI images (second to last columns). We use different MRI images for u_2.*

the MATLAB built-in function imnoise(·,'poisson') by scaling $Au_1 + c$ with a suitable factor before applying imnoise, and then scaling it back with the same factor. More precisely,

$$f = \text{factor} * \text{imnoise}((Au_1 + c)/\text{factor}, \text{'poisson'})$$

and we set factor to be $2 * 10^8$ in all cases. Note that the larger "factor" is, the larger noise level is. To generate g, we first generate $\mathcal{F}_p = R_\Lambda \mathcal{F}$, where R_Λ is the projection on the known frequency region Λ, and $\mathcal{F}$ is the unitary discrete Fourier transform. Here, we choose R_Λ to be the sampling along the 30 radial lines and the 10% random sampling described in [20]. The Gaussian noise with standard deviation 0.05 is also added to both the real and imaginary parts of $\mathcal{F}_p u_2$.

In all experiments, we choose the same initializations for a fair comparison; u_1^0 is obtained by the expectation-maximization algorithm [49], and u_2^0 is obtained by the inverse discrete Fourier transform with zero filling (see Figure 2). For the individual reconstruction models (26) and (27), the JAnal model (12) as well as our JSTF model (10), we use one level piecewise cubic B-spline wavelet frame transformation. For the JSDDTF model (11) as well as the DDTF individual reconstruction models (28) and (29), we use 8×8 undecimated discrete cosine transform filters [51] for the initial guess of (15). In all cases, we set $\kappa = 1$. In (10), (11), (28), and (29), we set $\mu_1 = 0.05$, $\mu_2 = 1$, $\alpha_1^k = \alpha_2^k = 0.001$, $\beta_1^k = \beta_2^k = \gamma^k = 0.00005$, and $\rho_1^j = \rho_2^j = 0.5$ in all cases. All parameters related to the (joint) sparsity are manually chosen so that we obtain the optimal restoration results, especially the optimal results in the PET image for the joint reconstruction models. In addition, as the wavelet frame systems consist of a low pass filter and high pass filters and the low pass filter coefficients are not sparse in general [13], we follow the convention that we do not penalize the frame coefficients corresponding to the low pass filter.

Table 1 summarizes relative errors, PSNR values, and correlations of the aforementioned five restoration models, and Figures 3, 4, 5, 6, 7, and 8 present visual comparisons of the results. We can see that both JSTF (10) and JSDDTF (11) model consistently outperform both the individual reconstruction models and the existing joint reconstruction models in [26]. This verifies that there exists a correlation on image singularities between the two modality images, and exploiting this correlation results in the better reconstruction results. In addition, compared to the existing joint reconstruction models, we can see that our proposed models

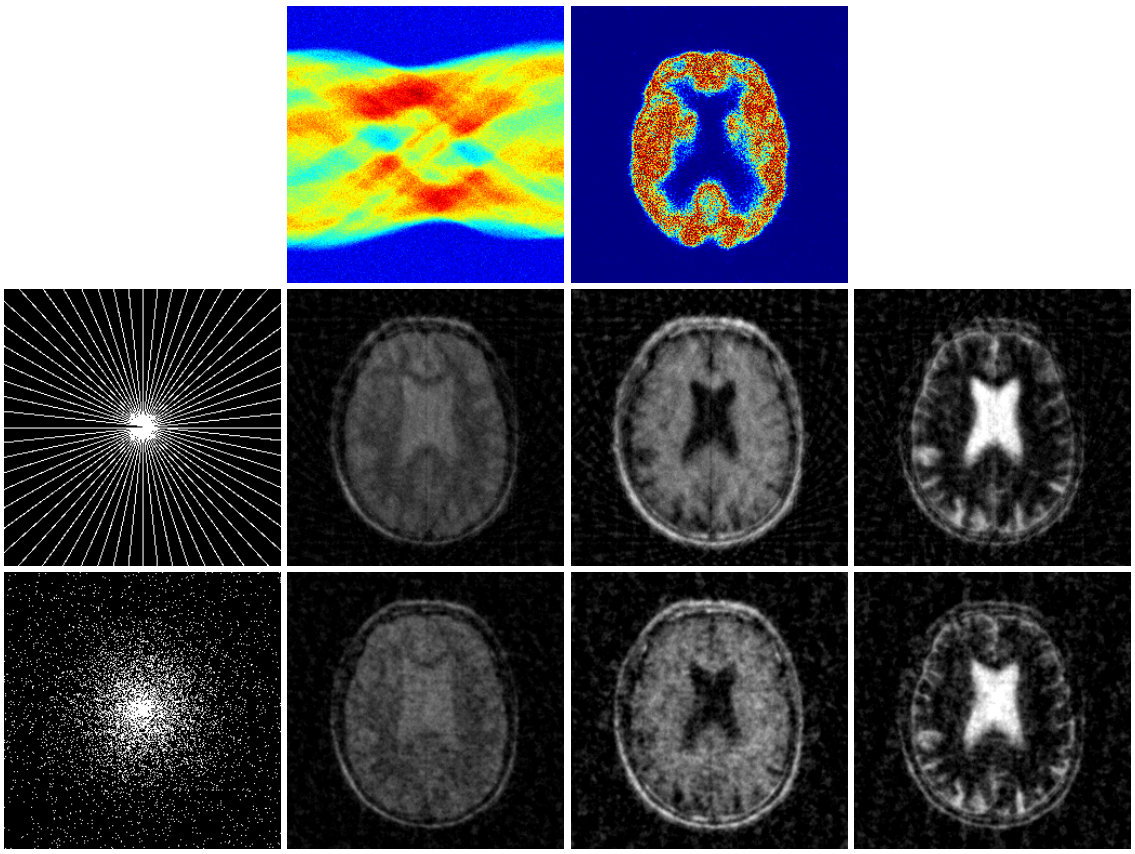

Figure 2. *Synthesized data and initializations. The first row describes the synthesized PET data f (left) and u_1^0 (right). The second row depicts the radial sampling projection R_Λ, followed by u_2^0 for PET-PD, PET-$T1$, and PET-$T2$, respectively. The third row shows the random sampling projection R_Λ, followed by the initializations of u_2^0 for PET-PD, PET-$T1$, and PET-$T2$, respectively.*

introduce fewer artifacts in both modality images, leading to the visual improvements over the existing joint reconstruction methods which are consistent with the improvements in indices. It is worth noting that our proposed model can improve the restoration qualities even using a static wavelet tight frame. This demonstrates that the improvements mainly come from simultaneously considering the different regularities of the different modality images and the joint sparsity. Meanwhile, we can see that the QPLS model and the JAnal model (12), which only take the structural correlation into account, show the degradations in MRI restoration results compared to the independent reconstruction methods due to the artifacts. Finally, we list some zoomed-in views in Figure 9 to illustrate that our models (10) and (11) restore structures better than the existing methods.

