

# RED-PSM: Regularization by Denoising of Partially Separable Models for Dynamic Imaging

Berk Iskender<sup>1</sup>

berki2@illinois.edu

Marc L. Klasky<sup>2</sup>

mklasky@lanl.gov

Yoram Bresler<sup>1</sup>

ybresler@illinois.edu

<sup>1</sup>University of Illinois at Urbana-Champaign, USA<sup>2</sup>Los Alamos National Laboratory, USA

## Abstract

*Dynamic imaging involves the recovery of a time-varying 2D or 3D object at each time instant using its undersampled measurements. In particular, in dynamic tomography, only a single projection at a single view angle may be available at a time, making the problem severely ill-posed. In this work, we propose an approach, RED-PSM, which combines for the first time two powerful techniques to address this challenging imaging problem. The first, are partially separable models, which have been used to introduce a low-rank prior for the spatio-temporal object. The second is the recent Regularization by Denoising (RED), which provides a flexible framework to exploit the impressive performance of state-of-the-art image denoising algorithms, for various inverse problems. We propose a partially separable objective with RED and an optimization scheme with variable splitting and ADMM. Our objective is proved to converge to a value corresponding to a stationary point satisfying the first-order optimality conditions. Convergence is accelerated by a particular projection-domain-based initialization. We demonstrate the performance and computational improvements of our proposed RED-PSM with a learned image denoiser by comparing it to a recent deep-prior-based method TD-DIP. Although the emphasis is on dynamic tomography, we also demonstrate the performance advantages of RED-PSM in a dynamic cardiac MRI setting.*

## 1. Introduction

Time-varying tomography is a challenging ill-posed inverse problem that involves reconstructing a dynamic object using its sequential projections at each time instant. Since the measurements are inconsistent due to the evolving object, traditional reconstruction algorithms lead to significant artifacts.

The problem arises in micro-CT [1], myocardial perfusion imaging [2], thoracic CT [3], imaging of fluid flow processes [4, 5] and dynamic imaging of material samples undergoing compression [6, 7]. Also, it is closely related to the dynamic MRI (dMRI) problem, which typically arises

in cardiac imaging [8].

Previous work [9, 10] treated the problem as time-sequential sampling of bandlimited signals and provided an optimal view angle sampling order and theoretical guarantees for unique and stable reconstruction. However, the approach is limited by its bandlimitedness assumptions.

Several methods [11, 12] recover the underlying motion field from projections using the Radon transform properties. However, these methods require the object information a priori. Other algorithms [13, 14] alternate between estimating the motion field and the time-varying object. These methods assume the total density to be preserved for a given object at different times which may be a limiting assumption. For instance, imaging a fixed slice of a time-varying object under compression may violate this assumption due to perpendicularly evolving parts.

In dynamic tomography and MRI, partially-separable models (PSM) have been used to represent the underlying object [15] and the time-sequential projections [16, 17]. The projection-domain PSM in [16, 17] carries the PSM to the projection domain using the harmonic representation of projections, and provides uniqueness and stability analysis for the problem. In spite of this advantage of the projection domain PSM over its object-domain counterparts, the performance of this approach is limited by the null space of the measurement operator.

A different PSM-based method for dMRI [18] imposes the low-rank structure as a soft constraint. It uses a hybrid objective with the PSM model used in the data fidelity term and the object in a simple  $l_1$ -norm penalty, with an additional penalty term to penalize the mismatch between the PSM and the object. The method does not use a guided initial guess for the object or for the PSM basis functions. A fast alternative low-rank (or PSM)-based method for dMRI [19] decomposes the object representation into three components: the mean signal, a low-rank PSM, and the residual, which is assumed to be sparse in the Fourier domain. Then, each component is estimated in order. Although on some of the data sets used for the comparison in [19] competing methods are either faster or more accurate, the method provides better results when averaged over all data sets.

Recently, object-domain deep image prior (DIP)-based algorithms [20,21] have been proposed for dMRI. Providing impressive results, DIP-based algorithms such as [20] suffer from overfitting and usually require handcrafted early stopping criteria during the optimization of generator network parameters. In an attempt to overcome the overfitting and the need for handcrafted early stopping, the authors of [21] regularize the problem by constraining the geodesic distances between generated objects in proportion to the distances between their latent representation. However, this requires the computation of the Jacobian of the generator network at each iteration of the update of the weights and significantly increases the computational load and run time.

A different approach [22] combines partially separable and generative models. It employs spatial and temporal regularization by penalizing the generator weights and a particular initialization scheme for the inputs to the generator networks. Although it combines the PSM with the recent DIP framework, this method has the following limitations, which are overcome by the proposed approach: (i) The spatial generator is an artifact removal network taking the full-sized spatial basis functions as input. Since this prevents a patch-based implementation, it may be difficult to scale to high-resolution 3D+temporal settings. (ii) The CNN prior for natural images may not be useful as a prior for the individual spatial basis functions, since the least-squares optimal spatial basis functions are the left singular vectors of the complete object, and as such may not have the structure of natural images. (iii) As with other DIP-based methods, if the additional penalty on the generator parameters is insufficient, this method can be prone to overfitting.

*Contributions*

1. To the best of our knowledge, RED-PSM is the first PSM-based approach to dynamic imaging, that uses a pre-trained, learned (RED-based [23]) spatial prior.
2. A novel and effective ADMM algorithm for the resulting new minimization problem enforcing PSM as a hard constraint and incorporating the learned RED spatial regularizer is proposed.
3. The method is supported by theoretical analysis, with a convergence guarantee.
4. Compared to a recent DIP-based [20] algorithm, RED-PSM achieves better reconstruction accuracy with orders of magnitude faster run times.
5. A version of the approach with a patch-based regularizer is shown to provide almost equivalent reconstruction accuracy. This makes the proposed method conveniently scalable to high-resolution 3D or 4D settings.

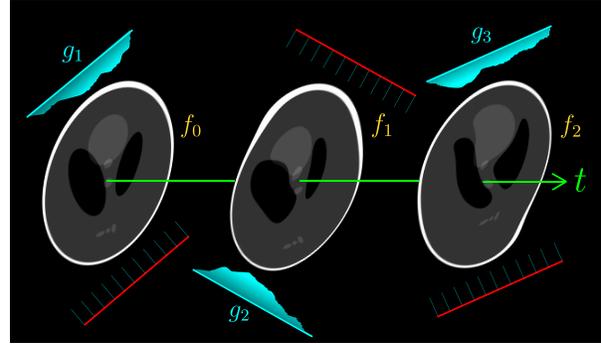


Figure 1: Projection acquisition geometry for time-varying tomography of the object  $f_t$  with single measurement  $g_t$  at each time instant for  $t \in \{0, 1, 2\}$ .

