Fully Decomposable Compressive Sampling With Joint Optimization for Multidimensional Sparse Representation

Wenrui Dai, *Member, IEEE*, Yong Li, *Student Member, IEEE*, Junni Zou, *Member, IEEE*, Hongkai Xiong, *Senior Member, IEEE*, and Yuan F. Zheng, *Fellow, IEEE*

***Abstract*—Conventional compressive sampling methods cannot efficiently exploit structured sparsity for sampling multidimensional signals like video sequences. In this paper, we propose a fully decomposable compressive sampling model that adopts the Kronecker product framework to exploit the structured sparsity spanning multidimensional signals. It enables efficient sampling in a progressive fashion by retaining the block-diagonal feature of Kronecker products. A synthetic sensing matrix is developed for joint optimization over sampling signals with multiple dimensions. Instead of adjusting global Gram matrix, separable minimization of mutual coherence in multiple dimensions is jointly formulated for a stable recovery with enhanced convergence rate. Sampling rate allocation is considered to improve recovery performance based on the decomposable compressive sampling. The proposed model is employed in video acquisition for temporal sparsity along motion trajectory. Experiment results show that the proposed model can improve the recovery performance with a reduced number of necessary samples in comparison to the state-of-the-art methods.**

***Index Terms*—Sparse representation, matrix optimization, compressive video sampling, video acquisition.**

I. Introduction

MULTIMEDIA signal acquisition usually requires a large number of sensors to sample the raw data. Shannon's celebrated theorem demonstrates that traditional sampling requires a minimum rate twice the highest frequency of the signal, namely the Nyquist rate, to achieve sufficient sampling without loss of any information. Therefore, traditional data acquisition systems commonly employ a *sample-then-compress* framework to acquire signals. However, most signals of interest have concise (sparse) linear representations with respect to some basis, e.g. Fourier or wavelet basis, which means that the transformed (compressed) coefficients equal or close to zero could be discarded. Recognizing the deficiency of directly sampling raw data, compressive sampling (CS [1], [2]) has been developed to make sub-Nyquist sampling by recording fewer projections into an incoherent set of measurement vectors.

The *compression-into-sampling* system can generate irregular and inadequate measurements conditioned on the actual sparsity of signals, so that the necessary sampling rate can be substantially reduced. Since exact reconstruction via nonlinear recovery regularized by ℓ_0 quasi-norm is an NP-hard problem, the ℓ_1 minimization [3] was commonly adopted as a tractable alternative for sparse recovery. Its recovery performance is demonstrated to be guaranteed under the constraints of the null space property (NSP [4]) and the restricted isometry property (RIP [5]) for measurement vectors. However, conventional CS [6], [7] relies on a vectorized representation for the whole signal of interest, which would obscure the intrinsic structure of multidimensional signals. As a result, these methods are restricted in applications for multidimensional signals, e.g. hyperspectral imaging [8]–[10], magnetic resonance imaging (MRI [11]), and image/video acquisition [12].

Recently, compressive sampling has been inspired to investigate high-speed spatial and temporal imaging resolution with a reversed-complexity manner, which would favor the ubiquitous and interactive multimedia access for increasing mobile communication. To facilitate a distributed CS protocol, the sparsifying basis and synthetic sensing matrix for sampling are optimized for compact representation of varying sparsities. Traced back to [13], compressive imaging was first employed into video representation and coding based on a single-pixel camera model, where 3-D wavelet transform was adopted to improve the frame-by-frame reconstruction with its 2-D counterpart. In [14], an iterative multiscale framework was developed to exploit motion estimation and compensation for enhanced sparsity at finer scales in the 3-D wavelet domain. Cossalter *et al.* [15] combined compressive sensing and the information obtained by tracking the motion objects. To improve sampling efficiency, each frame

Manuscript received December 13, 2016; revised August 1, 2017; accepted October 26, 2017. Date of publication November 13, 2017; date of current version December 26, 2017. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Hongbin Li. This work was supported in part by the National Natural Science Foundation of China under Grant 61501294, Grant 61622112, Grant 61472234, Grant 61720106001, Grant 61529101, and Grant 61425011, in part by China Postdoctoral Science Foundation under Grant 2015M581617, and in part by the Program of Shanghai Academic Research Leader under Grant 17XD1401900. *(Corresponding author: Hongkai Xiong.)*

W. Dai is with the Department of Biomedical Informatics, University of California, San Diego, San Diego, CA 92093, USA (e-mail: wed004@ucsd.edu).

Y. Li and H. Xiong are with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: marsleely@sjtu.edu.cn; xionghongkai@sjtu.edu.cn).

J. Zou is with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: zou-jn@cs.sjtu.edu.cn).

Y. F. Zheng is with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210 USA (e-mail: zheng@ece.osu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2017.2773427

was further segmented into non-overlapping blocks which were estimated with a linear combination of blocks in previous frames [16], [17]. For distributed compressive video sensing, an efficient modified gradient projection algorithm was developed for sparse reconstruction at the decoder [18]. To further reduce random measurements, Ma *et al.* [19] adopted a pseudorandom downsampling of the 2-D Fourier transform for fast online encoding of each frame and the approximate message passing approach for offline decoding with rapid convergence.

To improve the recovery performance, motion estimation and compensation were introduced into the block-based CS framework. In essence, a video sequence was decomposed into reference frames and the other CS frames reconstructed based on adaptive bases extracted from adjacent reconstructed reference frames. Liu *et al.* [20] developed an adaptive block-based CS framework for video, which classified blocks with their inter-frame correlations and texture region. Mum *et al.* [21] improved the quality of reconstructed frames by combining the motion-compensated prediction from the reference frame and the recovered residual with the smooth projected Landweber reconstruction. Later, this method was extended to multi-hypothesis prediction based on multiple reference frames [22]. In [23], Karhunen-Loève transform (KLT) bases were adopted to exploit long-term inter-frame correlations, which adaptively represented sparsities for each block based on reference frames. These methods achieved a potential gain over conventional video acquisition methods. However, they cannot provide theoretical tractability and performance guarantees, as explicit utilization of motion information hampers the derivation of an adaptive sensing matrix to ensure strict sparsity property for the spatio-temporal correlations.

Kronecker product framework was first introduced for multidimensional signals in Kronecker compressive sensing (KCS [24]), where spatio-temporal sparsity was jointly modeled with the Kronecker product of a 2-D spatial sparsifying basis and 1-D temporal sparsifying basis. KCS achieved the stable recovery for compressive video sampling with a block-diagonal sensing matrix synthesized by sensing matrix and full-rank identity matrix for spatial and temporal dimension. Thus, the recovery performance is degraded in KCS due to its redundant sampling in temporal dimension. To improve the recovery efficiency, multiway compressed sensing (MWCS [25]) developed a two-step process to utilize the Kronecker sensing structure, including fitting a low-rank model in compressed domain and per-mode decompression. However, MWCS is an NP-hard problem, as it heavily relies on tensor rank estimation. Recently, generalized tensor compressive sensing (GTCS [26]) was developed to employ the Kronecker structure for simultaneous acquisition and representation from all tensor modes, which considerably reduces the computational complexity for recovery by skipping the vectorization of video frames and tensor rank estimation. However, its recovery performance is degraded over the fixed basis (e.g. 3D DCT basis) due to the error accumulation with tensor mode continuing.

Furthermore, sensing matrix optimization is not considered in [24]–[26], which would substantially enhance the CS reconstruction with derived basis. However, existing optimization methods [27]–[30] cannot perform efficiently for synthetic sensing matrix generated by Kronecker products. Tremendous computation burden is led by iteratively optimizing the sensing matrix to force the Gram matrix to identity matrix. Considering the separability of Kronecker matrix, it would be desirable to design an efficient algorithm based on separable metrics like mutual coherence.

This paper proposes a fully decomposable compressive sampling model to exploit the structured sparsity in multidimensional signals. The contribution of this paper is two-fold. Firstly, the proposed model adopts a rank-deficient matrix to improve compressive sampling in multiple dimensions. Integrating with the Kronecker product framework, it can achieve a higher compression ratio with a guarantee of recovery performance. Secondly, the proposed model optimizes the synthetic sensing matrix by jointly minimizing the mutual coherence of the projection matrix. Decomposable optimization is allowed for sampling signals with multiple dimensions using sampling rate-distortion optimization. Given sparsifying basis, the optimized sensing matrix can achieve both reliable performance and improved convergence rate in reconstruction.

To validate the efficacy of the proposed model, we employed it into video acquisition to evaluate its recovery performance under various sampling rates. It improves the compression ratio with a guaranteed recovery performance by performing compressive sampling in both spatial and temporal dimensions. The proposed model is shown to outperform conventional CS schemes and latest Kronecker product based methods in terms of PSNR and visual quality. We also evaluated the sensing matrix optimization and sampling rate allocation scheme, which can effectively improve the recovery performance and computational complexity based on given number of measurements.

The reminder of this paper is organized as follows. In Section II, the relevant preliminaries and analysis on Kronecker compressive sensing (KCS) are introduced. Section III proposes the fully decomposable compressive sampling model, including its formulation, upper bound of mutual coherence, stable recovery and advantages to KCS. Sensing matrix optimization for fully decomposable CS is developed in Section IV with a well-designed algorithm for sampling rate allocation. In Section V, fully decomposable CS is employed into video acquisition with joint optimization for spatio-temporal sparsity. Section VI provides extensive experiment results to evaluate the proposed video acquisition schemes. Finally, we conclude this paper in Section VII.

II. Preliminaries and Motivation

In the rest of this paper, we reserve normal symbols to scalar variables and boldface symbols to vector variables for clarity. Table I specifies the frequently-used variables in this paper.

A. Compressive Sampling

A signal $\mathbf{x} \in \mathbb{R}^N$ is said to be K-sparse over the basis Ψ, if $\mathbf{x}$ can be decomposed by $\Psi\theta$ with $\|\theta\|_0 = K \ll N$. According to CS theory, the K-sparse signal $\mathbf{x}$ can be exactly reconstructed from $M = O(K \log(N/K))$ linear projections onto the

TABLE I
ABBREVIATION TABLE

SUMMARY OF ABBREVIATIONS AND MODEL PARAMETERS	
$\mathbf{X}$	original video cubic
$\mathbf{Y}$	CS measurements of $\mathbf{X}$
$\mathbf{F}_i$	the *i*-th frame of $\mathbf{X}$
θ	sparse/compressive representation of signal $\mathbf{x}$
Φ	the sensing matrix
Ψ	the sparsifying basis
$\mathbf{D}$	the projection matrix $\mathbf{D} = \Phi\Psi$
δ_K	the *K*-restricted isometry constant
ϕ^i	the *i*-th row of matrix Φ
ψ^j	the *j*-th column of matrix Ψ
$\mathbf{d}^i$	the *i*-th column of matrix $\mathbf{D}$
Φ_s	sensing matrix for the spatial dimension
Φ_t	sensing matrix for the temporal dimension
$r_t(r_s)$	sampling rate for the temporal (spatial) dimension
r	the overall sampling rate $r = r_t \cdot r_s$
μ	the mutual coherence
N	the number of samples
M	the number of measurements
K	the sparsity

sensing matrix Φ. Here, the nonadaptive sensing matrix $\Phi \in \mathbb{R}^{M\times N}$ is designed to satisfy the incoherent condition with fixed sparsifying basis Ψ [31]. It has been shown that the independent and identically distributed (i.i.d.) Gaussian or Bernoulli random matrices can facilitate a simple construction of nonadaptive sensing matrices [32]. The signal $\mathbf{x}$ is recovered by finding the sparsest vector θ consistent with the measurement vector $\mathbf{y} = \Phi\mathbf{x} = \Phi\Psi\theta$, which can be solved by the ℓ_0-minimization problem [33], [34] or the computationally tractable ℓ_1 optimization problem.

$$\hat{\theta} = \arg\min \|\theta\|_1 \;\; s.t. \;\; \|\mathbf{y} - \Phi\Psi\theta\|_2 \leq \sigma. \tag{1}$$

Commonly, (1) can be solved through Basis Pursuit [35]–[38] at the computational complexity of $\mathcal{O}(N^3)$.