5. Conclusions. In this paper, we proposed a joint sparsity based tight frame regularization PET-MRI joint reconstruction model, together with a PAM algorithm. The numerical experiments show that our models (10) and (11) both outperform the existing models in [26]. This performance gain of our proposed models mainly comes from taking the different regular-

Table 1
Comparison of relative errors, PSNR, and correlations.

Radial sampling									
Images		Indices	Initial	Analysis	DDTF	QPLS	JAnal	JSTF	JSDDTF
PET-PD	PET	RelErr	0.3156	0.1112	0.0939	0.1040	0.0945	0.0937	**0.0905**
		PSNR	18.8555	27.9170	29.3820	28.5011	29.3297	29.4054	**29.7059**
		Corr	0.9384	0.9918	0.9942	0.9928	0.9942	0.9943	**0.9946**
	MRI	RelErr	0.3448	0.2148	0.2129	0.2878	0.2732	0.2022	**0.1929**
		PSNR	24.0252	28.1347	28.2127	25.5958	26.0482	28.6616	**29.0692**
		Corr	0.9018	0.9633	0.9648	0.9350	0.9410	0.9678	**0.9707**
PET-T1	PET	RelErr	0.3156	0.1112	0.0939	0.1034	0.0952	0.0934	**0.0893**
		PSNR	18.8555	27.9170	29.3820	28.5453	29.2649	29.4358	**29.8250**
		Corr	0.9384	0.9918	0.9942	0.9929	0.9941	0.9943	**0.9948**
	MRI	RelErr	0.3571	0.1992	0.1809	0.2266	0.2146	0.1753	**0.1663**
		PSNR	19.9312	25.0007	25.8395	23.8804	24.3556	26.1104	**26.5667**
		Corr	0.8967	0.9695	0.9752	0.9601	0.9641	0.9764	**0.9787**
PET-T2	PET	RelErr	0.3156	0.1112	0.0939	0.1069	0.0964	0.0923	**0.0881**
		PSNR	18.8555	27.9170	29.3820	28.2628	29.1616	29.5351	**29.9369**
		Corr	0.9384	0.9918	0.9942	0.9924	0.9940	0.9944	**0.9949**
	MRI	RelErr	0.3907	0.2415	0.2285	0.2678	0.2497	0.2108	**0.2023**
		PSNR	19.9442	24.1204	24.6029	23.2245	23.8311	25.3016	**25.6593**
		Corr	0.9005	0.9635	0.9676	0.9551	0.9608	0.9723	**0.9745**
Random sampling									
Images		Indices	Initial	Analysis	DDTF	QPLS	JAnal	JSTF	JSDDTF
PET-PD	PET	RelErr	0.3156	0.1112	0.0939	0.1027	0.0943	0.0933	**0.0906**
		PSNR	18.8555	27.9170	29.3820	28.6088	29.3456	29.4376	**29.6987**
		Corr	0.9384	0.9918	0.9942	0.9930	0.9942	0.9943	**0.9946**
	MRI	RelErr	0.3366	0.2081	0.2082	0.2729	0.2618	0.2003	**0.1919**
		PSNR	24.2355	28.4134	28.4054	26.0581	26.4169	28.7440	**29.1145**
		Corr	0.9064	0.9657	0.9663	0.9408	0.9452	0.9683	**0.9709**
PET-T1	PET	RelErr	0.3156	0.1112	0.0939	0.1036	0.0950	0.0929	**0.0889**
		PSNR	18.8555	27.9170	29.3820	28.5293	29.2815	29.4772	**29.8588**
		Corr	0.9384	0.9918	0.9942	0.9929	0.9941	0.9943	**0.9948**
	MRI	RelErr	0.3712	0.1952	0.1785	0.2180	0.2093	0.1707	**0.1627**
		PSNR	19.5948	25.1765	25.9556	24.2185	24.5722	26.3422	**26.7584**
		Corr	0.8878	0.9709	0.9759	0.9629	0.9658	0.9777	**0.9797**
PET-T2	PET	RelErr	0.3156	0.1112	0.0939	0.1072	0.0964	0.0928	**0.0882**
		PSNR	18.8555	27.9170	29.3820	28.2321	29.1553	29.4865	**29.9256**
		Corr	0.9384	0.9918	0.9942	0.9924	0.9939	0.9943	**0.9949**
	MRI	RelErr	0.4052	0.2357	0.2270	0.2551	0.2397	0.2089	**0.2016**
		PSNR	19.6260	24.3331	24.6594	23.6451	24.1873	25.3805	**25.6890**
		Corr	0.8925	0.9654	0.9680	0.9591	0.9638	0.9728	**0.9746**

ities of the different modality images into consideration, as well as the structural correlation. Finally, our convergence analysis demonstrates that the sequence generated by our algorithm globally converges to a critical point of the proposed model.

Acknowledgment. We would like to thank the anonymous reviewers for their constructive suggestions and comments that helped tremendously with improving the presentations of this paper.

