SIAM J. IMAGING SCIENCES
Vol. 10, No. 4, pp. 2250–2274

PCM-TV-TFV: A Novel Two-Stage Framework for Image Reconstruction from Fourier Data*

Weihong Guo†, Guohui Song‡, and Yue Zhang†

Abstract. We propose in this paper a novel two-stage projection correction modeling (PCM) framework for image reconstruction from (nonuniform) Fourier measurements. PCM consists of a projection stage (P-stage) motivated by the multiscale Galerkin method and a correction stage (C-stage) with an edge guided regularity fusing together the advantages of total variation and total fractional variation. The P-stage allows for continuous modeling of the underlying image of interest. The given measurements are projected onto a space in which the image is well represented. We then enhance the reconstruction result at the C-stage that minimizes an energy functional consisting of a fidelity in the transformed domain and a novel edge guided regularity. We further develop efficient proximal algorithms to solve the corresponding optimization problem. Various numerical results in both one-dimensional signals and two-dimensional images have also been presented to demonstrate the superior performance of the proposed two-stage method to other classical one-stage methods.

Key words. edge guided reconstruction, Fourier measurements, total fractional order variation

AMS subject classifications. 35R11, 65K10, 65F22, 90C25

DOI. 10.1137/17M1130666

1. Introduction.

Image reconstruction from Fourier measurements has been a fundamental problem in various applications, such as magnetic resonance imaging (MRI) [4, 33, 36, 41, 42, 51], ultrasound imaging [19, 59], and synthetic radar imaging [3, 18]. The reconstruction methods in the literature can be roughly classified into two categories: the discrete models and the continuous ones. The discrete models view the underlying image as a discrete vector with certain *fixed* resolution and usually obtain its approximation through solving a discrete optimization problem consisting of a fidelity term and a regularity term. There are various regularity terms used in the literature, such as total variation (TV) [47, 53], total generalized variation [11, 38], and total fractional variation (TFV) [16, 17, 63]. Moreover, many efficient algorithms such as the alternating direction method of multipliers (ADMM) and primal-dual methods have been proposed [7, 8, 9, 31, 48, 49, 60, 66] to solve corresponding optimization problems. On the other hand, the continuous models consider the underlying image as a piecewise smooth function and recover the image from a function approximation point of view. One of its advantages is the flexibility in setting resolution and it has been

*Received by the editors May 17, 2017; accepted for publication (in revised form) August 8, 2017; published electronically November 27, 2017.
http://www.siam.org/journals/siims/10-4/M113066.html
Funding: The work of the authors was supported in part by grants NSF-DMS-1521582 and NSF-DMS-1521661.
†Department of Mathematics, Applied Mathematics and Statistics, Case Western Reserve University, Cleveland, OH 44106 (wxg49@case.edu, yxz772@case.edu).
‡Department of Mathematics, Clarkson University, Potsdam, NY 13699 (gsong@clarkson.edu).

successfully employed in the reconstruction of super-resolution images [13, 24]. It has also been shown to have superior performance in generalized/infinite-dimensional compressive sensing reconstruction [1, 2].

Image reconstruction has usually been formulated as an optimization problem that minimizes an energy functional in the following form:

$$\min \mathcal{L}(x) + \mathcal{R}(x),$$

where $\mathcal{L}(x)$ is a fidelity term depending on an empirical estimation of the distribution of noise, and $\mathcal{R}(x)$ is a regularity term with a prior estimation of the structure of the underlying image. For instance, $\mathcal{L}(x)$ is often a least squares term while the noise is assumed to be Gaussian, and other formulations could also be found in [39]. On the other hand, a widely used regularity term $\mathcal{R}(x)$ is the l_1 type constraint incorporating certain sparsity prior knowledge. Such sparsity may come from edge estimation [12, 35, 40], wavelet transformation [20, 26, 55, 64], different orders of TV [11, 15, 21, 63], etc.

Most of existing methods solve the above optimization problem with various fidelity terms and regularity terms in either continuous or discrete settings. We will refer to them as one-stage methods in this paper. We will leverage both discrete and continuous models to develop *two-stage* projection correction modeling (PCM), in which the first stage (P-stage) employs a continuous model and the second stage (C-stage) imposes a discrete regularization/penalty term on the model. In particular, for image reconstruction from Fourier measurements, we will later show in numerical experiments that the proposed two-stage PCM has a superior performance compared with other popular one-stage methods.

We demonstrate the idea of PCM with a reconstruction problem. The given data are some finite uniform or nonuniform Fourier measurements. The goal is to reconstruct the underlying image from these measurements. At the P-stage, we consider the underlying image as a function f in a processing domain that is usually a Hilbert space spanned by some basis such as polynomials or wavelets. We will find an "optimal" approximation (projection) in the processing domain by minimizing a certain data fidelity term. In other words, the P-stage projects the Fourier (k-space) measurements into another processing domain that has an accurate representation of the underlying function. We also point out that the P-stage could also be viewed as a dimension reduction step, since we often choose a projection on a much lower dimension subspace. It will also help to reduce the computational time at the second stage. Due to the noise in the measurements and/or imperfect selection of the basis, the approximation in the P-stage will also contain errors. To further improve the reconstruction, we will impose a discrete regularization at the C-stage.

At the C-stage, we will find a "corrected" approximation in the same processing domain by minimizing the sum of a date fidelity term and a regularity term. The data fidelity is the difference between the corrected approximation and the approximated function obtained at the P-stage. On the other hand, we will employ a regularity term on the discrete vector that is the evaluation of the corrected approximation function on discrete grids. In particular, we will consider a hybrid regularity combining TV and TFV. The TV regularization has very good performance in keeping edges in the reconstruction but suffers from the staircase artifact which causes oil-painted blocks. It is mainly due to the fact that TV is a local operator. On the other hand, the TFV is a recent proposed regularization term in image processing and has achieved

promising results [16, 17, 44, 58, 63, 65]. The TFV is based on fractional order derivatives which are generalizations of the integer derivatives. It can be derived from Abel's integral equation and more details can be found in [32, 45]. Imposing TFV could reduce the staircase artifact due to its nonlocal nature. However, the edges are damped in the reconstructed image based on TFV regularization. Therefore, we will adopt a hybrid regularity with TV on the "edges" and TFV on the "smooth" part.

We remark that edge detection becomes an important task in our method, since we will not know where the true edges are before we proceed. We would like to mention that there is another research pipeline in multimodality image reconstruction where computed tomography (CT) and MRI scanning would be run at the same time and CT images are used to enhance the performance of MRI as well as reduce the processing time [10, 22, 37, 40]. The set of edges will be given since the CT is much faster than MRI. However, we consider a more challenging task in this paper where the edges are not known and will be reconstructed recursively in our algorithm.

We briefly summarize our contributions below:

(1) We propose a novel two-stage PCM method of image reconstruction that leverages advantages of both discrete and continuous models.
(2) We introduce a novel regularity that uses TV and TFV adaptively.
(3) We employ a precise and efficient edge based variational constraint in the regularity term, while the edge is determined through combining techniques of image morphology and *thresholding* strategy.
(4) We develop an efficient proximal algorithm for solving the proposed model.

The rest of the paper is organized as follows. We introduce the proposed PCM framework for image reconstruction from Fourier measurements in section 2. We develop in section 3 an efficient proximal algorithm for the general PCM model and further employ it to derive an algorithm for a specific PCM-TV-TFV model. Numerical experiments and comparisons are presented in section 4. We finally make some concluding remarks in section 5.

2. Projection correction modeling.

We will present the proposed PCM framework for image reconstruction from Fourier measurements in this section. To this end, we first give a brief introduction of the image reconstruction problem.

Suppose the underlying image is a real-valued function $f : \Omega \to \mathbb{R}$, where $\Omega \subseteq \mathbb{R}^2$ is a bounded domain of the image. We are given its Fourier data $\hat{f}$ in the following form:

$$\hat{f} = \mathcal{S}\mathcal{F}f + \epsilon, \tag{1}$$

- $\mathcal{S}$ is the sampling operator (might be uniform or nonuniform),
- $\mathcal{F}$ is the continuous Fourier transform as

$$\mathcal{F}f(w) = \int_{\mathbb{R}^2} f(x) \exp(-2\pi i w \cdot x) dx,$$

- $\epsilon \in \mathbb{C}^m$ is random noise.

Our goal is to recover the underlying image f from the given Fourier data.

The challenges come from the nonuniformness of $\mathcal{S}$ as well as the appearance of the noise ϵ. It is well received that most images are piecewise smooth with potential jumps around the