Compressive sampling relies on the fundamental principle of incoherent sampling, which makes the sensing matrix Φ and the sparsifying basis Ψ to be as incoherent as possible. This principle can be guaranteed by the restricted isometry property (RIP [5], [39]) under constraints on Φ and Ψ. However, it requires a combinatorial computational complexity to verify a general matrix $D = \Phi\Psi$. Thus, mutual coherence is introduced for more concrete recovery based on Φ and Ψ [40].

Definition 1 (Mutual Coherence): The mutual coherence of the orthonormal bases $\Phi \in \mathbb{R}^{N\times N}$ and $\Psi \in \mathbb{R}^{N\times N}$ is the maximum absolute value for the inner product between elements of the two bases.

$$\mu(\Phi, \Psi) = \max_{1\leq i,j\leq N} \mid (\phi^i, \psi^j) \mid, \tag{2}$$

Under lower coherence between Φ and Ψ, the signal $\mathbf{x}$ can be recovered by a smaller set of CS samples with less mutual information.

B. Compressive Sampling for Multi-Dimensional Signals

Compressive sampling for multi-dimensional signals involves complex models for high-dimensional sparsity. Conventional (global) CS methods reshape the multi-dimensional signal into a single 1-D vector without considering the space $\mathbf{X}$ spans. Thus, $\mathbf{X}$ is sampled by directly multiplying the 1-D vector with a global dense matrix Φ.

For multi-dimensional signals, it is critical to find appropriate bases to jointly model the high-dimensional sparsity. One promising way is to combine the bases for sparsity in various dimensions with Kronecker product. Here, we define the Kronecker product $(\mathbf{A} \otimes \mathbf{B})_{PR\times QS}$ of two matrices $\mathbf{A}$ and $\mathbf{B}$ with sizes $P \times Q$ and $R \times S$.

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} A(1,1)\mathbf{B} & A(1,2)\mathbf{B} & \cdots & A(1,Q)\mathbf{B} \\ A(2,1)\mathbf{B} & A(2,2)\mathbf{B} & \cdots & A(2,Q)\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ A(P,1)\mathbf{B} & A(P,2)\mathbf{B} & \cdots & A(P,Q)\mathbf{B} \end{bmatrix}, \tag{3}$$

where $A(p,q)$ is the element at p-th row and q-th column in matrix $\mathbf{A}$. In the Kronecker product framework, a multi-dimensional sparsifying basis Ψ is constructed under the assumption that multi-dimensional signal $\mathbf{X}$ is sparse or compressible conditioned on basis Ψ_k along its k-th dimension. Thus, Ψ can be synthesized with the Kronecker product of Ψ_k, $1 \leq k \leq E$ for all the E dimensions of $\mathbf{X}$.

$$\Psi = \Psi_1 \otimes \Psi_2 \otimes \cdots \otimes \Psi_E. \tag{4}$$

Thus, the i-th column ψ^i of Ψ can be represented by

$$\psi^i = \psi_1^{i_1} \otimes \psi_2^{i_2} \otimes \cdots \otimes \psi_E^{i_E}, \tag{5}$$

where $\psi_k^{i_k}$ is the corresponding column in sparsifying basis Ψ_k for the k-th dimension. It is demonstrated in [41] and [42] that hyperbolic basis is optimal for (4) among bases derived from its universality along different dimensions. For example, a 2-D wavelet transform with basis Ψ_s and a 1-D wavelet transform with basis Ψ_t can sparsely represent an image and a smooth or piecewise smooth sequence of pixels, respectively. The Kronecker product of Ψ_s and Ψ_t would construct a hyperbolic wavelet basis with different degrees of smoothness for each dimension [43]. On the contrary, global CS would require multiplexing sensors to operate simultaneously along all dimensions, which are prohibitive and impractical for video acquisition due to the physical complexity and latency.

Without loss of generality, we consider the signal $\mathbf{X}$ of E dimensions. The mutual coherence of the sensing matrix $\hat{\Phi} = \Phi_1 \otimes \cdots \otimes \Phi_E$ and sparsifying basis $\hat{\Psi} = \Psi_1 \otimes \cdots \otimes \Psi_E$ can be formulated as

$$\mu(\hat{\Phi}, \hat{\Psi}) = \prod_{k=1}^{E} \mu(\Phi_k, \Psi_k) \leq \min_{1\leq k\leq E} \mu(\Phi_k, \Psi_k). \tag{6}$$

Equation(6) implies that KCS requires less number of samples than performing CS independently on each dimension (aka, independent CS) to recover the signals of interest, as $\mu(\hat{\Phi}, \hat{\Psi})$ will not exceed $\mu(\Phi_k, \Psi_k)$, $1 \leq k \leq E$.

To summarize, global CS directly performs compressed sampling and global reconstruction for the entire signals. However, it is computationally intensive in practice due to its exponentially increasing data volume and complexity. Independent CS employs individual compressed sampling and individual reconstruction for each dimension. For multi-dimensional signal $\mathbf{X} \in \mathbb{R}^{N_1\times N_2\times\cdots\times N_E}$, independent CS performs sampling and

recovery for $i = 1, \ldots, N_k$ along the k-th dimension in an isolated manner.

$$\text{vec}\left(\mathbf{Y}_i\right) = \Phi_k \text{vec}\left(\mathbf{X}_i\right) = \Phi_k \Psi_k \theta_i,$$

where $\mathbf{X}_i \in \mathbb{R}^{N_1 \times N_{k-1} \times N_{k+1} \times \cdots \times N_E}$ and $\Phi_k \in \mathbb{R}^{M_k \times (N_1 \cdot N_{k-1} \cdot N_{k+1} \cdots N_E)}$. For example, independent CS separately handles each frame of video sequences with Φ_t and Ψ_t. Thus, it degrades the sampling and recovery performance by neglecting the temporal sparsity. KCS provides a tractable alternative to balance the complexity and performance for sampling and recovery. It employs the Kronecker product framework to generate sensing matrix for joint recovery.

$$\mathbf{Y} = \left(\Phi_1 \otimes \cdots \otimes \Phi_E\right) \mathbf{X} = \left(\Phi_1 \otimes \cdots \otimes \Phi_E\right) \times \left(\Psi_1 \otimes \cdots \otimes \Psi_E\right) \theta.$$

For compact representation, Φ_k, $k = 2, \ldots, E$ are commonly set to identity matrices $\mathbf{I}$ in KCS. Thus, sampling is not performed along these dimensions. However, KCS cannot sufficiently capture the temporal sparsity in video sequences, as the identity sensing matrix is adopted for temporal sampling, as shown in (21). This fact means that the measurements $\mathbf{Y}$ individually sampled from frames $\mathbf{F}_1, \ldots, \mathbf{F}_T$ are redundant due to the temporal correlations in video sequences. The spatial measurements for adjacent frames could be further clustered with temporal sparsity to reduce the necessary samples for recovery.

III. Fully Decomposable Compressive Sampling (FDCS)

Recall the synthetic sensing matrix $\hat{\Phi} = \Phi_s \otimes \mathbf{I}$ in KCS, where $\mathbf{I}$ and Φ_s can be regarded as factorized sensing matrices for temporal and spatial dimension, respectively. The identity matrix $\mathbf{I}$ offers complete temporal components, while Φ_s determines the overall sampling rate by compressing the spatial measurements. In this section, we propose a general CS framework for multidimensional signals to perform decomposable sampling in each of its subspaces.

A. Formulation

Consider a multidimensional signal $\mathbf{X}$ represented by a union of E subspaces $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_E$, where $\mathcal{S}_i$ is spanned by the sparsifying basis Ψ_i. It is obvious that KCS cannot sufficiently exploit the sparsity in the multidimensional signals with $E \geq 2$, as its measurements sampled with the full-rank identity matrix are redundant in the temporal dimension. To develop an efficient sampling model, we redefine the overall sensing matrix $\hat{\Phi}$ based on the rank-deficient sensing matrices Φ_i for $\mathcal{S}_i$. The measurements $\mathbf{Y}$ are obtained by sampling $\mathbf{X}$ based on a Kronecker product framework.

$$\text{vec}(\mathbf{Y}) = \left(\Phi_1 \otimes \cdots \otimes \Phi_E\right) \cdot \text{vec}(\mathbf{X}), \tag{7}$$

where Φ_i is the $M_i \times N$ sensing matrix for $\mathcal{S}_i$ and $M = \prod_{i=1}^{E} M_i$ for all the E dimensions, respectively. According to

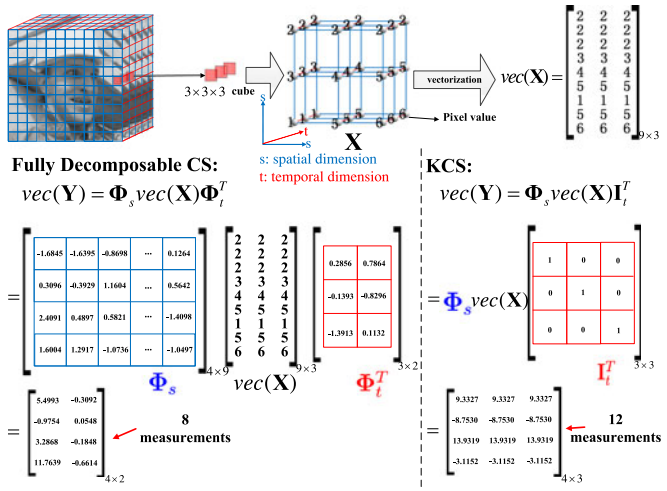


Fig. 1. An illustrative example for sampling with KCS and the proposed fully decomposable CS (FDCS) model, where FDCS and KCS are employed on a $3 \times 3 \times 3$ cube extracted from *Foremen* sequence. It shows that FDCS projects $\mathbf{X}$ into a vector with 8 elements, while KCS requires 4 more samples which can be clustered due to the temporal redundancy.

[44], (7) can be rewritten in matrix product from.

$$\mathbf{Y} = \prod_{i=1}^{E} \Phi_i \mathcal{V}_i\left(\mathbf{X}\right) \tag{8}$$

where $\mathcal{V}_i$ is the vectorization operator for the i-th subspace $\mathcal{S}_i$ that $\mathbf{X}$ lives in. Consequently, we define the fully decomposable compressive sampling for multidimensional signal $\mathbf{X}$.

Definition 2 (Fully Decomposable Compressive Sampling (FDCS)): A multidimensional signal $\mathbf{X}$ is called to be fully decomposed over a union of E subspaces $\mathcal{S}_i$, when there exists sensing matrices Φ_i and sparsifying bases Ψ_i, $i = 1, \ldots, E$ satisfying that

$$\mathbf{Y} = \otimes_{i=1}^{E} \Phi_i \Psi_i \theta_i, \tag{9}$$

where θ_i is the decomposable sparse representation for the projection of $\mathbf{X}$ onto subspace $\mathcal{S}_i$.