**2. Problem Statement**

In a 2D setting, the goal in the ill-posed dynamic tomography problem is to reconstruct a time-varying object  $f(\mathbf{x}, t)$ ,  $\mathbf{x} \in \mathbb{R}^2$  vanishing outside a disc of diameter  $L$ , from its projections

$$g(\cdot, \theta, t) = \mathcal{R}_\theta\{f(\mathbf{x}, t)\}$$

obtained using the Radon transform operator  $\mathcal{R}_\theta$  at angle  $\theta$ . Considering time-sequential sampling, in which only one projection is acquired at each time instant, and sampling uniform in time, the acquired measurements are

$$\{g(s, \theta_p, t_p)\}_{p=0}^{P-1}, \forall s, t_p = p\Delta_t, \quad (1)$$

where  $s$  is the offset of the line of integration from the origin (i.e., detector position), and  $P$  is the total number of projections (and temporal samples) acquired. The sampling of the  $s$  variable is assumed fine enough and is suppressed in the notation. The angular sampling scheme, the sequence  $\{\theta_p\}_{p=0}^{P-1}$ , with  $\theta_p \in [0, 2\pi]$ , is considered as a free design parameter. Figure 1 shows a simplified time-sequential sampling scheme for a time-varying object.

Our objective in dynamic tomography, and in dynamic imaging in general, is to reconstruct the underlying object with temporal variation  $\{f(\mathbf{x}, t_p)\}_{p=0}^{P-1}$  from the time-sequential measurements in (1). The challenge is that because each projection belongs to a different object, the projections in (1) are inconsistent. Therefore, a conventional, e.g., filtered backprojection (FBP) reconstruction as for a static object results in significant reconstruction artifacts.

**3. Partially Separable Models (PSM)**

For spatio-temporal inverse problems such as dynamic MRI and tomography, the underlying object can be accurately represented using a partially-separable model (PSM), which effectively introduces a low-rank prior to the problem. For dynamic tomography, a PSM can represent the

imaged object  $f$ , or its full set of projections  $g$ . In this paper we use an object-domain PSM.

The representation of a dynamic object  $f(\mathbf{x}, t)$  by a  $K$ -th order PSM is the series expansion

$$f(\mathbf{x}, t) = \sum_{k=0}^{K-1} \Lambda_k(\mathbf{x})\psi_k(t). \quad (2)$$

The model helps with interpretability by separating the spatial structure from the temporal dynamics. This expansion is dense in  $\mathcal{L}_2$  [24], meaning that any finite energy object can be approximated arbitrarily well by such a model of sufficiently high order. Empirically, modest values of  $K$  provide high accuracy in applications to MR cardiac imaging [15, 25, 26]. Theoretical analysis in [17] shows that for a spatially bandlimited object undergoing a time-varying affine transformation (i.e, combination of time-varying translation, scaling, and rotation) of bounded magnitude, a low order PSM provides a good approximation.

The PSM using a  $d$ -dimensional representation for the temporal functions leads to significant reduction of  $\approx PL^2$  representation parameters for an unstructured object of diameter  $L$  with  $P$  temporal samples to  $\approx KL^2 + Kd$  with  $K \ll P$  and  $d \ll P$ . Thus, it provides an effective parsimonious model for the spatio-temporal object  $f$ .

Also, by propagating the PSM object model to the projection domain, it enables quantitative analysis [17] of the choice of optimal sequential projection angular schedule.

## 4. Proposed Method: RED-PSM

### 4.1. Variational Formulation

We use a discretized version of the PSM (2) for the dynamic object, with the object  $f(\cdot, t)$  at each time instant  $t = 0, 1, \dots, P-1$ , represented by a  $J \times J$ -pixel image. Vectorizing these images to vectors  $f_t \in \mathbb{R}^{J^2}$ , the entire dynamic object is the  $J^2 \times P$  matrix  $f = [f_0 \dots f_{P-1}] = \Lambda\Psi^T \in \mathbb{R}^{J^2 \times P}$ . Denoting the  $t$ -th column of the  $P \times P$  identity matrix by  $e_t$ , we have  $f e_t = f_t$  for the  $t$ -th individual time frame of  $f$ . The columns  $\Lambda_k$  and  $\Psi_k$  of  $\Lambda \in \mathbb{R}^{J^2 \times K}$  and  $\Psi \in \mathbb{R}^{P \times K}$  are the discretized spatial and temporal basis functions for the PSM representation, respectively.

Assuming that the x-ray detector has  $J$  bins, the projection of the object at time  $t$  is  $g_t = g(\cdot, \theta_t, t) = R_{\theta_t} f_t \in \mathbb{R}^J$ , where the measurement matrix  $R_{\theta_t} \in \mathbb{R}^{J \times J^2}$  computes the projection at view angle  $\theta_t$ .

We formulate the recovery of  $f$  as the solution  $\hat{f} = \hat{\Lambda}\hat{\Psi}^T$  to the following variational problem

$$(\hat{\Lambda}, \hat{\Psi}) = \arg \min_{\Lambda, \Psi} \sum_{t=0}^{P-1} \|R_{\theta_t} \Lambda \Psi^T e_t - g_t\|_2^2 + \lambda \rho(\Lambda \Psi^T e_t) + \xi \|\Psi\|_F^2 + \xi \|\Lambda\|_F^2 \quad \text{where } \Psi = UZ. \quad (3)$$

The first term is the data fidelity term measuring the fit between available undersampled measurements  $g_t$  of the true object and the measurements obtained from the estimated object  $f = \Lambda\Psi^T \in \mathbb{R}^{J^2 \times P}$ . The second term with weight  $\lambda > 0$  is a *spatial* regularizer injecting relevant spatial prior to the problem. The last two terms with weight  $\xi > 0$  prevent the bilinear factors  $\Lambda$  or  $\Psi$  from growing without limit.

Finally, the constraint  $\Psi = UZ$  is an implicit *temporal* regularizer that restricts the temporal basis functions  $\Psi$  to a  $d$ -dimensional subspace of  $\mathbb{R}^P$  spanned by a fixed basis  $U \in \mathbb{R}^{P \times d}$ . In practice, we incorporate this constraint by explicit substitution (reparametrization of  $\Psi$  in terms of the free variable  $Z \in \mathbb{Z}^{d \times K}$ ) into the objective, and the minimization in (3) over  $\Psi$  is thus replaced by minimization over  $Z$ . This reduces the number of degrees of freedom in  $\Psi$  to a fixed number  $dK$ , independent of the number  $P$  of temporal sampling instants. For notational conciseness, We do not display this constraint/reparametrization in the sequel, but it is used throughout.

### 4.2. Incorporating Regularization by Denoising

For the spatial regularizer  $\rho(\cdot)$  we consider ‘‘Regularization by Denoising (RED)’’ [23]. RED proposes a recovery method using a denoiser in an explicit regularizer of the form

$$\rho(f_t) = \frac{1}{2} f_t^T (f_t - D(f_t)) \quad (4)$$

where  $D : \mathbb{R}^{J^2} \rightarrow \mathbb{R}^{J^2}$  is the denoising operator. Recent works using RED provide impressive results for various *static* reconstruction tasks, including high-dimensional cases [27]. Providing significant flexibility for the type of denoisers that can be used, RED still requires  $D$  to be differentiable and locally homogeneous, and to satisfy the passivity condition  $\|G(f)\| \leq \|f\|$ , for its theoretical analysis to apply.<sup>1</sup>

For the conventional variational formulation

$$\hat{f}_t = \arg \min_{f_t} \|R_{\theta_t} f_t - g_t\|_2^2 + \lambda \rho(f_t),$$

an efficient choice are iterative algorithms [23] that use the standard ‘‘early termination’’ approach [29], and only require a single use of the denoiser per iteration.