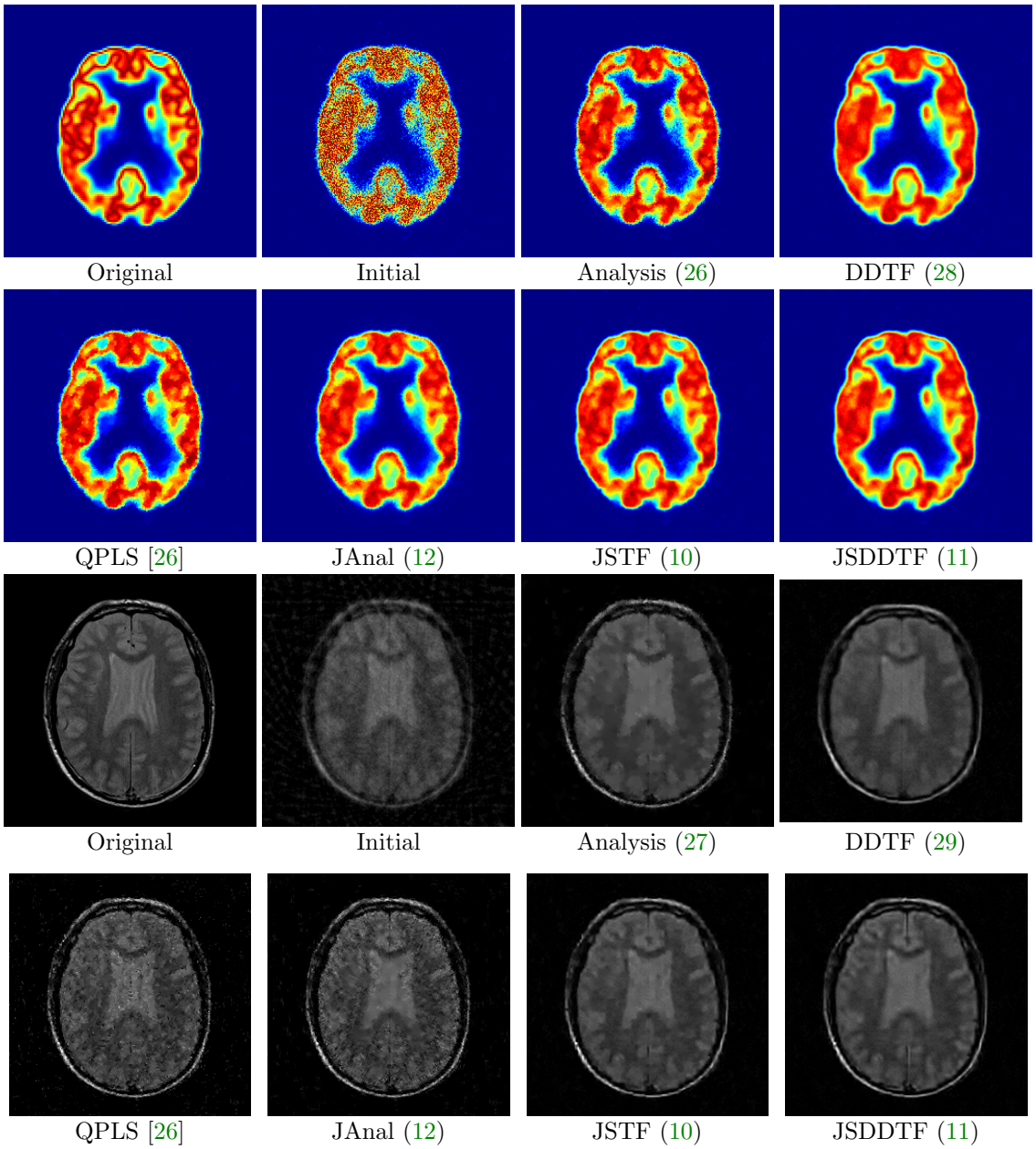

Figure 3. *Visual comparison of PET-PD radial joint reconstruction results. The first and second rows describe the PET images, and the third and fourth rows depict the MRI images.*

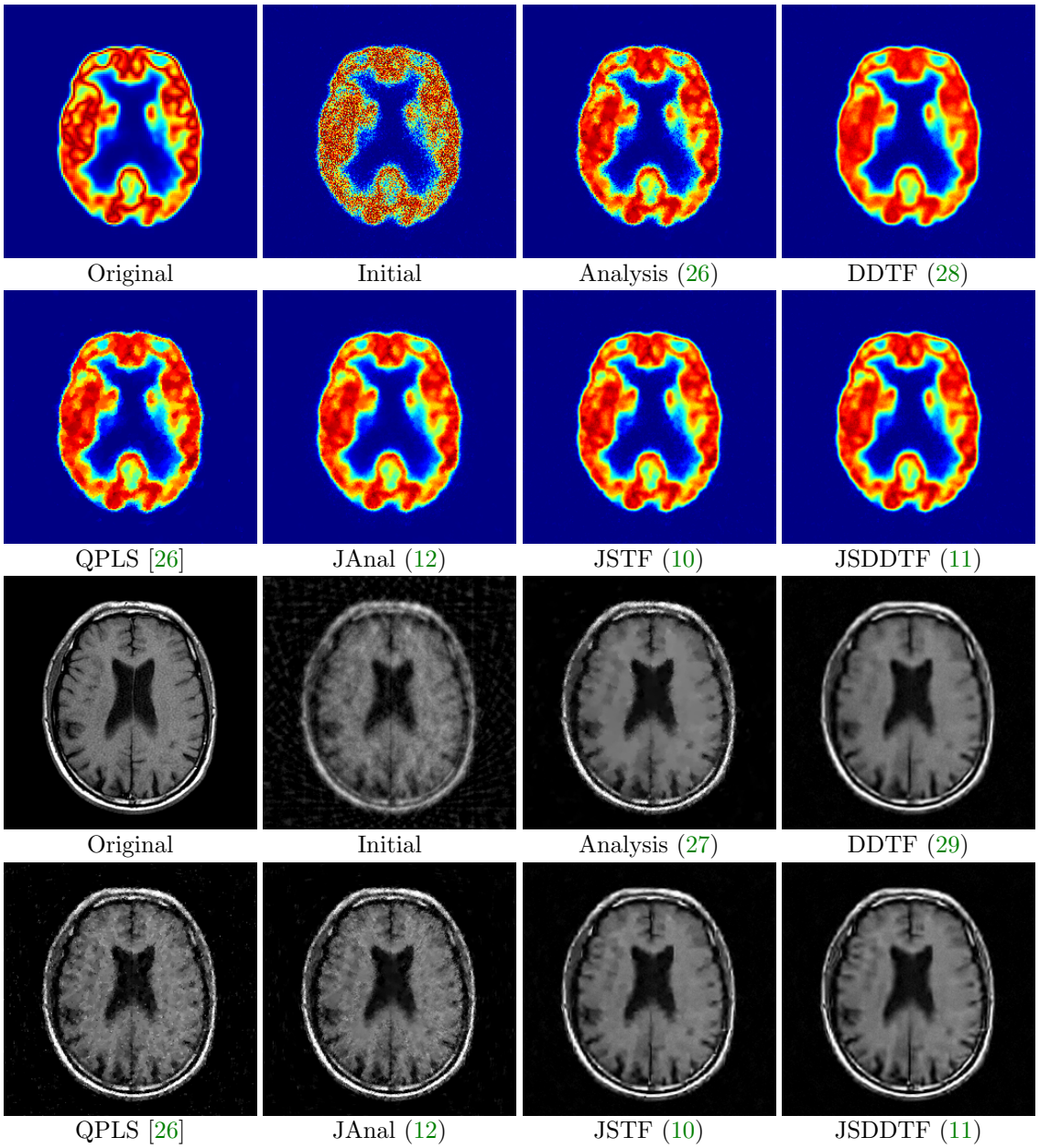

Figure 4. *Visual comparison of PET-T1 radial joint reconstruction results. The first and second rows describe the PET images, and the third and fourth rows depict the MRI images.*