Here, the transform $\Phi_i \Psi_i \theta_i$ refers to the measurements corresponding to the subspace $\mathcal{S}_i$ spanned by the sparsifying basis Ψ_i. To be concrete, $\Phi_i \in \mathbb{R}^{M_i \times N}$ in (8) compresses projection in subspace $\mathcal{S}_i$ into a vector of M_i measurements, as it is always underdetermined for $M_i < N$. When $\mathbf{X}$ is decomposable over the E subspaces, operation $\Phi\mathbf{X}$ obtains CS measurements of $\mathbf{X}$ in a progressive fashion based on Φ_i and Ψ_i for $\mathcal{S}_i$. When progressively performing compressive sampling from $i = 1$ to E, $\Phi_i \in \mathbb{R}^{M_i \times N}$ clusters projection of signals at corresponding position in a partial union $\mathcal{S}_1 \otimes \cdots \otimes \mathcal{S}_{i-1}$ of subspaces into an M_i-tuple vector.

In comparison to KCS, FDCS adopts an underdetermined matrix Φ_i instead of the identity matrix $\mathbf{I}$ for each subspace, so that it can simultaneously exploit the structured sparsity of these subspaces to obtain the compressed measurements $\mathbf{Y}$. Fig. 1 provides a supporting evidence by employing FDCS and KCS on a $3 \times 3 \times 3$ cube extracted from *Foremen* sequence, respectively. It shows that fully decomposable CS projects $\mathbf{X}$

into a matrix $\mathbf{Y}$ with 8 measurements, while KCS requires 12 measurements to recover the cube. FDCS simultaneously exploits spatial and temporal correlations to generates fewer measurements that are randomly distributed.

B. Upper Bound for Mutual Coherence

In this section, we estimate the redundancy of $\mathbf{X}$ in the sense of mutual coherence derived by the FDCS. Given sensing matrix Φ and sparsifying basis Ψ, the upper bound of mutual coherence is developed to guarantee the recovery performance with Kronecker product framework. Given sensing matrix $\Phi \in \mathbb{R}^{M\times N}$, the sampling rate is defined as $r = M/N$. In the Kronecker product framework, the sparsifying basis for $\mathbf{X}$ is $\hat{\Psi} = \Psi_1 \otimes \cdots \otimes \Psi_E$. The corresponding sensing matrices for fully decomposable CS is $\hat{\Phi}^{(F)} = \Phi_1 \otimes \cdots \otimes \Phi_E$, while Φ_i, $i = 2, \ldots, E$ are set to identity matrices $\mathbf{I}$ for KCS. In Theorem 1, the mutual coherence for the fully decomposable CS model is demonstrated to be upper-bounded by the one derived from KCS.

Theorem 1: Given corresponding sampling rates $r_i \in (0, 1)$ for Φ_i, $i = 1, \ldots, E$, in the fully decomposable CS model and the overall sampling rate $r = \prod_{i=1}^{E} r_i$, under Gaussian random sampling, with a high probability asymptotically approaching 1, the mutual coherence $\mu(\hat{\Phi}^{(F)}, \hat{\Psi})$ for FDCS is upper-bounded by the mutual coherence $\mu(\hat{\Phi}^{(K)}, \hat{\Psi})$ for KCS.

$$\mu(\hat{\Phi}^{(F)}, \hat{\Psi}) \leq \mu(\hat{\Phi}^{(K)}, \hat{\Psi}) \tag{10}$$

Proof: Please refer to Appendix A. ■

The proposed sampling scheme can be viewed as an E-th order CS, which makes FDCS suitable for the common reconstruction framework. It is well recognized that the mutual coherence should be made small enough to ensure stable CS recovery. As is known, the sampling rate of a sensing matrix determines the coherence with a universal basis, Theorem 1 demonstrates that, by fixing overall sampling rate, FDCS can guarantee a smaller mutual coherence in comparison to KCS. This fact implies that FDCS can achieve better recovery performance based on the same number of measurements.

On the other hand, rate allocation is another critical factor on the performance of fully decomposable CS. According to CS theory, the more sparse a signal is, the less measurements are required to achieve stable recovery. Thus, the structured sparsity in the multi-dimensional signals could be better approximated with measurements properly assigned by fully decomposable CS. For example, the video sequence is treated as a combination of spatial and temporal sparsity, where the structure along temporal dimensions varies with scene changes and object motions, and the detail information in each frame determines the structure of spatial dimension. Thus, the recovery performance can be substantially improved with these structured sparsities. To balance the stability and performance for reconstructing various structures, it is reasonable to adaptively select sampling rates for multidimensional signals based on the statistics of the subspace they live in.

C. Stable Recovery

In this section, we demonstrate that stable recovery can be achieved for fully decomposable CS. Given the multidimensional $\mathbf{X}$, denote $\mathbf{D} = \Phi\Psi$ and $\mathbf{D}_i = \Phi_i\Psi_i$ the projection dictionary for the multidimensional signals $\mathbf{X}$ and its i-th subspace $\mathcal{S}_i$. Assuming that the projection of $\mathbf{X}$ onto each subspace $\mathcal{S}_i$ is K_i-sparse, we show that there exists $\mathbf{D}$ stable for block sparse vector $\mathbf{u}$ with $2K_i$-sparse component $\mathbf{u}_i$ for $\mathcal{S}_i$.

Theorem 2 (Stable Recovery): Given projection dictionary $\mathbf{D} = \mathbf{D}_1 \otimes \cdots \otimes \mathbf{D}_E$, it is stable for arbitrary block-sparse vector $\mathbf{u}$ if and only if its $2K_i$-sparse component $\mathbf{u}_i$ for $\mathcal{S}_i$ satisfies

$$C_1^i \|\mathbf{u}_i\|_2^2 \leq \|\mathbf{D}_i\mathbf{u}_i\|_2^2 \leq C_2^i \|\mathbf{u}_i\|_2^2. \tag{11}$$

Here, $\mathbf{u}_i = \theta_1 - \theta_2$ is the difference between two K_i sparse vector θ_1 and θ_2.

Proof: Please refer to Appendix B. ■

Theorem 2 implies that the multidimensional signal $\mathbf{X}$ can be recovered when block RIP is maintained for $\mathbf{D}_i$ in the i-th subspace $\mathcal{S}_i$. For each subspace $\mathcal{S}_i$, its vectorized component $\mathcal{V}_i\mathbf{X}$ can be separately recovered with $\ell_{2,1}$ optimization, when the number of measurements is greater than the sparsity K_i. Thus, the multidimensional signals $\mathbf{X}$ can be recovered by progressively combining the vectorized projection from all the subspaces $\mathcal{S}_i$.

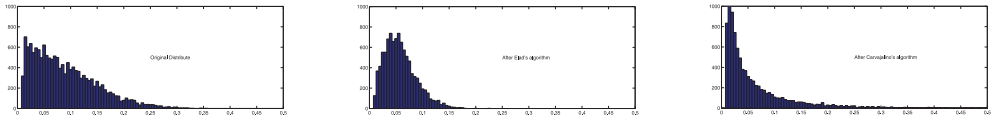
In the fully decomposable CS model, we use a series of underdetermined matrices $\Phi_1, \ldots, \Phi_E$ as the sensing matrices corresponding to its subspaces $\mathcal{S}_1, \ldots, \mathcal{S}_E$. Its measurements enable a high compression ratio which comes from the extra compressed temporal dimensions. The fully decomposable CS model can sufficiently exploit the structured sparsity in multiple subspaces to derive irregular measurements to eliminate their correlations. In comparison to KCS, the fully decomposable CS model would linearly increase the computational complexity for sampling, as (8) shows that the proposed scheme could progressively sample each subspace. The proposed sampling structure retains the feature of traditional CS with the exception of an extra cache to store temporary frame snapshot. This cache is utilized for the subsequent second-order CS. Once the samples of fully decomposable CS are received, an efficient joint recovery will help us reconstruct the original 3-D signal.

IV. Sensing Matrix Optimization for Multidimensional Signals

In compressive sampling, the reconstruction performance depends on the compressibility of the signal, the choice of the reconstruction algorithm, and the incoherence between sensing matrix and the sparsifying basis. In this section, we develop an algorithm to further optimize the synthetic sensing matrix for reconstruction in the fully decomposable CS model.

A. Minimization of Mutual Coherence

The randomly generated sensing matrix can enjoy additional benefit from universality, but it cannot perfectly satisfy the incoherence requirements. To improve the sampling efficiency and reconstruction performance, it is promising to develop an



(a) Without optimization (b) Elad's method with Threshold= 0.2 and Shrink Constant= 0.6 (c) Carvajalino's method

Fig. 2. Histogram of the absolute values of the off-diagonal entries of the Gram matrix G before the optimization and afterwards.

optimized algorithm to substantially reduce the coherence between sensing and sparsifying matrices. In the fully decomposable CS model, it is imperative to optimize the sensing matrix by considering the large dimensionality of Kronecker product matrix. It is shown that the optimized sensing matrix with smaller mutual coherence with a given basis would achieve efficient sampling and fast convergence.

Recall a related definition of mutual coherence for the projection dictionary $\mathbf{D} = \Phi\Psi$.

$$\mu(\mathbf{D}) = \max_{1\le i\ne j\le N} \frac{\mid (\mathbf{d}^i)^T \mathbf{d}^j \mid}{\|\mathbf{d}^i\| \cdot \|\mathbf{d}^j\|}, \tag{12}$$

where $\mathbf{d}^i$ denotes the i-th column of $\mathbf{D}$. The mutual coherence measures the worst similarity between columns of $\mathbf{D}$. Since pursuit techniques cannot distinguish two closely related columns, the mutual coherence plays an important role on the performance of recovery algorithms. Projection dictionary with small mutual coherence tends to achieve incoherent sampling and indicate a robust reconstruction, while the one with large coherence is more likely to fail. Suppose that the sparse representation for $\mathbf{x}_0$ is in the form of $\mathbf{D}\alpha_0$ with $\|\alpha_0\|_0 \le \frac{1}{2}(1 + \frac{1}{\mu(\Phi\Psi)})$. We have the following results for ℓ_0 and ℓ_1 optimization.

1) The vector α_0 is necessarily the sparsest one to describe $\mathbf{x}_0$ by solving the ℓ_0 optimization.

$$\min_{\alpha} \|\alpha\|_0 \quad s.t.\ \mathbf{x}_0 = \Phi\Psi\alpha \tag{13}$$

2) The Basis Pursuit (BP) algorithm for approximating α_0 guarantees an exact solution to the linear programming problem.

$$\min_{\alpha} \|\alpha\|_1 \quad s.t.\ \mathbf{x}_0 = \Phi\Psi\alpha \tag{14}$$

Thus, the effect of mutual coherence on the measurements can be evaluated. The two metrics imply that the upper bound of $\|\alpha\|_0$ will be relaxed when $\mu(\Phi\Psi)$ decreases. This relaxed upper bound allows a wider set of candidate signals to reside under the umbrella of successful CS. It benefits the reconstruction in two aspects. First, more accurate recovery can be obtained with the same number of measurements. Second, a fast convergence to the ℓ_1 convex optimization can be achieved. Therefore, the optimization for sensing matrix can be transferred to minimization of the mutual coherence of projection dictionary $\mathbf{D}$.

In practice, to ensure a fair comparison to actual behavior of sparse representations and pursuit algorithm's performance, we replace the worst-case stand-point with an "average" measure of mutual coherence. To average the significant mutual coherence between columns of $\mathbf{D}$, the Gram matrix $\mathbf{G} = \mathbf{D}^T\mathbf{D}$ is considered for the normalized columns of $\mathbf{D}$.