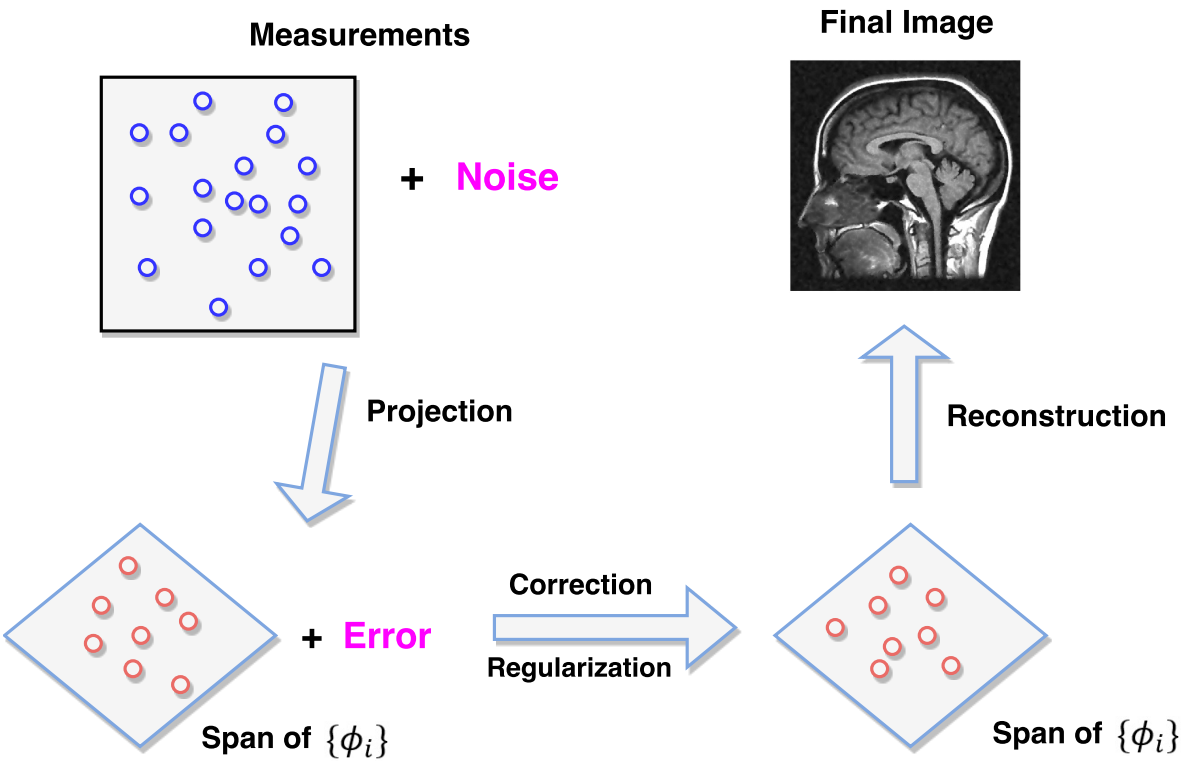


Figure 1. *General PCM framework.*

edges. The inverse Fourier transform will not work directly here. The nonuniformness of the samples in the frequency domain (k-space) will make the inverse process unstable and might bring extra approximation errors when the sampled data contain noise. Moreover, the Fourier basis is amenable to Gibbs oscillations artifacts in representing a piecewise smooth function. To overcome such challenges, we will propose a two-stage PCM framework below.

We point out that the PCM framework can be viewed as a generalization of the classical multiscale Galerkin method in finite element analysis. In contrast to the classical triangle or polyhedron mesh segmentation in a single scale, the region has elements in different scales. The advantage of doing so is to increase the numerical stability of the solver. More details can be found in chapter 13 of [30].

We present the general idea of such a PCM framework in Figure 1. We shall next introduce the P-stage and the C-stage in the PCM framework with more details.

2.1. The P-stage. Suppose $\{\phi_1, \phi_2, \ldots, \phi_n\}$ is a basis of a subspace of the processing domain and we will find an "optimal" approximation

$$\tilde{f} = \sum_{j=1}^{n} (\boldsymbol{c}_f)_j \phi_j,$$

where its coefficient $\boldsymbol{c}_f$ is obtained through solving the following least squares problem:

$$\boldsymbol{c}_f = \operatorname*{argmin}_{\boldsymbol{c}} \|\mathcal{S}\mathcal{F}\Phi\boldsymbol{c} - \hat{f}\|_2^2, \tag{2}$$

where $\Phi\boldsymbol{c} = \sum_{j=1}^{n} c_j\phi_j$. The above least squares problem could be solved efficiently by conjugate gradient solvers.

The performance of the P-stage depends on the selection of the basis $\{\phi_i\}_{i=1}^n$ according to the prior knowledge of the property of the image. We point out that the above framework has also been used in [29, 55] for approximating the inverse frame operator with admissible frames. Analysis of the approximation error has also been discussed there. In particular, the Fourier basis has been used to obtain a stable and efficient numerical approximation of a smooth function from its nonuniform Fourier measurements. However, the Fourier basis might not be appropriate to represent a piecewise smooth function. Instead, we will consider wavelets including classical ones such as the Haar or Daubechies wavelet or more recent ones such as curvelet [43, 56] or shearlet [27, 34], which have a more accurate representation for piecewise smooth functions. In this regard, it could also be viewed as a generalization of the admissible frame method in [29, 55]. In this paper, one-dimensional (1D) and 2D Haar wavelets are used in the numerical experiments, but the general PCM framework could also work with other wavelets.

We remark that even with a reasonable selection of basis $\{\phi_i\}_{i=1}^n$, $\boldsymbol{c}_f$ from (2) might not be accurate due to the noise in the data. We proceed to improve it in the correction stage to alleviate those effects.

2.2. The C-stage. We will find a "corrected" approximation

$$g = \sum_{j=1}^{n} (\boldsymbol{c}_g)_j \phi_j$$

in the same processing domain through solving the following regularization optimization problem:

$$\min_{\boldsymbol{c}_g} \frac{1}{2}\|\boldsymbol{c}_g - \boldsymbol{c}_f\|_2^2 + \mathcal{R}(\boldsymbol{c}_g),$$

where $\mathcal{R}(c_g)$ imposes prior knowledge based on the understanding of the processing domain. The idea is to find a new set of coefficients that are close to $\boldsymbol{c}_f$ found in the P-stage such that the new approximation satisfies some regularities. In terms of optimization, it provides the proximal guidance during the algorithm implementation.

2.2.1. TV-TFV regularity. We next discuss the choice of the regularity term $\mathcal{R}(\boldsymbol{c}_g)$. One popular choice in the literature uses $\|c_g\|_1$ by assuming the underlying function has a sparse representation in the processing domain. However, it might not be the best choice for the problem of image reconstruction from Fourier measurements. It does not incorporate the property of images that are piecewise smooth and contain many textures/features. Moreover, it might have the bias issue in statistics literature that involves the modeling error brought by imperfect selection of basis $\{\phi_i\}_{i=1}^n$; e.g., see [28, chapters 5 and 7]. That is, a regularity term

on the coefficients $\|c_g\|_1$ might not be the best choice for alleviating the bias issue. Instead, we impose regularization in image/function values on a discrete grid through some prior knowledge about the underlying images such as the piecewise smoothness or textures/features. In particular, we shall employ a regularity that combines TV and TFV. To this end, we first review some definitions and notation of fractional order derivatives.

We point out that there are several definitions of fractional order derivatives, such as Riemann–Liouville (RL), Grunwald–Letnikov, and Caputo. We will employ the RL definition [32, 45] in this paper. The left, right, and central RL derivatives of order $\alpha \in (n-1, n)$, $n \in \mathbb{N}$, for a function $f(x)$ supported on an interval $[a, b]$, are defined by

$$ {}_aD_x^\alpha f(x) = \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dx^n} \int_a^x (x-\tau)^{n-\alpha-1} f(\tau) d\tau,$$

$$ {}_xD_b^\alpha f(x) = \frac{(-1)^n}{\Gamma(n-\alpha)} \frac{d^n}{dx^n} \int_x^b (\tau-x)^{n-\alpha-1} f(\tau) d\tau,$$

and

$$ {}_aD_b^\alpha f(x) = \frac{1}{2}({}_aD_x^\alpha f(x) + \ {}_xD_b^\alpha f(x) \cdot (-1)^n),$$

where $\Gamma(\alpha)$ is the Euler's Gamma function

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt.$$

We next introduce the discretization of the RL fractional derivative. For simplicity of presentation, we will show the 1D discretization over an interval $[a, b]$. The discretization over a 2D regular domain will be a direct generalization of this procedure along horizontal and vertical directions. We consider the following n equidistant nodes on $[a, b]$:

$$x_i = \frac{(i-1)(b-a)}{n} + a, \quad i = 1, 2, \ldots, n.$$

Let $L^{(\alpha)}$ and $R^{(\alpha)}$ be the matrix approximations of the left- and right-sided RL α-order derivative operator ${}_aD_x^\alpha$ and ${}_xD_b^\alpha$ accordingly. With Dirichlet boundary conditions that $f(a) = f(b) = 0$, it follows [50] that $L^{(\alpha)}$ and $R^{(\alpha)}$ are two triangular strip matrices with the following structure:

$$L_n^{(\alpha)} = n \begin{pmatrix} w_0^\alpha & w_1^\alpha & \cdots & \cdots & w_{n-1}^\alpha & w_n^\alpha \\ 0 & w_0^\alpha & w_1^\alpha & \ddots & \ddots & w_{n-1}^\alpha \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & w_0^\alpha & w_1^\alpha & \vdots \\ 0 & \ddots & \ddots & 0 & w_0^\alpha & w_1^\alpha \\ 0 & 0 & \cdots & 0 & 0 & w_0^\alpha \end{pmatrix}$$

and

$$R_n^{(\alpha)} = n \begin{pmatrix} w_0^\alpha & 0 & 0 & \cdots & 0 & 0 \\ w_1^\alpha & w_0^\alpha & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & w_0^\alpha & \ddots & 0 \\ w_{n-1}^\alpha & \ddots & \ddots & w_1^\alpha & w_0^\alpha & 0 \\ w_n^\alpha & w_{n-1}^\alpha & \cdots & \cdots & w_1^\alpha & w_0^\alpha \end{pmatrix},$$

where $w_0^\alpha = 1$, $w_j^\alpha = (-1)^j \binom{\alpha}{j}$. These coefficients then can be constructed iteratively:

$$w_j^\alpha = \left(1 - \frac{1+\alpha}{j}\right) w_{j-1}^\alpha, \quad j = 1, 2, \ldots, n.$$

Furthermore, when $\alpha \in (1, 2)$, the matrix approximation of the central RL derivative $C^{(\alpha)}$ will then become

$$C_n^{(\alpha)} = \frac{1}{2}(L^{(\alpha)} + R^{(\alpha)}).$$

Let $U \in \mathbb{R}^N$ be the given discretized image u under an ordinary xy-coordinate with pixel values lexicographically ordered in a column vector. For simplicity, we assume that the original image u is square with $n = \sqrt{N}$ rows and columns. Let $\otimes$ denote the Kronecker product. It follows from applying the central RL derivative $C^{(\alpha)}$ to the images along the x-direction that

$$u_x^{(\alpha)} = (I_n \otimes C_n^{(\alpha)})U,$$

where $I_n \in \mathbb{R}^{n\times n}$ is the identity matrix. Similarly, along the y-direction we will have

$$u_y^{(\alpha)} = (C_n^{(\alpha)} \otimes I_n)U.$$

The procedure will be the same for other choices of $L_n^{(\alpha)}$ and $R_n^{(\alpha)}$. The l_1 norm regularization over $u_x^{(\alpha)}$ and $u_y^{(\alpha)}$ leads to the TFV models [63].