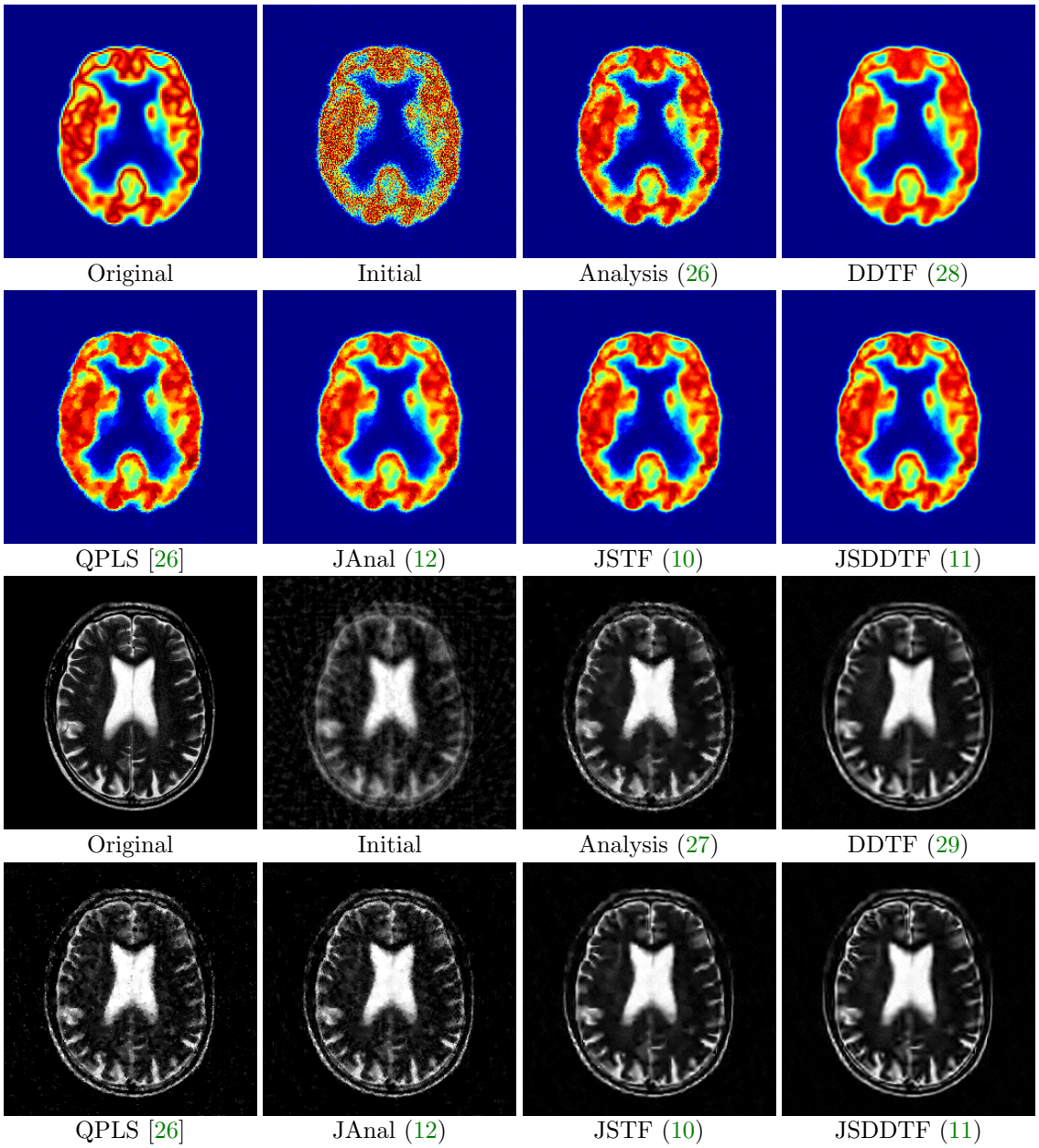

Figure 5. *Visual comparison of PET-T2 radial joint reconstruction results. The first and second rows describe the PET images, and the third and fourth rows depict the MRI images.*

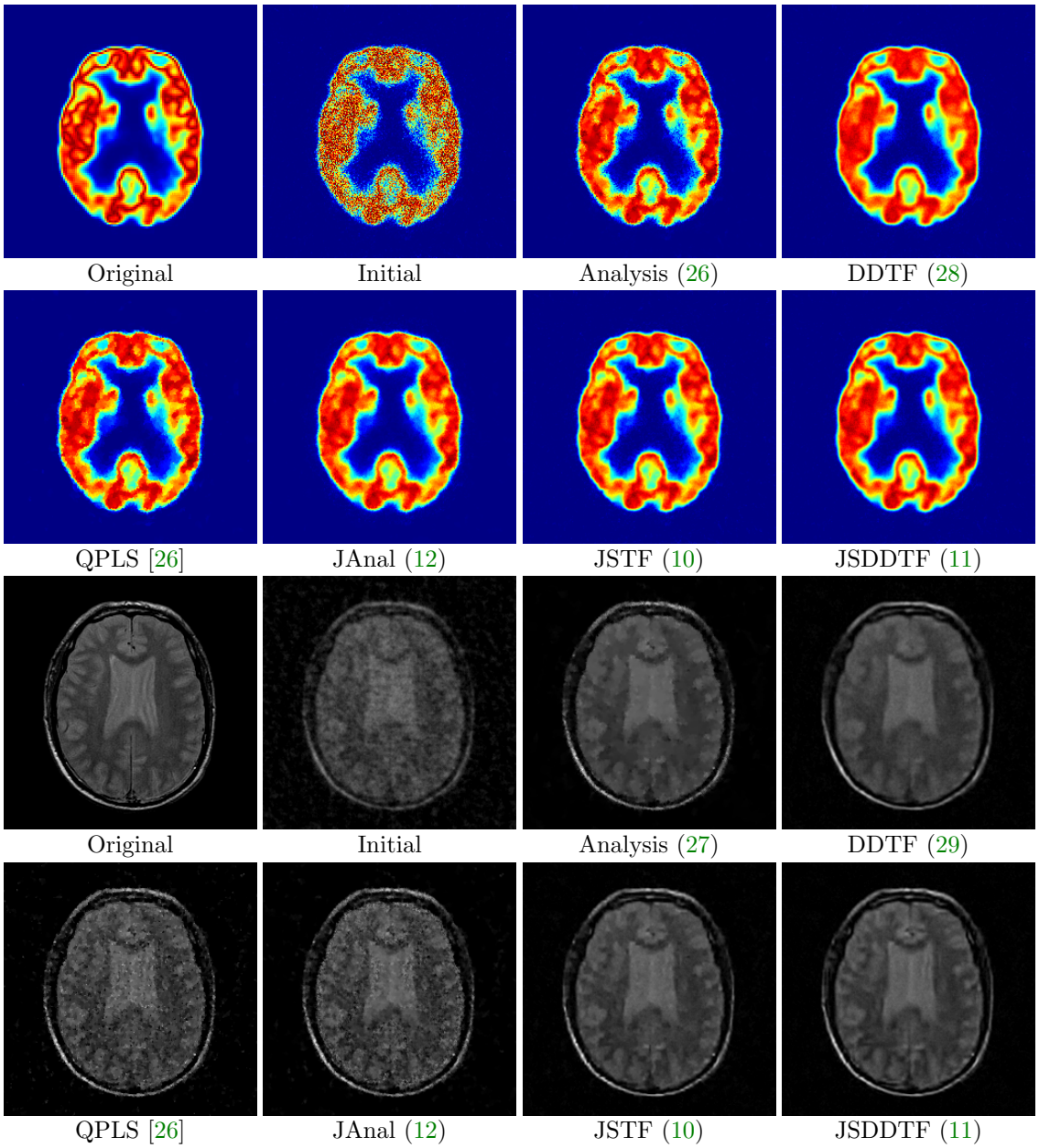

Figure 6. *Visual comparison of PET-PD random joint reconstruction results. The first and second rows describe the PET images, and the third and fourth rows depict the MRI images.*