Definition 3: The t-average mutual coherence of $\mathbf{D}$ is

$$\mu_t(\mathbf{D}) = \frac{\sum_{1\le i\ne j\le k}(\mid g_{ij} \mid \ge t)\cdot \mid g_{ij} \mid}{\sum_{1\le i\ne j\le k}(\mid g_{ij} \mid \ge t)}, \tag{15}$$

where g_{ij} is the off-diagonal entry of Gram matrix $\mathbf{G}$ that represents the inner product between the normalized columns $\mathbf{d}^i/\|\mathbf{d}^i\|$ and $\mathbf{d}^j/\|\mathbf{d}^j\|$.

In [27], *Elad* developed an algorithm to minimize the $\mu_t(\mathbf{D})$ where off-diagonal entries in $\mathbf{G}$ that are larger than a threshold γ will be reduced with a shrunk constant and the rank of $\mathbf{G}$ would be subsequentially retained. In practice, the value of $\mu_t(\mathbf{D})$ would be reduced iteratively. Obviously, its shortcoming lies in that some small off-diagonal entries in $\mathbf{G}$ will be increased in the process of "artificial" rank retaining and will affect the RIP of $\mathbf{D}$. To overcome it, a learning method was targeted to simultaneously optimize sensing matrix and sparsifying basis by finding Φ that makes the corresponding Gram matrix $\mathbf{G}$ approximate the identity matrix [29].

$$\Psi^T\Phi^T\Phi\Psi \approx \mathbf{I}_n. \tag{16}$$

In Fig. 2, we provide the histograms of the absolute values of off-diagonal entries in $\mathbf{G}$ before and after the optimization. Fig. 2(b) shows that Elad's algorithm [27] preserves the ordering of the absolute entries in $\mathbf{G}$ and leads to better distribution, while it creates some artificial large values entries that are not present in the original matrix. These artificial values would completely ruin the worst-case guarantees of the pursuit algorithms. On the contrary, the learning based algorithm [29] improves the distribution of off-diagonal entries by orthogonalizing arbitrary subset of columns in $\mathbf{D}$. It makes all the off-diagonal entries in $\mathbf{G}$ shrink to zero and the diagonal entries approach one. Fig. 2(c) suggests that [29] provides the best mutual coherence and RIP of $\mathbf{D}$, as the histogram decreases monotonically. Consequently, the mutual coherence of corresponding synthetic sensing matrix is optimized for the Kronecker product in the fully decomposable CS model.

B. Separable Optimization for Kronecker Product

The separability of Kronecker product enables us to design a specific algorithm to simplify both the design procedure and the implementation of the synthetic sensing matrix to reduce complexity. In this section, we demonstrate the optimality of the separable minimization for the mutual coherence in the Kronecker product framework. Thus, the optimized synthetic sensing matrix can be obtained by making Kronecker products of the ones optimized separately. Moreover, the optimization process is also separable, which preserves the block feature of

Kronecker product matrix and enables fast low-scale matrix computation.

The mutual coherence of the projection dictionary in fully decomposable CS is formulated as $\mu(\hat{\Phi}\hat{\Psi}) = \mu(\Phi_1\Psi_1)\cdots\mu(\Phi_E\Psi_E)$ for the E subspaces. According to the theorem, we can obtain the optimal mutual coherence $\mu(\hat{\Phi}\hat{\Psi})$ by calculating the product of E optimized mutual coherence $\mu(\Phi_1\Psi_1),\ldots,\mu(\Phi_E\Psi_E)$. For each subspace $\mathcal{S}_i$, Φ_i is optimized for wavelet basis Ψ_i. For arbitrary $1 \le i, j \le E$, $\mu(\Phi_i\Psi_i)$ and $\mu(\Phi_j\Psi_j)$ are independent with each other. Thus, Φ_i can be determined separately to minimize $\mu(\Phi_i\Psi_i)$.

Algorithm 1 is developed to optimize matrix Φ by separately minimizing mutual coherence for spatial and temporal sensing matrices based on a fixed sparsifying basis Ψ. Taking the temporal sensing matrix as example, the optimization starts with a randomly generated Gaussian matrix $\Phi_t \in \mathbb{R}^{M\times N}$. The optimal sensing matrix $\hat{\Phi}_t$ is obtained under the constraint that the Gram matrix $\mathbf{G} = \Psi^T\Phi_t^T\Phi_t\Psi$ is close to an $N \times N$ identity matrix. Rewriting (16) by multiplying both side of $\mathbf{G}$ with Ψ and Ψ^T, we obtain

$$\Psi\Psi^T\Phi_t^T\Phi_t\Psi\Psi^T - \Psi\Psi^T \approx 0. \tag{17}$$

Thus, the optimized sensing matrix $\hat{\Phi}_t$ minimizes $\|\Psi\Psi^T\Phi_t^T\Phi_t\Psi\Psi^T - \Psi\Psi^T\|_F^2$ with the fixed basis Ψ. Let us define the eigendecomposition of $\Psi\Psi^T$ as $\Psi\Psi^T = \mathbf{Q}\mathbf{\Sigma}\mathbf{Q}^T$, where $\mathbf{\Sigma} = \mathrm{diag}(\lambda_1,\ldots,\lambda_N)$ is the diagonal matrix whose entries are the eigenvalues of $\Psi\Psi^T$ and $\mathbf{Q}$ is the orthogonal matrix composed of corresponding eigenvectors. Therefore, $\|\mathbf{\Sigma}\mathbf{Q}^T\Phi_t^T\Phi_t\mathbf{Q}\mathbf{\Sigma} - \mathbf{\Sigma}\|_F^2$ can be considered as an equivalent substitution of (17). Substituting $\mathbf{W}$ with $\Phi\mathbf{Q}$, we formulate the equivalent minimization problem for the residual.

$$\hat{\mathbf{W}} = \arg\min_{\mathbf{W}} \|\mathbf{\Sigma} - \mathbf{\Sigma}\mathbf{W}^T\mathbf{W}\mathbf{\Sigma}\|_F^2 \tag{18}$$

Thus, $\hat{\Phi}_t = \hat{\mathbf{W}}\mathbf{Q}^T$ minimizes (17).

Denote $\mathbf{v}_i = [\lambda_1 W_{i,1},\ldots,\lambda_N W_{i,N}]$ the i-th row of matrix $\mathbf{W}\mathbf{\Sigma}$. We define $\mathbf{R} = \mathbf{\Sigma} - \sum_{i=1}^{M}\mathbf{v}_i^T\mathbf{v}_i$ and $\mathbf{R}_j = \mathbf{\Sigma} - \sum_{i:i\neq j}\mathbf{v}_i^T\mathbf{v}_i$. In the optimization, $\mathbf{v}_j$ is obtained by minimizing the residuals $\|\mathbf{R}_j - \mathbf{v}_j^T\mathbf{v}_j\|_F^2$. Provided that the eigendecomposition of $\mathbf{R}_j$ is $\mathbf{R}_j = \mathbf{U}_j\mathbf{\Delta}_j\mathbf{U}_j^T$, $\mathbf{v}_j$ is set to $\sqrt{\delta_{1,j}}\mathbf{u}_{1,j}$ to eliminate the largest item of residuals. Here, $\delta_{1,j}$ is the largest eigenvalue of $\mathbf{\Delta}_j$ and $\mathbf{u}_{1,j}$ is its corresponding eigenvector in $\mathbf{U}_j$. Therefore, $\hat{\mathbf{W}}$ and $\hat{\Phi}_t$ can be solved from $\{\mathbf{v}_i\}$ and $\{\lambda_i\}$. Similarly, we can find the optimal spatial sensing matrix $\hat{\Phi}_s$, and consequently, generate the synthetic sensing matrix $\hat{\Phi} = \hat{\Phi}_s \otimes \hat{\Phi}_t$.

C. Sampling Rate-Distortion Optimization

Since compressive sensing is separably performed on both spatial and temporal dimension, sampling rate-distortion performance is optimized for fully decomposable CS. Considering that Algorithm 1 minimizes $\|\Psi\Psi^T\Phi^T\Phi\Psi\Psi^T - \Psi\Psi^T\|_F^2$ for the corresponding dimension, we introduce Lagrangian cost function $\mathcal{L}(\Phi_s, \Phi_t)$ for optimized sampling rate allocation.

Algorithm 1: Separable Sensing Matrix Optimization.

1: **Task:** Find the optimal sensing matrix $\hat{\Phi}_i$ that minimizes $\mu(\Phi_i\Psi_i)$.
2: **Initialization:** Sparsifying basis Ψ_i and a randomly generated Gaussian matrix $\Phi_i \in \mathbb{R}^{M\times N}$
3: Eigen-decomposition: $\Psi_i\Psi_i^T = \mathbf{Q}\mathbf{\Sigma}\mathbf{Q}^T$
4: Set $\mathbf{W} = \Phi\mathbf{Q}$
5: Set $j = 1$;
6: **Repeat** M **times:**
7: Compute $\mathbf{R}_j = \mathbf{\Sigma} - \sum_{1\le i\le M, i\neq j}\mathbf{v}_i^T\mathbf{v}_i$
8: Eigen-decomposition: $\mathbf{R}_j = \mathbf{U}_j\mathbf{\Delta}_j\mathbf{U}_j^T$
9: Sort the largest diagonal elements $\delta_{1,j}$ and its eigenvector $\mathbf{u}_{1,j}$
10: Update $\mathbf{w}_j$ using $[\lambda_1\mathbf{w}_{j,1},\ldots,\lambda_N\mathbf{w}_{j,N}] = \sqrt{\delta_{1,j}}\mathbf{u}_{1,j}$
11: Set $j = j + 1$
12: Update M components of $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_M]$
13: Compute the optimal $\hat{\Phi}_i = \mathbf{W}\mathbf{Q}^T$

$$\mathcal{L}(\Phi_1,\ldots,\Phi_E) = \sum_{i=1}^{E}\eta_i\|\Psi_i^T\Phi_i^T\Phi_i\Psi_i - \mathbf{I}\|_F^2, \tag{19}$$

where η_i is the Lagrangian multiplier for the i-th subspace $\mathcal{S}_i$. Given the total sampling rate r, we minimize $\mathcal{L}(\Phi_1,\ldots,\Phi_E)$ based on their partial sampling rates $r_1,\ldots,r_E$.

Theorem 3: Given arbitrary smooth wavelet bases Ψ_i for the i-th subspace $\mathcal{S}_i$, with a high probability, there exists optimal sampling rate allocation r_i, $1 \le i \le E$ to minimize the upper bound of recovery error.

Proof: Please refer to Appendix C. ■

Theorem 3 implies that there exists an optimal sampling rate allocation $(r_1,\ldots,r_E)$ for the multidimensional signals. It shows that the sampling rate allocation is related to the spatial and temporal sparsifying basis, which affects the sparsity K_i, $1 \le i \le E$ of multidimensional signals.

In practice, the heuristic optimization would be prohibitive to determine sampling rate. To balance the efficiency and recovery performance, a set of L candidate pairs of spatial and temporal sampling rates $\{(r_1^1,\ldots,r_E^1),\ldots,(r_1^L,\ldots,r_E^L)\}$ are predetermined for selection. For each pair $(r_1^l,\ldots,r_E^l)$, the Lagrangian cost function is evaluated for comparison, so that the optimal synthetic sensing matrix can be derived based on the optimized allocation scheme. In Section V, we elaborate the sampling rate allocation scheme for compressive video sampling.