We remark that even though $C_n^{(\alpha)}$ is a dense matrix, the matrix $I_n \otimes C_n^{(\alpha)}$ will be sparse. Furthermore, we observe that w_j^α decays very fast. For example, take $\alpha = 1.3$, the first few w_j's,

$$w_0 = 1, \quad w_1 = -1.3, \quad w_2 = 0.195, \quad w_3 = 0.0455,$$
$$w_4 = 0.0493, \quad w_5 = 0.01, \quad w_6 = 0.006 \cdots .$$

To further enhance the sparsity of the operation matrix, one can truncate w_j's at a certain level to improve the efficiency of the program.

We are now ready to introduce the C-stage in the following form:

$$\min_{\boldsymbol{c}_g} \frac{1}{2}\|\boldsymbol{c}_g - \boldsymbol{c}_f\|_2^2 + \mu_t\|\nabla g|_\Gamma\|_1 + \mu_f\|\nabla^\alpha g|_{\Gamma^c}\|_1 \tag{3}$$

$$\text{s.t.} \quad g = \Phi \boldsymbol{c}_g. \tag{4}$$

where Γ is an open domain centered around the edge part of the image and ∇^α is the discretized α-order fractional differential operator.

We point out that an important question in the above model (3) is to select the Γ set. It denotes an open region centered around the edges rather than the edges themselves. The reason is that the set of edges has zero measure in $\mathbb{R}^2$ and it will not be very meaningful to consider the 2D total (fractional) variation on that.

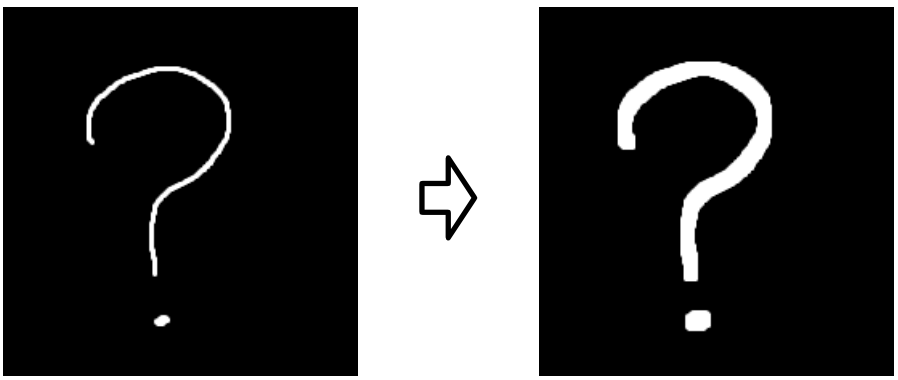
2.2.2. Construction of the Γ set. We shall next discuss the construction of the Γ set. We first detect the edges of the image reconstructed from the P-stage. In order to obtain accurate edges, our approach is to find a rough estimation at the initial step and then update it iteratively with the reconstructed image during the implementation of the reconstruction algorithm. Specifically, we use the first few iterations of the TV model [53] as a warm start to obtain an initial estimate of the edges. Meanwhile, we take an early termination in such an iterative method to avoid staircase artifacts. We then apply these filter edge detectors such as Sobel, Canny, and Prewitt filters [14, 23, 57] on the initial result to improve the accuracy of the edges.

Once the edges are obtained, an initial Γ set can be easily constructed via morphologic dilation [54]; see Figure 2 for a visual illustration. The original image is a binary valued image. This operation converts the neighborhood of 0's of the question mark to all 1's. For all the numerical experiments, the converted neighborhood width is set to be 5 pixels.

We then update the Γ set through a few iterations of the previous results. We next discuss the update of the Γ set in more detail. We denote the Γ set obtained at the ith iteration by Γ_i. When estimating Γ_k, we use the previous Γ_i's,

$$\Gamma_k = \texttt{round}\left(\frac{1}{k-1}\sum_{i=1}^{k-1}\Gamma_i\right),$$

where $\texttt{round}(\cdot)$ is the standard round function that returns the closest integer. This is to ensure that Γ_k will be a binary matrix. A more general thresholding scheme with a given distribution $\mathcal{D}$ at a certain confidence level $t \in [0, 1)$ will be



(a) Original Image (b) Dilated Image

Figure 2. *Illustration of image dilation.*

$$\Gamma_k = \left(\sum_{i=q}^{k-1} w_i \Gamma_i > t \right)$$

for some $1 \le q \le k-1, q \in N_+$, $w_i \sim \mathcal{D}$, $w_q \le w_{q+1} \le \cdots \le w_{k-1}$, $\sum_{i=q}^{k-1} w_i = 1$.

We could further reduce the computational cost by thresholding on the number of iterations. We point out that the reconstructed image will become more accurate during the iterations of the algorithm and there will be little variance in the detected edges after certain iterations. That is, it is reasonable to use a small number of iterations in edge detection.

We summarize the procedure for reconstructing the initial Γ set as follows:

(1) Warm up and edge detection.
   (a) Run the TV model for a few iterations (in our experiment $3 \sim 5$ will be enough). This depends on a rough estimate of the noise level. A quick noise variance estimation can be found at [52].
   (b) Use an implement filter based edge detector such as Canny or Sobel filters to detect the edge set.

(2) Obtain Γ through image dilation.

The Γ set is updated as the underlying image is iteratively updated.

2.3. PCM-TV-TFV in image denoising. We shall summarize the P-stage and the C-stage to present the complete process of the two-stage PCM-TV-TFV model as follows:

$$\begin{aligned} \mathbf{P}: \ & \boldsymbol{c}_f = \operatorname*{argmin}_{\boldsymbol{c}} \|\mathcal{S}\mathcal{F}\Phi\boldsymbol{c} - \hat{f}\|_2^2, \\ \mathbf{C}: \ & \min_{\boldsymbol{c}_g} \frac{1}{2}\|\boldsymbol{c}_g - \boldsymbol{c}_f\|_2^2 + \mu_t \|\nabla g|_\Gamma\|_1 + \mu_f \|\nabla^\alpha g|_{\Gamma^c}\|_1 \\ & \text{s.t.} \quad g = \Phi \boldsymbol{c}_g, \end{aligned} \tag{5}$$

where Γ is an open domain centered around the edges.

We will demonstrate the advantages of the above proposed two-stage model through numerical comparisons with other popular models later.

3. Algorithms. In this section we shall develop a proximal algorithm scheme for solving the general PCM optimization problem. Moreover, we will also introduce a split Bregman scheme [31] for solving the TV-TFV regularity problem. We will then combine them to derive a specific algorithm for the PCM-TV-TFV model.

3.1. General PCM model solver. We shall first consider the following general PCM model with a general regularity term $\mathcal{R}$:

$$\begin{aligned} \mathbf{P}: & c_f = \operatorname*{argmin}_{c} \|\mathcal{S}\mathcal{F}\Phi\boldsymbol{c} - \hat{f}\|_2^2, \\ \mathbf{C}: & \min_{\boldsymbol{c}_g} \frac{1}{2}\|\boldsymbol{c}_g - \boldsymbol{c}_f\|_2^2 + \mathcal{R}(\boldsymbol{c}_g). \end{aligned}$$

We will introduce a general proximal algorithm for solving the above PCM optimization problem. We remark that the proximal algorithms refer to a class of algorithms that are widely used in modern convex optimization literature; for a comprehensive survey, see [49].

To this end, we first review the definition of the proximal operator: for a convex function $\mathcal{R}$, the proximal operator **prox** [49] is defined as

$$\mathbf{prox}_{\lambda\mathcal{R}}(\boldsymbol{v}) = \underset{\boldsymbol{c}}{\operatorname{argmin}}\left\{\mathcal{R}(\boldsymbol{c}) + \frac{1}{2\lambda}\|\boldsymbol{c}-\boldsymbol{v}\|_2^2\right\}.$$

We point out that computing the proximal operator is equivalent to solving a trust region problem [25, 49]. Moreover, we could derive closed forms of such proximal operators for many popular functions. For example, if $\mathcal{R}(\boldsymbol{c}) = \boldsymbol{c}^T A\boldsymbol{c}$ for some symmetric positive semidefinite matrix A, then

$$\mathbf{prox}_{\lambda\mathcal{R}}(\boldsymbol{v}) = (A + 1/\lambda I)^{-1}\boldsymbol{v}/\lambda.$$

If $\mathcal{R}(\boldsymbol{c}) = \|\boldsymbol{c}\|_1$, we have

$$(\mathbf{prox}_{\lambda\mathcal{R}}(\boldsymbol{v}))_i = \begin{cases} v_i - \lambda & \text{if } v_i > \lambda, \\ v_i + \lambda & \text{if } v_i < -\lambda, \\ 0 & |v_i| \le \lambda, \end{cases}$$

which is the componentwise soft thresholding operator. Throughout the rest of the paper, we will use $\mathbf{prox}_{\lambda}(\boldsymbol{v})$ to denote this soft thresholding operation.

Consequently, we present an efficient proximal algorithm in Algorithm 1 for solving the general PCM optimization problem, based on Nesterov's accelerated gradient method as well as the FISTA algorithm [6, 9, 46].

We observe that Algorithm 1 involves a linear program and an iterative proximal operator evaluation problem. Both of them can be computed very efficiently. The update step for λ^k is taken from the FISTA algorithm [6, 9, 46], which will accelerate the convergence.

Algorithm 1. Accelerated proximal algorithm.