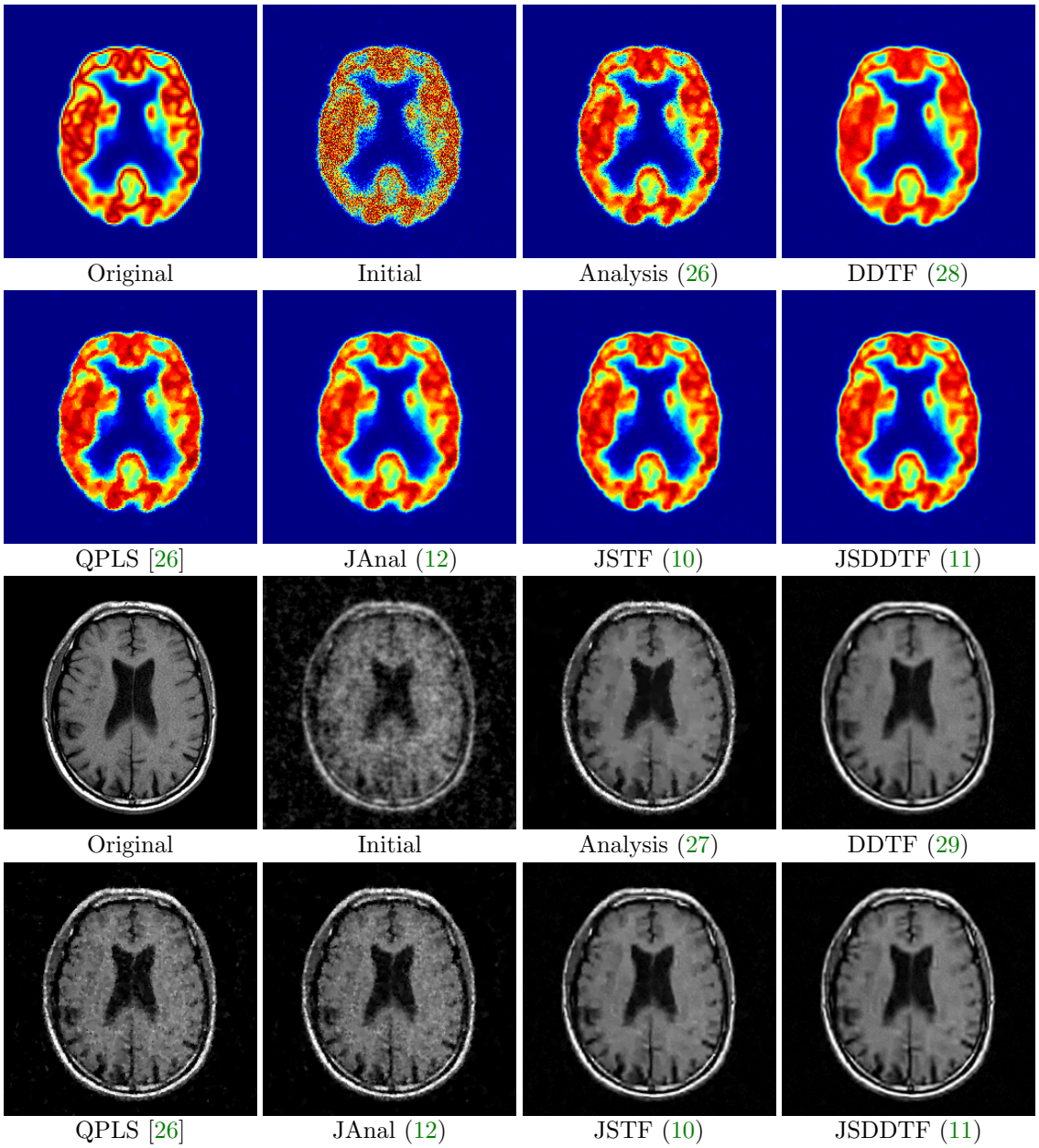

Figure 7. *Visual comparison of PET-T1 random joint reconstruction results. The first and second rows describe the PET images, and the third and fourth rows depict the MRI images.*

