

Video Compressive Sensing for Dynamic MRI

Jianing V. Shi^{1,2,*}, Wotao Yin², Aswin C. Sankaranarayanan³, and Richard G. Baraniuk¹

¹ Department of Electrical and Computer Engineering, Rice University

² Department of Mathematics, UCLA

³ Department of Electrical and Computer Engineering, Carnegie Mellon University

January 28, 2014

Abstract

We present a video compressive sensing framework, termed kt-CSLDS, to accelerate the image acquisition process of dynamic magnetic resonance imaging (MRI). We are inspired by a state-of-the-art model for video compressive sensing that utilizes a linear dynamical system (LDS) to model the motion manifold. Given compressive measurements, the state sequence of an LDS can be first estimated using system identification techniques. We then reconstruct the observation matrix using a joint structured sparsity assumption. In particular, we minimize an objective function with a mixture of wavelet sparsity and joint sparsity within the observation matrix. We derive an efficient convex optimization algorithm through alternating direction method of multipliers (ADMM), and provide a theoretical guarantee for global convergence. We demonstrate the performance of our approach for video compressive sensing, in terms of reconstruction accuracy. We also investigate the impact of various sampling strategies. We apply this framework to accelerate the acquisition process of dynamic MRI and show it achieves the best reconstruction accuracy with the least computational time compared with existing algorithms in the literature.

1 Introduction

Our fascination with detail has lead to sensors of ever increasing capabilities. In many modalities, the traditional sampling theory associated with Nyquist sampling theorem fails to deliver high spatio-temporal resolution or at best, delivers this at prohibitive costs. Compressive sensing [9, 8, 17] has recently arisen as a paradigm to revolutionize the design of sensors and signal processing [5], and is promised to deliver better sensors. The key insight of compressive sensing is that one can design a sensing system that only acquires a few linear measurements and recover the signals via convex optimization or greedy pursuit. Two assumptions are essential to the success of compressive sensing: a) the signal can be approximated using sparse representation under a suitable basis or dictionary; b) linear measurements are suitably incoherent with the basis or dictionary in which the signal is represented. Compressive sensing enables one to acquire highly undersampled data during the acquisition process, with the sampling rate way below the Nyquist sampling frequency.

*Corresponding author's email address: jianing@math.ucla.edu

A prominent application of compressive sensing is to accelerate the acquisition process of magnetic resonance imaging (MRI). In fact, the discovery of compressive sensing was largely motivated by the MRI problem [9], where one wishes to reconstruct an object based on incomplete Fourier samples. The physical mechanism of MRI [14, 28, 34] requires scanning in the Fourier space in order to reconstruct an object. The speed of imaging is fundamentally limited by physical constraints such as gradient amplitude and slew rate, as well as physiological constraints [31]. Compressive sensing has proven to be very successful in accelerating the acquisition process of MRI and has opened up many possibilities for new clinical applications [31, 32]. Dynamic MRI reconstructs a dynamic sequence of images based on measurements in spatial frequency versus time (k-t) domain, roughly speaking, video acquired in the Fourier space. As pointed out in [32], dynamic MRI is challenging due to the time-varying nature of the imaging object and the spatio-temporal tradeoff. Moreover, the discrepancy between the dynamic nature of the moving object and a static scene assumption for sensing creates a spatio-temporal recovery error [45, 39]. To tackle such a challenge, we leverage ideas from state-of-the-art video compressive sensing frameworks.

Video compressive sensing is nontrivial due to its high-dimensional representation and the ephemeral nature of videos. In order to achieve recovery using as few samples as possible, it requires one to exploit the redundancy in the ambient space and go beyond a frame-by-frame reconstruction [52]. The general philosophy is to identify a model in which signal can be represented parsimoniously, and identify the basis or dictionary to sparsify the signal under the signal model. Several signal models have been employed to perform video compressive sensing. One early approach considered sparse representation in both spatial and temporal domains by treating video as a three-dimensional matrix, and employed 3D wavelet transform to sparsify the video [52]. A later approach took advantage of the small inter-frame differences together with spatial 2D wavelet transform within each frame, which was implemented in the compressive coded aperture video camera [35]. Further work along the video coding ideas sought to reconstruct individual frames based on wavelet sparsity in the spatial domain, while modeling temporal dependencies between frames using motion compensation methods, such as lifting based wavelet sparsity [48] and optical flow [44]. Multi-scale recovery algorithms along with various motion compensation mechanisms were investigated in [38, 45, 39]. Work based on the separation of background and moving objects were investigated in [10]. Video compressive sensing models based on the entire image typically involve a dense measurement matrix, which is computationally expensive, therefore block-based video compressive sensing [20] divided each frame into small blocks in order to accelerate computation. Dictionary learning-based methods were proposed to identify task specific basis for compressive sensing reconstruction [41, 13, 42]. Another approach of exploiting the redundancy was to consider motion manifold and build a global model for the video cube. The key idea is to project the original video cube onto a motion manifold, and perform reconstruction within a low-dimensional space. Motion manifold models arise in many computer vision and machine learning problems, such as dynamical textures modeling [18, 11], human activity tracking [6, 50], video-based face recognition [2], data-driven motion synthesis [29], video compressive sensing [46], coded strobing photography [49]. A key promise of the motion manifold model is to obtain a compact representation of high-dimensional data by exploring the spatio-temporal structures, hence enabling computation on the low-dimensional manifold.

Prior work on compressively sensed dynamic MRI include k-t SPARSE [33], k-t FOCUSS [25], Modified CS [30], MASTeR [4], Subtraction Sparsity [43], and L+S reconstruction [36]. k-t SPARSE and k-t FOCUSS both use the wavelet transform to model sparsity in the temporal domain. Modified CS identifies signal support in the first frame to facilitate reconstruction of the rest video frames using compressive sensing. MASTeR uses motion adaptive spatio-temporal regularization to perform reconstruction based on compressive k-t data. Subtraction Sparsity is designed for contrast-enhanced magnetic resonance angiography, which subtracts a pre-contrast mask from all post-contrast frames to promote sparsity in the resulting difference images. L+S reconstruction uses a low rank and sparse matrix decomposition to separate background and dynamic components.

In this paper, we have chosen to build upon CS-LDS [46], a video compressive sensing approach for time varying signals modeled as a linear dynamical system (LDS). Encouraged by its high-fidelity reconstruction quality for a wide arrange of videos and its ability to achieve high compression rates, we extend the CS-LDS model to the k-t domain. We propose a novel and efficient algorithm, which we call kt-CSLDS. Our proposed algorithm takes advantage of the orthonormal property of the Fourier operator and uses a number of numerical techniques to achieve high computational efficiency. We provide theoretical guarantee for its global convergence. We use the kt-CSLDS model to accelerate the image acquisition process of dynamic MRI. Finally, we investigate the impact of sampling strategies on the reconstruction quality of dynamic MRI.

2 Compressive Sensing Dynamic MRI Model

Learning a low-dimensional signal model based on video data is an important topic in computer vision, signal processing and machine learning. Linear dynamical systems (LDSs) are a particularly useful model which builds a compact representation of the spatial and temporal variations in image sequences. This arises in a range of applications including object recognition, video segmentation, and video synthesis. We are primarily motivated by video synthesis, since it aligns perfectly with goal of compressive sensing. Our video compressive sensing model is largely inspired by CS-LDS [46], which couples compressive sensing with linear dynamical systems to perform video compressive sensing. Since we are interested in accelerating the image acquisition process of MRI, we extend the CS-LDS model from the spatial domain to the Fourier domain.

2.1 Notations

We clarify the notations in this section. A video can be denoted by a 3D tensor $Y^3 \in \mathbb{R}^{n_x \times n_y \times l}$. For ease of notation, we vectorize each frame of the video and represent it using 2D matrices; hence, for the rest of the paper, we represent videos as $Y \in \mathbb{R}^{n \times l}$, where $n = n_x n_y$. We use $X \in \mathbb{R}^{d \times l}$ to denote the state sequence over time, where at each time t the state vector is \mathbf{x}_t . We denote the compressive sensed k-t video cube as $Z \in \mathbb{R}^{m \times l}$, where m is the number of measurements.

We use different notations for row space and column space of matrices. In particular, for the observation matrix $C \in \mathbb{R}^{n \times d}$, we define the row vector and column vector as follows. We denote each row of C using a row vector $\mathbf{c}_{(i)} := \mathbf{e}_{(i)} C$, where $\mathbf{e}_{(i)} = (0, \dots, 1, \dots, 0)^\top$ is a

row selector with the i th element being 1, $i = 1, 2, \dots, n$. Similarly, we denote each column of C using a column vector $\mathbf{c}_j := C\mathbf{e}_j$, where $\mathbf{e}_j = (0, \dots, 1, \dots, 0)$ is a column selector with the j th element being 1, $j = 1, 2, \dots, d$.

We use ψ_2 to denote the 2D wavelet transform. We define the following operator to denote frame-by-frame wavelet transform for each frame of the video cube $C \in \mathbb{R}^{n \times d}$ as

$$\Psi\mathbf{c}_j = \Psi(\mathbf{c}_j) := \text{vec}(\psi_2\mathbf{c}_j^2), \quad (1)$$

where \mathbf{c}_j^2 denotes matricization of the 3D tensor \mathbf{c}_j by collapsing the first two dimension, resulting in a 2D matrix,

$$\mathbf{c}_j^2 = \text{mat}(\mathbf{c}_j), \quad (2)$$

for $j = 1, 2, \dots, d$.

With such a notation, the frame-by-frame wavelet transform for the entire video cube can be represented by

$$\Psi(C) = \begin{pmatrix} \left| \begin{array}{c} \text{vec}(\psi_2\mathbf{c}_1^2) \\ \vdots \end{array} \right| & \left| \begin{array}{c} \text{vec}(\psi_2\mathbf{c}_2^2) \\ \vdots \end{array} \right| & \dots & \left| \begin{array}{c} \text{vec}(\psi_2\mathbf{c}_d^2) \\ \vdots \end{array} \right| \end{pmatrix}. \quad (3)$$

2.2 Compressive Sensed Dynamic MRI

We first introduce the signal model for compressive sensing dynamic MRI. Dynamic MRI imaging typically takes measurements of a moving object $\mathbf{y}_t \in \mathbb{R}^n$ in the Fourier space, which result in a sequence of Fourier measurements $\mathbf{z}_t \in \mathbb{R}^n$,

$$\mathbf{z}_t = \mathcal{F}\mathbf{y}_t + \xi_t, \quad (4)$$

where $\xi_t \in \mathbb{R}^n$ is the measurement noise. Traditional imaging system takes the full Fourier space samples and reconstructs a video using inverse Fourier transform,

$$\hat{\mathbf{y}}_t = \mathcal{F}^{-1}\mathbf{z}_t. \quad (5)$$

With the traditional imaging sequence, the sampling rate needs to satisfy the Nyquist sampling theorem, which fundamentally limits the temporal resolution of the dynamic MRI imaging.

Now with compressive sensing, we take partial measurements in the k-t space and increase the temporal resolution of the imaging acquisition process,

$$\mathbf{z}_t = \Phi_t\mathcal{F}\mathbf{y}_t + \xi_t, \quad (6)$$

where $\Phi_t \in \mathbb{R}^{m \times n}$ is the measurement matrix, and $\mathbf{z}_t \in \mathbb{R}^m$ represents partial Fourier measurements. Note Φ_t is a row selector in the Fourier space, and takes the form of a subsampled identity matrix. Our goal is recover \mathbf{y}_t based on measurement matrix and partial Fourier measurements.

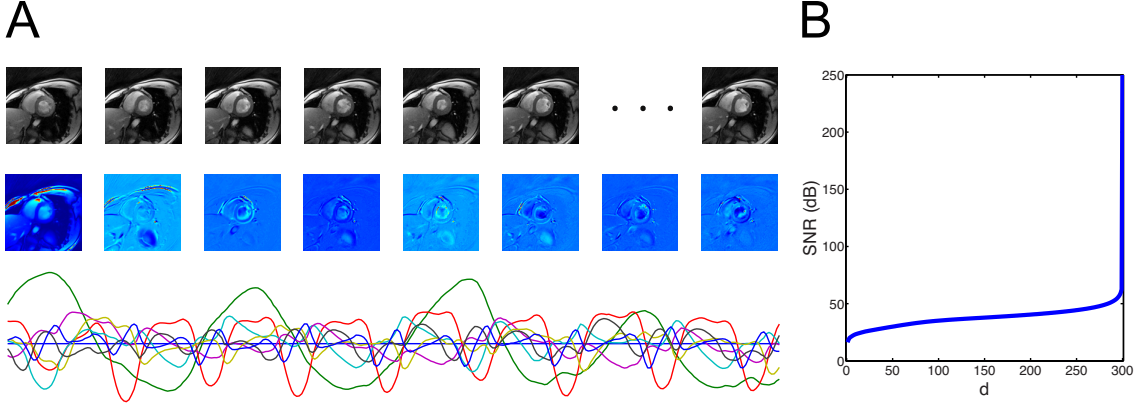


Figure 1: (A) Illustration of the LDS model for a sample heart video: (top) sequence of individual frames of a video Y , (middle) observation matrix C , (bottom) state sequence X . (B) LDS as a good approximation of the original video, where we plot SNR as a function of d .