Input: $\hat{f}$, Φ, $\boldsymbol{c}_g^0$, $\boldsymbol{d}^0$, λ^0

Output: Reconstructed image g

P-stage: Construct $\mathcal{SF}\Phi$ from (2), solve $\boldsymbol{c}_f = (\mathcal{SF}\Phi)^\dagger \hat{f}$.
C-stage:
Initialize $c_g^0 = c_f$, $k = 0$.
while rela_err $> \epsilon$ **do**
  $k = k + 1$;
  $\lambda^k = \frac{k}{k+3}$;
  $\boldsymbol{c}_g^k = \mathbf{prox}_{\lambda^k\mathcal{R}}(\boldsymbol{d}^k - \lambda^k(\boldsymbol{d}^k - \boldsymbol{c}_f))$;
  $\boldsymbol{d}^{k+1} = \boldsymbol{c}_g^k + \lambda_k(\boldsymbol{c}_g^k - \boldsymbol{c}_g^{k-1})$;
  rela_err $= \|c_g^k - c_g^{k-1}\|_2/\|c_g^{k-1}\|_2$;
end while
Construct $g = \Phi\boldsymbol{c}_g^k$.

3.2. TV-TFV regularity solver. In this subsection we shall consider the following TV-TFV regularity problem:

$$\min_f \frac{1}{2}\|f-\tilde{f}\|^2+\mu_t\|\nabla f|_\Gamma\|_1+\mu_f\|\nabla^\alpha f|_{\Gamma^c}\|_1, \tag{6}$$

where Γ again is an open domain centered around the "edge" set. If μ_f is set to be 0 and the width of Γ is large enough, the above model will reduce to the classical TV denoising model,

$$\min_f \frac{1}{2}\|f-\tilde{f}\|^2+\mu_t\|\nabla f\|_1.$$

Similarly, if Γ is set to be $\emptyset$, it will become the TFV model,

$$\min_f \frac{1}{2}\|f-\tilde{f}\|^2+\mu_f\|\nabla^\alpha f\|_1.$$

We remark that this model is used in the image denoising problem. The workflow of the image denoising problem with TV-TFV regularization is summarized in Figure 3.

We shall introduce an algorithm for solving the above TV-TFV regularization problem. We will start with a warm-up procedure to get an estimate of the edge regions. We will then employ the split Bregman method [31, 62] to derive the algorithm. We present it in Algorithm 2.

We will demonstrate that the proposed TV-TFV regularization at the C-stage has a superior performance by numerical comparisons with the TV denoising model and the TFV denoising model in section 4.

3.3. PCM-TV-TFV solver. We now present an algorithm for solving the PCM-TV-TFV model (5) for image reconstruction in Algorithm 3. It follows from a direct application of ADMM [9].

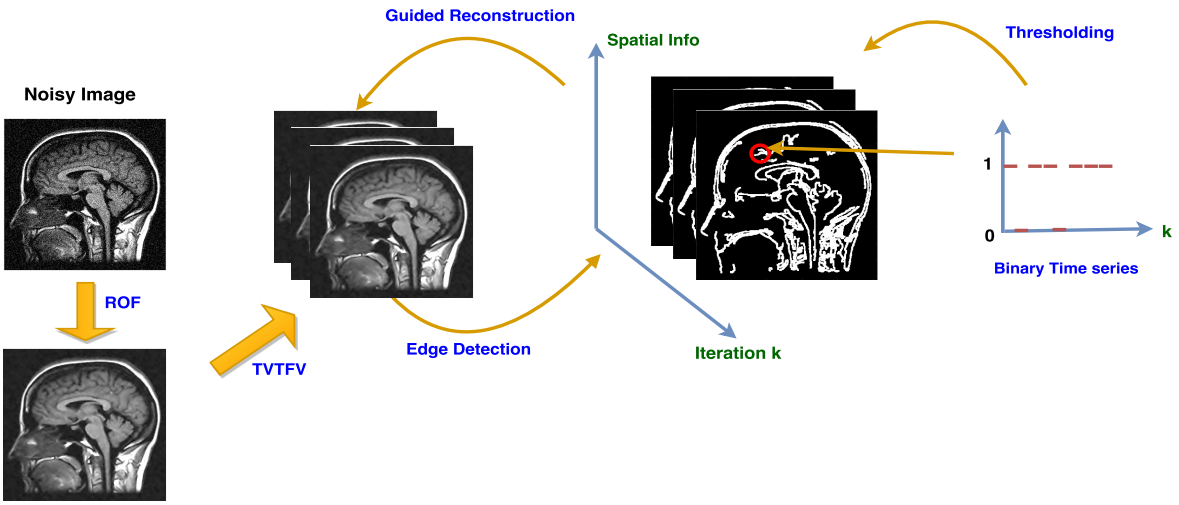


Figure 3. *Illustration of the TV-TFV regularization scheme in the image denoising problem.*

Algorithm 2. TV-TFV denoising algorithm.

Initialize $\bar{\Gamma}_0 = \Omega$, $f^1 = \tilde{f}$, confidence level t, length of stored sequence n.
Warm up with $3 \sim 5$ TV iterations and update $\bar{\Gamma}_0$ to $\bar{\Gamma}_1$.
for $k = 1, 2, \ldots$ **do**
  if $k > n$ **then**
    $\bar{\Gamma}_k = \left(\sum_{i=k-n+1}^{k} \bar{\Gamma}_i\right) > n * t;$
  end if
  Update $\Gamma_k = \text{dilate}(\bar{\Gamma}_k)$;
  Update $d^k = \nabla f|_{\Gamma_k}$, $e^k = \nabla^\alpha f|_{\Gamma_k^c}$.
  Solve f-subproblem

$$f^{k+1} = \underset{f}{\text{argmin}} \frac{1}{2}\|f - \hat{f}\|_2^2 + \mu_t \|\nabla f|_{\Gamma_k} - d^k + dd^k\|_2^2 + \mu_f \|\nabla^\alpha f|_{\Gamma_k} - e^k + ee^k\|_2^2$$

  d-subproblem

$$d^{k+1} = \mathbf{prox}_{\mu_t/\lambda_t}\left(\nabla f^{k+1}|_{\Gamma_k} + dd^k\right)$$

  dd-subproblem

$$dd^{k+1} = dd^k + \gamma_1\left(\nabla f^{k+1}|_{\Gamma_k} - d^{k+1}\right)$$

  e-subproblem

$$e^{k+1} = \mathbf{prox}_{\mu_f/\lambda_f}\left(\nabla^\alpha f|_{\Gamma_k} + ee^k\right)$$

  ee-subproblem

$$ee^{k+1} = ee^k + \gamma_2\left(\nabla^\alpha f|_{\Gamma_k} - e^{k+1}\right)$$

  if $\|f^{k+1} - f^k\|_2 / \|f^k\|_2 < \epsilon$ **then**
    break
  end if
end for

4. Numerical experiments. In this section we demonstrate the superior performance of our proposed two-stage PCM framework for image reconstruction from Fourier measurements. In particular, we will use numerical experiments to show that

(1) the projection step itself can achieve accurate recovery in the case without noise and bias error;
(2) the projection correction with TV regularity (PCM-TV) has a better performance than many of the state-of-the-art continuous models;

Algorithm 3. PCM-TV-TFV algorithm.

Input: $\hat{f}$, Φ, $\boldsymbol{d}^0$, λ^0
Output: Reconstructed image g
P-stage: Construct $\mathcal{SF}\Phi$ from (2), solve $\boldsymbol{c}_f = (\mathcal{SF}\Phi)^\dagger \hat{f}$.
C-stage: Initialize $\bar{\Gamma}_0 = \Omega$, confidence level t, length of stored sequence n.
Warm up with $3 \sim 5$ TV iterations and update $\bar{\Gamma}_0$ to $\bar{\Gamma}_1$.
for $k = 1, 2, \ldots$ **do**
  if $k > n$ **then**
    $\bar{\Gamma}_k = \left(\sum_{i=k-n+1}^{k} \bar{\Gamma}_i\right) > n * t$;
  end if
  Update $\Gamma_k = \text{dilate}(\bar{\Gamma}_k)$;
  Update $d^k = \nabla(\Phi \boldsymbol{c}_g^k)|_{\Gamma_k}$, $e^k = \nabla^\alpha(\Phi \boldsymbol{c}_g^k)|_{\Gamma_k^c}$.
  Solve $\boldsymbol{c}_g$-subproblem

$$\boldsymbol{c}_g^{k+1} = \underset{\boldsymbol{c}_g}{\text{argmin}} \frac{1}{2}\|\boldsymbol{c}_g - \boldsymbol{c}_f\|_2^2 + \mu_t \|\nabla(\Phi \boldsymbol{c}_g)|_{\Gamma_k} - d^k + dd^k\|_2^2 + \mu_f \|\nabla^\alpha(\Phi \boldsymbol{c}_g)|_{\Gamma_k} - e^k + ee^k\|_2^2$$

  d-subproblem

$$d^{k+1} = \mathbf{prox}_{\mu_t/\lambda_t}\left(\nabla\left(\Phi \boldsymbol{c}_g^{k+1}\right)|_{\Gamma_k} + dd^k\right)$$

  dd-subproblem

$$dd^{k+1} = dd^k + \gamma_1\left(\nabla\left(\Phi \boldsymbol{c}_g^{k+1}\right)|_{\Gamma_k} - d^{k+1}\right)$$

  e-subproblem

$$e^{k+1} = \mathbf{prox}_{\mu_f/\lambda_f}\left(\nabla^\alpha\left(\Phi \boldsymbol{c}_g^{k+1}\right)|_{\Gamma_k} + ee^k\right)$$

  ee-subproblem

$$ee^{k+1} = ee^k + \gamma_2\left(\nabla^\alpha\left(\Phi \boldsymbol{c}_g^{k+1}\right)|_{\Gamma_k} - e^{k+1}\right)$$

  if $\|c_g^{k+1} - c_g^k\|_2 / \|c_g^k\|_2 < \epsilon$ **then**
    break
  end if
end for

(3) the TV-TFV regularity leads to better results in image denoising;
(4) the projection correction with TV-TFV regularity (PCM-TV-TFV) model further improves the results of the PCM-TV.