However, the regularized PSM objective in (3) does not allow to propagate the RED updates on  $f$  to the respective basis functions in an efficient manner. To overcome this difficulty, we perform a bilinear variable splitting  $f = \Lambda\Psi^T$

<sup>1</sup>While many powerful denoisers satisfy these conditions [23], recent work [28] provides an alternative framework to explain the good performance of RED with denoisers not satisfying conditions in [23].

and obtain our final formulation

$$\begin{aligned} \min_{f, \Lambda, \Psi} \sum_t \|R_{\theta_t} \Lambda \Psi_k^T e_t - g_t\|_2^2 + \lambda \rho(f e_t) \quad (5) \\ + \xi \|\Psi\|_F^2 + \xi \|\Lambda\|_F^2 \quad \text{s.t.} \quad f = \Lambda \Psi^T. \end{aligned}$$

Since the PSM is enforced as a hard constraint, the estimated object  $f$  is constrained to have  $\text{rank}(f) \leq K$ . We propose an algorithm based on ADMM to solve (5).

To cast the problem with the PSM constraint into the ADMM framework, we form the augmented Lagrangian in the scaled form, [30, 31]

$$\begin{aligned} \mathcal{L}_\beta[\Lambda, \Psi, f; \gamma] = \sum_t \left\| R_{\theta_t} \Lambda \Psi^T e_t - g_t \right\|_2^2 + \lambda \rho(f e_t) \\ + \xi \|\Psi\|_F^2 + \xi \|\Lambda\|_F^2 - \frac{\beta}{2} \|\gamma\|_F^2 + \frac{\beta}{2} \|\Lambda \Psi^T - f + \gamma\|_F^2, \quad (6) \end{aligned}$$

where  $\gamma \in \mathbb{R}^{P \times J^2}$  represents the scaled dual variable and  $\beta > 0$  is the penalty parameter.

Then, ADMM can be used to solve (6) as in Algorithm 1.

---

#### Algorithm 1 RED-PSM

---

**input:**  $\Lambda^{(0)}, \Psi^{(0)}, \gamma^{(0)}, f^{(0)} = \Lambda^{(0)} \Psi^{(0)T}, \beta > 0, \lambda > 0, \xi > 0$

- 1: **for**  $i \in \{1, \dots, I\}$  **do**
  - 2:  $\Lambda^{(i)} = \arg \min_{\Lambda} \left\{ \sum_t \|R_{\theta_t} \Lambda \Psi^{(i-1)T} e_t - g_t\|_2^2 + \frac{\beta}{2} \|\Lambda \Psi^{(i-1)T} - f^{(i-1)} + \gamma^{(i-1)}\|_F^2 + \xi \|\Lambda\|_F^2 \right\}$
  - 3:  $\Psi^{(i)} = \arg \min_{\Psi} \left\{ \sum_t \|R_{\theta_t} \Lambda^{(i)} \Psi^T e_t - g_t\|_2^2 + \frac{\beta}{2} \|\Lambda^{(i)} \Psi^T + \gamma^{(i-1)} - f^{(i-1)}\|_F^2 + \xi \|\Psi\|_F^2 \right\}$
  - 4:  $\forall t: f_t^{(i)} = \arg \min_{f_t} \left\{ \lambda \rho(f_t) + \frac{\beta}{2} \|(\Lambda^{(i)} \Psi^{(i)T} + \gamma^{(i-1)}) e_t - f_t\|_2^2 \right\}$
  - 5:  $\gamma^{(i)} = \gamma^{(i-1)} + \Lambda^{(i)} \Psi^{(i)T} - f^{(i)}$
  - 6: **end for**
- 

Line 4 in Algorithm 1 corresponds to the variational denoising of  $(\Lambda^{(i)} \Psi^{(i)T} + \gamma^{(i-1)}) e_t$  with regularization  $\lambda \rho(f_t)$ . Algorithm 1 is reformulated into Algorithm 2 to use RED as in [23] and [28], where  $D_\phi$  is the denoiser for which we have the gradient rule

$$\nabla \rho(f_t) = f_t - D_\phi(f_t) \quad (7)$$

and each  $f_t$  update in Step 4 is a single fixed-point iteration step using the approach of early stopping [29] (Sec. 4.3.2). In contrast to an iterative method to perform Line 4 in Algorithm 1, Step 4 in Algorithm 2 requires only single use of the denoiser per iteration of ADMM.

### 4.3. Regularization Denoiser

The regularization denoiser  $D_\phi$  has a DnCNN [32] architecture and was trained in a supervised manner on a training

---

#### Algorithm 2 RED-PSM with efficient $f$ step

---

**Note:** Inputs, and Lines 1-3 & 5-6 are the same as Algorithm 1.

$$4: \forall t: f_t^{(i)} = \frac{\lambda}{\lambda + \beta} D_\phi(f_t^{(i-1)}) + \frac{\beta}{\lambda + \beta} (\Lambda^{(i)} \Psi^{(i)T} + \gamma^{(i-1)}) e_t$$


---

set of *static* 2D slices  $f_i \in \mathbb{R}^{J^2}, i = 1, \dots, N$  assuming that such data will be available in the settings of interest. Thus, the RED steps are agnostic to the specific motion type. The training objective for the denoiser is

$$\min_{\phi} \sum_i \|f_i - D_\phi(\tilde{f}_i)\|_F^2 \quad (8)$$

where  $\tilde{f}_i = f_i + \eta_i$ , and the injected noise  $\eta_i \sim \mathcal{N}(0, \sigma_i^2 I)$  has noise level  $\sigma_i \sim U[0, \sigma_{\max}]$  spanning a range of values, so that the denoiser learns to denoise data with various noise levels.

### 4.4. Convergence Analysis

In [33](Section V.), we follow an approach along the lines of [34] to analyze convergence. We show that under mild technical conditions, the objective in Algorithm 1 is guaranteed to converge (with increasing number  $I$  of iterations) to a value corresponding to a stationary point of the Lagrangian, that is, satisfying the necessary conditions for first order optimality. In practice, Algorithm 2 with the efficient  $f$  step version, which we implemented and used in the experiments reported in Section 5 has better run times, and rapid empirical convergence. However, its analysis requires additional steps, which are not particularly illuminating. Thus, following the tradition in related analyses [29, 35], we focus on the analysis of the nominal Algorithm 1.

## 5. Experiments

### 5.1. Datasets

Three categories of data sets are used in this work.