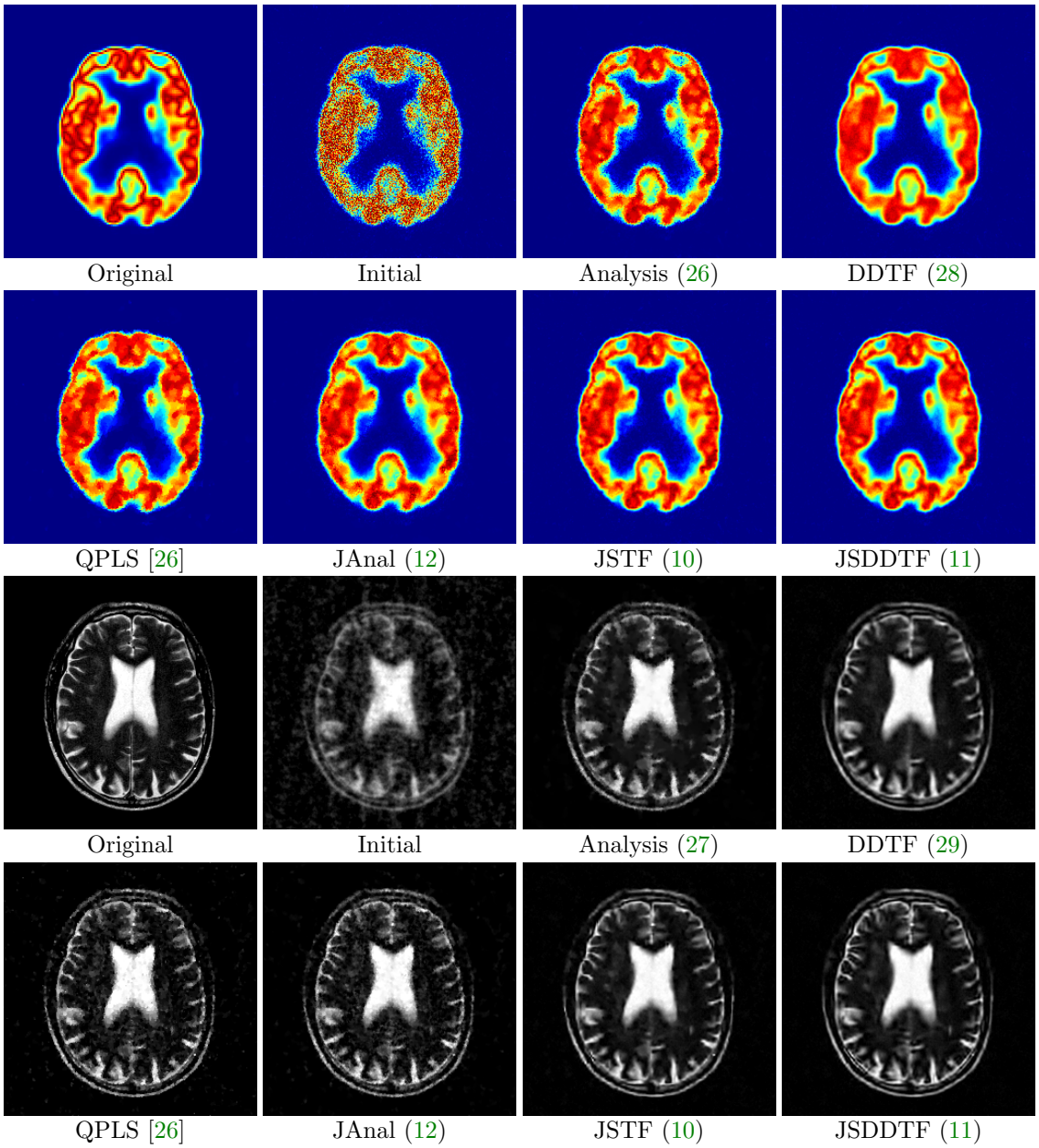

Figure 8. *Visual comparison of PET-T2 random joint reconstruction results. The first and second rows describe the PET images, and the third and fourth rows depict the MRI images.*

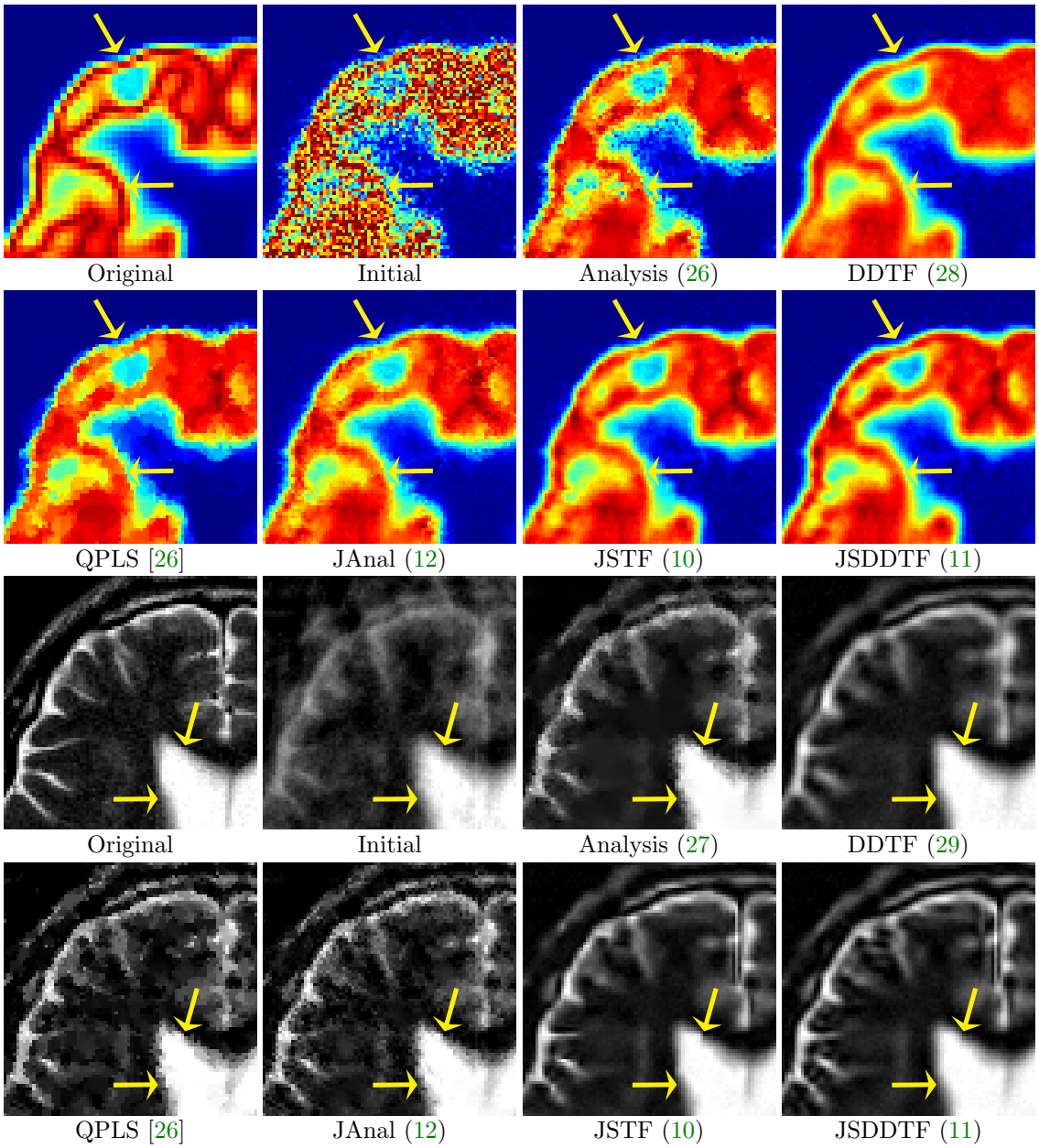

Figure 9. *Zoomed-in views of Figure 5 The first and second rows describe the PET images, and the third and fourth rows depict the MRI images. The yellow arrows indicate the regions worth noticing.*

REFERENCES

[1] M. Aharon, M. Elad, and A. Bruckstein, K-*SVD: An algorithm for designing overcomplete dictionaries for sparse representation*, IEEE Trans. Signal Process., 54 (2006), pp. 4311–4322, https://doi.org/10.1109/TSP.2006.881199.

[2] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, *Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Lojasiewicz inequality*, Math. Oper. Res., 35 (2010), pp. 438–457, https://doi.org/10.1287/moor.1100.0449.

[3] C. Bao, H. Ji, Y. Quan, and Z. Shen, *Dictionary learning for sparse coding: Algorithms and convergence analysis*, IEEE Trans. Pattern Anal. Mach. Intell., 38 (2016), pp. 1356–1369, https://doi.org/10.1109/TPAMI.2015.2487966.

[4] C. Bao, H. Ji, and Z. Shen, *Convergence analysis for iterative data-driven tight frame construction scheme*, Appl. Comput. Harmon. Anal., 38 (2015), pp. 510–523, https://doi.org/10.1016/j.acha.2014.06.007.