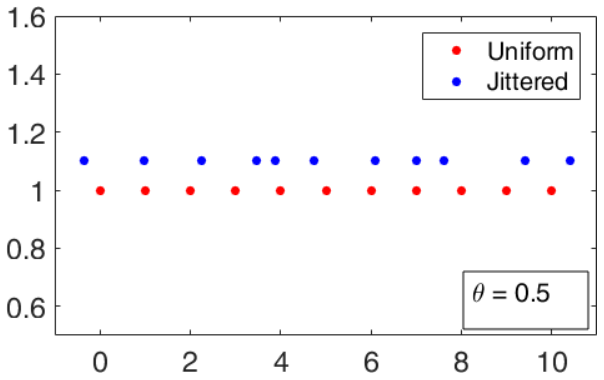


Figure 4. *Illustration of jittered and uniform sampling schemes.*

We will focus on the function reconstruction from nonuniform measurements. In particular, we consider the jittered sampling in the frequency domain. That is, we assume the sampling is taken at the following frequencies:

$$w_k = k + \eta_k, \quad \eta_k \sim U[-\theta, \theta], k = -\frac{m}{2}, \ldots, \frac{m}{2} - 1.$$

We display an example of the jittered sampling in Figure 4. All the experiments were run with MATLAB R2015b on a PC with an Intel i7-3820 3.60 GHz CPU and 16.0 GB RAM.

4.1. Accurate recovery of the P-stage. We will use a simple example to demonstrate the accurate recovery of the P-stage when there is no noise and no bias error in the model. For a selected processing domain $\mathcal{H}_n = \text{span}\{\phi_i\}_{i=1}^n$, we say it has no bias error if the underlying function $f \in \mathcal{H}_n$.

We will use the classical Haar wavelets in the processing domain. In the 1D case, the mother wavelet $\psi(x)$ is

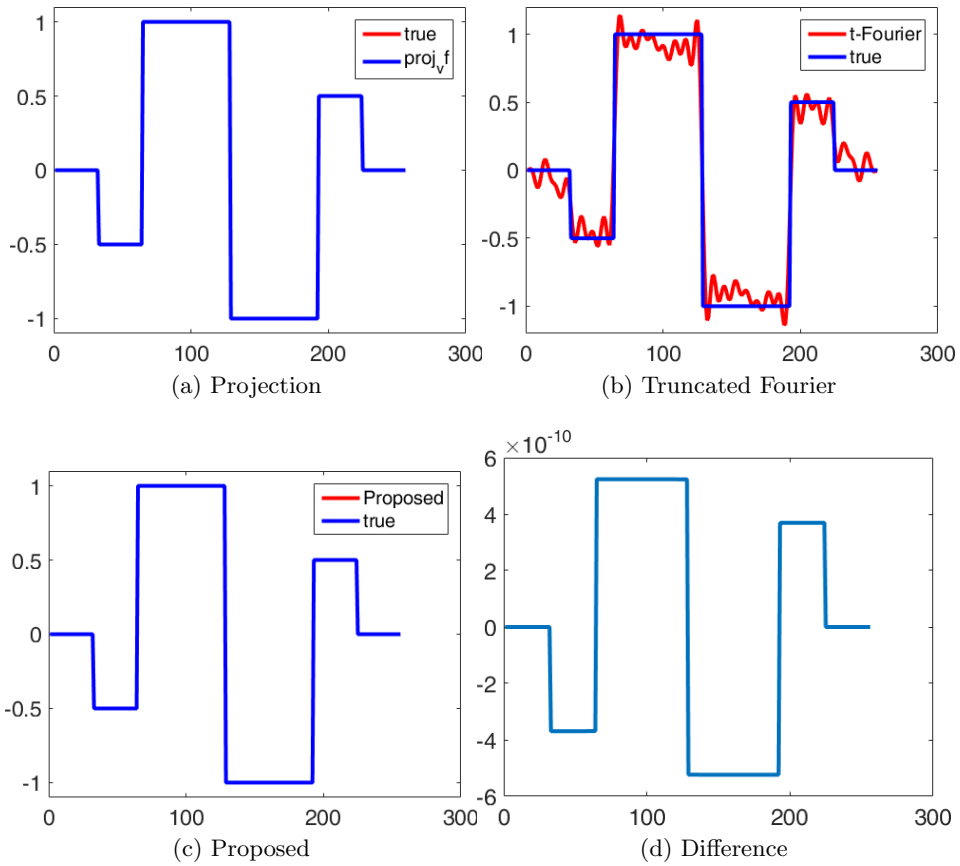
$$\psi(x) = \begin{cases} 1 & \text{if } 0 \le x < 1/2, \\ -1 & \text{if } 1/2 \le x < 1, \\ 0 & \text{otherwise,} \end{cases}$$

and its descendants are $\psi_{n,k} = 2^{n/2}\psi(2^n x - k)$, $k = \mathbb{Z}, x \in [0, 1)$. The 2D Haar wavelet is formulated by simple cross-product.

We consider the following piecewise constant test function:

$$f(x) = \begin{cases} -1/2 & \text{if } -1/8 \le x < 1/4, \\ 1 & \text{if } 1/4 \le x < 1/2, \\ -1 & \text{if } 1/2 \le x < 5/8, \\ 1/2 & \text{if } 5/8 \le x < 3/4, \\ 0 & \text{otherwise.} \end{cases}$$

One can verify that the support of this test function is a subset of that of the selected Haar wavelets. Figure 5 shows that our proposed projection stage achieves a very accurate recovery for noiseless case.



(a) Projection (b) Truncated Fourier

(c) Proposed (d) Difference

Figure 5. *Bias-free reconstruction for piecewise constant signal without noise:* (a) *Ground truth signal and its projection on* $\mathcal{H}_n$. (b) *Reconstruction from truncated Fourier series which suffers from Gibbs oscillation.* (c) *Reconstruction from projection stage. Here* $m = 128$, $n = 32$, $resol = 1/256$, $\theta = 0.25$. (d) *The difference between the reconstruction and the ground truth.*

4.2. PCM-TV model. We consider in this subsection the case where the reconstruction contains bias error. In other words, the underlying function does not lie in the finite-dimensional subspace spanned by the chosen basis. We will consider the following piecewise linear test function $f(x)$:

$$f(x) = \begin{cases} 1 & \text{if } 1/16 \le x < 1/8, \\ -1/2 & \text{if } 1/8 \le x < 1/4, \\ 1 & \text{if } 1/4 \le x < 1/2, \\ -\frac{8}{3}x + \frac{7}{3} & \text{if } 1/2 \le x < 7/8, \\ 0 & \text{otherwise.} \end{cases}$$

We will use the same Haar wavelets to construct the processing domain. Figure 6 displays the bias error.

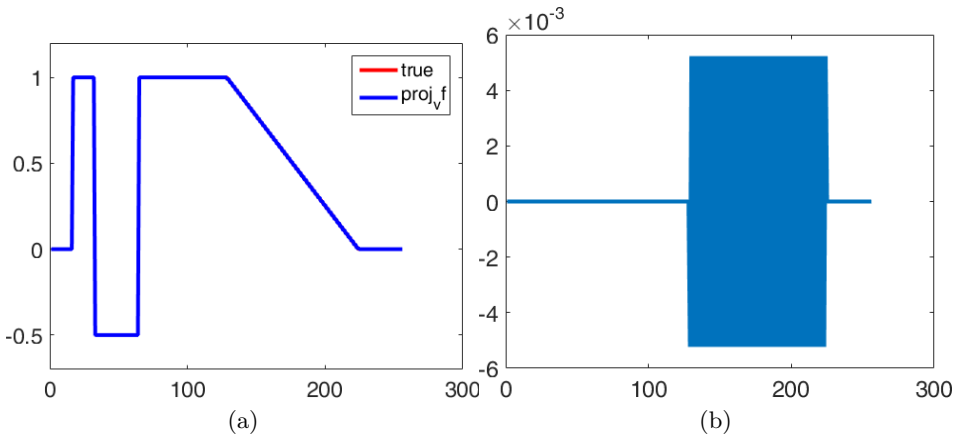


Figure 6. (a) *Piecewise linear function (red) and its projection onto* $\mathcal{H}_n$, $n = 128$ *(blue).* (b) *The difference between the two.*

Suppose we are given $m = 256$ nonuniform Fourier measurements with some added Gaussian noise ϵ at various noise levels $\sigma = \|\epsilon\|_2/\|\hat{f}\|_\infty$.