*Walnut Dataset:* We use the CT reconstructions of two different (static) walnut objects from the publicly available 3D walnut CT dataset [36]. We create a dynamic test object by synthetically warping the central axial slice of one of the walnut objects using a sinusoidal piecewise-affine time-varying warp [37]. To be precise, the image is divided into a  $J \times J$  uniformly spaced rectangular grid, and the following vertical displacement is applied on each row separately to drive the temporally varying warp  $\Delta_{j,t} = -C(t) \sin(3\pi j/J), j \in \{0, \dots, J-1\}$ , where  $C(t)$  is a linearly increasing function of  $t$  and  $C(0) = 0$ . *Static* axial, coronal, and sagittal slices of the other walnut object are used to train the denoiser  $D_\phi$ .

*Compressed Object Dataset:* This data set is obtained from a materials science experiment [38] with a sequence

of nine increasing compression steps applied to an object, with a full set of radiographic projections collected (using Carl Zeiss Xradia 520 Versa) and reconstructed by the instrument’s software at each step. Using this quasi-static data set, a fixed axial slice is extracted from each reconstruction. Nine extracted slices are interpolated to  $P$  time frames using a recent deep learning-based video interpolation algorithm [39]. The denoiser  $D_\phi$  for our experiments on this data set was trained using the axial slices of the *static* pre-compression and post-compression objects, which would be available in actual dynamic loading experiments.

We note that all algorithms compared in this paper are agnostic to both the synthetic warp applied on the static walnut slice and to the data-driven interpolation method used for the compressed object.

Spatio-temporal projection data for each dataset is simulated by a parallel-beam projection with  $J = 128$  detector bins of the dynamic phantoms, a single projection at each of the  $P$  time instants. The sequence of projection angles  $\{\theta_t\}_{t=0}^{P-1}$  (a free experimental design parameter) was chosen to follow the bit-reversed view angle sampling scheme, which has been shown [16] to provide an especially favorable conditioning of the recovery problem. The simulated measurements are corrupted using AWGN with standard deviation  $\sigma = 5 \cdot 10^{-3}$ . This noise level leads to the FBP (with Ram-Lak filter) of the full set of  $P = 512$  *projections at each time instant* having a PSNR of approximately 46 dB. When, in the actual experiments with sequentially sampled data, only  $1/P$  of this data is used, the PSNR of the reconstruction may be expected to be lower.

Ground-truth frames for  $P = 4$  are shown in Figure 2.

*Cardiac dMRI Dataset:* For a more direct comparison with the setting and data used in dMRI works, we also test RED-PSM on the “retrospective” cardiac dMRI data in [20]. The data includes 23 distinct time frames for one cardiac cycle. Details of the data and experiments are in Section 5.4.3.

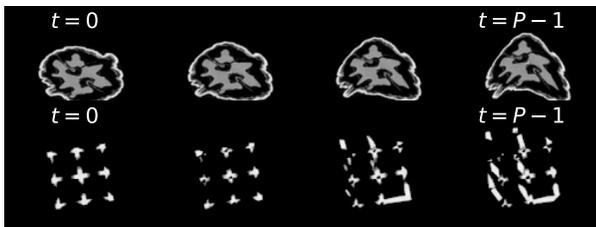


Figure 2: Ground-truth frames for the time-varying walnut (top) and compressed object (bottom) datasets, uniformly sampled in time,  $t \in \{0, \dots, P - 1\}$ , for  $P = 4$ .

## 5.2. Baseline Methods

**PSM-TV:** Similar to the proposed approach, this algorithm also uses a partially separable model to represent the object with a different regularization type. For PSM-TV,

regularization penalizes the  $\ell_1$ -norm of the discrete 2D total variation of the temporal slices of  $f$  at each time instant. To this end, the constraint  $f = \Lambda\Psi^T$  is implemented by substitution into the objective in (5), and the definition of  $\rho$  is changed to  $\rho(\Lambda\Psi^T e_t) = \text{TV}(\Lambda\Psi^T e_t)$ , and the rest of the objective is kept the same. The unconstrained problem is then solved for  $\{\hat{\Lambda}, \hat{\Psi}\}$  (using, as a convenient implementation, the Adam optimizer in Pytorch). Finally, the estimated object is obtained as  $\hat{f} = \hat{\Lambda}\hat{\Psi}^T$ .

**TD-DIP [20]:** TD-DIP is a recent method based on the Deep Image Prior (DIP) approach.<sup>2</sup> It uses a mapping network  $M_\alpha$  and a generative model  $N_\beta$  sequentially to obtain the estimated object at each time instant  $f_t$  from fixed and handcrafted latent representations  $\chi_t$ . Because TD-DIP was originally proposed for dynamic MRI, we modified the objective minimally for dynamic tomography as

$$\min_{\alpha, \beta} \sum_t \|g_t - R_{\theta(t)}((N_\beta \circ M_\alpha)(\chi_t))\|^2. \quad (9)$$

For the comparisons in this work, identical mapping network and generator architectures, latent representation dimensionality, optimizer, learning rate, and decay schemes are used as in the available online implementation [40]. The original work focuses on the beating heart problem and thus proposes a helix-shaped latent manifold for  $\chi_t$  with cycles equal to the number of heartbeats during measurement acquisition. Since we do not have a repetition assumption for the motion types included in this paper, we use a linear manifold as explained in the original paper [20]. Thus, the method is sometimes denoted as “TD-DIP (L)” in Section 5.4 for clarification.

## 5.3. Experimental Settings

All methods are run on a workstation with an Intel(R) Xeon(R) Gold 5320 CPU and NVIDIA RTX A6000 GPU. In practice, we used a minor variation of Algorithm 2, where we combined the subproblems in  $\Lambda$  and  $\Psi$ , and minimized with respect to both basis functions simultaneously using gradient descent with Adam [41] optimizer.

**Denoiser training.** The upper limit for noise level used in training the denoiser was set to  $\sigma_{\max} = 5 \cdot 10^{-2}$ . For the dynamic walnut object, the denoiser  $D_\phi$  is trained on the central 200 axial, 200 sagittal, and 200 coronal slices of another walnut CT reconstruction downsampled to size  $128 \times 128$ . For the compressed object, axial slices of pre-compression and post-compression *static* objects, containing 462 slices in total, are used to train  $D_\phi$ . For the cardiac MRI setting, the denoiser was trained on the static MRI

<sup>2</sup>It would be interesting to include yet another baseline method – the DIP-based PSM approach [22] (also developed for MRI). However we were unable to do so as an implementation of this method was not available, and because of potential issues with replicating its performance and adapting to our CT problem.