[5] E. M. Bardsley, D. Calvetti, and E. Somersalo, *Hierarchical regularization for edge-preserving reconstruction of PET images*, Inverse Problems, 26 (2010), 035010, https://doi.org/10.1088/0266-5611/26/3/035010.

[6] J. Bolte, S. Sabach, and M. Teboulle, *Proximal alternating linearized minimization for nonconvex and nonsmooth problems*, Math. Program., 146 (2014), pp. 459–494, https://doi.org/10.1007/s10107-013-0701-9.

[7] L. Borup, R. Gribonval, and M. Nielsen, *Bi-framelet systems with few vanishing moments characterize Besov spaces*, Appl. Comput. Harmon. Anal., 17 (2004), pp. 3–28, https://doi.org/10.1016/j.acha.2004.01.004.

[8] X. Bresson and T. F. Chan, *Fast dual minimization of the vectorial total variation norm and applications to color image processing*, Inverse Probl. Imaging, 2 (2008), pp. 455–484, https://doi.org/10.3934/ipi.2008.2.455.

[9] M. Burger, J. Müller, E. Papoutsellis, and C. B. Schönlieb, *Total variation regularization in measurement and image space for PET reconstruction*, Inverse Problems, 30 (2014), 105003, https://doi.org/10.1088/0266-5611/30/10/105003.

[10] J. F. Cai, R. H. Chan, L. Shen, and Z. Shen, *Convergence analysis of tight framelet approach for missing data recovery*, Adv. Comput. Math., 31 (2009), pp. 87–113, https://doi.org/10.1007/s10444-008-9084-5.

[11] J. F. Cai, R. H. Chan, and Z. Shen, *A framelet-based image inpainting algorithm*, Appl. Comput. Harmon. Anal., 24 (2008), pp. 131–149, https://doi.org/10.1016/j.acha.2007.10.002.

[12] J. F. Cai, R. H. Chan, and Z. Shen, *Simultaneous cartoon and texture inpainting*, Inverse Probl. Imaging, 4 (2010), pp. 379–395, https://doi.org/10.3934/ipi.2010.4.379.

[13] J. F. Cai, B. Dong, S. Osher, and Z. Shen, *Image restoration: Total variation, wavelet frames, and beyond*, J. Amer. Math. Soc., 25 (2012), pp. 1033–1089, https://doi.org/10.1090/S0894-0347-2012-00740-1.

[14] J. F. Cai, H. Ji, Z. Shen, and G. B. Ye, *Data-driven tight frame construction and image denoising*, Appl. Comput. Harmon. Anal., 37 (2014), pp. 89–105, https://doi.org/10.1016/j.acha.2013.10.001.

[15] J.-F. Cai, S. Osher, and Z. Shen, *Split Bregman methods and frame based image restoration*, Multiscale Model. Simul., 8 (2010), pp. 337–369, https://doi.org/10.1137/090753504.

[16] C. Catana, A. R. Guimaraes, and B. R. Rosen, *PET and MR imaging: The odd couple or a match made in heaven?*, J. Nucl. Med., 54 (2013), pp. 815–824, https://doi.org/10.2967/jnumed.112.112771.

[17] R. H. Chan, T. F. Chan, L. Shen, and Z. Shen, *Wavelet algorithms for high-resolution image reconstruction*, SIAM J. Sci. Comput., 24 (2003), pp. 1408–1432, https://doi.org/10.1137/S1064827500383123.

[18] S. R. Cherry, *Multimodality in vivo imaging systems: Twice the power or double the trouble?*, Annu. Rev. Biomed. Eng., 8 (2006), pp. 35–62, https://doi.org/10.1146/annurev.bioeng.8.061505.095728.

[19] J. Choi, B. Dong, and X. Zhang, *Limited tomography reconstruction via tight frame and simultaneous sinogram extrapolation*, J. Comput. Math., 34 (2016), pp. 575–589, https://doi.org/10.4208/jcm.1605-m2016-0535.

[20] B. Deka and S. Datta, *A practical under-sampling pattern for compressed sensing MRI*, in Advances in Communication and Computing, Springer India, New Delhi, 2015, pp. 115–125, https://doi.org/10.1007/978-81-322-2464-8_9.

[21] B. Dong, J. Li, and Z. Shen, *X-Ray CT image reconstruction via wavelet frame based regularization and radon domain inpainting*, J. Sci. Comput., 54 (2013), pp. 333–349, https://doi.org/10.1007/s10915-012-9579-6.

[22] B. Dong and Z. Shen, *MRA-based wavelet frames and Aapplications*, in Mathematics in Image Processing, IAS/Park City Math. Ser. 19, AMS, Providence, RI, 2013, pp. 9–158.

[23] B. Dong and Z. Shen, *Image restoration: A data-driven perspective*, in Proceedings of the 8th International Congress on Industrial and Applied Mathematics, Higher Education Press, Beijing, 2015, pp. 65–108.

[24] B. Dong, Z. Shen, and P. Xie, *Image restoration: A general wavelet frame based model and its asymptotic analysis*, SIAM J. Math. Anal., 49 (2017), pp. 421–445, https://doi.org/10.1137/16M1064969.

[25] M. J. Ehrhardt and S. R. Arridge, *Vector-valued image processing by parallel level sets*, IEEE Trans. Image Process., 23 (2014), pp. 9–18, https://doi.org/10.1109/TIP.2013.2277775.

[26] M. J. Ehrhardt, K. Thielemans, L. Pizarro, D. Atkinson, S. Ourselin, B. F. Hutton, and S. R. Arridge, *Joint reconstruction of PET-MRI by exploiting structural similarity*, Inverse Problems, 31 (2015), 015001, https://doi.org/10.1088/0266-5611/31/1/015001.

[27] M. A. T. Figueiredo and J. M. Bioucas-Dias, *Restoration of Poissonian images using alternating direction optimization*, IEEE Trans. Image Process., 19 (2010), pp. 3133–3145, https://doi.org/10.1109/TIP.2010.2053941.

[28] M. Fornasier and H. Rauhut, *Recovery algorithms for vector-valued data with joint sparsity constraints*, SIAM J. Numer. Anal., 46 (2008), pp. 577–613, https://doi.org/10.1137/0606668909.

[29] B. Goldluecke, E. Strekalovskiy, and D. Cremers, *The natural vectorial total variation which arises from geometric measure theory*, SIAM J. Imaging Sci., 5 (2012), pp. 537–563, https://doi.org/10.1137/110823766.

[30] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan, *Magnetic Resonance Imaging: Physical Principles and Sequence Design*, Wiley, New York, 1999, https://doi/org/10.1002/978118633953.

[31] E. Haber and M. H. Gazit, *Model fusion and joint inversion*, Surv. Geophys., 34 (2013), pp. 675–695, https://doi.org/10.1007/s10712-013-9232-4.

[32] B. Han, G. Kutyniok, and Z. Shen, *Adaptive multiresolution analysis structures and shearlet systems*, SIAM J. Numer. Anal., 49 (2011), pp. 1921–1946, https://doi.org/10.1137/090780912.