We will use TV as the regularity at the C-stage of our two-stage PCM method and call it PCM-TV. Moreover, we will compare the proposed two-stage PCM-TV with the following popular one-stage method with different regularities:

- ℓ^1 regularization model,

$$\min_{\boldsymbol{c}} \frac{1}{2}\|\mathcal{S}\mathcal{F}\Phi\boldsymbol{c} - \hat{f}\|_2^2 + \lambda\|\boldsymbol{c}\|_1;$$

- Tikhonov regularization model,

$$\min_{\boldsymbol{c}} \frac{1}{2}\|\mathcal{S}\mathcal{F}\Phi\boldsymbol{c} - \hat{f}\|_2^2 + \lambda\|\boldsymbol{c}\|_2^2;$$

- single-stage TV (SS-TV) model,

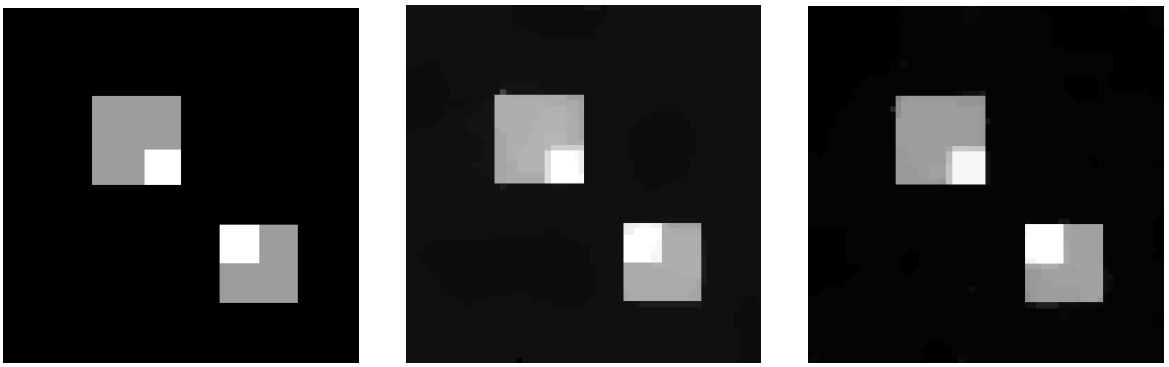
$$\min_{\boldsymbol{c}} \frac{1}{2}\|\mathcal{S}\mathcal{F}\Phi\boldsymbol{c} - \hat{f}\|_2^2 + \lambda\|\nabla\Phi\boldsymbol{c}\|_1.$$

In particular, we will compare them in terms of a few different performance measurements including structural similarity (ssim) [61], peak signal-to-noise ratio (psnr), signal-to-noise ratio (snr), and relative error (rela_err) defined as follows:

$$psnr = 10\log_{10}\frac{d_1 d_2(\max_{i,j} u_{ij})^2}{\|u-\hat{u}\|_F^2},$$
$$snr = 10\log_{10}\frac{\|mean(\hat{u}) - u\|_F^2}{\|\hat{u}-u\|_F^2},$$
$$rela\_err = \frac{\|u-\hat{u}\|_2}{\|u\|_2},$$

Table 1
Numerical comparison of the different models on different noise level.

σ	Model	ssim	psnr	snr	rela_err
0.1	Proposed	**0.9283**	**39.14**	**33.58**	**0.0161**
	ℓ^1 regularization	0.8478	36.45	30.88	0.0220
	SS-TV	0.8466	35.72	30.16	0.0239
	Tikhonov	0.7640	32.17	26.60	0.0360
0.4	Proposed	**0.8505**	**29.41**	**23.84**	**0.0495**
	ℓ^1 regularization	0.7092	26.21	20.64	0.0715
	SS-TV	0.6830	23.81	18.24	0.0942
	Tikhonov	0.3041	21.09	15.52	0.1289
0.7	Proposed	**0.7096**	**24.87**	**19.30**	**0.0834**
	ℓ^1 regularization	0.6221	21.93	16.36	0.1170
	SS-TV	0.6816	22.60	17.03	0.1083
	Tikhonov	0.1414	15.51	9.951	0.2445



(a) Ground Truth (b) SS-TV (c) PCM-TV

Figure 7. *Reconstruction comparison with bias error for the 2D image. Here* $m = 128 \times 128$, $n = 64 \times 64$, $imsize = 256 \times 256$, $\sigma = 0.6$.

where $u \in \mathbb{R}^{d_1 \times d_2}$ is the true image and $\hat{u} \in \mathbb{R}^{d_1 \times d_2}$ is the reconstruction. We present the numerical results in Table 1.

We can observe from Table 1 that our proposed PCM-TV has a better performance than all three other one-stage methods.

We next consider the 2D case. We will consider the 2D function f with a randomly chosen square support of $[0.25, 0.5]^2 \cup [0.61, 0.83]^2$ under where the entire image region is defined as $[0, 1]^2$. We display it in Figure 7. The 2D Haar wavelet used here is a direct product of the 1D Haar wavelet leading to $\mathcal{H}_n^2 := \mathcal{H}_n \otimes \mathcal{H}_n$.

In particular, we will use a piecewise constant test function whose support does not lie in $\mathcal{H}_n^2$. It consists of four squares and the white ones do not lie in $\mathcal{H}_n$, which will cause the bias error in reconstruction.

We will compare the proposed PCM-TV model with the SS-TV method. We point out that TV regularization usually yields better results than ℓ^1 regularization and ℓ^2 regularization in

Table 2
Numerical details for the 2D reconstruction in Figure 7 with bias error, $\sigma = 0.6$.

Model	psnr	snr	rela_err	time(s)
SS-TV	33.8196	20.87	0.086	**10.8**
PCM-TV	**35.6589**	**22.7092**	**0.069**	11.0

Table 3
Numerical results for brain image denoising with different regularities.

Model	psnr	snr	rela_err
Noisy	20.9199	6.0853	0.3914
TV	31.0007	16.1661	0.1226
TFV	31.6910	16.8565	0.1133
TV-TFV	**32.1948**	**17.3602**	**0.1069**

the 2D imaging problem. We present the corresponding numerical results and computational time of the split Bregman iterations in Table 2.

We could observe from the above numerical results that the two-stage PCM-TV has a better performance than the SS-TV regularization method. We next explain the differences between them. We point out that the two-stage PCM-TV model is equivalent to the following optimization problem:

$$\min_{\boldsymbol{c}_g} \frac{1}{2}\|\boldsymbol{c}_g - A^\dagger \hat{f}\|^2 + \lambda \mathcal{R}(\boldsymbol{c}_g),$$

where $A = \mathcal{S}\mathcal{F}\Phi$ and $\mathcal{R}$ is the TV operator. The corresponding first order optimality condition (from Fermat's rule [5]) implies that

$$\mathbf{0} \in \boldsymbol{c}_g - A^\dagger \hat{f} + \lambda \partial \mathcal{R}(\boldsymbol{c}_g), \tag{7}$$

where $\partial\mathcal{R}$ denotes the subdifferential of $\mathcal{R}$. On the other hand, for the SS-TV model

$$\min_{\boldsymbol{c}_g} \frac{1}{2}\|A\boldsymbol{c}_g - \hat{f}\|^2 + \lambda \mathcal{R}(\boldsymbol{c}_g),$$

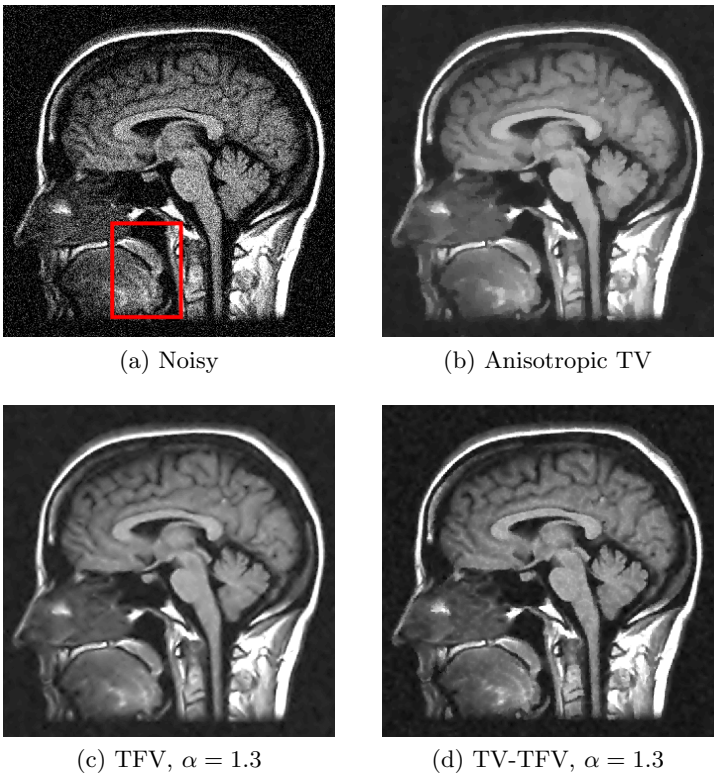
the first order condition implies that

$$\mathbf{0} \in \boldsymbol{c}_g - A^\dagger \hat{f} + \lambda (A^*A)^{-1} \partial \mathcal{R}(\boldsymbol{c}_g). \tag{8}$$

Compared with (7), the descent direction in (8) is distorted by the factor $(A^*A)^{-1}$. It might not only cause extra computational cost but also bring numerical instability and extra errors in computing the inverse of A^*A.

4.3. Performance of TV-TFV regularity. In this subsection we demonstrate the advantages of the proposed TV-TFV regularity. We consider the following general image denoising problem:

$$\min_{u} \frac{1}{2}\|u - \hat{u}\|_2^2 + \lambda \mathcal{R}(u),$$



(a) Noisy (b) Anisotropic TV

(c) TFV, $\alpha = 1.3$ (d) TV-TFV, $\alpha = 1.3$

Figure 8. *Denoising results for models with different regularities. Here α denotes the fractional order of the derivative in the TFV model.*

where $\hat{u}$ is the given noisy image. We will compare three different regularities in the above model: TV, TFV, and TV-TFV.

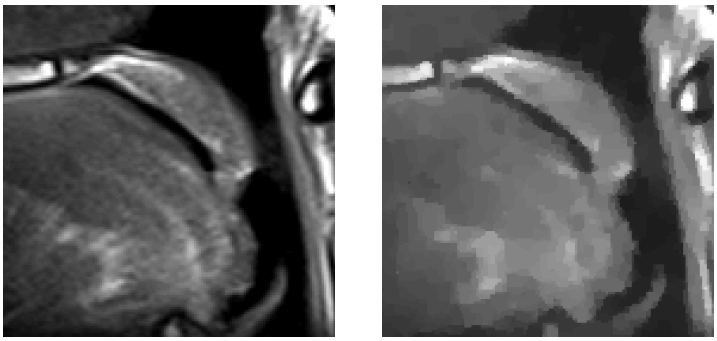
We display the noisy image and the reconstructed images from these three denoising methods in Figure 8.[1] To better understand the performance, we zoom in on the selected parts of the image and display them in Figure 9.