V. Application Into Video Acquisition

In this section, we employ the proposed FDCS into compressive video sampling. Consider that a video sequence $\mathcal{F}$ consists of T consecutive frames $\mathbf{F}_1,\ldots,\mathbf{F}_T$ with a same resolution of $N = N_R \times N_C$. Let us denote $\mathbf{X}$ the $N \times T$ matrix for the video sequence.

$$\mathbf{X} = (\mathrm{vec}(\mathbf{F}_1)\ \mathrm{vec}(\mathbf{F}_2)\cdots\mathrm{vec}(\mathbf{F}_T)) \in \mathbb{R}^{N\times T}. \tag{20}$$

Algorithm 2: Sampling Rate-distortion Optimization for Fully Decomposable CS (FDCS).

Task: Determine the optimal sampling rate allocation from the L candidate pairs $\{(r_s^1, r_t^1), \ldots, (r_s^L, r_t^L)\}$
Initialization: Sparsifying basis Ψ_t and a randomly generated Gaussian matrix $\Phi_t \in \mathbb{R}^{M\times N}$
Eigen-decomposition: $\Psi\Psi^T = \mathbf{Q}\mathbf{\Sigma}\mathbf{Q}^T$
for $l = 1, \ldots, L$ **do**
    Estimate $\hat{\Phi}_s^l$ and $\hat{\Phi}_t^l$ based on the sampling rates r_s and r_t.
    Derive the cost $\mathcal{L}(\hat{\Phi}_s^l, \hat{\Phi}_t^l)$;
    Set $l = l + 1$
end for
Find the minimum in $\mathcal{L}(\hat{\Phi}_s^l, \hat{\Phi}_t^l)$, $l = 1, \ldots, L$
Generate synthetic sensing matrix $\hat{\Phi} = \hat{\Phi}_s^l \otimes \hat{\Phi}_t^l$

Here, the i-th column vector $\text{vec}(\mathbf{F}_i)$ of $\mathbf{X}$ is the 1-D vectorization of $N_R \times N_C$ frame $\mathbf{F}_i$. In progressive sampling for video acquisition [24], the sensing matrix for $\mathbf{X}$ is a block diagonal matrix generated by Kronecker product $\hat{\Phi} = \Phi_s \otimes \mathbf{I}$, where the $M \times N$ diagonal sub-block Φ_s enables sampling of vectorized frames $\text{vec}(\mathbf{F}_1), \ldots, \text{vec}(\mathbf{F}_T)$.

$$\text{vec}(\mathbf{Y}) = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_T \end{bmatrix} = \begin{bmatrix} \Phi_s & 0 & \cdots & 0 \\ 0 & \Phi_s & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Phi_s \end{bmatrix} \cdot \text{vec}(\mathbf{X}), \quad (21)$$

where $\mathbf{Y} = (\mathbf{y}_1\ \mathbf{y}_2 \cdots \mathbf{y}_T)$ is the $M \times T$ matrix with $\mathbf{y}_i$ indicating the $M \times 1$ measurements for $\mathbf{F}_i$. According to (21), KCS utilizes the identity matrix $\mathbf{I}$ to combine measurements from each frame, so that the block diagonal matrix $\hat{\Phi}$ samples each frame $\mathbf{F}_i$ into corresponding measurements $\mathbf{y}_i$ with Φ_s. By contrast, fully decomposable CS considers the synthetic sparsifying basis defined in (4) to generate $\mathbf{Y}$ by jointly optimizing over $\hat{\Phi}$ and $\hat{\Psi}$ for spatial and temporal dimension.

FDCS simultaneously performs compressive sampling in spatial and temporal dimension for efficient representation of video sequences. In comparison to KCS, temporal sensing matrix Φ_t is adopted for structured sparsity along the motion trajectory. Sampling rate allocation scheme is developed to assign sampling rates to spatial and temporal dimensions. Algorithm 2 is adopted to adaptively estimate spatial and temporal sensing matrices and the cost is derived based on the estimated synthetic sensing matrix. Sensing matrix for each dimension is optimized according to Algorithm 1.

Remarkably, the block-diagonal structure is reserved in the product $\Phi_s \mathbf{X} \Phi_t^T$, which means that the fully decomposable CS model satisfies the progressive sampling for video acquisition. Furthermore, fully decomposable CS can also obtain measurements separably in spatial and temporal domain with existing hardware implementations of CS imaging systems [45], [46] with separable sensing devices. According to (8), the temporal sensing can be implemented similarly to the spatial sensing. Once the 2-D spatial frames are compressively sensed and temporally stored, the DMD (digital micromirror device) is used to linearly combine spatial measurements at corresponding positions like a second-order CS. In Fig. 3, FDCS and KCS are evaluated over various video sequences. The measurements obtained by KCS are redundant on the temporal dimension (horizontally), as they are similar in the same row.

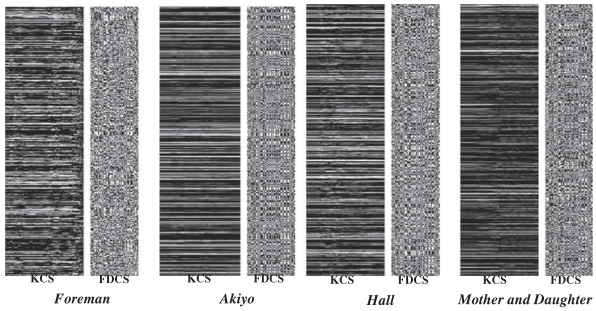


Fig. 3. The measurements comparison between KCS and fully decomposable CS (FDCS) in different video sequences. The measurements obtained by KCS are redundant along the horizontal direction (temporal dimension), as pixels with similar gray-scale values are randomly distributed in the compressed columns (spatial dimension) but located in the same row. By contrast, FDCS generates less measurements randomly distributed in both the horizontal and vertical directions, which implies that the spatial and temporal correlations are simultaneously exploited during the sampling.

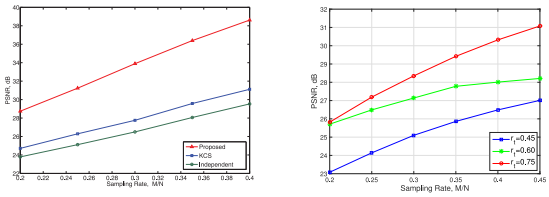
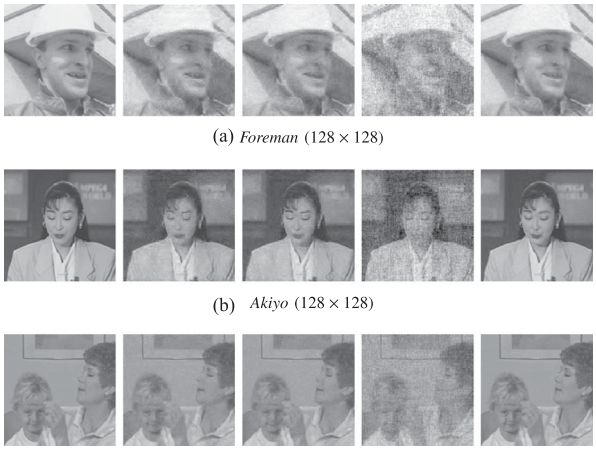


Fig. 4. (a) The reconstruction performance (dB) obtained by independent CS, KCS and fully decomposable CS under $r \in [0.20, 0.40]$ and $r_t = 0.6$, respectively; (b) The reconstruction performance (dB) obtained by fully decomposable CS under $r \in [0.20, 0.45]$ and $r_t = 0.45$, 0.60, and 0.75, respectively.

In Fig. 4(a) and 4(b), we provide examples for sampling *Akiyo* and *Foreman* sequences with various allocation of spatial and temporal sampling rates. Fig. 4(a) shows the sampling-rate-distortion performance for *Akiyo* sequence with a resolution of $128 \times 128 \times 128$ with the overall sampling rate r varying in the interval $[0.20, 0.40]$. The temporal sampling rate r_t is fixed at 0.6. It shows that FDCS outperforms KCS and independent CS by a gain of 3–8 dB in PSNR. In Fig. 4(b), we fix the overall sampling rate and compare the recovery performance of the proposed FDCS model with various temporal sampling rates. The *Foreman* sequence features scene movement, which is reflected in sharp changes in the value of each pixel across frames, so that it requires more temporal measurements. The best recovery is achieved with $r_t = 0.75$.

VI. Experimental Results

To evaluate the proposed fully decomposable CS (FDCS) model in video acquisition, we compare it with the state-of-the-art CS schemes in terms of recovery performance and analyze the influence of sensing matrix optimization. We also discuss the optimal configuration of rate allocation for various sequences.

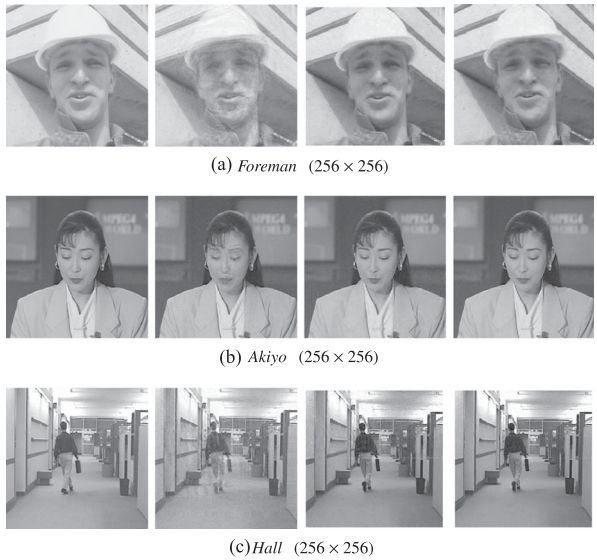


(a) *Foreman* (128 × 128)

(b) *Akiyo* (128 × 128)

(c) *Mother and Daughter* (128 × 128)

Fig. 5. Visual quality for recovered frames obtained by various CS schemes, including independent CS, KCS, GTCS and the proposed fully decomposable CS (FDCS). From left to right: Original, Independent CS, KCS with $r = 0.35$; GTCS with $r = 0.35$ and $r_t = 0.75$; FDCS with $r = 0.35$ and $r_t = 0.75$.



(a) *Foreman* (256 × 256)

(b) *Akiyo* (256 × 256)

(c) *Hall* (256 × 256)

Fig. 6. Visual quality for recovered frames obtained by various Kronecker product based schemes, including KCS, GTCS and the proposed fully decomposable CS (FDCS). From left to right: original, GTCS, KCS, and FDCS.

The Daubechies-8 wavelet was used to form sparsifying basis, and appropriate random matrices were set as the initial temporal and spatial sensing matrices. Referring to the matrix structure, we crop the standard video sequence around the center to have frames of size 128×128 pixels (approximate QCIF resolution) and 256×256 pixels (approximate CIF resolution). To appropriately utilize the temporal sparsity within hardware capacity, 128 frames are selected to form a test sequence. In the experiments, the basis pursuit (BP) solver is used to recover the original video signals. All the experiments are executed on a workstation with 3.3-GHz CPU and 12-GB RAM. It is noted that we develop the proposed algorithm over the MATLAB toolbox at *http://dsp.rice.edu/kcs*.