2.3 Linear Dynamical Systems

Linear time-invariant dynamical systems (LDS) can be expressed using three components: 1) an observation model that defines the state space and the observation matrix linking observations to state space, 2) state transition model that captures dynamics on the state space, and 3) an initial condition.

More specifically, the discrete form of LDS can be expressed as

$$\mathbf{y}_t = C\mathbf{x}_t + \omega_t \quad \omega_t \sim \mathcal{N}(0, Q) \quad (7a)$$

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + \nu_t \quad \nu_t \sim \mathcal{N}(0, R) \quad (7b)$$

at each time instant ($t = 1, 2, \dots, l$), together with the initial condition x_0 . In the above, $\mathbf{y}_t \in \mathbb{R}^n$ represents the observation (in our case, the original video frames), $\mathbf{x}_t \in \mathbb{R}^d$ are the hidden states, $C \in \mathbb{R}^{n \times d}$ is the observation matrix, $A \in \mathbb{R}^{d \times d}$ is the transition matrix. $\omega_t \in \mathbb{R}^n$ is the process noise, which include excitation driving the stochastic process and error in the Markov model. $\nu_t \in \mathbb{R}^n$ is the observation noise, modeling inaccuracies in the LDS model.

We first illustrate the concept of LDS model using a sample video, see Figure 1(A). With such a model, observations \mathbf{y}_t can be represented as linear transformation of the state \mathbf{x}_t , corrupted by observation noise, whereas the states \mathbf{x}_t evolve according to a first-order Markov process corrupted by process noise. The noise terms ω_t and ν_t are assumed to be temporally white, independent of each other, the states and the observations. If the noises are assumed to be zero-mean Gaussian spatially, then the LDS model corresponds to a first-order Gaussian Markov random process. We focus on the Gaussian noise case throughout this paper, with $\omega_t \sim \mathcal{N}(0, Q)$ and $\nu_t \sim \mathcal{N}(0, R)$.

In the case where $d \ll n$, the motion manifold is a good model to reduce the high-dimensional video cube into a low-dimensional representation. We illustrate such a concept in Figure 1(B). The key promise of using LDS relies on the assumption that high-dimensional

signal $\mathbf{y}_t \in \mathbb{R}^n$ can be faithfully represented using low-dimensional state sequence $\mathbf{x}_t \in \mathbb{R}^d$, with $d \ll n$.

Given $Y = (y_1, y_2, \dots, y_l)$, we can obtain a d -dimensional LDS approximation of the original video cube Y through SVD, and measure the accuracy of such an approximation. When we have an estimate of Y , $\hat{Y} = C(d)X(d)$, the reconstruction SNR of \hat{Y} is given as

$$\text{SNR} = 10 \log_{10} \frac{\|Y\|_F^2}{\|\hat{Y} - Y\|_F^2}, \quad (8)$$

which is a function of d . We obtain reasonably good SNR even at low d , as shown in Figure 1(B).

2.4 Compressive Measurement Model

Now with compressive measurements in the Fourier space, we have

$$\mathbf{z}_t = \Phi_t \mathcal{F} C \mathbf{x}_t + \omega_t \quad \omega_t \sim \mathcal{N}(0, Q) \quad (9a)$$

$$\mathbf{x}_{t+1} = A \mathbf{x}_t + \nu_t \quad \nu_t \sim \mathcal{N}(0, R), \quad (9b)$$

where \mathcal{F} is the multi-dimensional Fourier transform in the image domain Ω . $\Phi_t \in \mathbb{R}^{m \times n}$ is the measurement matrix, essentially the row selector which stipulates where we sample in the k space during the image acquisition process,

$$\Phi_t = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Our measurement model is comprised of two components, a time-invariant component and a time-variant component, as proposed in [46, 47],

$$\mathbf{z}_t = \begin{pmatrix} \bar{\mathbf{z}}_t \\ \tilde{\mathbf{z}}_t \end{pmatrix} = \begin{pmatrix} \bar{\Phi} \mathcal{F} \\ \tilde{\Phi}_t \mathcal{F} \end{pmatrix} \mathbf{y}_t. \quad (10)$$

Note that $\bar{\Phi}$ is the time-invariant component, which is designed to facilitate estimation for the state sequence. $\tilde{\Phi}_t$ is the time-variant component to allow innovation during the acquisition process, due to the ephemeral nature of videos.

The overall architecture of kt-CSLDS is shown in Figure 2, where we take compressive measurements of the original video signal. Each frame of the video is a high-dimensional signal is projected to a low-dimensional representation that follows a Markov process.

The design of the measurement model is critical to the success of video compressive sensing. The time-invariant component needs to satisfy the observability condition for LDS. In the Fourier domain, the low-frequency content of a video changes marginally, while the high-frequency content changes more drastically from frame to frame. We therefore sample the low-frequency content more densely. Note in both kt-FOCUSS and MASTeR, low-frequency domain is sampled in a similar fashion. The time-variant component is designed to satisfy the incoherence assumptions for compressive sensing, promoting randomness of the Fourier samples in the temporal domain. Within each frame, we sample the Fourier space according to the theoretical results obtained by [26].

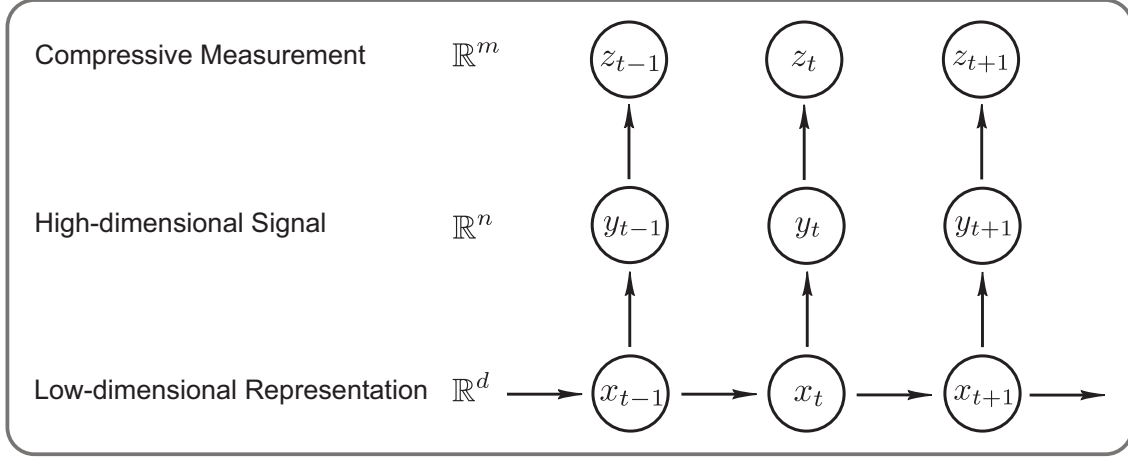


Figure 2: Architecture of the kt-CSLDS model. High-dimensional signal is projected to a low-dimensional representation that follows a Markov process. The compressive measurements are observations drawn from the high-dimensional signal.

3 Compressive Sensing Reconstruction Algorithm

3.1 Algorithm Flow

The overall architecture of imaging acquisition and video reconstruction process is summarized in Figure 3. Upon acquiring Fourier samples using the aforementioned sampling strategy, we reconstruct the video frames for dynamic MRI via a two-step procedure.

We first estimate the state sequence \hat{X} based on the time-invariant samples using system identification. We then reconstruct the observation matrix \hat{C} based on both time-invariant and time-variant measurements by exploring certain sparsity assumptions. The final video reconstruction can be obtained by

$$\hat{Y} = \hat{C}\hat{X}. \quad (11)$$

Here we set up the general framework for algorithm flow, whose details will be discussed in the following two sections. In Section 3.2 we discuss two different methods for state sequence estimation. In Section 3.3, we derive an efficient numerical algorithm for observation matrix reconstruction and prove its theoretical convergence.

3.2 State Sequence Estimation

State sequence \mathbf{x}_t can be estimated based on the time-invariant component,

$$\bar{\mathbf{z}}_t = \bar{\Phi}\mathcal{F}C\mathbf{x}_t + \omega_t \quad \omega_t \sim \mathcal{N}(0, Q) \quad (12a)$$

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + \nu_t \quad \nu_t \sim \mathcal{N}(0, R). \quad (12b)$$

Since that Fourier transform is a linear operator, this sub-system is a linear time-invariant system. Given that there is no input data, the system is driven purely by stochastic noise.

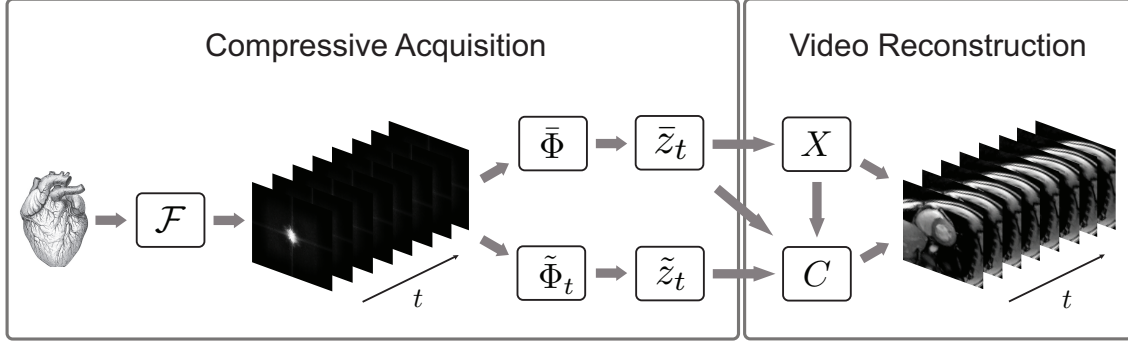


Figure 3: Architecture of the kt-CSLDS imaging and reconstruction processes. During the imaging acquisition process of dynamic MRI, compressive samples are obtained in the Fourier space. We then perform video reconstruction based on the compressive measurements.

Given compressive measurements of video data $\bar{Z} = (\bar{z}_1, \bar{z}_2, \dots, \bar{z}_t, \dots, \bar{z}_l)$, we assume they are generated by the time-invariant LDS described in Eq. (12). Our goal is to determine the order of the unknown system d , and the forward Kalman filter state sequence $\hat{X} = (x_1, x_2, \dots, x_t, \dots, x_l)$ up to a similarity transformation.

There exist many ways to perform state inference and system identification for the LDS model. State inference refers to the process of estimating hidden states over time $\hat{X} := (x_1, \dots, x_t, \dots, x_l)$ given the observations $\bar{Z} := (\bar{z}_1, \dots, \bar{z}_t, \dots, \bar{z}_l)$ and parameters $\theta := \{A, C, Q, R\}$. On the other hand, system identification involves finding the parameters θ and the distribution over hidden states $p(X|\bar{Z}, \theta)$ that maximizes the likelihood of the observed data \bar{Z} .

When one uses the recursive formulation as stated in (12), it has a connection to the Kalman filter. The stochastic LDS models the distribution of outputs $p(\bar{z}_{1:l})$, and the inference problem for LDS aims to estimate the distribution over hidden states $p(x_t|\bar{z}_{1:l})$. The inference can be carried out recursively, by combining a forward pass and a backward pass. The forward pass takes the initial state x_0 together with a collection of observation states $y_{1:t}$, and computes x_t recursively, resulting in the Kalman filter. The backward pass takes the observations from \bar{z}_l to \bar{z}_{t+1} , and corrects the results from the forward pass by evaluating the influence of future observations, which is also known as the Rauch-Tung-Striebel (RTS) equation.