We observe that the anisotropic TV suffers from the staircase artifact due to the fact that the TV is a local operator. On the other hand, the reconstruction with TFV regularity has a blurry effect on the edges. This is not surprising because the TFV is a nonlocal method and it is less edge sensitive than TV. Instead, the TV-TFV regularity avoids such artifacts and has a better reconstruction of both the edges and the overall image.

We also present the numerical results of different performance measurements in Table 3. The TV-TFV regularity shows better results in such measurements as well.

4.4. PCM-TV-TFV versus PCM-TV. Finally, we will combine the projection stage and the correction stage with TV-TFV regularity instead of TV regularity to further improve the performance.

[1]Image retrieved from http://radiopaedia.org/ by Frank Gaillard.



(a) Ground Truth (b) Anisotropic TV

(c) TFV (d) TV-TFV

Figure 9. *Detailed comparison between different models.*

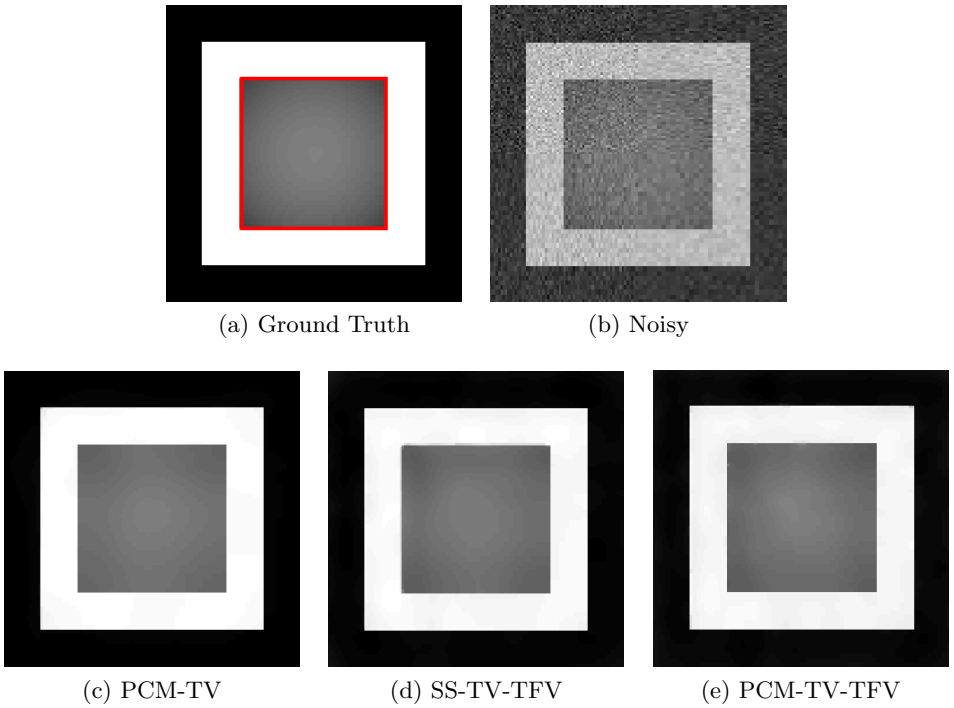
Table 4
Numerical results for denoising with different regularities.

Model	psnr	snr	rela_err	time(s)
PCM-TV	23.1515	20.4278	0.0952	**12.3**
SS-TV-TFV	24.1506	21.1566	0.0875	37.1
PCM-TV-TFV	**25.3128**	**21.9563**	**0.0798**	24.4

We display the ground truth image and the reconstructed images from various methods in Figure 10, including the inverse Fourier method, the PCM-TV method, the PCM-TV-TFV method, and the SS-TV-TFV method. We point out that the optimization parameters λ, μ_t, and μ_f of the above models are all selected to achieve the best performance empirically.

We point out that the noisy image in Figure 10 is obtained directly by the inverse Fourier transform and we can see that the noise level is quite high in this case. All the PCM-TV, SS-TV-TFV, and PCM-TV-TFV models are able to produce more reasonable visual results. To see a deep comparison, we zoom in on the red square part of Figure 10 and present the approximation errors in Figure 11.

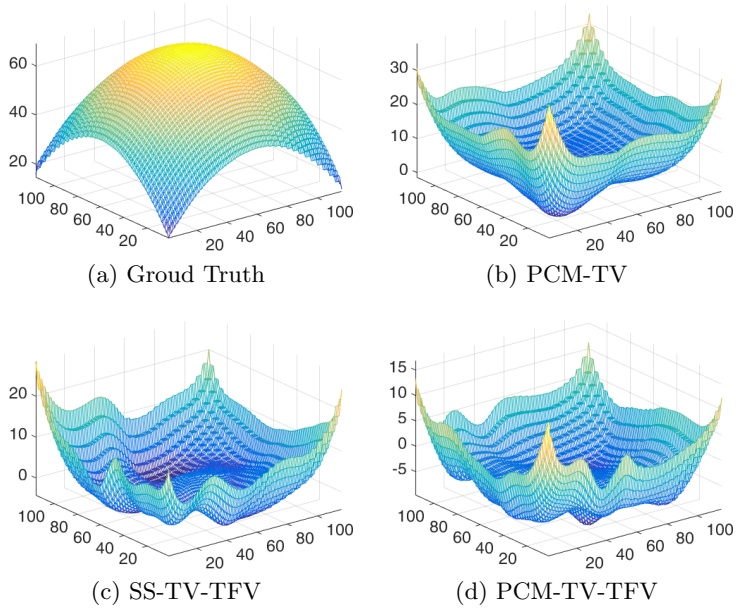
The numerical results are shown in Table 4. We can observe that the proposed PCM-TV-TFV has a better performance than SS-TV-TFV in both accuracy and efficiency. Due to the extra effort of computing the TFV term, it takes more time than the PCM-TV method, but it obtains a much better accuracy.



(a) Ground Truth (b) Noisy

(c) PCM-TV (d) SS-TV-TFV (e) PCM-TV-TFV

Figure 10. *Reconstruction comparison between PCM-TV, SS-TV-TFV, and PCM-TV-TFV. Noisy image* (b) *is obtained from inverse Fourier transform. Here* $m = 128 \times 128$, $n = 96 \times 96$, $imsize = 256 \times 256$, $\sigma = 0.4$.



(a) Groud Truth (b) PCM-TV

(c) SS-TV-TFV (d) PCM-TV-TFV

Figure 11. *Zoomed-in comparison of the red square in Figure* 10. *Figure* (a) *is the true surface. Figure* (b) *shows the difference between the truth and the one reconstructed by PCM-TV. Relative error* $\approx 62\%$. *Figure* (c) *shows the difference between the truth and the reconstruction by SS-TV-TFV. Relative error* $\approx 38\%$. *Figure* (d) *shows the difference between the truth and the one reconstructed by PCM-TV-TFV. Relative error* $\approx 32\%$.

5. Conclusions. In this paper we propose a general two-stage PCM framework for image reconstruction from Fourier measurements. The projection step alleviates the instability from the nonuniformness of the Fourier measurements and the correction step further reduces the noise and the bias effects in the previous step. A precise edge guided TV-TFV regularity shows its own advantages over the models with single TV or TFV regularity. The numerical experiments demonstrate that such a combination enhances the reconstruction and reduces the drawbacks of the TV and TFV themselves. Furthermore, we also show that the proposed PCM-TV-TFV has a superior performance even when the measurements have considerably heavy noise.

We assume that we have some a priori information of an appropriate basis for representing the underlying image. For example, wavelets might be good choices for approximating natural images. A poor choice of basis would lead to a large error in the P-stage. However, the regularization process at the C-stage would alleviate this error to some extent instead of aggravating it. Throughout the paper we only use Haar wavelets to demonstrate the advantages of the proposed two-stage PCM method in the numerical experiments. However, other bases such as shearlets, curvelets, or adaptive wavelets could also be incorporated to further improve the results. Moreover, the proposed two-stage framework for Fourier measurements could also be extended to other linear measurements such as the Radon transform, etc.

Acknowledgments. The authors would like to thank Yiqiu Dong for discussions. The authors would also like to thank the anonymous referees and the associate editor, Stanley Osher, for their effort and valuable comments.

REFERENCES

[1] B. Adcock and A. C. Hansen, *Generalized sampling and infinite-dimensional compressed sensing*, Found. Comput. Math., 2015, pp. 1–61.

[2] B. Adcock, A. C. Hansen, C. Poon, and B. Roman, *Breaking the Coherence Barrier: A New Theory for Compressed Sensing*, preprint, arXiv:1302.0561, 2013.

[3] M. T. Alonso, P. Lopez-Dekker, and J. J. Mallorqui, *A novel strategy for radar imaging based on compressive sensing*, IEEE Trans. Geosci. Remote Sensing, 48 (2010), pp. 4285–4295.

[4] R. Archibald, A. Gelb, and R. B. Platte, *Image reconstruction from undersampled Fourier data using the polynomial annihilation transform*, J. Sci. Comput., 67 (2016), pp. 432–452.

[5] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books Math./Ouvrages Math. SMC, Springer, New York, 2011.

[6] A. Beck and M. Teboulle, *A fast iterative shrinkage-thresholding algorithm for linear inverse problems*, SIAM J. Imaging Sci., 2 (2009), pp. 183–202.

[7] S. Becker, J. Bobin, and E. J. Cands, *NESTA: A fast and accurate first-order method for sparse recovery*, SIAM J. Imaging Sci., 4 (2011), pp. 1–39.

[8] S. Becker, E. J. Candès, and M. C. Grant, *Templates for convex cone problems with applications to sparse signal recovery*, Math. Program. Comput., 3 (2011), pp. 165–218.

[9] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, *Distributed optimization and statistical learning via the alternating direction method of multipliers*, Found. Trends Mach. Learn., 3 (2011), pp. 1–122.