A. Recovery Performance

Firstly, we evaluate various CS-based sampling schemes on the cropped video cubes, including the independent CS recovery [9], the GTCS [26], the KCS [24] and the proposed fully decomposable CS (FDCS) model. For the independent recovery, CS independently obtains and recovers measurements for each individual frame with the sparsifying basis Ψ_s. The 3D-DCT basis is used for GTCS. For KCS and FDCS, the same Kronecker product (hyperbolic) wavelet basis is adopted. It is worth mentioning that KCS only utilizes a spatial-compressed synthetics matrix to reconstruct signals from the entire measurements, while the FDCS takes advantage of optimized holo-compressed sensing matrix $\hat{\Phi} = \hat{\Phi}_s \otimes \hat{\Phi}_t$ to sufficiently exploit the spatio-temporal correlations in video sequences. The total sampling rate (ratio of measurements and pixels) is set as $r = 0.35$. For the FDCS model, the temporal sampling rate r_t is selected as 0.75 to maximally envelop the temporal structures of all test sequences. Figs. 5 and 6 show the reconstructed frames for independent CS, GTCS, KCS, and the proposed FDCS model over test sequences with various resolutions, including *Foreman*, *Akiyo*, *Mother and Daughter*, and *Hall*. In comparison to KCS, GTCS and CS, the proposed FDCS model achieves best visual quality for recovered frames over all the test sequences. Its improvement on recovery owes to the availability of extra temporal compression, which tends to preserve more details with the same number of measurements.

B. Sensing Matrix Optimization

This section validates the optimized sensing matrix against the randomly generated matrix in FDCS. The sensing matrix optimization algorithm is evaluated in terms of the PSNR of reconstructed frames and time cost. Gaussian random matrix Φ and optimized sensing matrix $\hat{\Phi}$ are adopted to collect fully decomposable CS measurements under various sampling rates, respectively. Here, the optimized matrix $\hat{\Phi}$ is calculated from Φ with Algorithm 1. The mutual coherence of the random and optimized matrices is 0.092 and 0.055, respectively. Table II gives the detailed recovery quality and computational complexity of FDCS with both the optimized and the randomly generated sensing matrix. Remarkably, the optimized matrix achieves better recovery performance as well as reduces about 80% of the computational complexity in comparison to the random matrix. This fact coincides with the conclusions in Section IV.

C. Sampling Rate Allocation

Finally, we discuss the sampling rate allocation for the proposed FDCS model. Instead of full sampling along the temporal dimension in KCS, FDCS sufficiently exploits the temporal correlations to further compress the temporal components and achieve higher sampling efficiency. Therefore, it is reasonable to consider rate allocation for the spatial and temporal sampling to obtain optimal overall performance. As mentioned in Section II, the hyperbolic basis outperforms any other basis for its universality in various dimensions, which means that we can also construct appropriate sensing matrix according to the sparsity along both spatial and temporal dimensions.

TABLE II
RECOVERY PERFORMANCE FOR FULLY DECOMPOSABLE CS (FDCS) WITH OPTIMIZED AND RANDOM SENSING MATRIX UNDER VARIOUS SAMPLING RATES

Sequence	Model	r=0.20		r=0.25		r=0.30	
		PSNR (dB)	Time (S)	PSNR (dB)	Time (S)	PSNR (dB)	Time (S)
Foreman	Random	25.82	2.9e+04	27.19	3.4e+04	28.35	3.3e+04
	Optimized	**25.88**	**5.5e+03**	**27.22**	**5.6e+03**	**28.45**	**5.9e+03**
Akiyo	Random	26.60	3.1e+04	28.82	2.9e+04	31.00	3.1e+04
	Optimized	26.60	**6.1e+03**	**28.93**	**6.5e+03**	30.97	**4.6e+03**
Hall	Random	22.95	3.1e+04	24.57	2.7e+04	26.28	3.5e+04
	Optimized	**22.97**	**5.2e+03**	**24.59**	**4.5e+03**	**26.38**	**7.6e+03**
MD	Random	28.56	3.9e+04	30.31	4.5e+04	32.04	4.4e+04
	Optimized	**33.38**	**9.8e+03**	**35.93**	**8.6e+03**	**38.78**	**1.3e+04**
Sequence	Model	r=0.35		r=0.40		r=0.45	
		PSNR (dB)	Time (S)	PSNR (dB)	Time (S)	PSNR (dB)	Time (S)
Foreman	Random	29.42	3.3e+04	30.32	3.7e+04	31.08	4.1e+04
	Optimized	**29.52**	**5.1e+03**	**30.40**	**6.5e+03**	**31.20**	**5.4e+03**
Akiyo	Random	33.10	2.9e+04	35.30	3.0e+04	37.77	3.3e+04
	Optimized	**33.16**	**1.1e+04**	**35.48**	**7.2e+03**	**37.80**	**7.7e+03**
Hall	Random	27.98	3.2e+04	29.77	2.7e+04	31.92	2.8e+04
	Optimized	**28.05**	**5.9e+03**	**29.87**	**5.4e+03**	**32.02**	**4.3e+03**
MD	Random	33.73	3.7e+04	35.34	4.3e+04	36.88	4.2e+04
	Optimized	**41.58**	**8.6e+03**	**44.12**	**7.3e+03**	**45.59**	**4.9e+03**

Fig. 7. Recovery performance for fully decomposable CS (FDCS) under various sampling rates. The overall sampling rates for the first and second rows are 0.25 and 0.45, respectively. (a): Original; (e): KCS; (b)(f): FDCS with $r_t = 0.45$; (c)(g): FDCS with $r_t = 0.6$; (d)(h): FDCS with $r_t = 0.75$.

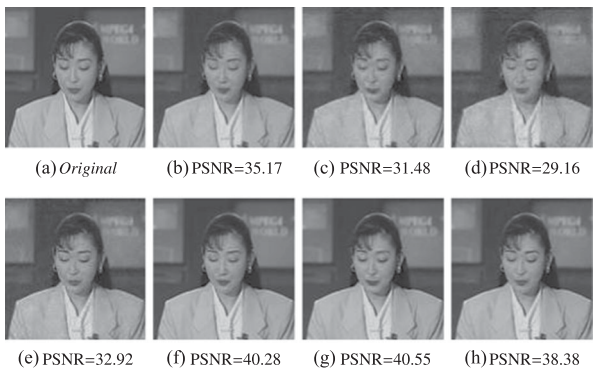


Fig. 8. Recovery performance for full decomposable CS (FDCS) under various sampling rates. The overall sampling rates for the first and second rows are 0.25 and 0.45, respectively. (a): Original; (e): KCS; (b)(f): FDCS with $r_t = 0.45$; (c)(g): FDCS with $r_t = 0.6$; (d)(h): FDCS with $r_t = 0.75$.

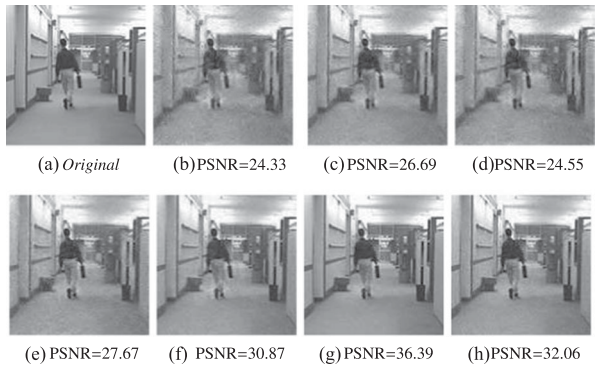
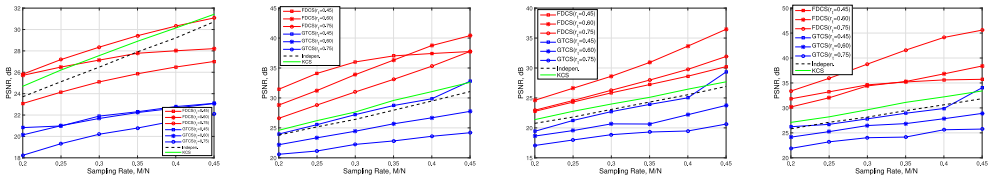


Fig. 9. Recovery performance for fully decomposable CS (FDCS) under various sampling rates. The overall sampling rates for the first and second rows are 0.25 and 0.45, respectively. (a): Original; (e): KCS; (b)(f): FDCS with $r_t = 0.45$; (c)(g): FDCS with $r_t = 0.6$; (d)(h): FDCS with $r_t = 0.75$.

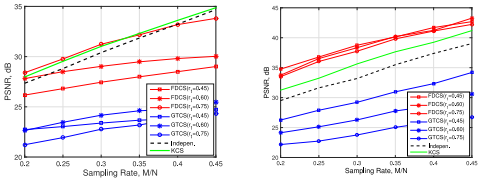
Fig. 10. Recovery performance for fully decomposable CS (FDCS) under various sampling rates. The overall sampling rates for the first and second rows are 0.25 and 0.45, respectively. (a): Original; (e): KCS; (b)(f): FDCS with $r_t = 0.45$; (c)(g): FDCS with $r_t = 0.6$; (d)(h): FDCS with $r_t = 0.75$.

Intuitively, the number of measurements would determine the quality of reconstruction in CS. However, we can notice in Fig. 7(d) and (f) that lower sampling rate could help achieve better reconstruction. To achieve the best performance in FDCS, the allocation of sampling rates for different dimensions should follow the distribution of sparsity, as shown in Figs. 7–10. For example, more measurements are required for the temporal dimension, as there exist sharp changes across frames in *Foreman* sequence. In Fig. 7, the reconstruction quality in (c)(g) with larger r_t is better than (b)(f) with the same overall sampling rate r. In contrast to the *Foreman* sequence, the *Akiyo* sequence is much more smooth along the temporal dimension. Thus, small



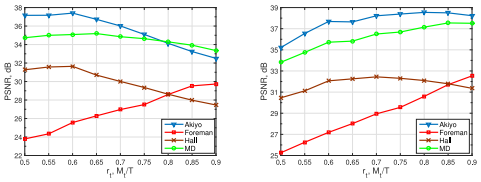
(a) *Foreman* (b) *Akiyo* (c) *Hall* (d) *Mother and Daughter*

Fig. 11. Reconstruction performance for fully decomposable CS (FDCS), GTCS, KCS, and independent CS over four sequences with sizes of $128 \times 128 \times 128$.



(a) *Foreman* (b) *Akiyo*

Fig. 12. Reconstruction performance for fully decomposable CS (FDCS), GTCS, KCS, and independent CS over two sequences with sizes of $256 \times 256 \times 128$.



(a) $128 \times 128 \times 128$ (b) $256 \times 256 \times 128$

Fig. 13. Sampling-rate-distortion curves for test sequences *Akiyo*, *Foreman*, *Hat* and *Mother and Daughter* under overall sampling rate 0.4.

r_t can also guarantee an accurate reconstruction, as shown in Fig. 8.

Figs. 11 and 12 provide the sampling rate-distortion curves obtained by FDCS, GTCS, KCS, and the independent CS for test sequences with various resolutions. These curves are obtained with total sampling rates ranging from 0.2 to 0.45. For fully decomposable CS and GTS, the temporal sampling rates are set to 0.45, 0.6 and 0.75, respectively. To sufficiently exploit temporal correlations, sensing matrix optimization is also employed into fully decomposable CS. Figs. 11 and 12 show that FDCS outperforms the state-of-the-art methods, especially in the regions of low sampling rates ($r < 0.3$). Remarkably, FDCS performs better under the situations that temporal sampling is sufficient. In comparison to the Kronecker product-based method KCS, FDCS achieves a noticeable gain (up to 8 dB) in PSNR. Fig. 13 provides the sampling-rate-distortion curves for various test sequences with sizes $128 \times 128 \times 128$ and $256 \times 256 \times 128$ under the overall sampling rate 0.4. It shows that the optimal sampling rate allocation for sequences would vary due to the various statistics for spatial and temporal dimension, i.e. sparsity. For example, recovery performance of *Foreman* raises with the growth of temporal sampling rate, as local motion of textures (i.e. the region of face) would require more measurements for perfect recovery.