3.2.1 Review of System Identification

Generally speaking, there are two types of methods for system identification. One approach is to obtain parameters θ and distribution $p(X|\bar{Z}, \theta)$ using maximum likelihood solution through iterative techniques such as expectation maximization (EM). The EM approach utilizes the Kalman filter and Kalman smoother, which requires the entire observation sequences. EM guarantees convergence to a local maximum in the likelihood surface, and is sensitive to initial condition. Another approach for system identification is to use subspace methods to obtain solutions, which is known as subspace identification.

In subspace identification, it is desirable to find the minimal model order for the state space representation under the constraint that the reduced model approximates the output data. It is well-known that the minimal order is equal to the rank of the block Hankel matrix, defined as

$$\mathcal{H}_d(z) = \begin{pmatrix} z_1 & z_2 & \cdots & z_{l-d+1} \\ z_2 & z_3 & \cdots & z_{l-d+2} \\ \vdots & \vdots & \ddots & \vdots \\ z_d & z_{d+1} & \cdots & z_l \end{pmatrix}. \quad (13)$$

Therefore, subspace identification methods typically exploit the rank of block Hankel matrix, and relies on matrix decomposition to obtain state sequence estimates \hat{X} as well as the realization of the state space model parameterized by $\{A, C, Q, R\}$. Algorithm details vary among different subspace identification methods: a) construction of the block Hankel matrices differ depending on whether it is covariance driven or data driven, b) the matrix decomposition methods vary using different user defined weighting matrices in the projection methods. In several classical subspace identification algorithms, oblique projection is employed when there exist both output data and input data, which reduces to orthogonal projection in the case of stochastic identification. These algorithms include principle component analysis (PCA), unweighted principle component analysis (UPCA) [3], canonical variate analysis (CVA) [27], numerical algorithms for subspace identification (N4SID) [37], and multivariate output error state space (MOESP) [51]. These methods can determine the order of the system for the state space model, and estimate system matrices $\{A, C\}$ up to a similarity transformation. Note that subspace identification methods mentioned above use orthogonal projections and can be computationally expensive. It is noteworthy that N4SID provides asymptotically optimal solution for the forward Kalman filter state sequence \hat{X} , in the sense of maximum likelihood. However, the memory storage and computation requirement of N4SID and other subspace methods are prohibitively expensive for video data. As pointed out in [18], under mild conditions, one can obtain a closed-form solution for a canonical model realization. We adopt such a strategy in this paper in favor of its algebraic simplicity and computational efficiency.

3.2.2 Canonical Model Realization

In stochastic identification, the goal is typically to determine the system matrices $\theta = \{A, C, Q, R\}$ up to a similarity transformation. Obtaining $\{A, C, Q, R\}$ is also called a *realization* of the system. The ambiguity of the system identification is well-known, in a sense that there is no unique choice of system matrices which can generate the same sample path given suitable initial condition. As long as $T \in \mathbb{R}^{d \times d}$ is invertible, one can generate the same dynamics by substituting A with TAT^{-1} , C with CT^{-1} , Q with TQT^{-1} , and initial condition x_0 with Tx_0 . Given the bilinear product between C and x_t , it immediately follows that any estimate of the forward Kalman filter state sequence \hat{X} is accurate only up to a similarity transformation.

In order to obtain a unique realization for the LDS, one chooses a representative from these equivalent solutions, which results in the so-called *canonical model realization* [18]. This can be achieved by imposing additional constraints or imposing regularization on the solution.

Algorithm 1 State Sequence Estimate with SVD

Input: Fourier measurements \bar{Z} acquired by time-invariant measurement matrix $\bar{\Phi}$ and system order d

1° Formulate block Hankel matrix $\mathcal{H}_d(\bar{z})$

2° Perform SVD $\mathcal{H}_d(z) = U\Sigma V^\top$

3° Keep the first d eigensystems U_d, Σ_d, V_d

4° Calculate state sequence estimate $\hat{X} = \Sigma_d V_d^\top$

3.2.3 Canonical Model Realization with Truncated SVD

Corollary 3.1 (Canonical model realization, [18]) *Suppose one has access to the full video data in the spatial domain $Y = (y_1, y_2, \dots, y_t, \dots, y_l)$. Assuming the observation matrix $C \in \mathbb{R}^{n \times d}$ of the canonical model has orthonormal columns, i.e., $C^\top C = I$, then one can obtain a closed-form solution for the forward Kalman filter state sequence \hat{X} and canonical model realization \hat{C} based on the simplest form of Hankel matrix $\mathcal{H}_{1,l}$.*

Proof. Once formulating the simplest form of block Hankel matrix $\mathcal{H}_1(y)$, we note the following relationship with state sequence $X = (x_1, x_2, \dots, x_t, \dots, x_l)$ and noise term $W = (\omega_1, \omega_2, \dots, \omega_t, \dots, \omega_l)$,

$$\mathcal{H}_1(y) = CX + W.$$

The estimation for observation matrix C and state sequence X can be formulated as

$$\hat{C}, \hat{X} = \arg \min_{C, X} \|\mathcal{H}_{1,l}(y) - CX\|_F^2.$$

Given the bilinear product between C and X , one can immediately inspect that the solution is not unique. Since we have imposed additional constraint $C^\top C = I$, one can obtain a canonical model realization. It follows from the fixed rank property of SVD [24] that a unique closed-form solution can be obtained through the SVD:

$$\mathcal{H}_1(y) = U\Sigma V^\top \quad \text{where} \quad U^\top U = I \quad V^\top V = I.$$

Note the system order d can be determined from the rank of the block Hankel matrix, which leads to the following

$$\hat{C} = U \quad \hat{X} = \Sigma V^\top, \tag{14}$$

where both estimates are closed-form solutions. ■

The above result sheds light on a simpler path of estimating state sequence \hat{X} , without the computational burden of system identification methods. We remark that canonical model realization based on the simplest form of Hankel matrix [18] essentially does not exploit the structure of LDS. Once we consider a higher degree Hankel matrix, the structure and observability of LDS comes into consideration.

Formally, an LDS is said to be observable if, for any possible state sequence, the current state can be determined in finite time using only the outputs. Less formally, observability refers to the idea that it is possible to determine the behavior of the entire system based on merely the system's outputs. Conversely, an LDS is said to be unobservable if the current values of some states cannot be determined through output sensors. There exists a convenient test for observability.

Lemma 3.2 (Observability, [7]) *For an LDS, equipped with system matrices (C, A) and state space dimension d , the system is observable if the observability matrix*

$$\mathcal{O}(C, A) = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{d-1} \end{pmatrix} \quad (15)$$

is full rank.

The rationale for this test is that if $\mathcal{O}(C, A)$ is rank d , then each of the d states is viewable through linear combinations of the system output $Y = (y_1, y_2, \dots, y_t, \dots, y_l)$.

Theorem 3.3 *Given time-invariant compressive measurements $\bar{\Phi}$ and Fourier video data $\bar{Z} = (\bar{z}_1, \bar{z}_2, \dots, \bar{z}_t, \dots, \bar{z}_l)$, suppose the observability matrix $\mathcal{O}(\bar{\Phi}FC, A)$ is full rank, then there exists a closed-form solution for the forward filter state sequence \hat{X} .*

Proof. Given the video compressive sensing model, we use the time-invariant component to estimate the state sequence. We formulate the block Hankel matrix based on the Fourier measurements \mathcal{H}_d , and by denoting $\bar{C} = \bar{\Phi}FC$,

$$\begin{aligned} \mathcal{H}_d(\bar{z}) &= \begin{pmatrix} \bar{z}_1 & \bar{z}_2 & \dots & \bar{z}_{l-d+1} \\ \bar{z}_2 & \bar{z}_3 & \dots & \bar{z}_{l-d+2} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{z}_d & \bar{z}_{d+1} & \dots & \bar{z}_l \end{pmatrix} \\ &= \begin{pmatrix} \bar{C}x_1 & \bar{C}x_2 & \dots & \bar{C}x_{l-d+1} \\ \bar{C}Ax_1 & \bar{C}Ax_2 & \dots & \bar{C}Ax_{l-d+1} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{C}A^{d-1}x_1 & \bar{C}A^{d-1}x_2 & \dots & \bar{C}A^{d-1}x_{l-d+1} \end{pmatrix} \\ &= \begin{pmatrix} \bar{C} \\ \bar{C}A \\ \vdots \\ \bar{C}A^{d-1} \end{pmatrix} \begin{pmatrix} x_1 & x_2 & \dots & x_{l-d+1} \end{pmatrix} \\ &= \mathcal{O}(\bar{\Phi}FC, A) \begin{pmatrix} x_1 & x_2 & \dots & x_{l-d+1} \end{pmatrix}. \end{aligned}$$

Under the assumption that $\mathcal{O}(\bar{\Phi}FC, A)$ is full-rank, the LDS is observable, according to Lemma 3.4. Therefore, one can obtain a canonical model realization through the SVD

$$\mathcal{H}_d(\bar{z}) = \mathcal{O}(\bar{\Phi}FC, A) \begin{pmatrix} x_1 & x_2 & \dots & x_{l-d+1} \end{pmatrix} = \tilde{U} \tilde{\Sigma} \tilde{V}^\top. \quad (16)$$

This leads to the estimate for the state sequence:

$$\hat{X} = \tilde{\Sigma} \tilde{V}^\top, \quad (17)$$

as a modified closed-form solution. ■

The above result leads to Algorithm 1, where one can estimate the state sequence based on a very simple procedure using SVD.

Algorithm 2 State Sequence Estimate without SVD

Input: video data \bar{Z} , system order d

1° Formulate simplest block Hankel matrix $\mathcal{H}_{1,1,l}(z) = [\bar{z}_1, \bar{z}_2, \dots, \bar{z}_t, \dots, \bar{z}_l]$

2° Initialize $k = 0$, $C^0 \in \mathbb{R}^{n \times d}$, $X^0 \in \mathbb{R}^{d \times l}$, $\omega = 1$, $\tilde{\omega} > 1$, $\delta > 0$, $\gamma_1 \in (0, 1)$

while not converged **do**

 3° Set $(C, X, M) = (C^k, X^k, M^k)$

 4° Compute $M_\omega \leftarrow \omega M^k + (1 - \omega)C^k X^k$

 5° Compute $C_+(\omega) \leftarrow M_\omega X^\top (X X^\top)^\dagger$

 6° Compute $X_+(\omega) \leftarrow (C_+(\omega)^\top C_+(\omega))^\dagger (C_+(\omega)^\top M_\omega)$

 7° Compute $M_+(\omega) \leftarrow \mathcal{H}_{1,1,l}(z)$

 8° Compute residual ratio $\gamma(\omega) \leftarrow \|M - C_+(\omega)X_+(\omega)\|_F / \|M - CX\|_F$

if $\gamma(\omega) \geq 1$ **then**

 Set $\omega = 1$ and go to 4°

end if

 9° Update $(C^{k+1}, X^{k+1}, M^{k+1}) = (C_+(\omega), X_+(\omega), M_+(\omega))$

 Update $k \leftarrow k + 1$

if $\gamma(\omega) \geq \gamma_1$ **then**

 Set $\delta = \max(\delta, 0.25(\omega - 1))$ and $\omega = \min(\omega + \delta, \tilde{\omega})$

end if

end while

3.2.4 Canonical Model Realization without SVD

Algorithm 1 exploits the full rank of the block Hankel matrix, which represents the model complexity. It is well-known the rank of block Hankel matrix can be corrupted when there is noise in the data. Moreover, it is often desirable in system identification to reduce the model complexity. In the view of video compressive sensing, it is favorable to obtain a most compact representation of the video data and perform computation on the corresponding low-dimensional manifold. We thus extend Algorithm 1 to the low rank case.

We formulate the system identification problem as follows:

$$\hat{C}, \hat{X} = \arg \min_{C, X} \|CX - M\|_F^2 \quad \text{s.t.} \quad M = \mathcal{H}_{1,1,l}(z) \quad \text{rank}(CX) = d \quad (18)$$

where a low-rank factorization is sought, resulting in spatial factor $C \in \mathbb{R}^{n \times d}$ and temporal factor $X \in \mathbb{R}^{d \times l}$. Such an approach is designed to alleviate possible corruption of noise, which can increase the rank of block Hankel matrix. We adopt a low-rank factorization algorithm based on nonlinear successive over-relaxation (SOR) [53]. This results in Algorithm 2, which avoids the computation burden of SVD and obtains estimation for both the observation matrix \hat{C} and state sequence \hat{X} .