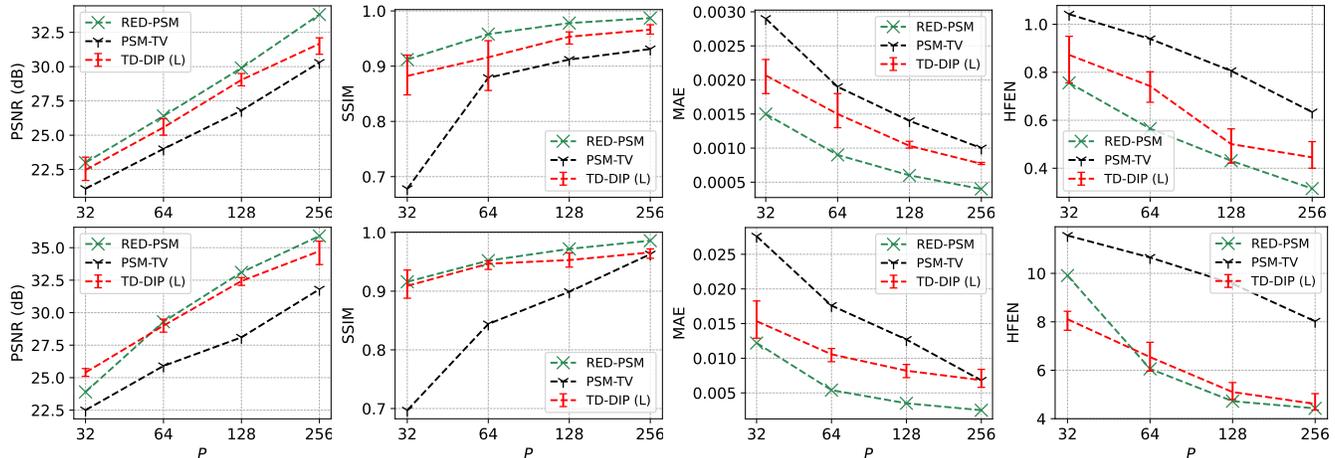


Figure 3: Reconstruction PSNR, SSIM, MAE, and HFEN for the time-varying walnut object (top) and compressed material (bottom) vs.  $P$  using different methods. For TD-DIP, accuracy statistics for the best PSNR reconstruction during optimization are reported assuming such an oracle is available. The minimum and maximum values are indicated with error bars and the mean values are connected by the dashed line for different  $P$  for three different runs.

		(a) Walnut				(b) Compressed Material			
$P$	Method	PSNR (dB)	SSIM	MAE (1e-3)	HFEN	PSNR (dB)	SSIM	MAE (1e-2)	HFEN
32	PSM-TV (R)	21.1	0.678	2.9	1.04	22.5	0.697	2.9	11.6
	TD-DIP (L)	22.5	0.882	2.1	0.87	<b>25.4</b>	0.909	1.5	<b>8.1</b>
	RED-PSM (Pr)	<b>22.7</b>	<b>0.914</b>	<b>1.5</b>	<b>0.79</b>	23.9	<b>0.916</b>	<b>1.2</b>	9.9
64	PSM-TV (R)	24.0	0.879	1.9	0.94	25.6	0.845	1.8	10.7
	TD-DIP (L)	25.6	0.916	1.5	0.74	29.0	0.947	1.1	6.5
	RED-PSM (Pr)	<b>26.4</b>	<b>0.958</b>	<b>0.9</b>	<b>0.57</b>	<b>29.3</b>	<b>0.952</b>	<b>0.5</b>	<b>6.0</b>
128	PSM-TV (R)	26.8	0.912	1.4	0.78	29.2	0.907	1.1	9.6
	TD-DIP (L)	29.0	0.953	1.0	0.50	32.4	0.953	0.8	5.1
	RED-PSM (Pr)	<b>29.9</b>	<b>0.978</b>	<b>0.6</b>	<b>0.43</b>	<b>33.1</b>	<b>0.972</b>	<b>0.4</b>	<b>4.7</b>
256	PSM-TV (R)	30.1	0.934	1.0	0.63	31.9	0.963	0.7	8.0
	TD-DIP (L)	31.7	0.966	0.8	0.45	34.7	0.966	0.7	4.6
	RED-PSM (Pr)	<b>33.8</b>	<b>0.987</b>	<b>0.4</b>	<b>0.32</b>	<b>35.9</b>	<b>0.986</b>	<b>0.3</b>	<b>4.4</b>

Table 1: Reconstruction accuracies for different  $P$ . Random (R) and ProSep (Pr) initialization for the PSM methods. For TD-DIP, the reported accuracies are for the best PSNR using a “stopping oracle”, averaged over three runs with random initial conditions.

training slices of the ACDC dataset [42]. For all datasets,  $D_\phi$  is trained for 500 epochs using the Adam optimizer with a learning rate of  $5 \cdot 10^{-3}$ . Each convolutional layer is followed by a ReLU nonlinearity except for the final layer, which has a single-channel output. We test both direct and residual DnCNN denoisers where the former predicts the denoised image and the latter estimates the noise from the noisy input.

Architectural details for denoisers used in our experiments are provided in Table 6 in Supplementary Material C. We use a fixed pre-trained denoiser for all  $P$  for the same object type.

**Temporal Basis.** In all experiments, we use a fixed basis  $U$  as a cubic spline interpolator to interpolate the low-dimensional temporal representation  $Z$  to  $\Psi$ .

**Initialization.** Unless stated otherwise, spatial and temporal basis functions are initialized using the SVD truncated to order  $K$  the rank of the dynamic object estimate pro-

duced by a recent projection-domain PSM-based method “ProSep” [17]. If the ProSep estimate has rank smaller than  $K$ , the remaining basis functions are initialized as 0. Otherwise, all spatial basis functions are initialized as 0 and the latent representations  $z_k$  of the temporal basis functions are initialized randomly as  $z_k \sim \mathcal{N}(0, I)$ .

**Tomographic Acquisition Scheme.** All methods mentioned in this paper use the bit-reversed angular sampling scheme, over the range  $[0, \pi]$ .

**Run Times.** For  $P = 256$  and using the specified computational resources and parameter settings, To achieve the peak PSNR during optimization, RED-PSM with ProSep initialization requires  $50 < \text{iterations} < 150$  taking about 2 to 6 minutes whereas TD-DIP with batch size  $P$  typically requires  $> 30k$  steps, taking about 3.5 hours to complete. Hence, RED-PSM provides a speedup over TD-DIP by a factor of 35 to 105 to reach its peak performance. We note that based on the parameter configuration, the speedup factor may vary. However, the proposed method provides a significant run time reduction in all cases.

**Evaluation Metrics.** Four quantitative metrics were implemented for evaluating reconstruction accuracy (i) the *peak signal-to-noise ratio* (PSNR) in dB; (ii) the *structural similarity index* (SSIM) [43]; (iii) the *mean absolute error* (MAE); and (iv) the *high-frequency error norm* (HFEN) [44] defined as  $\text{HFEN}(f, f_r) = \|\text{LoG}(f) - \text{LoG}(f_r)\|_2$  where LoG is a rotationally symmetric Laplacian of Gaussian filter with a standard deviation of 1.5 pixels.