VII. Conclusion

This paper proposes a novel fully decomposable CS (FDCS) for highly compressed multidimensional signals. With the structured sparsity, FDCS preserves a higher sampling efficiency by further compressing projected signals lived in multiple subspaces. Due to the block feature of Kronecker product, sampling rate allocation is developed to derive the overall sensing matrix for optimized reconstruction performance over the multiple subspaces. With the knowledge that projection matrix with low mutual coherence would lead to lower reconstruction errors, mutual coherence for each subspace is minimized by optimizing its sensing matrix based on the projection dictionary of FDCS. The mutual coherence is proved to be divisible, which makes it possible for obtaining the optimal synthetic sensing matrix from each dimension in a fast low-scale matrix computation. For validation, the proposed FDCS is employed in compressive video sampling. Experiments results show that the proposed scheme substantially improves the reconstruction accuracy and further reduces the necessary number of samples.

Appendix A
Proof of Theorem 1

We begin with $E = 2$. In the Kronecker product framework, the mutual coherence for synthetic sensing matrix $\Phi_s \otimes \Phi_t$ and the sparsifying basis $\Psi_s \otimes \Psi_t$ can be decomposed by

$$\mu(\Phi_s \otimes \Phi_t, \Psi_s \otimes \Psi_t) = \mu(\Phi_s, \Psi_s) \cdot \mu(\Phi_t, \Psi_t). \tag{22}$$

Thus, we consider the mutual coherence for spatial and temporal dimension, respectively.

Without loss of generality, we assume that the corresponding spatial sensing matrices $\Phi_s^{(F)}$ and $\Phi_s^{(K)}$ for the fully decomposable CS and KCS models are the selected sub-matrices of $N \times N$ orthonormal bases $\tilde{\Phi}_s$. Given $M \times N$ and $M_s \times N$ spatial sensing matrix $\Phi_s^{(K)}$ and $\Phi_s^{(F)}$, $\mu(\Phi_s^{(K)}, \Psi_s^{(K)})$ and $\mu(\Phi_s^{(F)}, \Psi_s^{(F)})$ for the fully decomposable CS and KCS model satisfy with probability at least $1 - 5e^{-\tau}$ that [24], [31], [47]

$$\frac{\mu(\Phi_s^{(F)}, \Psi_s)}{\mu(\Phi_s^{(K)}, \Psi_s)} \leq \sqrt{\frac{rN(N-1)}{1-r}} \cdot r_s C K \tau \log(\tau K \log N) \times (\log K)^2.$$

Here, C is a fixed constant. For simplicity, we denote $\Delta(K, \tau, N) = \sqrt{N(N-1)} \cdot K\tau \log(\tau K \log N)(\log K)^2$.

Similarly, we assume that fully decomposable CS select M_t column vectors from the $T \times T$ orthonormal bases $\tilde{\Phi}_t$ to

generate the temporal sensing matrix $\Phi_t^{(F)}$. Thus, the temporal mutual coherence for fully decomposable CS is derived as

$$\mu(\tilde{\Phi}_t, \Psi_t) = \max_{1\le i,j\le T} \left|\langle \tilde{\phi}_t^i, \psi_t^j\rangle\right| = \max_{1\le i,j\le T} \left|\sum_{1\le k\le T} \tilde{\phi}_t(i,k)\psi_t(k,j)\right|.$$

Given Ψ_t, we denote $\bar{\mu} = \max_{1\le i,j\le T} |\psi_t(i,j)|$ the mutual coherence $\mu(\mathbf{I}, \Psi_t)$ for KCS. When the elements of $\tilde{\Phi}_t$ obey a normal distribution having mean zero and variance $1/T$, we can find that $\langle \tilde{\phi}_t, \psi_t\rangle \sim \mathcal{N}(0, 1/T)$, as ψ_t is an orthonormal vector. According to [48], we can obtain that

$$Pr\left\{ \left|\langle \tilde{\phi}_t, \psi_t\rangle\right| \le \sqrt{\frac{1-r}{r}} \cdot \frac{\bar{\mu}}{r_s \Delta(K,\tau,N)} \right\} \ge 1 - 2\exp\left(-\frac{(T-2)(1-r)\bar{\mu}^2}{2r r_s^2 \triangle^2(K,\tau,N)}\right).$$

As a result, when T grows, fully decomposable CS would yield a smaller mutual coherence in comparison to KCS with a high probability asymptotically approaching 1.

Subsequently, we extend this result to the cases with $E \ge 2$. Considering that KCS adopts identity sampling matrix $\mathbf{I}$ for each dimension with $i = 2, \ldots, E$. We can obtain similar results to $E = 2$. Supposing that fully decomposable CS and KCS leverage sampling matrices $\Phi_1^{(F)}$ and $\Phi_1^{(K)}$ selected from the $N \times N$ orthonormal bases $\tilde{\Phi}_1$, we can find with probability at least $1 - 5e^{-\tau}$ that

$$\frac{\mu(\Phi_1^{(F)}, \Psi_1)}{\mu(\Phi_1^{(K)}, \Psi_1)} \le \sqrt{\frac{rN(N-1)}{1-r}} \cdot r_1 C K \tau \log(\tau K \log N) \times (\log K)^2,$$

where r_i is the sampling rate for the i-th dimension and the overall sampling rate $r = \prod_{i=1}^E r_i$. Without loss of generality, we suppose the elements of $T_i \times T_i$ orthonormal bases $\tilde{\Phi}_i$ for the i-th dimension, $i = 2, \ldots, E$, obey a normal distribution with zero mean and variance $1/T_i$. Thus, we can obtain that

$$Pr\left\{ \left|\langle \tilde{\phi}_i, \psi_i\rangle\right| \le \left[\sqrt{\frac{1-r}{r}} \cdot \frac{\bar{\mu}}{r_1 \Delta(K,\tau,N)}\right]^{\frac{1}{E-1}} \right\} \ge 1 - 2\exp\left(-(T_i - 2)\left[\frac{(1-r)\bar{\mu}^2}{2r r_1^2 \triangle^2(K,\tau,N)}\right]^{\frac{1}{E-1}}\right).$$

Therefore, provided arbitrary $E \in \mathbb{N}$, with the growth of T_i, $i = 2, \ldots, E$, fully decomposable CS would yield a smaller mutual coherence in comparison to KCS with a high probability asymptotically approaching 1.

Appendix B
Proof of Theorem 2

Considering $\mathbf{D}_i = \Phi_i \Psi_i$ for the i-th subspace $\mathcal{S}_i$, we rewrite the Kronecker product for the multidimensional signals $\mathbf{X}$ of E dimensions.

$$\text{vec}(\mathbf{Y}) = (\mathbf{D}_1 \otimes \cdots \otimes \mathbf{D}_E)\,\text{vec}(\mathbf{X}). \tag{23}$$

Since $(\mathbf{A} \otimes \mathbf{B}) \cdot \text{vec}(\mathbf{X}) = \text{vec}(\mathbf{Y})$ is equivalent to $\mathbf{AXB}^T = \mathbf{Y}$, $\mathbf{Y}$ can be obtained by multiplying $\mathbf{X}$ with $\mathbf{D}_1, \ldots, \mathbf{D}_E$ in an arbitrary order. Without loss of generality, the sequential order is $1, \ldots, E$ for proof. For the i-th subspace $\mathcal{S}_i$, we find for arbitrary $2K_i$-sparse vector $\mathbf{u}_i$ derived from the $\ell_{2,1}$ optimization that $C_1^i \|\mathbf{u}_i\|_2^2 \le \|\mathbf{D}_i \mathbf{u}_i\|_2^2 \le C_2^i \|\mathbf{u}_i\|_2^2$. For the E subspaces,

$$\prod_{i=1}^E C_1^i \|\mathbf{u}_i\|_2^2 \le \prod_{i=1}^E \|\mathbf{D}_i \mathbf{u}_i\|_2^2 \le \prod_{i=1}^E C_2^i \|\mathbf{u}_i\|_2^2. \tag{24}$$

Thus, there exists $\mathbf{u} = \mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_E$ satisfying Equation (11).

Appendix C
Proof of Theorem 3

When optimizing the Lagrangian cost function for sampling rate, the upper bound for recover error would be minimized. Given smooth wavelet bases Ψ_i, since sensing matrix optimization is separately performed for the E dimension, we can obtain for the i-th dimension with $i = 1, \ldots, E$ [24]

$$\|\mathbf{x}^{(i)} - \hat{\mathbf{x}}^{(i)}\|_2 \le C_i \left(\frac{M_i}{\sqrt{N_i} \cdot \mu(\Phi_i, \Psi_i)}\right)^{-\left(\frac{1}{2K_i} + \frac{1}{4}\right)}.$$

Here, we assume that the multidimensional signal $\mathbf{X}$ is K_i-sparse for the i-th subspace $\mathcal{S}_i$ in the fully decomposable CS model. Considering that $\|\mathbf{X} - \hat{\mathbf{X}}\|_2 = \otimes_{i=1}^E \|\mathbf{x}^{(i)} - \mathbf{x}^{(i)}\|_2$, we have

$$\|\mathbf{X} - \hat{\mathbf{X}}\|_2 \le \prod_{i=1}^E C_i \left(\frac{M_i}{\sqrt{N_i} \cdot \mu(\Phi_i, \Psi_i)}\right)^{-\left(\frac{1}{2K_i} + \frac{1}{4}\right)}.$$

Since the sparsity $K = \max(k_1, \ldots, k_E)$, we can obtain for constant $C = \prod_{i=1}^E C_i$,

$$\|\mathbf{X} - \hat{\mathbf{X}}\|_2 \le C \left(\frac{M}{\sqrt{N} \prod_{i=1}^E \mu(\Phi_i, \Psi_i)}\right)^{-\left(\frac{1}{2K} + \frac{1}{4}\right)}.$$

Considering mutual coherence for $\mathcal{S}_i$, we can find that

$$\mu(\Phi_i, \Psi_i) \ge \sqrt{\frac{N_i - M_i}{M_i(N_i - 1)}} = \sqrt{\frac{1 - r_i}{r_i(N_i - 1)}}.$$

Therefore, we can obtain that

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 \le C \left(N \prod_{i=1}^E r_i \cdot \sqrt{\frac{(N_i - 1)}{N_i} \cdot \frac{r_i}{1 - r_i}}\right)^{-2\sum_{i=1}^E \frac{1}{K_i}}. \tag{25}$$

When fixing the overall sampling rate by $r = \prod_{i=1}^E r_i$, there exists r_i, $i = 1, \ldots, E$ to minimize the recovery error.