Corollary 3.4 *There exists at least a subsequence $\{(C^k, X^k, M^k)\}$ generated by Algorithm 2 that satisfies the first-order optimality conditions of (18) in the limit.*

Proof. We omit the proof here, since it is an immediate result of Theorem 3.5 in [53]. ■

The estimated state sequence by factorizing the Hankel matrix is only accurate up to a linear transformation. In other words, there is no unique solution for the state sequence

estimation. Such an ambiguity poses difficulty for reconstruction of the observation matrix. This difficulty is resolved by a joint sparsity assumption, as we will describe in the next session.

3.3 Observation Matrix Reconstruction

Once we have estimated the state sequence $\hat{\mathbf{x}}_t$, the next component of the algorithm is to reconstruct the observation matrix C .

3.3.1 Joint Structured Sparsity

Denoting observation matrix as $C \in \mathbb{R}^{n \times d}$, we formulate the reconstruction model as follows:

$$\min_C \alpha \sum_{i=1}^n \|\mathbf{e}_{(i)} \Psi(C)\|_2 + \beta \sum_{j=1}^d \|\Psi(C) \mathbf{e}_j\|_1 + \sum_{t=1}^l \frac{1}{2} \|\mathbf{z}_t - \Phi_t \mathcal{F}(C \hat{\mathbf{x}}_t)\|_2^2, \quad (19)$$

where Ψ denotes the frame-by-frame wavelet transform operator, defined in (3). The first two regularization terms concern structured sparsity for the observation matrix. The first term is the joint sparsity regularization, $\sum_{i=1}^n \|\mathbf{e}_{(i)} \Psi(C)\|_2$, which encourages all the columns $\Psi(C)$ to share a common yet small support. The second term, $\sum_{j=1}^d \|\Psi(C) \mathbf{e}_j\|_1$, demands $\Psi(C) \mathbf{e}_j$ to be sparse under wavelet transform, based on the assumption that each frame of the observation matrix is image-like. The joint sparsity is critical to the success of reconstruction, due to the ambiguity introduced by the non-uniqueness within the state sequence estimation.

Computationally one immediately notices the first two terms are non-smooth and both involve C . Moreover, the first term operates on the row space of matrix C while the second term operates on the column space of matrix C . In addition, the amount of data and the number of variables are large in our application. For these reasons, it is difficult to solve the optimization problem by off-the-shell algorithms for ℓ_1 minimization. We propose to apply the alternating direction method of multipliers (ADMM) [21] in such a way that all subproblems are easy to solve and can handle a large amount of data in a short time.

3.3.2 Alternating Direction Method of Multipliers

ADMM combines variable splitting techniques with the augmented Lagrangian method for solving constrained optimization with separable objective functions. It is also referred to as the alternating direction augmented Lagrangian method by several groups in the community. Alternating direction methods originated from solving PDEs [19, 40] and were later extended to solving variational problems associated with PDEs [21]. Recently, there has been a surge of interest in fast optimization algorithms using ADMM methodology for solving ℓ_1 and TV regularized problems [23, 1, 22, 54]. A state-of-art algorithm solved for group sparsity problems, which include joint sparsity [16]. Our new algorithm is based on ADMM and is optimized for the best computational efficiency.

Algorithm 3 Joint Structured Sparsity Optimization

Initialize $U \in \mathbb{R}^{n \times d}$, $V \in \mathbb{R}^{n \times d}$, $C \in \mathbb{R}^{n \times d}$, $k = 0$.

while convergence criteria not met **do**

U -subproblem:

$$U^{k+1} = \arg \min_U \sum_{j=1}^n \alpha \|\mathbf{u}_{(i)}\|_2 + \frac{\alpha\mu}{2} \|\mathbf{u}_{(i)} - \mathbf{e}_{(i)}\Psi(C^k) - \kappa_{(i)}^k\|_2^2$$

V -subproblem:

$$V^{k+1} = \arg \min_V \sum_{j=1}^d \beta \|\mathbf{v}_j\|_1 + \frac{\beta\mu}{2} \|\mathbf{v}_j - \Psi(C^k)\mathbf{e}_j - \lambda_j^k\|_2^2$$

C -subproblem:

$$C^{k+1} = \arg \min_C \frac{\alpha\mu}{2} \sum_{i=1}^n \|\mathbf{u}_{(i)}^{k+1} - \mathbf{e}_{(i)}\Psi(C) - \kappa_{(i)}^k\|_2^2 + \frac{\beta\mu}{2} \sum_{j=1}^d \|\mathbf{v}_j^{k+1} - \Psi(C)\mathbf{e}_j - \lambda_j^k\|_2^2 \\ + \sum_{t=1}^l \frac{1}{2} \|\mathbf{z}_t - \Phi_t \mathcal{F}(C\hat{\mathbf{x}}_t)\|_2^2$$

Multipliers update:

$$\kappa_{(i)}^{k+1} = \kappa_{(i)}^k - \gamma(\mathbf{u}_{(i)}^{k+1} - \Psi\mathbf{c}_{(i)}^{k+1})$$

$$\lambda_j^{k+1} = \lambda_j^k - \gamma(\mathbf{v}_j^{k+1} - \Psi\mathbf{c}_j^{k+1})$$

$$k = k + 1$$

end while

We first introduce additional variables to split the energy between different regularization terms, which results in the following constrained optimization,

$$\begin{aligned} \min_{U, V, C} \quad & \alpha \sum_{i=1}^n \|\mathbf{u}_{(i)}\|_2 + \beta \sum_{j=1}^d \|\mathbf{v}_j\|_1 + \sum_{t=1}^l \frac{1}{2} \|\mathbf{z}_t - \Phi_t \mathcal{F}(C\hat{\mathbf{x}}_t)\|_2^2 \\ \text{s.t.} \quad & \mathbf{u}_{(i)} = \mathbf{e}_{(i)}\Psi(C) \quad \mathbf{v}_j = \Psi(C)\mathbf{e}_j, \end{aligned} \quad (20)$$

where we have introduced variables U and V . The rows of $U \in \mathbb{R}^{n \times d}$ are $\mathbf{u}_{(i)}$, $i = 1, 2, \dots, n$, and the columns of $V \in \mathbb{R}^{n \times d}$ are \mathbf{v}_j , $j = 1, 2, \dots, d$.

To apply ADMM, we introduce the augmented Lagrangian of problem (20),

$$\begin{aligned} \min_{U, V, C} \quad & \alpha \sum_{i=1}^n \|\mathbf{u}_{(i)}\|_2 + \beta \sum_{j=1}^d \|\mathbf{v}_j\|_1 + \sum_{t=1}^l \frac{1}{2} \|\mathbf{z}_t - \Phi_t \mathcal{F}(C\hat{\mathbf{x}}_t)\|_2^2 \\ & + \frac{\alpha\mu}{2} \sum_{i=1}^n \|\mathbf{u}_{(i)} - \mathbf{e}_{(i)}\Psi(C) - \kappa_{(i)}\|_2^2 + \frac{\beta\mu}{2} \sum_{j=1}^d \|\mathbf{v}_j - \Psi(C)\mathbf{e}_j - \lambda_j\|_2^2, \end{aligned} \quad (21)$$

where we use $\mathcal{L}(U, V, C)$ to denote the augmented Lagrangian, and use $\kappa_{(i)}$ and λ_j to denote the scaled Lagrange multipliers. When C is fixed, minimizing over U and V are independent. Therefore, we apply ADMM to (21) in which we alternate between minimizing its objective function over U, V with fixed C and minimizing it over C with fixed U, V , along with the updates to λ_j , $\kappa_{(i)}$. We summarize the algorithm flow in Algorithm 3 and explain in the next few subsections how to efficiently solve each of its subproblems.

3.3.3 Joint Sparsity

The U -subproblem models the joint sparsity between the different spatial factors within the observation matrix. Noticing the optimization is independent with respect to each row $\mathbf{u}_{(i)}$

of U , we therefore solve for $i = 1, 2, \dots, n$,

$$\mathbf{u}_{(i)}^{k+1} = \arg \min_{\mathbf{u}_{(i)}} \alpha \|\mathbf{u}_{(i)}\|_2 + \frac{\alpha\mu}{2} \|\mathbf{u}_{(i)} - \mathbf{e}_{(i)}\Psi(C^k) - \kappa_{(i)}^k\|_2^2. \quad (22)$$

Lemma 3.5 (Shrinkage for ℓ_2 norm, [16]) *For any $\lambda, \mu > 0$ and $x, y \in \mathbb{R}^n$, the minimizer to*

$$\min_y \lambda \|y\|_2 + \frac{\mu}{2} \|y - x\|_2^2$$

is given by

$$y = \mathcal{S}_2\left(x, \frac{\lambda}{\mu}\right) := \max\left\{\|x\|_2 - \frac{\lambda}{\mu}, 0\right\} \odot \frac{x}{\|x\|_2}, \quad (23)$$

where \odot denotes component-wise product and \mathcal{S}_2 stands for ℓ_2 -shrinkage.

One can derive closed-form solution to the U-subproblem (22),

$$\mathbf{u}_{(i)}^{k+1} = \mathcal{S}_2\left(\mathbf{e}_{(i)}\Psi(C^k) + \kappa_{(i)}^k, \frac{1}{\mu}\right), \quad i = 1, 2, \dots, n. \quad (24)$$

3.3.4 Wavelet Sparsity

The V -subproblem concerns the wavelet sparsity, and is reduced to a sequence of the same ℓ_1 minimization problems with different data. Since computation on each column of V matrix \mathbf{v}_j is completely decoupled, we solve for each column independently. For $j = 1, 2, \dots, d$,

$$\mathbf{v}_j^{k+1} = \arg \min_{\mathbf{v}_j} \beta \|\mathbf{v}_j\|_1 + \frac{\beta\mu}{2} \|\mathbf{v}_j - \Psi(C^k)\mathbf{e}_j - \lambda_j^k\|_2^2. \quad (25)$$

For any $\lambda, \mu > 0$ and $x, y \in \mathbb{R}^n$, the minimizer to

$$\min_y \lambda \|y\|_1 + \frac{\mu}{2} \|y - x\|_2^2$$

is given by

$$y = \mathcal{S}_1\left(x, \frac{\lambda}{\mu}\right) := \max\left\{|x| - \frac{\lambda}{\mu}, 0\right\} \odot \text{sgn}(x). \quad (26)$$

where \odot denotes component-wise product and \mathcal{S}_1 stands for ℓ_1 -shrinkage.

The closed-form solution to the V -subproblem (25) is

$$\mathbf{v}_j^{k+1} = \mathcal{S}_1\left(\Psi(C^k)\mathbf{e}_j + \lambda_j^k, \frac{1}{\mu}\right), \quad j = 1, 2, \dots, d. \quad (27)$$

3.3.5 Reconstruction Fidelity

The C -subproblem, as it involves multiple terms and all the input data, is the most time consuming to solve. Specifically, it is

$$\begin{aligned} C^{k+1} = \arg \min_C & \frac{\alpha\mu}{2} \sum_{i=1}^n \|\mathbf{u}_{(i)}^{k+1} - \mathbf{e}_{(i)}\Psi(C) - \kappa_{(i)}^k\|_2^2 + \frac{\beta\mu}{2} \sum_{j=1}^d \|\mathbf{v}_j^{k+1} - \Psi(C)\mathbf{e}_j - \lambda_j^k\|_2^2 \\ & + \sum_{t=1}^l \frac{1}{2} \|\mathbf{z}_t - \Phi_t \mathcal{F}(C\hat{\mathbf{x}}_t)\|_2^2. \end{aligned}$$

Once we further write out the fidelity term, and utilizing the linearity property of the discrete Fourier transform, we have

$$\begin{aligned}
H(C) &:= \sum_{t=1}^l \frac{1}{2} \|\mathbf{z}_t - \Phi_t \mathcal{F}(C \hat{\mathbf{x}}_t)\|_2^2 \\
&= \sum_{t=1}^l \frac{1}{2} \|\mathbf{z}_t - \Phi_t \mathcal{F} \sum_{j=1}^d \mathbf{c}_j \hat{x}_{t,j}\|_2^2 \\
&= \sum_{t=1}^l \frac{1}{2} \|\mathbf{z}_t - \sum_{j=1}^d \hat{x}_{t,j} \Phi_t \mathcal{F} \mathbf{c}_j\|_2^2.
\end{aligned}$$