## 5.4. Results

### 5.4.1 Reconstruction accuracies for different $P$

In Figure 3 and Table 1, the performance of RED-PSM is compared with PSM-TV and with TD-DIP that uses a linear latent representation for both objects. While Figure 3

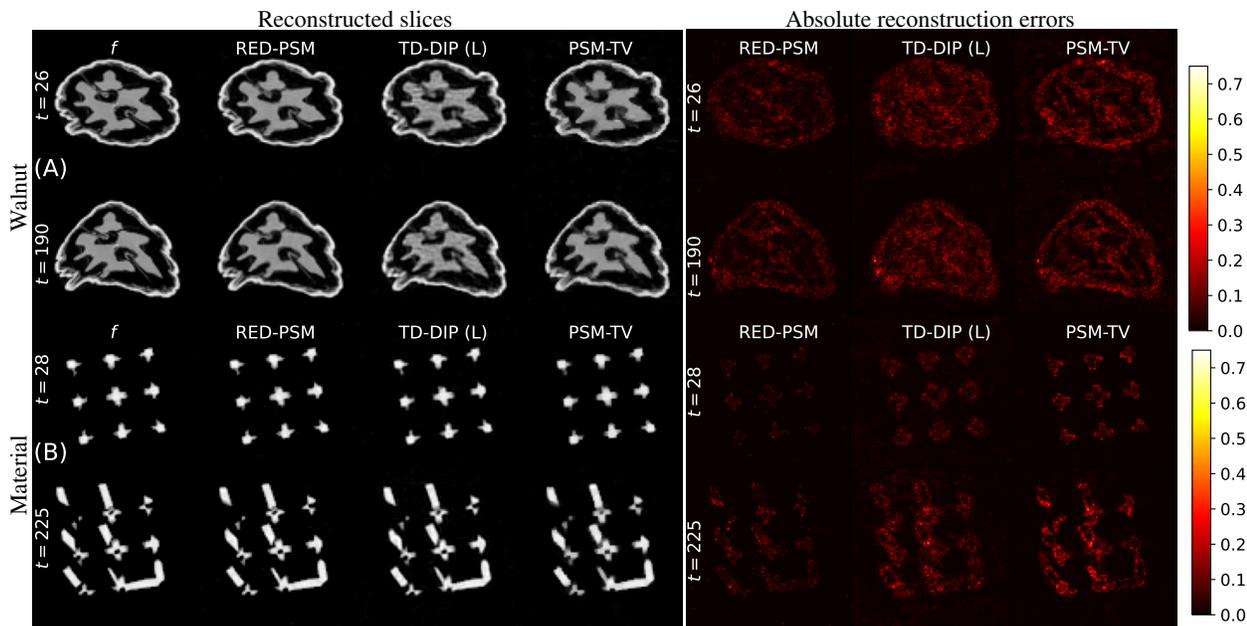


Figure 4: Comparison of reconstructed object frames at two time instants using different methods for  $P=256$ , and the corresponding normalized absolute reconstruction errors for (A) the time-varying walnut, and (B) compressed object.

facilitates the comparison of metrics for different  $P$  for the various methods, since the range of metrics is very large for varying  $P$ , Table 1 emphasizes important differences between methods for the same  $P$ , and provides detail for a more precise quantitative comparison. As expected, the estimates improve with increasing  $P$  for all methods. In terms of PSNR, the proposed algorithm performs on par with or better than *the best PSNR* TD-DIP estimate for various  $P$ , with improvement for RED-PSM enhancing with the increase in  $P$ . Moreover, in terms of SSIM, MAE, and HFEN, the improvements with RED-PSM are even more significant. Although these observations are valid for both objects, for the walnut, RED-PSM provides slightly greater improvement over TD-DIP.

Figure 4 compares the reconstructions for both objects at two different  $t$  for  $P = 256$ . As expected, providing blurry reconstructions lacking finer details for both objects, PSM-TV performs the worst among the compared methods. The TD-DIP reconstructions contain visible noise-like artifacts on the piecewise constant regions of the walnut object, which are alleviated by RED-PSM. This observation manifests itself also in the absolute difference figures, with error for TD-DIP distributed throughout the interior regions of the walnut. Furthermore, RED-PSM preserves sharp details around the shell of the walnut better. Throughout the compressed material object for the given time frames, compared to TD-DIP, RED-PSM has an almost consistently smaller absolute error. This difference is more prominent around the highly dynamic regions of the object, emphasizing the advantage of the proposed method.

In Figure 5, the reconstructed  $x$ - $t$  cross-sections of the

dynamic walnut are shown for different methods. The location of the cross-section is highlighted by a yellow line on the static  $x$ - $y$  frame at  $t = 0$ . Consistent with the comparison in Figure 4, RED-PSM provides reduced absolute error values throughout the respective cross-section. Also, as more apparent in the error figures, TD-DIP produces higher background errors.

#### 5.4.2 Effect of Initialization

The initialization of  $\Lambda$ ,  $\Psi$ , and  $f$  plays an important role in the performance and convergence speed of RED-PSM. We observe significant speed-up when rather than a random initialization, we initialize the algorithm with ProSep [16] estimated reconstruction. Figure 7 shows PSNR vs. iterations comparison for different initialization techniques for the dynamic walnut object with  $P = 256$ . The rest of the parameters were selected identically as indicated in Table 5 in Supplementary Material. This experiment highlights the advantages of initializing with ProSep estimated basis functions: eliminating the need for multiple runs for a best-case result; and speeding up convergence considerably.

#### 5.4.3 Cardiac dMRI data experiments

In this setting, different to the previous experiments, we used 4  $k$ -space radial lines (“spokes”) per frame at the bit-reversed angles. We used 1.4 and 2.8 cardiac cycles, with  $23 \times 4 = 92$  spokes/cycle, for a total of  $P=128$  and  $P=256$  spokes. The problem is still severely undersampled compared to the experiment in [20] where 13 spokes are used

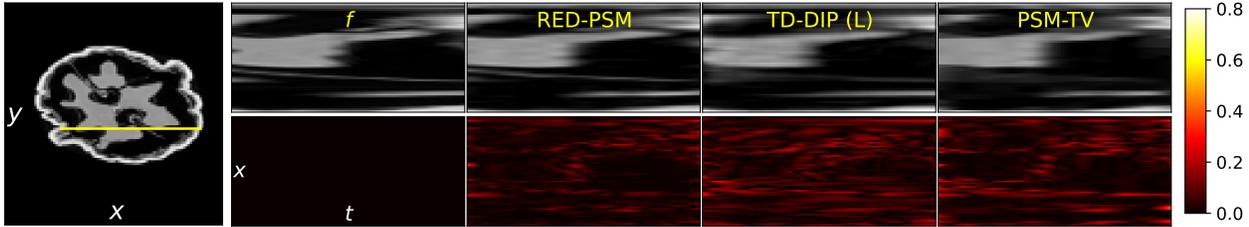


Figure 5: Comparison of reconstructed  $x$ - $t$  cross-sections using different methods for  $P=256$  (top), with corresponding normalized absolute error counterparts (bottom). The location of the cross-section is indicated with a yellow line on the static object at time  $t=0$  (left). The  $x$ - $y$ - $t$  coordinates are indicated on the static object and bottom left absolute error figure in white text.

per frame for 13 cycles, for a total of  $13 \times 23 \times 13 = 3,887$  spokes. Since the data is periodic, we also tested the helix latent scheme (H) for TD-DIP.