[33] R. B. Innis, V. J. Cunningham, J. Delforge, M. Fujita, A. Gjedde, R. N. Gunn, J. Holden, S. Houle, S. C. Huang, M. Ichise, H. Iida, H. Ito, Y. Kimura, R. A. Koeppe, G. M. Knudsen, J. Knuuti, A. A. Lammertsma, M. Laruelle, J. Logan, R. P. Maguire, M. A. Mintun, E. D. Morris, R. Parsey, J. C. Price, M. Slifstein, V. Sossi, T. Suhara, J. R. Votaw, D. F. Wong, and R. E. Carson, *Consensus nomenclature for in vivo imaging of reversibly binding radioligands*, J. Cereb. Blood Flow Metab., 27 (2007), pp. 1533–1539, https://doi.org/10.1038/sj.jcbfm.9600493.

[34] W. Jin, Y. Censor, and M. Jiang, *Bounded perturbation resilience of projected scaled gradient methods*, Comput. Optim. Appl., 63 (2016), pp. 365–392, https://doi.org/10.1007/s10589-015-9777-x.

[35] D. Kazantsev, S. Ourselin, B. F. Hutton, K. J. Dobson, A. P. Kaestner, W. R. B. Lionheart, P. J. Withers, P. D. Lee, and S. R. Arridge, *A novel technique to incorporate structural prior information into multi-modal tomographic reconstruction*, Inverse Problems, 30 (2014), 065004, https://doi.org/10.1088/0266-5611/30/6/065004.

[36] R. Kimmel, R. Malladi, and N. Sochen, *Images as embedded maps and minimal surfaces: Movies, color, texture, and volumetric medical images*, Int. J. Comput. Vis., 39 (2000), pp. 111–129, https://doi.org/10.1023/A:1008171026419.

[37] A. Krol, S. Li, L. Shen, and Y. Xu, *Preconditioned alternating projection algorithms for maximum* a posteriori *ECT reconstruction*, Inverse Problems, 28 (2012), 115005, https://doi.org/10.1088/0266-5611/28/11/115005.

[38] K. Kurdyka, *On gradients of functions definable in o-minimal structures*, Ann. Inst. Fourier (Grenoble), 48 (1998), pp. 769–783, http://www.numdam.org/item?id=AIF_1998__48_3_769_0.

[39] J. Liang, J. Ma, and X. Zhang, *Seismic data restoration via data-driven tight frame*, Geophysics, 79 (2014), pp. V65–V74, https://doi.org/10.1190/geo2013-0252.1.
[40] Y. Liu, J. F. Cai, Z. Zhan, D. Guo, J. Ye, Z. Chen, and X. Qu, *Balanced sparse model for tight frames in compressed sensing magnetic resonance imaging*, PLOS ONE, 10 (2015), 80119584, https://doi.org/10.1371/journal.pone.0119584.
[41] S. Lojasiewicz, *On semi-analytic and subanalytic geometry*, in Panoramas of Mathematics, Warsaw, 1992/1994, Banach Center Publ. 34, Institute of Mathematics of the Polish Academy of Sciences, Warsaw, 1995, pp. 89–104.
[42] D. G. Luenberger and Y. Ye, *Linear and Nonlinear Programming*, Internat. Ser. Oper. Res. Management Sci. 228, 4th ed., Springer, Cham, Switzerland, 2016, https://doi.org/10.1007/978-3-319-18842-3.
[43] M. Lustig, D. Donoho, and J. M. Pauly, *Sparse MRI: The application of compressed sensing for rapid MR imaging*, Magn. Reson. Med., 58 (2007), pp. 1182–1195, https://doi.org/10.1002/mrm.21391.
[44] S. Mallat, *A Wavelet Tour of Signal Processing, The Sparse Way*, 3rd ed., Academic Press, London, 2008.
[45] A. Rangarajan, I. T. Hsiao, and G. Gindi, *A Bayesian joint mixture framework for the integration of anatomical information in functional image reconstruction*, J. Math. Imaging Vision, 12 (2000), pp. 199–217, https://doi.org/10.1023/A:1008314015446.
[46] A. Ron and Z. Shen, *Affine systems in $L_2(\mathbf{R}^d)$: The analysis of the analysis operator*, J. Funct. Anal., 148 (1997), pp. 408–447, https://doi.org/10.1006/jfan.1996.3079.
[47] G. Sapiro and D. L. Ringach, *Anisotropic diffusion of multivalued images with applications to color filtering*, IEEE Trans. Image Process., 5 (1996), pp. 1582–1586, https://doi.org/10.1109/83.541429.
[48] Z. Shen, *Wavelet frames and image restorations*, in Proceedings of the International Congress of Mathematicians. Volume IV, Hindustan Book Agency, New Delhi, 2010, pp. 2834–2863.
[49] L. A. Shepp and Y. Vardi, *Maximum likelihood reconstruction for emission tomography*, IEEE Trans. Med. Imag., 1 (1982), pp. 113–122, https://doi.org/10.1109/TMI.1982.4307558.
[50] N. Sochen, R. Kimmel, and R. Malladi, *A general framework for low level vision*, IEEE Trans. Image Process., 7 (1998), pp. 310–318, https://doi.org/10.1109/83.661181.
[51] G. Strang, *The discrete cosine transform*, SIAM Rev., 41 (1999), pp. 135–147, https://doi.org/10.1137/S0036144598336745.
[52] D. W. Townsend, *Multimodality imaging of structure and function*, Phys. Med. Biol., 53 (2008), pp. R1–R39, http://stacks.iop.org/0031-9155/53/i=4/a=R01.
[53] D. Tschumperle and R. Deriche, *Vector-valued image regularization with PDEs: A common framework for different applications*, IEEE Trans. Pattern Anal. Mach. Intell., 27 (2005), pp. 506–517, https://doi.org/10.1109/TPAMI.2005.87.
[54] J. Wang and J. F. Cai, *Data-driven tight frame for multi-channel images and its application to joint color-depth image reconstruction*, J. Oper. Res. Soc. China, 3 (2015), pp. 99–115, https://doi.org/10.1007/s40305-015-0074-2.
[55] Y. Xu and W. Yin, *A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion*, SIAM J. Imaging Sci., 6 (2013), pp. 1758–1789, https://doi.org/10.1137/120887795.
[56] R. Zhan and B. Dong, *CT image reconstruction by spatial-Radon domain data-driven tight frame regularization*, SIAM J. Imaging Sci., 9 (2016), pp. 1063–1083, https://doi.org/10.1137/16M105928X.
[57] W. Zhou, J. F. Cai, and H. Gao, *Adaptive tight frame based medical image reconstruction: A proof-of-concept study for computed tomography*, Inverse Problems, 29 (2013), 125006, https://doi.org/10.1088/0266-5611/29/12/125006.