[10] J. G. Brankov, Y. Yang, R. M. Leahy, and M. N. Wernick, *Multi-modality tomographic image reconstruction using mesh modeling*, in Proceedings of the International Symposium on Biomedical Imaging, IEEE, 2002, pp. 405–408.

[11] K. Bredies, K. Kunisch, and T. Pock, *Total generalized variation*, SIAM J. Imaging Sci., 3 (2010), pp. 492–526.

[12] J.-F. Cai, B. Dong, and Z. Shen, *Image restoration: A wavelet frame based model for piecewise smooth functions and beyond*, Appl. Comput. Harmon. Anal., 41 (2016), pp. 94–138.
[13] E. J. Cands and C. Fernandez-Granda, *Towards a mathematical theory of super-resolution*, Comm. Pure Appl. Math., 67 (2014), pp. 906–956.
[14] J. Canny, *A computational approach to edge detection*, IEEE Trans. Pattern Anal. Mach. Intell., 1986, pp. 679–698.
[15] T. Chan, A. Marquina, and P. Mulet, *High-order total variation-based image restoration*, SIAM J. Sci. Comput., 22 (2000), pp. 503–516.
[16] D. Chen, Y. Chen, and D. Xue, *Fractional-order total variation image restoration based on primal-dual algorithm*, Abstract Appl. Anal., 2013 (2013), 585310.
[17] D. Chen, S. Sun, C. Zhang, Y. Chen, and D. Xue, *Fractional-order TV-l2 model for image denoising*, Cent. Eur. J. Phys., 11 (2013), pp. 1414–1422.
[18] V. C. Chen and H. Ling, *Time-Frequency Transforms for Radar Imaging and Signal Analysis*, Artech House, Norwood, MA, 2002.
[19] T. Chernyakova and Y. C. Eldar, *Fourier-domain beamforming: The path to compressed ultrasound imaging*, IEEE Trans. Ultrasonics Ferroelectrics Frequency Control, 61 (2014), pp. 1252–1267.
[20] J. K. Choi, B. Dong, and X. Zhang, *An Edge Driven Wavelet Frame Model for Image Restoration*, preprint, arXiv:1701.07158, 2017.
[21] N. Chumchob, K. Chen, and C. Brito-Loeza, *A fourth-order variational image registration model and its fast multigrid algorithm*, Multiscale Model. Simul., 9 (2011), pp. 89–128.
[22] X. Cui, H. Yu, G. Wang, and L. Mili, *Total variation minimization-based multimodality medical image reconstruction*, in SPIE Optical Engineering+Applications, International Society for Optics and Photonics, 2014, pp. 92121D–92121D.
[23] E. R. Davies, *Computer and Machine Vision: Theory, Algorithms, Practicalities*, Academic Press, New York, 2012.
[24] L.-J. Deng, W. Guo, and T.-Z. Huang, *Single image super-resolution via an iterative reproducing kernel Hilbert space method*, IEEE Trans. Circuits Syst. Video Techonol., 26 (2015).
[25] J. E. Dennis, Jr., and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Classics in Appl. Math., SIAM, Philadelphia, 1996.
[26] B. Dong and Z. Shen, *MRA-based wavelet frames and applications: Image segmentation and surface reconstruction*, in SPIE Defense, Security, and Sensing, International Society for Optics and Photonics, 2012, pp. 840102–840102.
[27] G. Easley, D. Labate, and W.-Q. Lim, *Sparse directional image representations using the discrete shearlet transform*, Appl. Comput. Harmon. Anal., 25 (2008), pp. 25–46.
[28] J. Friedman, T. Hastie, and R. Tibshirani, *The Elements of Statistical Learning*, Vol. 1, Springer Ser. Statist., Springer, New York, 2001.
[29] A. Gelb and G. Song, *A frame theoretic approach to the nonuniform fast Fourier transform*, SIAM J. Numer. Anal., 52 (2014), pp. 1222–1242.
[30] M. S. Gockenbach, *Understanding and Implementing the Finite Element Method*, SIAM, Philadelphia, 2006.
[31] T. Goldstein and S. Osher, *The split bregman method for l1-regularized problems*, SIAM J. Imaging Sci., 2 (2009), pp. 323–343.
[32] R. Gorenflo and F. Mainardi, *Fractional Calculus*, Springer, New York, 1997.
[33] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, *Generalized autocalibrating partially parallel acquisitions (GRAPPA)*, Magnetic Resonance Medicine, 47 (2002), pp. 1202–1210.
[34] W. Guo, J. Qin, and W. Yin, *A new detail-preserving regularization scheme*, SIAM J. Imaging Sci., 7 (2014), pp. 1309–1334.
[35] W. Guo and W. Yin, *Edge guided reconstruction for compressive imaging*, SIAM J. Imaging Sci., 5 (2012), pp. 809–834.
[36] L. He, T.-C. Chang, S. Osher, T. Fang, and P. Speier, *MR image reconstruction from undersampled data by using the iterative refinement procedure*, PAMM, 7 (2007), pp. 1011207–1011208.
[37] A. P. James and B. V. Dasarathy, *Medical image fusion: A survey of the state of the art*, Inform. Fusion, 19 (2014), pp. 4–19.

[38] F. Knoll, K. Bredies, T. Pock, and R. Stollberger, *Second order total generalized variation (TGV) for MRI*, Magnetic Resonance Medicine, 65 (2011), pp. 480–491.
[39] F. Li, S. Osher, J. Qin, and M. Yan, *A multiphase image segmentation based on fuzzy membership functions and l1-norm fidelity*, J. Sci. Comput., (2015), pp. 1–25.
[40] Y. Lu, J. Zhao, and G. Wang, *Edge-guided dual-modality image reconstruction*, IEEE Access, 2 (2014), pp. 1359–1363.
[41] M. Lustig, D. Donoho, and J. M. Pauly, *Sparse MRI: The application of compressed sensing for rapid MR imaging*, Magnetic Resonance Medicine, 58 (2007), pp. 1182–1195.
[42] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, *Compressed sensing MRI*, IEEE Signal Process. Mag., 25 (2008), pp. 72–82.
[43] J. Ma and G. Plonka, *The curvelet transform*, IEEE Signal Process. Mag., 27 (2010), pp. 118–133.
[44] A. Melbourne, N. Cahill, C. Tanner, M. Modat, D. Hawkes, and S. Ourselin, *Using fractional gradient information in non-rigid image registration: Application to breast MRI*, in SPIE Medical Imaging, International Society for Optics and Photonics, 2012, pp. 83141Z–83141Z.
[45] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, Wiley, New York, 1993.
[46] Y. Nesterov, *A method of solving a convex programming problem with convergence rate o* (1/k2), Soviet Math. Dokl., 27 (1983), pp. 372–376.
[47] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, *An iterative regularization method for total variation-based image restoration*, Multiscale Model. Simul., 4 (2005), pp. 460–489.
[48] S. Osher, Y. Mao, B. Dong, and W. Yin, *Fast linearized Bregman iteration for compressive sensing and sparse denoising*, Commun. Math. Sci., 8 (2010), pp. 93–111.
[49] N. Parikh, S. P. Boyd, et al., *Proximal algorithms*, Found. Trends Optim., 1 (2014), pp. 127–239.
[50] I. Podlubny, *Matrix approach to discrete fractional calculus*, Fract. Calc. Appl. Anal., 3 (2000), pp. 359–386.
[51] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, P. Boesiger, et al., *Sense: Sensitivity encoding for fast MRI*, Magnetic Resonance Medicine, 42 (1999), pp. 952–962.
[52] K. Rank, M. Lendl, and R. Unbehauen, *Estimation of image noise variance*, IEE Proc. Vision, Image Signal Process., 146 (1999), pp. 80–84.
[53] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.
[54] J. Serra, *Image Analysis and Mathematical Morphology*, Vol. 1, Academic Press, New York, 1982.
[55] G. Song and A. Gelb, *Approximating the inverse frame operator from localized frames*, Appl. Comput. Harmon. Anal., 35 (2013), pp. 94–110.
[56] J.-L. Starck, E. J. Candès, and D. L. Donoho, *The curvelet transform for image denoising*, IEEE Trans. Image Process., 11 (2002), pp. 670–684.
[57] R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer Science & Business Media, New York, 2010.
[58] R. Verdú-Monedero, J. Larrey-Ruiz, J. Morales-Sánchez, and J. L. Sancho-Gómez, *Fractional regularization term for variational image registration*, Math. Problems Engrg., 2009 (2009), 707026.
[59] N. Wagner, Y. C. Eldar, and Z. Friedman, *Compressed beamforming in ultrasound imaging*, IEEE Trans. Signal Process., 60 (2012), pp. 4643–4657.
[60] B. Wahlberg, S. Boyd, M. Annergren, and Y. Wang, *An ADMM algorithm for a class of total variation regularized estimation problems*, IFAC Proc. Vol., 45 (2012), pp. 83–88.
[61] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, *Image quality assessment: from error visibility to structural similarity*, IEEE Trans. Image Process., 13 (2004), pp. 600–612.
[62] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, *Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing*, SIAM J. Imaging Sci., 1 (2008), pp. 143–168.
[63] J. Zhang and K. Chen, *A total fractional-order variation model for image restoration with non-homogeneous boundary conditions and its numerical solution*, SIAM J. Imaging Sci., 8 (2015), pp. 2487–2518.

[64] Y. Zhang, B. Dong, and Z. Lu, *Minimization for wavelet frame based image restoration*, Math. Comp., 82 (2013), pp. 995–1015.
[65] Y. Zhang, Y. Pu, J. Hu, and J. Zhou, *A class of fractional-order variational image inpainting models*, Appl. Math. Inform. Sci., 6 (2012), pp. 299–306.
[66] M. Z. A. Bhotto, M. O. Ahmad, and M. N. S. Swamy, *An improved fast iterative shrinkage thresholding algorithm for image deblurring*, SIAM J. Imaging Sci., 8 (2015), pp. 1640–1657.