The metrics in Table 2 and the qualitative comparison in Figure 6 with zoomed-in reconstructions and absolute error maps, show that RED-PSM performs better than both versions of TD-DIP.

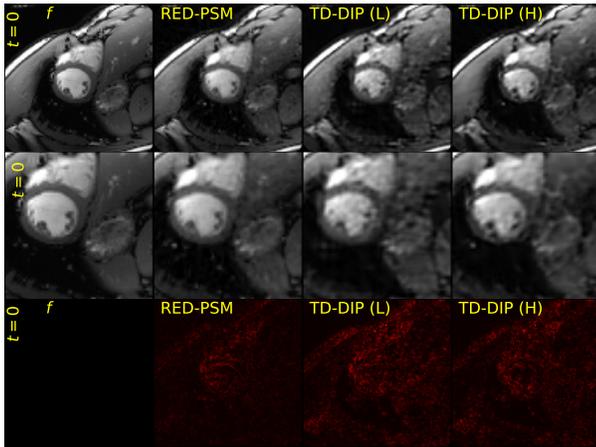


Figure 6: Reconstructed frames for  $P=256$  for retrospective dMRI data [20] with zoomed-in frames (middle row), and absolute reconstruction errors (last row).

$P$	Method	PSNR (dB)	SSIM	MAE (1e-2)	HFEN
128	TD-DIP (L)	34.6	0.923	1.7	3.38
	TD-DIP (H)	34.6	0.928	1.7	3.32
	<b>RED-PSM</b>	<b>36.6</b>	<b>0.939</b>	<b>1.4</b>	<b>2.73</b>
256	TD-DIP (L)	36.2	0.948	1.4	3.47
	TD-DIP (H)	36.4	0.947	1.4	3.44
	<b>RED-PSM</b>	<b>38.4</b>	<b>0.962</b>	<b>1.1</b>	<b>2.99</b>

Table 2: Reconstruction accuracies for RED-PSM and TD-DIP for the retrospective dMRI data [20].

### 5.5. PSNR vs. $t$ , and a Patch-Based Scheme

In Supplementary Material A, we compare the PSNR of the reconstructed frames for each  $t$  for all methods for the walnut object with  $P = 256$  and show the consistently better performance of RED-PSM.

Also, in Supplementary Material B, to improve scalability to high-resolution and high-dimensional settings, we propose and evaluate a patch-based version of RED-PSM with a patch-based denoiser  $D_\phi$ . In the experiments,

the patch-based scheme provided reconstruction accuracies similar to those reported here for the original scheme.

## 6. Conclusions

RED-PSM is the first PSM-based approach to dynamic imaging using a pre-trained and learned (RED-based) spatial prior. The objective in the proposed variational formulation is optimized using a novel and effective bilinear ADMM algorithm, which enforces the PSM as a hard constraint. Unlike existing PSM-based techniques, RED-PSM is supported by theoretical analysis, with a convergence guarantee to a stationary point of the objective. The results of the numerical experiments show better reconstruction accuracy and considerably faster run times compared to a recent DIP-based algorithm. A patch-based regularizer version of RED-PSM provides almost equivalent performance with a massive reduction of storage requirements, indicating the potential of our framework for dynamic high-resolution 2D or 3D settings.

Possible directions for future work include the application of RED-PSM to various imaging scenarios other than tomography and MRI, and robust denoiser training for the RED framework, since the deep denoisers encounter varying artifact distributions during optimization. This could also improve the generalizability of the framework to different input types.

**Acknowledgements.** This research was supported in part by Los Alamos National Labs under Subcontract No. 599416/CW13995. We thank Dr. Brian M. Patterson for providing the data for the compressed material experiments.

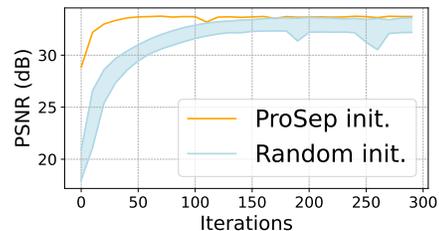


Figure 7: PSNR vs. iteration comparison for RED-PSM as described in Section 5.4.2. The area between the best and the worst PSNR for each iteration is shown in blue to highlight the varying performances of five different runs with random initialization.