This leads to the following equivalent problem:

$$\begin{aligned}
C^{k+1} = \arg \min_C & \frac{\alpha\mu}{2} \sum_{i=1}^n \|\mathbf{u}_{(i)}^{k+1} - \mathbf{e}_{(i)} \Psi(C) - \kappa_{(i)}^k\|_2^2 + \frac{\beta\mu}{2} \sum_{j=1}^d \|\mathbf{v}_j^{k+1} - \Psi(C) \mathbf{e}_j - \lambda_j^k\|_2^2 \\
& + \sum_{t=1}^l \frac{1}{2} \|\mathbf{z}_t - \sum_{j=1}^d \hat{x}_{t,j} \Phi_t \mathcal{F} \mathbf{c}_j\|_2^2.
\end{aligned} \tag{28}$$

In the C -subproblem, the first term operates in the row space while the second and third terms operate in the column space of the observation matrix C ; this is undesirable computationally. However, it is easy to see

$$\sum_{i=1}^n \|\mathbf{u}_{(i)}^{k+1} - \mathbf{e}_{(i)} \Psi(C) - \kappa_{(i)}^k\|_2^2 = \|U^{k+1} - \Psi(C) - \Upsilon^k\|_F^2 = \sum_{j=1}^d \|\mathbf{u}_j^{k+1} - \Psi(C) \mathbf{e}_j - \kappa_j^k\|_2^2,$$

which allows us to rewrite the C -subproblem as follows:

$$\begin{aligned}
C^{k+1} = \arg \min_C & \frac{\alpha\mu}{2} \sum_{j=1}^d \|\mathbf{u}_j^{k+1} - \Psi \mathbf{c}_j - \kappa_j^k\|_2^2 + \frac{\beta\mu}{2} \sum_{j=1}^d \|\mathbf{v}_j^{k+1} - \Psi \mathbf{c}_j - \lambda_j^k\|_2^2 \\
& + \sum_{t=1}^l \frac{1}{2} \|\mathbf{z}_t - \sum_{j=1}^d \hat{x}_{t,j} \Phi_t \mathcal{F} \mathbf{c}_j\|_2^2.
\end{aligned} \tag{29}$$

We rewrite the above objective function as

$$\begin{aligned}
& \frac{\alpha\mu}{2} \sum_{j=1}^d (\Psi \mathbf{c}_j)^\top (\Psi \mathbf{c}_j) - 2(\Psi \mathbf{c}_j)^\top (\mathbf{u}_j^{k+1} - \kappa_j^k) + (\mathbf{u}_j^{k+1} - \kappa_j^k)^\top (\mathbf{u}_j^{k+1} - \kappa_j^k) \\
& + \frac{\beta\mu}{2} \sum_{j=1}^d (\Psi \mathbf{c}_j)^\top (\Psi \mathbf{c}_j) - 2(\Psi \mathbf{c}_j)^\top (\mathbf{v}_j^{k+1} - \lambda_j^k) + (\mathbf{v}_j^{k+1} - \lambda_j^k)^\top (\mathbf{v}_j^{k+1} - \lambda_j^k) \\
& + \frac{1}{2} \sum_{t=1}^l \sum_{j=1}^d (\hat{x}_{t,j} \Phi_t \mathcal{F} \mathbf{c}_j)^\top (\hat{x}_{t,j} \Phi_t \mathcal{F} \mathbf{c}_j) + \frac{1}{2} \sum_{t=1}^l \sum_{j \neq j'} (\hat{x}_{t,j} \Phi_t \mathcal{F} \mathbf{c}_j)^\top (\hat{x}_{t,j'} \Phi_t \mathcal{F} \mathbf{c}_{j'}) \\
& - \sum_{t=1}^l \sum_{i=1}^d (\hat{x}_{t,j} \Phi_t \mathcal{F} \mathbf{c}_j)^\top \mathbf{z}_t + \frac{1}{2} \sum_{t=1}^T \mathbf{z}_t^\top \mathbf{z}_t.
\end{aligned} \tag{30}$$

By taking the first derivative of the objective function and setting it to zero, we derive the normal equation to problem (29). Note we use Ψ to denote the wavelet operator, and Ψ^\dagger to denote its adjoint operator. We use Φ_t to denote the row selector operator, and Φ_t^\dagger to denote its adjoint operator. Similarly, we use \mathcal{F} to denote the Fourier operator, and \mathcal{F}^\dagger to denote its adjoint operator. The normal equation is as follows:

$$\begin{aligned}
& \frac{\alpha\mu}{2} \begin{pmatrix} \Psi^\dagger\Psi & 0 & \cdots & 0 \\ 0 & \Psi^\dagger\Psi & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Psi^\dagger\Psi \end{pmatrix} \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \vdots \\ \mathbf{c}_d \end{pmatrix} + \frac{\beta\mu}{2} \begin{pmatrix} \Psi^\dagger\Psi & 0 & \cdots & 0 \\ 0 & \Psi^\dagger\Psi & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Psi^\dagger\Psi \end{pmatrix} \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \vdots \\ \mathbf{c}_d \end{pmatrix} \\
& + \frac{1}{2} \begin{pmatrix} \sum_t \hat{x}_{t,1}^2 (\Phi_t \mathcal{F})^\dagger \Phi_t \mathcal{F} & \sum_t \hat{x}_{t,1} \hat{x}_{t,2} (\Phi_t \mathcal{F})^\dagger \Phi_t \mathcal{F} & \cdots & \sum_t \hat{x}_{t,1} \hat{x}_{t,d} (\Phi_t \mathcal{F})^\dagger \Phi_t \mathcal{F} \\ \sum_t \hat{x}_{t,2} \hat{x}_{t,1} (\Phi_t \mathcal{F})^\dagger \Phi_t \mathcal{F} & \sum_t \hat{x}_{t,2}^2 (\Phi_t \mathcal{F})^\dagger \Phi_t \mathcal{F} & \cdots & \sum_t \hat{x}_{t,2} \hat{x}_{t,d} (\Phi_t \mathcal{F})^\dagger \Phi_t \mathcal{F} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_t \hat{x}_{t,d} \hat{x}_{t,1} (\Phi_t \mathcal{F})^\dagger \Phi_t \mathcal{F} & \sum_t \hat{x}_{t,d} \hat{x}_{t,2} (\Phi_t \mathcal{F})^\dagger \Phi_t \mathcal{F} & \cdots & \sum_t \hat{x}_{t,d}^2 (\Phi_t \mathcal{F})^\dagger \Phi_t \mathcal{F} \end{pmatrix} \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \vdots \\ \mathbf{c}_d \end{pmatrix} \\
& = \alpha\mu \begin{pmatrix} \Psi^\dagger(\mathbf{u}_1^{k+1} - \kappa_1^k) \\ \Psi^\dagger(\mathbf{u}_2^{k+1} - \kappa_2^k) \\ \vdots \\ \Psi^\dagger(\mathbf{u}_d^{k+1} - \kappa_d^k) \end{pmatrix} + \beta\mu \begin{pmatrix} \Psi^\dagger(\mathbf{v}_1^{k+1} - \lambda_1^k) \\ \Psi^\dagger(\mathbf{v}_2^{k+1} - \lambda_2^k) \\ \vdots \\ \Psi^\dagger(\mathbf{v}_d^{k+1} - \lambda_d^k) \end{pmatrix} + \begin{pmatrix} \sum_t \hat{x}_{t,1} (\Phi_t \mathcal{F})^\dagger \mathbf{z}_t \\ \sum_t \hat{x}_{t,2} (\Phi_t \mathcal{F})^\dagger \mathbf{z}_t \\ \vdots \\ \sum_t \hat{x}_{t,d} (\Phi_t \mathcal{F})^\dagger \mathbf{z}_t \end{pmatrix}.
\end{aligned} \tag{31}$$

We simplify the notation of the normal equation (31) as

$$\mathcal{LHS} \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \vdots \\ \mathbf{c}_d \end{pmatrix} = \mathcal{RHS},$$

and one can immediately notice that the normal equation (31) is not a diagonal system. In other words, \mathbf{c}_j 's are coupled. Solving for linear system (31) directly can be computationally expensive.

We use the prox-linear method [12] to decouple the system. In stead of solving for Eqn. (29) directly, we solve for the following problem using the prox-linear method that decouples all the \mathbf{c}_j 's:

$$\mathbf{c}_j^{k+1} = \arg \min_{\mathbf{c}_j} q(\mathbf{c}_j^k)^\top (\mathbf{c}_j - \mathbf{c}_j^k) + \frac{1}{2\delta} \|\mathbf{c}_j - \mathbf{c}_j^k\|_2^2, \tag{32}$$

where $q(\mathbf{c}_j^k) = \nabla_C \mathcal{L}(U, V, C)$. This allows us to solve for each \mathbf{c}_j using block coordinate descent in the Jacobian fashion,

$$\mathbf{c}_j^{k+1} = \mathbf{c}_j^k - \delta q(\mathbf{c}_j^k). \tag{33}$$

More careful inspection on $q(\mathbf{c}_j^k)$ reveals

$$\begin{aligned}
q(\mathbf{c}_j^k) &= 2\alpha\mu \Psi^\dagger(\Psi \mathbf{c}_j^k - (\mathbf{u}_j^{k+1} - \kappa_j^k)) + 2\beta\mu \Psi^\dagger(\Psi \mathbf{c}_j^k - (\mathbf{v}_j^{k+1} - \lambda_j^k)) \\
&+ \sum_t \hat{x}_{t,j}^2 (\Phi_t \mathcal{F})^\dagger \Phi_t \mathcal{F} \mathbf{c}_j^k - \sum_t \hat{x}_{t,j} (\Phi_t \mathcal{F})^\dagger \mathbf{z}_t.
\end{aligned} \tag{34}$$

Algorithm 4 ADMM for Constrained Optimization

Initialize $B \in \mathbb{R}^{n \times d}$, $C \in \mathbb{R}^{n \times d}$, $k = 0$.

while convergence criteria not met **do**

B-subproblem:

$$B^{k+1} \leftarrow \min_B \mathcal{L}(B, C^k, \lambda^k)$$

C-subproblem:

$$C^{k+1} \leftarrow \min_C \mathcal{L}(B^{k+1}, C, \lambda^k)$$

Multiplier update:

$$\Lambda^{k+1} \leftarrow \min_{\Lambda} \Lambda^k - \gamma(B^{k+1} - C^{k+1})$$

$$k = k + 1$$

end while

3.3.6 Theoretical Convergence

We now establish the convergence of Algorithm 2. We first rewrite the objective function of joint structured sparsity, in the unconstrained optimization form

$$\min_C \alpha \sum_{i=1}^n \|\mathbf{e}_{(i)} \Psi(C)\|_2 + \beta \sum_{j=1}^d \|\Psi(C) \mathbf{e}_j\|_1 + \sum_{t=1}^l \frac{1}{2} \|\mathbf{z}_t - \Phi_t \mathcal{F}(C \hat{\mathbf{x}}_t)\|_2^2,$$

as the constrained optimization

$$\begin{aligned} \min_{B, C} \quad & \alpha \sum_{i=1}^n \|\mathbf{e}_{(i)} \Psi(B)\|_2 + \beta \sum_{j=1}^d \|\Psi(B) \mathbf{e}_j\|_1 + \sum_{t=1}^l \frac{1}{2} \|\mathbf{z}_t - \Phi_t \mathcal{F}(C \hat{\mathbf{x}}_t)\|_2^2 \\ \text{s.t.} \quad & B - C = 0. \end{aligned}$$

We can further group the first two terms in the above constrained optimization together,

$$\begin{aligned} \min_{B, C} \quad & f(B) + g(C) \\ \text{s.t.} \quad & B - C = \mathbf{0}, \end{aligned}$$

with $f(B) = \alpha \sum_{i=1}^n \|\mathbf{e}_{(i)} \Psi(B)\|_2 + \beta \sum_{j=1}^d \|\Psi(B) \mathbf{e}_j\|_1$ and $g(C) = \sum_{t=1}^l \frac{1}{2} \|\mathbf{z}_t - \Phi_t \mathcal{F}(C \hat{\mathbf{x}}_t)\|_2^2$.

Consider the augmented Lagrangian function,

$$\mathcal{L}(B, C, \lambda) = f(B) + g(C) + \frac{\mu}{2} \|B - C - \Lambda\|_F^2, \quad (37)$$

where $\Lambda \in \mathbb{R}^{n \times d}$ is the scaled Lagrangian multiplier and $\mu > 0$ is a penalty parameter. With the above reformulation, we can consolidate Algorithm 2 into a simplified version, see Algorithm 4. Global and linear convergence for generalized ADMM was analyzed in [15] for constrained convex optimization problems. We extend those theoretical results from vector case to matrix case below.