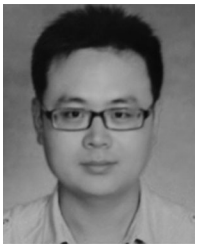
References

[1] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[3] J. Tropp and A. C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
[4] A. Cohen, W. Dahmen, and R. Devore, "Compressed sensing and best k-term approximation," *J. Amer. Math. Soc.*, vol. 22, no. 1, pp. 211–231, Jan. 2009.
[5] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Math.*, vol. 346, no. 9–10, pp. 589–592, May 2008.
[6] R. F. Marcia and R. M. Willett, "Compressive coded aperture video reconstruction," in *Proc. Eur. Signal Process. Conf.*, Lausanne, Switzerland, Aug. 2008.
[7] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, Apr. 2010.
[8] M. F. Duarte *et al.*, "Single pixel imaging via compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 83–91, Mar. 2008.
[9] T. Sun and K. F. Kelly, "Compressive sensing hyperspectral imager," in *Proc. Comput. Opt. Sens. Imag.*, San Jose, CA, USA, Oct. 2009, Paper CTuA.
[10] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Appl. Opt.*, vol. 47, no. 10, pp. B44–51, May 2008.
[11] M. Lustig, D. L. Donoho, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 72–82, Mar. 2008.
[12] G. Shi, D. Gao, X. Song, X. Xie, X. Chen, and D. Liu, "High-resolution imaging via moving random exposure and its simulation," *IEEE Trans. Image Process.*, vol. 20, no. 1, pp. 276–282, Jan. 2011.
[13] M. Wakin *et al.*, "Compressive imaging for video representation and coding," in *Proc. Picture Coding Symp.*, Beijing, China, Apr. 2006, pp. 1–6.
[14] J. Park and M. Wakin, "A multiscale framework for compressive sensing of video," in *Proc. Picture Coding Symp.*, Chicago, IL, USA, May. 2009, pp. 1–4.
[15] M. Cossalter, M. Tagliasacchi, G. Valenzise, and S. Tubaro, "Joint compressive video coding and analysis," *IEEE Trans. Multimedia*, vol. 12, no. 3, pp. 168–183, Apr. 2010.
[16] V. Stankovic, L. Stankovic, and S. Cheng, "Compressive video sampling," in *Proc. Eur. Signal Process. Conf.*, Lausanne, Switzerland, Aug. 2008, pp. 1–5.
[17] J. Prades-Nebot , Y. Ma, and T. Huang, "Distributed video coding using compressive sampling," in *Proc. Picture Coding Symp.*, Chicago, IL, USA, May. 2009, pp. 1–4.
[18] L. W. Kang and C. S. Lu, "Distributed compressed video sensing," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Taipei, Taiwan, Apr. 2009, pp. 1169–1172.
[19] J. Ma, G. Plonka, and M. Y. Hussaini, "Compressive video sampling with approximate message passing decoding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 9, pp. 1354–1364, Sep. 2012.
[20] Z. Liu, A. Y. Elezzabi, and H. V. Zhao, "Maximum frame rate video acquisition using adaptive compressed sensing," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 11, pp. 1704–1718, Nov. 2011.
[21] S. Mun and J. E. Fowler, "Residual reconstruction for block-based compressed sensing of video," in *Proc. IEEE Data Compression Conf.*, Snowbird, UT, USA, Mar. 2011, pp. 183–192.
[22] C. Chen, E. W. Tramel, and J. E. Fowler, "Compressed-sensing recovery of images and video using multihypothesis predictions," in *Proc. 45th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2011, pp. 1193–1198.
[23] Y. Liu, M. Li, and D. A. Pados, "Motion-aware decoding of compressed-sensed video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 3, pp. 438–444, Mar. 2013.
[24] M. F. Duarte and R. G. Baraniuk, "Kronecker compressive sensing," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 494–504, Feb. 2012.
[25] N. D. Sidiropoulos and A. Kyrillidis, "Multi-way compressed sensing for sparse low-rank tensors," *IEEE Signal Process. Lett.*, vol. 19, no. 11, pp. 757–760, Nov. 2012.
[26] S. Friedland, Q. Li, and D. Schonfeld, "Compressive sensing of sparse tensors," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4438–4447, Oct. 2014.
[27] M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5695–5702, Dec. 2007.
[28] Y. Weiss, H. S. Chang, and W. Freeman, "Learning compressed sensing," in *Proc. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Sep. 2007, vol. 1, pp. 535–541.
[29] J. M. Carvajalino and G. Sapiro, "Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1395–1408, Jul. 2009.
[30] J. Xu, Y. Pi, and Z. Cao, "Optimized projection matrix for compressive sensing," *EURASIP J. Adv. Signal Process.*, vol. 2010, no., Dec. 20101, Art. no. 560349.
[31] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Probl.*, vol. 23, no. 3, pp. 969–985, Apr. 2007.
[32] E. J. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
[33] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.
[34] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, vol. 24, no. 2, pp. 227–234, Apr. 1995.
[35] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.
[36] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6–18, Jan. 2006.
[37] D. L. Donoho, "For most large underdetermined systems of linear equations, the minimal l1-norm solution is also the sparsest solution," *Commun. Pure Appl. Math.*, vol. 59, no. 6, pp. 797–829, Jun. 2006.
[38] S. J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale l1 regularized least squares," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 606–617, Dec. 2007.
[39] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
[40] E. J. Candès, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
[41] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd ed. San Diego, CA: Academic, Dec. 2008.
[42] R. Hochmuth, "Wavelet characterizations for anisotropic Besov spaces," *Appl. Comput. Harmon. Anal.*, vol. 12, no. 2, pp. 179–208, Mar. 2002.
[43] M. F. Duarte and R. G. Baraniuk, "Kronecker product matrices for compressive sensing," in *Proc. IEEE Int Conf. Acoust., Speech, Signal Process.*, Princeton, NJ, USA, Mar. 2010, pp. 3650–3653.
[44] S. Jokar, "Sparse recovery and Kronecker products," in *Proc. 44th Annu. Conf. Inf. Sci. Syst.*, Princeton, NJ, USA, Mar. 2010.
[45] R. Robucci, J. D. Gray, L. K. Chiu. J. Romberg, and P. Hasler, "Compressive sensing on a CMOS separable-transform image sensor," *Proc. IEEE*, vol. 98, no. 6, pp. 1089–1101, May 2010.
[46] Y. August, C. Vachman, Y. Rivenson, and A. Stern, "Compressive hyperspectral imaging by random separable projections in both the spatial and the spectral domains," *Appl. Opt.*, vol. 52, no. 10, pp. D46-D54, Apr. 2013.
[47] L. Welch, "Lower bounds on the maximum cross correlation of signals," *IEEE Trans. Inf. Theory*, vol. 20, no. 3, pp. 397–399, May 1974.
[48] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845–2862, Jul. 2001.

Wenrui Dai (M'15) received the B.S., M.S., and Ph.D. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China in 2005, 2008, and 2014, respectively. He is currently a Postdoctoral Scholar with the Department of Biomedical Informatics, University of California, San Diego, San Diego, CA, USA. His research interests include learning-based image/video coding, image/signal processing, and predictive modeling.



Yong Li (S'15) received the B.S. degree from Xuzhou Normal University, Xuzhou, China, in 2009, and the M.S. degree from China University of Mining and Technology, Xuzhou, China, in 2012. He is currently working toward the Ph.D. degree with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China. From November 2014 to May 2015, November 2015 to May 2016, he was with the Department of Biomedical Informatics, University of California, San Diego, as a Visiting Scholar. His current research interests include compressive sensing, sparse representation, and image/signal processing.

Junni Zou (M'07) received the M.S. and the Ph.D. degrees in communication and information system from Shanghai University, Shanghai, China, in 2004 and 2006, respectively.

She is currently a Full Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. From 2006 to 2016, she was with the School of Communication and Information Engineering, Shanghai University, and became a Full Professor in 2014. From June 2011 to June 2012, she was with the Department of Electrical and Computer Engineering, University of California, San Diego, as a Visiting Professor. She has authored/co-authored more than 80 IEEE journal/conference papers, and 2 book chapters, including 16 IEEE Transactions journal papers. She holds 12 patents and has 10+ under reviewing patents. She has served on some technical program committees for the IEEE and other international conferences.

Her research interests include multimedia communication, network resource optimization, wireless communication, and network information theory

Dr. Zou was granted National Science Fund for Outstanding Young Scholar in 2016. She was a recipient of Shanghai Yong Rising Star Scientist Award in 2011, the First Prize of the Shanghai Technological Innovation Award in 2011, and the First Prize of the Shanghai Science and Technology Advancement Award in 2008.

Hongkai Xiong (M'01–SM'10) received the Ph.D. degree in communication and information system from Shanghai Jiao Tong University (SJTU), Shanghai, China, in 2003.

Since 2003, he has been with the Department of Electronic Engineering, SJTU, where he is currently a full Professor. From December 2007 to December 2008, he was a Research Scholar with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA. From 2011 to 2012, he was a Scientist with the Division of Biomedical Informatics, University of California, San Diego, San Diego, CA, USA. Since 2012, he has been a member of Innovative Research Groups, National Natural Science. His research interests include source coding/network information theory, signal processing, computer vision, and machine learning. He has published more than 190 refereed journal/conference papers.

Dr. Xiong was the recipient of the Top 10% Paper Award at the 2016 IEEE Visual Communication and Image Processing, the Best Student Paper Award at the 2014 IEEE Visual Communication and Image Processing, the Best Paper Award at the 2013 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, and the Top 10% Paper Award at the 2011 IEEE International Workshop on Multimedia Signal Processing. In 2016, he was granted Yangtze River Scholar Distinguished Professor from the Ministry of Education of China, and the Youth Science and Technology Innovation Leader from the Ministry of Science and Technology of China. He was also awarded Shanghai Academic Research Leader. In 2014, he was granted National Science Fund for Distinguished Young Scholar and Shanghai Youth Science and Technology Talent as well. In 2013, he was awarded Shanghai Shu Guang Scholar. In 2011, he was the recipient of the First Prize of the Shanghai Technological Innovation Award for Network-Oriented Video Processing and Dissemination: Theory and Technology. In 2010 and 2013, he was the recipient of the SMC-A Excellent Young Faculty Award of Shanghai Jiao Tong University. In 2009, he was awarded New Century Excellent Talents in University, Ministry of Education of China. He served as a TPC member for prestigious conferences, such as ACM Multimedia, ICIP, ICME, and ISCAS.

Yuan F. Zheng (F'97) received the undergraduate degree from Tsinghua University, Beijing, China, in 1970, and the M.S. and Ph.D. degrees in electrical engineering from The Ohio State University, Columbus, OH, USA, in 1980 and 1984, respectively.

From 1984 to 1989, he was with the Department of Electrical and Computer Engineering, Clemson University, Clemson, SC, USA. Since August 1989, he has been with The Ohio State University, where he is currently a Winbigler Designated Chair Professor, and was the Chairman of the Department of Electrical and Computer Engineering from 1993 to 2004. From 2004 to 2005, he spent a sabbatical year with Shanghai Jiao Tong University, Shanghai, China, and continued to be involved as the Dean of the School of Electronic, Information and Electrical Engineering until 2008. His research interests include two aspects. One is in wavelet transform for image and video, radar waveform and signal, and object tracking. The other is in robotics, which includes robotics for life science applications, multiple robots coordination, legged walking robots, and service robots. He has been on the editorial board of five international journals. He was the recipient of the Presidential Young Investigator Award from President Ronald Reagan in 1986, and the Research Awards from the College of Engineering, The Ohio State University, in 1993, 1997, and 2007, respectively, for his research contributions. He along with his students were the recipients of the Best Conference and Best Student Paper Award in 2000, 2002, 2006, 2009, and 2010, and the Fred Diamond for Best Technical Paper Award from the Air Force Research Laboratory, Rome, NY, USA, in 2006. In 2004 and 2005, he was appointed to the International Robotics Assessment Panel by the NSF, NASA, and NIH to assess the robotics technologies worldwide. In 2017, he was the recipient of the Innovator of the Year Award from The Ohio State University.