## References

- [1] E. Maire, C. Le Bourlot, J. Adrien, A. Mortensen, and R. Mokso, "20 Hz X-ray tomography during an in situ tensile test," *International Journal of Fracture*, vol. 200, no. 1, pp. 3–12, 2016.
- [2] I. Danad, J. Szymonifka, J. Schulman-Marcus, and J. K. Min, "Static and dynamic assessment of myocardial perfusion by computed tomography," *European Heart Journal - Cardiovascular Imaging*, vol. 17, no. 8, pp. 836–844, 03 2016. [Online]. Available: <https://doi.org/10.1093/ehjci/jew044>
- [3] P. J. Keall *et al.*, "Acquiring 4D thoracic CT scans using a multislice helical method," *Physics in Medicine & Biology*, vol. 49, no. 10, p. 2053, 2004.
- [4] A. Scanziani, K. Singh, T. Bultreys, B. Bijeljic, and M. J. Blunt, "In situ characterization of immiscible three-phase flow at the pore scale for a water-wet carbonate rock," *Advances in Water Resources*, vol. 121, pp. 446–455, 2018.
- [5] J. O'Connor, P. Tofts, K. Miles, L. Parkes, G. Thompson, and A. Jackson, "Dynamic contrast-enhanced imaging techniques: CT and MRI," *The British journal of radiology*, vol. 84, pp. S112–S120, 2011.
- [6] B. M. Patterson, N. L. Cordes, K. Henderson, J. J. Williams, T. Stannard, S. S. Singh, A. R. Ovejero, X. Xiao, M. Robinson, and N. Chawla, "In situ x-ray synchrotron tomographic imaging during the compression of hyper-elastic polymeric materials," *Journal of materials science*, vol. 51, no. 1, pp. 171–187, 2016.
- [7] B. M. Patterson, L. Kuettner, T. Shear, K. Henderson, M. J. Herman, A. Ionita, N. Chawla, J. Williams, T. Sun, K. Fezzaa *et al.*, "Synchrotron CT imaging of lattice structures with engineered defects," *Journal of Materials Science*, vol. 55, no. 25, pp. 11 353–11 366, 2020.
- [8] M. Salerno, B. Sharif, H. Arheden, A. Kumar, L. Axel, D. Li, and S. Neubauer, "Recent advances in cardiovascular magnetic resonance: techniques and applications," *Circulation: Cardiovascular Imaging*, vol. 10, no. 6, p. e003951, 2017.
- [9] P. Willis and Y. Bresler, "Optimal scan for time-varying tomography. I. Theoretical analysis and fundamental limitations," *IEEE Transactions on Image Processing*, vol. 4, no. 5, pp. 642–653, 1995.
- [10] N. Willis and Y. Bresler, "Optimal scan for time-varying tomography. II. Efficient design and experimental validation," *IEEE Transactions on Image Processing*, vol. 4, no. 5, pp. 654–666, 1995.
- [11] D. Robinson and P. Milanfar, "Fast local and global projection-based methods for affine motion estimation," *Journal of Mathematical Imaging and Vision*, vol. 18, no. 1, pp. 35–54, 2003.
- [12] X. Xiong and K. Qin, "Linearly estimating all parameters of affine motion using radon transform," *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4311–4321, 2014.
- [13] C. Jailin and S. Roux, "Dynamic tomographic reconstruction of deforming volumes," *Materials*, vol. 11, no. 8, 2018. [Online]. Available: <https://www.mdpi.com/1996-1944/11/8/1395>
- [14] C. Jailin, S. Roux, D. Sarrut, and S. Rit, "Projection-based dynamic tomography," *Physics in Medicine & Biology*, vol. 66, no. 21, p. 215018, 2021.
- [15] J. P. Haldar and Z.-P. Liang, "Spatiotemporal imaging with partially separable functions: A matrix recovery approach," in *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, April 2010, pp. 716–719.
- [16] B. Iskender, Y. Bresler, and M. L. Klasky, "Dynamic tomography reconstruction by projection-domain separable modeling," in *2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*. IEEE, 2022, pp. 1–5.
- [17] B. Iskender, M. L. Klasky, and Y. Bresler, "Dynamic tomography reconstruction by projection-domain separable modeling," *arXiv:2204.09935*, 2022.
- [18] B. Zhao, J. P. Haldar, A. G. Christodoulou, and Z.-P. Liang, "Image reconstruction from highly undersampled (k, t)-space data with joint partial separability and sparsity constraints," *IEEE Transactions on Medical Imaging*, vol. 31, no. 9, pp. 1809–1820, 2012.
- [19] S. Babu *et al.*, "Fast low rank column-wise compressive sensing for accelerated dynamic MRI," *IEEE Transactions on Computational Imaging*, 2023.
- [20] J. Yoo, K. H. Jin, H. Gupta, J. Yerly, M. Stuber, and M. Unser, "Time-dependent deep image prior for dynamic MRI," *IEEE Transactions on Medical Imaging*, vol. 40, no. 12, pp. 3337–3348, 2021.
- [21] Q. Zou, A. H. Ahmed, P. Nagpal, S. Kruger, and M. Jacob, "Dynamic imaging using a deep generative SToRM (GenSToRM) model," *IEEE Transactions on Medical Imaging*, vol. 40, no. 11, pp. 3102–3112, 2021.
- [22] A. H. Ahmed, Q. Zou, P. Nagpal, and M. Jacob, "Dynamic imaging using deep bi-linear unsupervised representation (DEBLUR)," *IEEE Transactions on Medical Imaging*, vol. 41, no. 10, pp. 2693–2703, 2022.
- [23] Y. Romano, M. Elad, and P. Milanfar, "The little engine that could: Regularization by denoising (red)," *SIAM Journal on Imaging Sciences*, vol. 10, no. 4, pp. 1804–1844, 2017.
- [24] in *Methods of Modern Mathematical Physics*, M. Reed and B. Simon, Eds. Academic Press, 1972.
- [25] S. Ma, H. Du, Q. Wu, and W. Mei, "Dynamic MRI reconstruction exploiting partial separability and t-SVD," in *2019 IEEE 7th International Conference on Bioinformatics and Computational Biology (ICBCB)*. IEEE, 2019, pp. 179–184.
- [26] S. Ma, Y. Fan, and Z. Li, "Dynamic MRI exploiting partial separability and shift invariant discrete wavelet transform," in *2021 6th International Conference on Image, Vision and Computing (ICIVC)*. IEEE, 2021, pp. 242–246.
- [27] Y. Sun, J. Liu, and U. Kamilov, "Block coordinate regularization by denoising," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

- [28] E. T. Reehorst and P. Schniter, “Regularization by denoising: Clarifications and new interpretations,” *IEEE transactions on computational imaging*, vol. 5, no. 1, pp. 52–67, 2018.
- [29] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [30] D. Gabay and B. Mercier, “A dual algorithm for the solution of nonlinear variational problems via finite element approximation,” *Computers & mathematics with applications*, vol. 2, no. 1, pp. 17–40, 1976.
- [31] J. Eckstein and D. P. Bertsekas, “On the douglas—rachford splitting method and the proximal point algorithm for maximal monotone operators,” *Mathematical Programming*, vol. 55, no. 1, pp. 293–318, 1992.
- [32] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising,” *IEEE transactions on image processing*, vol. 26, no. 7, pp. 3142–3155, 2017.
- [33] B. Iskender, M. L. Klasky, and Y. Bresler, “RED-PSM: Regularization by denoising of partially separable models for dynamic imaging,” *arXiv preprint arXiv:2304.03483*, 2023.
- [34] D. Hajinezhad and Q. Shi, “Alternating direction method of multipliers for a class of nonconvex bilinear optimization: convergence analysis and applications,” *Journal of Global Optimization*, vol. 70, pp. 261–288, 2018.
- [35] M. Udell *et al.*, “Generalized low rank models,” *Foundations and Trends® in Machine Learning*, vol. 9, no. 1, pp. 1–118, 2016.
- [36] H. Der Sarkissian, F. Lucka, M. van Eijnatten, G. Colacicco, S. B. Coban, and K. J. Batenburg, “A cone-beam X-ray computed tomography data collection designed for machine learning,” *Scientific data*, vol. 6, no. 1, p. 215, 2019.
- [37] “Piecewise Affine Transformation,” [https://scikit-image.org/docs/dev/auto\\_examples/transform/plot\\_piecewise\\_affine.html](https://scikit-image.org/docs/dev/auto_examples/transform/plot_piecewise_affine.html).
- [38] B. M. Patterson, “Personal communication,” 2022.
- [39] F. Reda, J. Kontkanen, E. Tabellion, D. Sun, C. Pantofaru, and B. Curless, “FILM: Frame interpolation for large motion,” *arXiv preprint arXiv:2202.04901*, 2022.
- [40] J. Yoo, “Time-Dependent Deep Image Prior for Dynamic MRI,” <https://github.com/jaejun-yoo/TDDIP>, 2021.
- [41] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [42] O. Bernard *et al.*, “Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved?” *IEEE Transactions on Medical Imaging*, 2018.
- [43] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [44] S. Ravishankar and Y. Bresler, “MR image reconstruction from highly undersampled k-space data by dictionary learning,” *IEEE Transactions on Medical Imaging*, vol. 30, no. 5, pp. 1028–1041, 2010.