Theorem 3.6 (Global Convergence) *The sequence $\{W^k\} := \{B^k, C^k, \Lambda^k\}$ generated by Algorithm 4 is guaranteed to be bounded. Moreover, if we assume there exists a saddle*

point $W^* := (B^*, C^*, \Lambda^*)$ to problem (35), and Hessian $H_g = \nabla^2 g$ satisfies the following condition,

$$\frac{\max(\alpha\mu, \beta\mu)}{\frac{1}{\mu} - \|H_g\|} + \gamma < 2, \quad (38)$$

then the sequence $\{W^k\}$ generated by Algorithm 3 converges to a KKT point of W^* ,

$$\lim_{k \rightarrow \infty} \|W^k - W^*\|_F^2 = 0.$$

Proof. First, we assume there exists a saddle point W^* , which means (B^*, C^*, Λ^*) satisfies the KKT conditions of problem (35):

$$\Lambda^* \in \partial f(B^*), \quad \Lambda^* \in \partial g(C^*), \quad B^* - C^* = \mathbf{0}.$$

Second, we verify that both functions $f(\cdot)$ and $g(\cdot)$ are convex. We have $f(B) = \alpha \sum_{i=1}^n \|\mathbf{e}_{(i)} \Psi(B)\|_2 + \beta \sum_{j=1}^d \|\Psi(B) \mathbf{e}_j\|_1$ and $g(C) = \sum_{t=1}^l \frac{1}{2} \|\mathbf{z}_t - \Phi_t \mathcal{F}(C \hat{\mathbf{x}}_t)\|_2^2$. Since $\|\cdot\|_p$ is convex when $p \geq 1$, and the fact sum of convex functions is also convex, we can easily verify that both functions are convex.

Third, based on Theorem 2.3 remark 3 condition (i) in [15], we know that the sequence $\{W^k\}$ is bounded.

Having obtained these assumptions, it follows from Theorem 2.3 in [15] that $\{W^k\}$ has a converging subsequence $\{W^{k'}\}$, whose limit is $W^* := \lim_{k' \rightarrow \infty} W^{k'}$. Hence we have global convergence. ■

4 Dynamic MRI Reconstruction Quality

4.1 Impact of Sampling Strategies

We now apply our algorithm to accelerate the acquisition process of dynamic MRI. Since the sampling strategy in the k-space has an impact on the reconstruction quality, we test three types of sampling strategies following the work of [26].

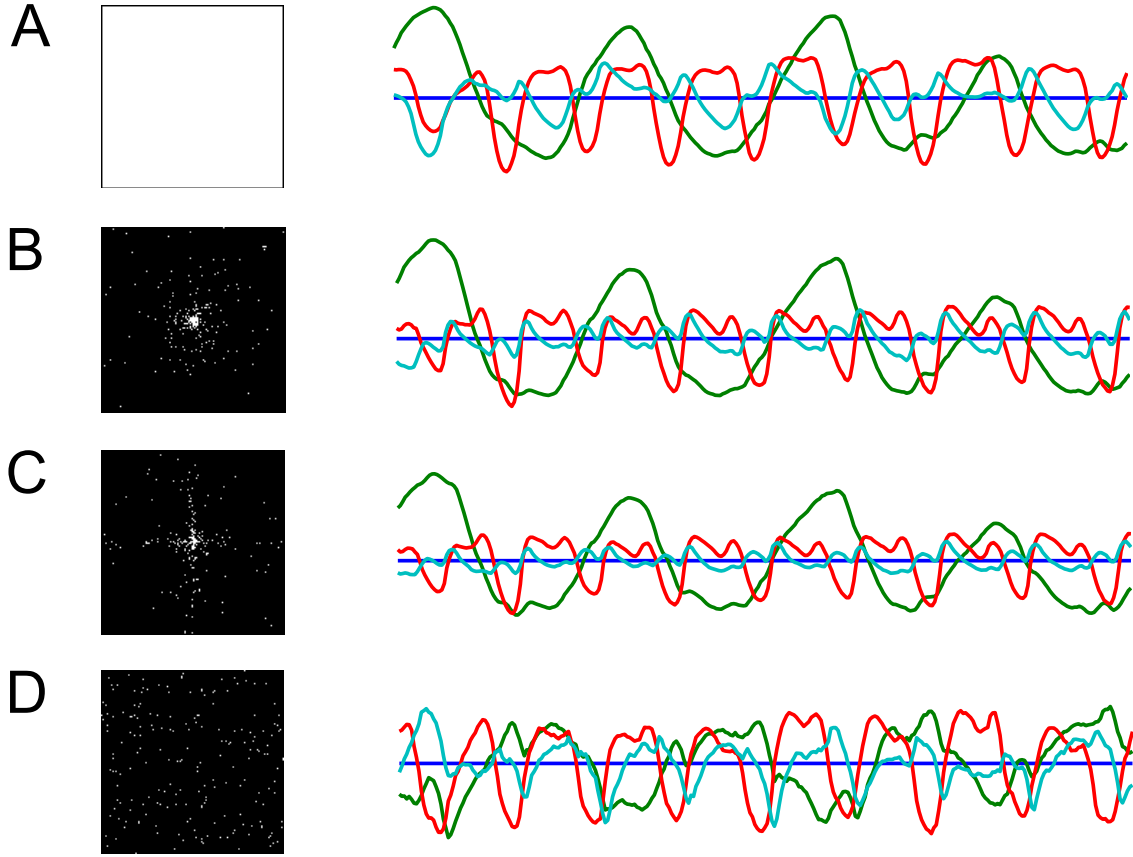


Figure 4: Impact of measurement matrix on the state sequence estimate, sampling distribution. Illustrated are four measurement matrices design for the time-invariant component $\bar{\Phi}$, with all compressive sensing cases using 200 samples: (A) measurement matrix covering all samples in k-space, (B) measurement matrix following *distance* distribution, (C) measurement matrix with *hyperbolic* distribution, (D) measurement matrix with *uniform* distribution. Left column shows the measurement matrix, right column shows the estimated state sequence using Algorithm 1.

We first illustrate the impact of measurement matrix on the state sequence estimate. Figure 4 shows three measurement matrices that cover all range of frequencies, however follow different probability distributions:

- *distance*:
probability of sampling falls over as inverse of squared distance to the k-space center.
- *hyperbolic*:
probability of sampling falls over as a hyperbolic function in the k-space.
- *uniform*:
probability of sampling is uniform in the k-space.

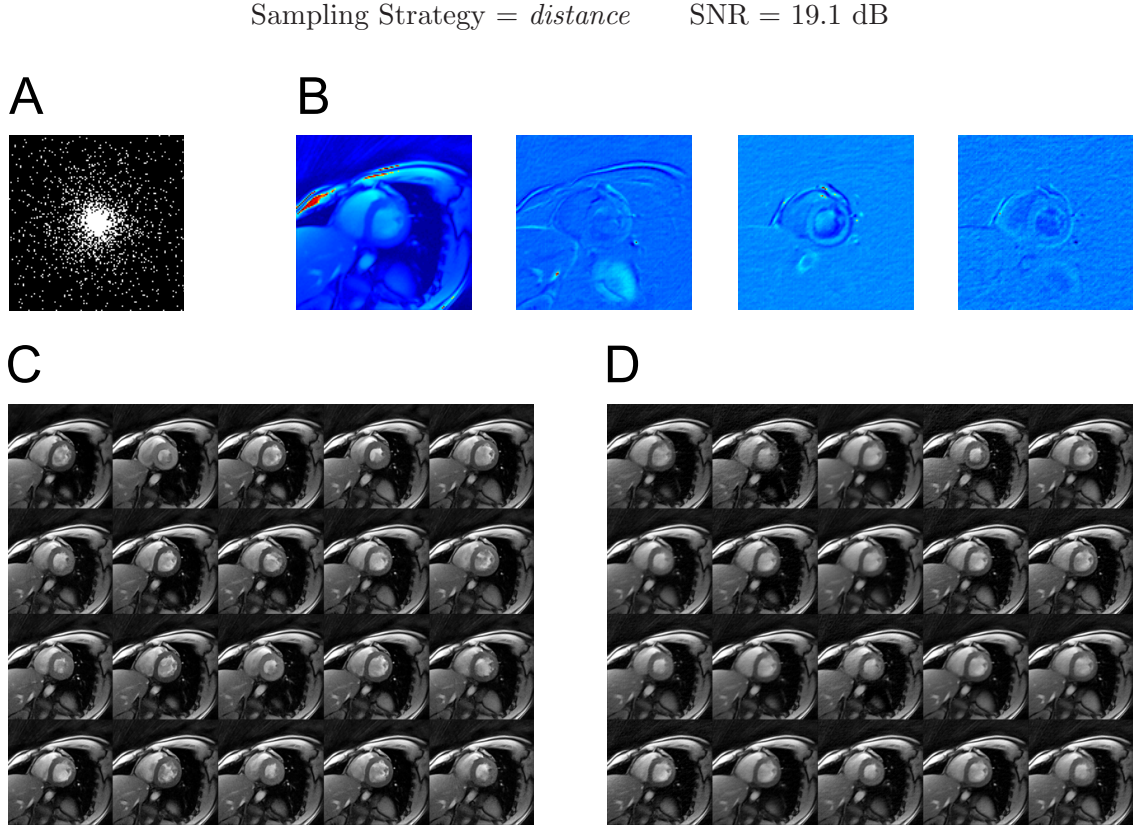


Figure 5: Reconstruction result for dynamic MRI data, based on 10% k-space data. Sampling strategy adopts the *distance* distribution. Shown are (A) measurement matrix Φ_t at $t = 1$, (B) observation matrix C with $d = 4$, (C) original video frames \mathbf{y}_t (D) reconstructed video frames $\hat{\mathbf{y}}_t$. The reconstruction SNR is 19.1 dB.

Our numerical results indicate the best measurement matrix design is to sample k-space according to the *distance* strategy, where one samples the k-space in a density that falls off as 1 over the squared distance to the center of k-space.

We show the reconstruction quality of dynamic MRI using different sampling strategies. The cardiac MRI dataset used in this experiment was described in [55]. We obtained the reconstructed video for real-time MRI of a human heart, whose spatial resolution is subsampled at 128×128 and temporal resolution is 33 ms, with 300 frames in total.

Note we only simulated a single coil with a homogeneous coil sensitivity map. We simulated k-t data by taking the Fourier transform and performing subsampling. Define the samples in Fourier space as

$$\Omega = \{(\omega_1^k, \omega_2^k)\}_{k=1}^m \subset \left\{-\frac{n_x}{2} + 1, \dots, \frac{n_x}{2}, -\frac{n_y}{2} + 1, \dots, \frac{n_y}{2}\right\} \quad (39)$$

assuming a uniform Cartesian grid.

Sampling Strategy = *hyperbolic* SNR = 15.3 dB

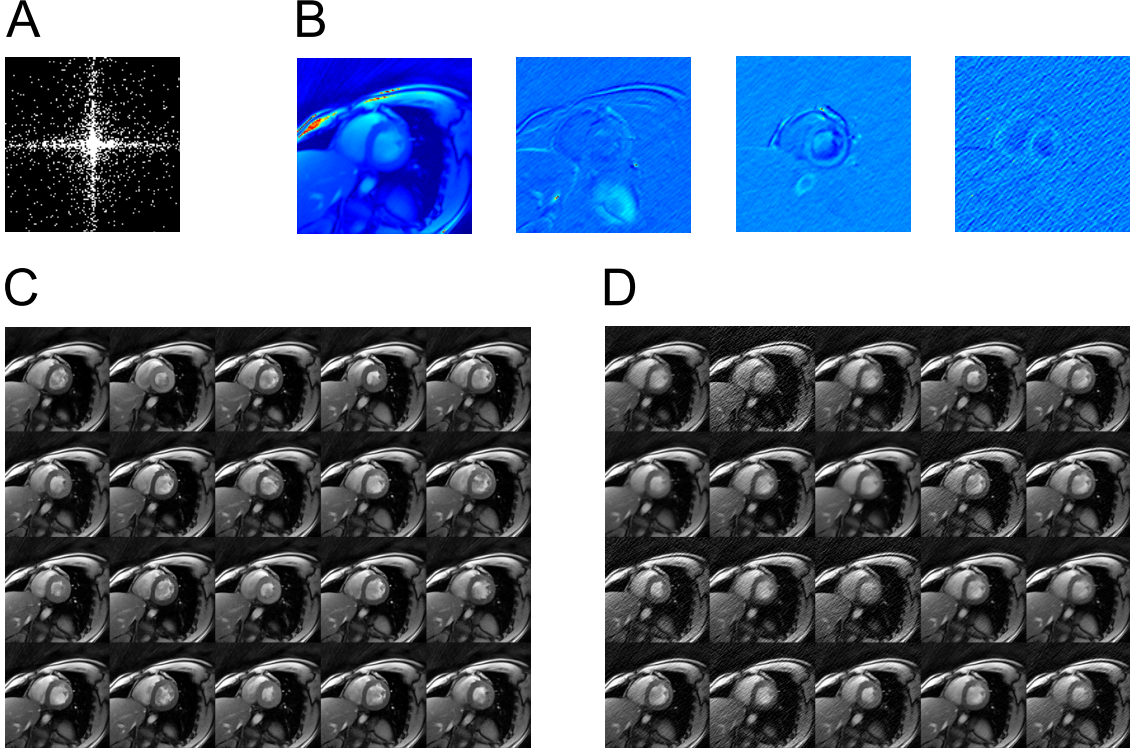


Figure 6: Reconstruction result for dynamic MRI data, based on 10% k-space data. Sampling strategy adopts the *hyperbolic* distribution. Shown are (A) measurement matrix Φ_t at $t = 1$, (B) observation matrix C with $d = 4$, (C) original video frames \mathbf{y}_t (D) reconstructed video frames $\hat{\mathbf{y}}_t$. The reconstruction SNR is 15.3 dB.

We show the reconstruction result for dynamic MRI with 10% k-t data, using the *distance* sampling strategy in Figure 5. We construct Ω by subsampling the Fourier space i.i.d. according to density

$$\eta(\omega_1, \omega_2) \propto (\omega_1^2 + \omega_2^2 + 1)^{-1}. \quad (40)$$

We attain a SNR of 19.1 dB in the reconstruction using the *distance* sampling strategy.

We show the reconstruction result for the dynamic MRI with 10% k-t data, using the *hyperbolic* sampling strategy in Figure 6. We construct Ω by subsampling the Fourier space i.i.d. according to density

$$\eta(\omega_1, \omega_2) \propto (\omega_1^2 + \omega_2^2 + 1)^{-3/2}. \quad (41)$$

We attain a SNR of 15.3 dB in the reconstruction using the *hyperbolic* sampling strategy.

Sampling Strategy = *uniform* SNR = 7.1 dB

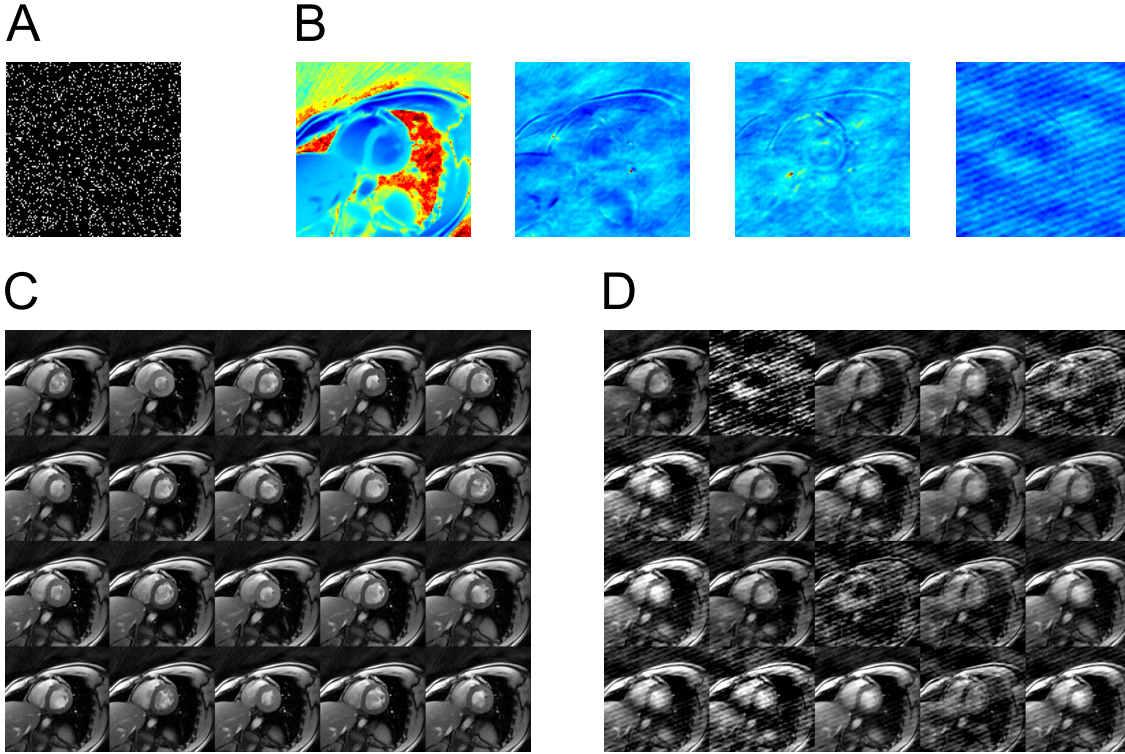


Figure 7: Reconstruction result for dynamic MRI data, based on 10% k-space data. Sampling strategy adopts the *uniform* distribution. Shown are (A) measurement matrix Φ_t at $t = 1$, (B) observation matrix C with $d = 4$, (C) original video frames \mathbf{y}_t (D) reconstructed video frames $\hat{\mathbf{y}}_t$. The reconstruction SNR is 7.1 dB.

We show the reconstruction result for the dynamic MRI with 10% k-t data, using the *uniform* sampling strategy in Figure 7. We construct Ω by subsampling the Fourier space i.i.d. according to density

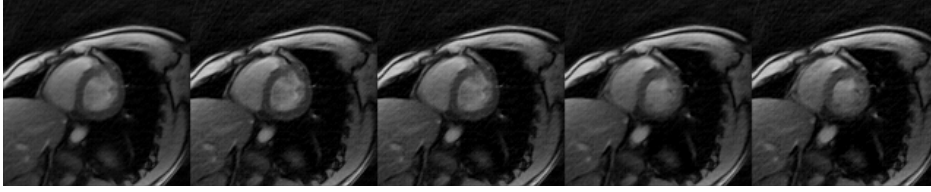
$$\eta(\omega_1, \omega_2) \propto 1. \quad (42)$$

We attain a SNR of 7.1 dB in the reconstruction using the *uniform* sampling strategy.

4.2 Comparison with Prior Art

We compare kt-CSLDS with prior art in the literature, which includes kt-SPARSE, MAS-TeR, and L+S. Figure 8 and Figure 9 show the numerical results on two datasets described in [55]. Both datasets can be downloaded from the paper website provided by the authors. Our numerical results show that kt-CSLDS achieves excellent reconstruction quality.

A

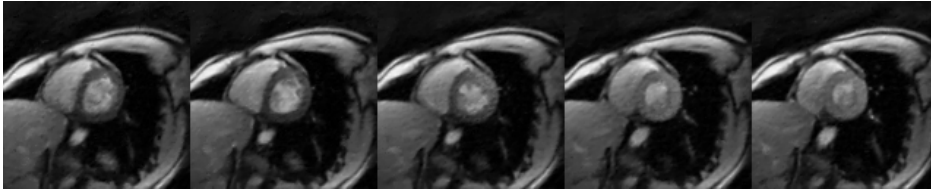


Original

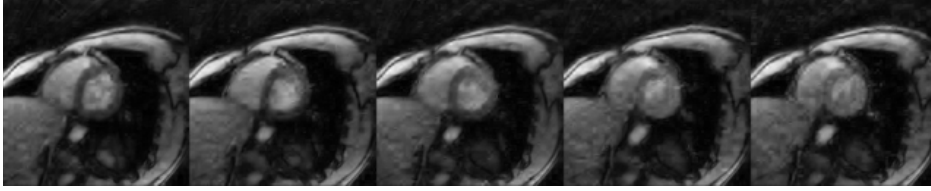
B



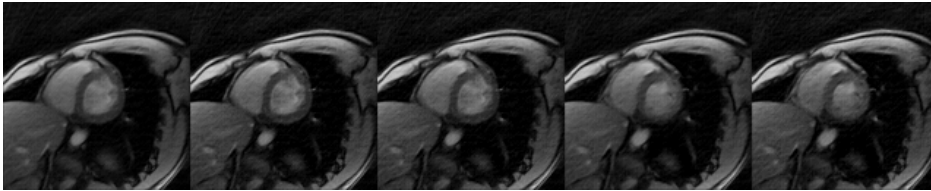
kt-SPARSE



MASTeR



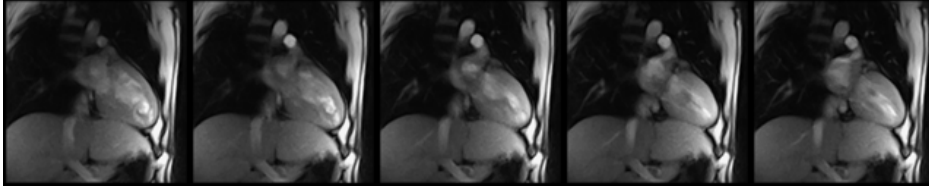
L+S



kt-CSLDS

Figure 8: Comparison of kt-CSLDS with prior art in the literature. Numerical results are based on one dataset for dynamic heart imaging, with 1.5 mm resolution, 8 mm section thickness. The original dynamic MRI is acquired at 30 ms acquisition time with 300 frames. We downsample the heart video to 128×128 spatial resolution and simulate a single coil acquisition. We use $10\times$ compression rate for this experiment, and employ the *distance* sampling strategy for compressive measurement. (A) Sample frames from the original heart video. (B) Reconstructed frames based on different video compressive sensing algorithms. Their respective reconstruction SNRs are as follows: kt-SPARSE (13.0 dB), MASTeR (18.8 dB), L+S (15.8 dB), kt-CSLDS (19.1 dB).

A

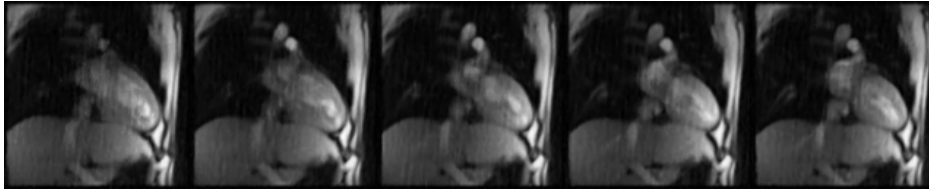


Original

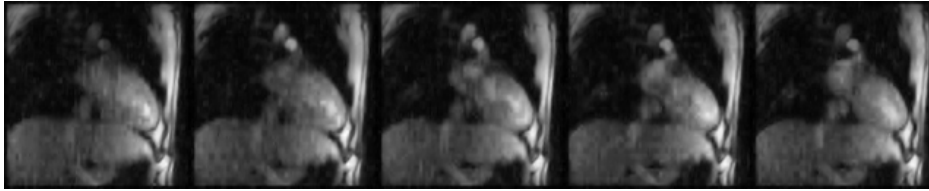
B



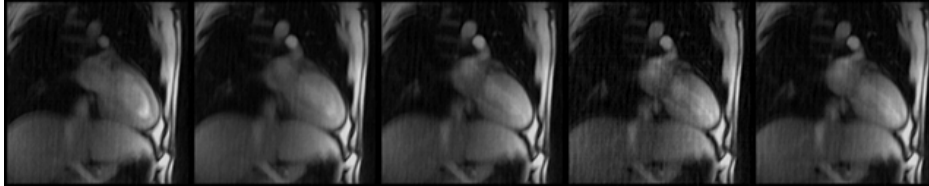
kt-SPARSE



MASTeR



L+S



kt-CSLDS

Figure 9: Comparison of kt-CSLDS with prior art in the literature. Numerical results are based on another dataset for dynamic heart imaging, with 2.0 mm resolution, 8 mm section thickness. The original dynamic MRI is acquired at 22 ms acquisition time with 360 frames. We downsample the heart video to 128×128 spatial resolution and simulate a single coil acquisition. We use $10\times$ compression rate for this experiment and employ the *distance* sampling strategy for compressive measurement. (A) Sample frames from the original heart video. (B) Reconstructed frames based on different video compressive sensing algorithms. Their respective reconstruction SNRs are as follows: kt-SPARSE (14.0 dB), MASTeR (19.4 dB), L+S (16.3 dB), kt-CSLDS (20.3 dB).

Table 1: Comparison of reconstruction SNR

Model	10×	20×	30×	40×	50×
kt-SPARSE	13.0 dB	11.3 dB	10.5 dB	9.9 dB	9.5 dB
MASTeR	18.8 dB	14.6 dB	12.7 dB	11.6 dB	11.0 dB
L+S	15.8 dB	12.4 dB	10.7 dB	9.7 dB	9.3 dB
kt-CSLDS	19.1 dB	16.3 dB	15.0 dB	13.3 dB	12.8 dB

Table 2: Comparison of computation time

Model	10×	20×	30×	40×	50×
kt-SPARSE	371.7 s	401.9 s	455.0 s	491.1 s	585.2 s
MASTeR	422.6 s	430.9 s	425.7 s	423.6 s	433.9 s
L+S	1390.2 s	1403.6 s	1402.8 s	1406.4 s	1406.4 s
kt-CSLDS	7.6 s	5.0 s	4.9 s	4.5 s	4.4 s

Table 1 compares the reconstruction SNR of different video compressive sensing models under various compression rate. Table 2 shows their respective computation time. In comparison, kt-CSLDS achieves the best reconstruction quality while consuming the least computational time.

5 Conclusions

In this paper, we built upon video compressive sensing ideas to accelerate the imaging acquisition process of dynamic MRI. We extended CS-LDS model to the Fourier-time space, resulting in so-called kt-CSLDS. Efficient numerical algorithm was derived based on ADMM. Theoretical analysis was carried out to ensure global convergence. Numerical results show that kt-CSLDS achieves favorable reconstruction quality while being computationally efficient, in comparison with state-of-the-art dynamic MRI compressive sensing literature.

LDS provides a compact model for video sequences, which approximates high-dimensional signal using low-dimensional representation. Therefore, kt-CSLDS benefits from such a compact representation, since the number of unknowns are much smaller compared with the original video cube. This explains why our model achieves high-fidelity reconstruction results given compressive measurements. The computational speed we gain is a result of both smaller dimensionality of the optimization problem and customized algorithm based on ADMM.

There are many ways to build upon the current kt-CSLDS framework. Our current methodology takes all the video data and performs batch process. Future work seeks online version of the current reconstruction algorithm. Regarding the measurement strategy, we have shown empirically that the best strategy is to sample the k-space according to the *distance* strategy. Such a result is consistent with theory for Fourier compressive sensing

for static MR imaging [26]. It remains an open theoretical question why such a strategy is optimal for video compressive sensing.

Acknowledgments

We thank Rachel Ward for generously providing the code to generate the sampling patterns in the k-space.

References

- [1] M. Afonso, J. Bioucas-Dias, and M. Figueiredo. Fast image recovery using variable splitting and constrained optimization. *IEEE Trans. Image Process.*, 19(9):2345–2356, 2010.
- [2] G. Aggarwal, A. Roy-Chowdhury, and R. Chellappa. A system identification approach for video-based face recognition. In *Proc. Int. Conf. Pattern Recognition (ICPR)*, pages 175–178, 2004.
- [3] K. S. Arun and S. Y. Kung. Balanced approximation of stochastic systems. *SIAM J. Matrix Anal. Appl.*, 11:42–68, 1990.
- [4] M. S. Asif, L. Hamilton, M. Brummer, and J. Romberg. Motion-adaptive spatio-temporal regularization for accelerated dynamic MRI. *Magn. Reson. Med.*, 2012. doi: 10.1002/mrm.24524.
- [5] R. G. Baraniuk. More is less: Signal processing and the data deluge. *Science*, 331(6018):717–719, 2011.
- [6] A. Bissacco, A. Chiuso, Y. Ma, and S. Soatto. Recognition of human gaits. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2001.
- [7] R. W. Brockett. *Finite Dimensional Linear Systems*. Wiley, 1970.
- [8] E. J. Candès and J. Romberg. Quantitative robust uncertainty principles and optimally sparse decompositions. *Foundations of Comput. Math.*, 6(2):227–254, 2006.
- [9] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory*, 52(2):489–509, 2006.
- [10] V. Cevher, A. C. Sankaranarayanan, M. F. Duarte, D. Reddy, R. G. Baraniuk, and R. Chellappa. Compressive sensing for background subtraction. In *Proc. European Conf. Computer Vision (ECCV)*, 2008.
- [11] A. B. Chan and N. M. Vasconcelos. Modeling, clustering, and segmenting video with mixtures of dynamic textures. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(5):909–926, 2008.

- [12] G. Chen and M. Teboulle. A proximal-based decomposition method for convex minimization problems. *Math. Programming*, 64(1):81–101, 1994.
- [13] H.-W. Chen, K. Li-Wei, and L. Chun-Shien. Dictionary learning-based distributed compressive video sensing. In *Proc. Picture Coding Symposium (PCS)*, 2010.
- [14] R. Damadian. Tumor detection by nuclear magnetic resonance. *Science*, 171(3976):1151–1153, 1971.
- [15] W. Deng and W. Yin. On the global and linear convergence of the generalized alternating direction method of multipliers. Technical report, Rice University, 2012.
- [16] W. Deng, W. Yin, and Y. Zhang. Group sparse optimization by alternating direction method. Technical report, Rice University, 2011.
- [17] D. L. Donoho. Compressed sensing. *IEEE Trans. Inf. Theory*, 52(4):1289–1306, 2006.
- [18] G. Doretto, A. Chiuso, Y. N. Wu, and S. Soatto. Dynamic textures. *Int. J. Computer Vision*, 51(2):91–109, 2003.
- [19] J. Douglas and H. H. Rachford. On the numerical solution of the heat conduction problem in 2 and 3 space variables. *Trans. American Mathematical Society*, 82:421–439, 1956.
- [20] J. E. Fowler, S. Mun, and E. W. Tramel. Block-based compressive sensing of image and video. *Foundations and Trends in Signal Processing*, 4(4):297–416, 2012.
- [21] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite-element approximations. *Computers & Mathematics with Applications*, 2:17–40, 1976.
- [22] D. Goldfarb, S. Ma, and K. Scheinberg. Fast alternating linearization methods for minimizing the sum of two convex functions. Technical report, Columbia University, 2010.
- [23] T. Goldstein and S. Osher. The split Bregman algorithm for L1 regularized problems. *SIAM J. Imaging Sci.*, 2(2):323–343, 2008.
- [24] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, MD, 2nd edition, 1989.
- [25] H. Jung, K. Sung, K. Nayak, E. Kim, and J. Ye. k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI. *Magn. Reson. Med.*, 61(1):103–116, 2009.
- [26] F. Krahmer and R. Ward. Beyond incoherence: stable and robust sampling strategies for compressive imaging. Technical report, University of Texas at Austin, <http://arxiv.org/pdf/1210.2380v2.pdf>, 2012.

- [27] W. E. Larimore. Canonical variete analysis in identification, filtering and adaptive control. In *Proc. IEEE Conf. Decision and Control*, pages 596–604, 1990.
- [28] P. C. Lauterbur. Image formation by induced local interactions: examples of employing nuclear magnetic resonance. *Nature*, 242(5394):190–191, 1973.
- [29] C. S. Lee and A. Elgammal. Human motion synthesis by motion manifold learning and motion primitive segmentation. In *Proc. Int. Conf. Articulated Motion and Deformable Objects*, pages 464–473, 2006.
- [30] W. Lu and N. Vaswani. Modified compressive sensing for real-time dynamic MR imaging. In *Proc. IEEE Int. Conf. Image Process.*, 2009.
- [31] M. Lustig, D. L. Donoho, and J. M. Pauly. Sparse MRI: the application of comprssed sensing for rapid mr imaging. *Magn. Reson. Med.*, 58:1182–1195, 2007.
- [32] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly. Compressed sensing MRI. *IEEE Signal Process. Mag.*, 25:72–82, 2008.
- [33] M. Lustig, M. Santos, D. Donoho, and J. Pauly. k-t SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity. In *Proc. Int. Soc. Magn. Reson. Med. (ISMRM)*, page 2420, 2006.
- [34] P. Mansfield. Multi-planar imaging formation using NMR spin echoes. *J. Physics C. Solid State Phys.*, 10:55–58, 1977.
- [35] R. Marcia and R. Willet. Compressive coded aperture video reconstruction. In *Proc. European Signal Processing Conf. (EUSIPCO)*, 2008.
- [36] R. Otazo, E. J. Candès, and D. Sodickson. Low-rank and sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components. Technical report, Stanford University, 2013.
- [37] P. V. Overschee and B. D. Moor. N4SID: subspace algorthims for the identification of combined deterministic-stochastic systems. *Automatica*, 30(1):75–93, 1994.
- [38] J. Y. Park and M. B. Wakin. A multiscale framework for compressive sensing of video. In *Proc. Picture Coding Symposium (PCS)*, May 2009.
- [39] J. Y. Park and M. B. Wakin. Multiscale algorithm for reconstructing videos from streaming compressive measurements. *J. Electron. Imaging*, 22(2):021001, 2013.
- [40] D. H. Peaceman and H. H. Rachford. The numerical solution of parabolic elliptic differential equations. *SIAM J. Appl. Math.*, 3:28–41, 1955.
- [41] J. Prades-Nebot, M. Yi, and T. Huang. Distributed video coding using compressive sampling. In *Proc. Picture Coding Symposium (PCS)*, 2009.
- [42] A. Rajwade, D. Kittle, T.-H. Tsai, D. Brady, and L. Carin. Coded hyperspectral imaging and blind compressive sensing. Technical report, Duke University, 2012.

- [43] S. Rapacchi, F. Han, Y. Natsuaki, R. Kroeker, A. Plotnik, E. Lehman, J. Sayre, G. Laub, J. P. Finn, and P. Hu. High spatial and temporal resolution dynamic contrast-enhanced magnetic resonance angiography using compressed sensing with magnitude image subtraction. *Magn. Reson. Med.*, page doi: 10.1002/mrm.24842, 2013.
- [44] D. Reddy, A. Veeraraghavan, and R. Chellappa. P2C2: programmable pixel compressive camera for high speed imaging. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2011.
- [45] A. C. Sankaranarayanan, C. Studer, and R. G. Baraniuk. CS-MUVI: Video compressive sensing for spatial-multiplexing cameras. In *Proc. IEEE Intl. Conf. Comp. Photography (ICCP)*, 2012.
- [46] A. C. Sankaranarayanan, P. Turaga, R. G. Baraniuk, and R. Chellappa. Compressive acquisition of dynamic scenes. In *Proc. European Conf. Computer Vision (ECCV)*, volume 6311, pages 129–142, 2010.
- [47] A. C. Sankaranarayanan, P. K. Turaga, R. Chellappa, and R. G. Baraniuk. Compressive acquisition of linear dynamical systems. *SIAM J. Imaging Sci.*, 6(4):2109–2133, 2013.
- [48] A. Secker and D. Taubman. Lifting-based invertible motion adaptive transform framework for highly scalable video compression. *IEEE Trans. Image Process.*, 12(12):1530–1542, 2003.
- [49] A. Veeraraghavan, D. Reddy, and R. Raskar. Coded strobing photography: Compressive sensing of high speed periodic videos. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(4):671–686, 2011.
- [50] A. Veeraraghavan, A. Roy-Chowdhury, and R. Chellappa. Matching shape sequences in video with an application to human movement analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(12):1896–1909, 2005.
- [51] M. Verhaegen. Identification of the deterministic part of MIMO state space model given innovations form from input-output data. *Automatica*, 30(1):61–74, 1994.
- [52] M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, and R. G. Baraniuk. Compressive imaging for video representation and coding. In *Proc. Picture Coding Symposium (PCS)*, Apr 2006.
- [53] Z. Wen, W. Yin, and Y. Zhang. Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm. *Math. Prog. Comp.*, 4:333–361, 2012.
- [54] J. Yang and Y. Zhang. Alternating direction algorithms for L1-problems in compressive sensing. *SIAM J. Sci. Comput.*, 33(1):250–278, 2011.
- [55] S. Zhang, M. Uecker, D. Voit, K. D. Merboldt, and J. Frahm. Real-time cardiovascular magnetic resonance at high temporal resolution: radial flash with nonlinear inverse reconstruction. *J. Cardiovasc. Magn. Reson.*, 12:39, 2010.