Spatiotemporal Maps for Dynamic MRI Reconstruction

Rodrigo A. Lobos, *Member, IEEE*, Xiaokai Wang, Rex T. L. Fung, Yongli He, David Frey, Dinank Gupta, Zhongming Liu, *Senior Member, IEEE*, Jeffrey A. Fessler, *Fellow, IEEE*, and Douglas C. Noll, *Fellow, IEEE*

***Abstract*—The partially separable functions (PSF) model is commonly adopted in dynamic MRI reconstruction, as is the underlying signal model in many reconstruction methods including the ones relying on low-rank assumptions. Even though the PSF model offers a parsimonious representation of the dynamic MRI signal in several applications, its representation capabilities tend to decrease in scenarios where voxels present different temporal/spectral characteristics at different spatial locations. In this work we account for this limitation by proposing a new model, called spatiotemporal maps (STMs), that leverages autoregressive properties of (k, t)-space. The STM model decomposes the spatiotemporal MRI signal into a sum of components, each one consisting of a product between a spatial function and a temporal function that depends on the spatial location. The proposed model can be interpreted as an extension of the PSF model whose temporal functions are independent of the spatial location. We show that spatiotemporal maps can be efficiently computed from autocalibration data by using advanced signal processing and randomized linear algebra techniques, enabling STMs to be used as part of many reconstruction frameworks for accelerated dynamic MRI. As proof-of-concept illustrations, we show that STMs can be used to reconstruct both 2D single-channel animal gastrointestinal MRI data and 3D multichannel human functional MRI data.**

***Index Terms*—Dynamic MRI reconstruction, partially separable functions, autoregression, gastrointestinal MRI, functional MRI.**

I. INTRODUCTION

DATA-ACQUISITION below the Nyquist rate is desired in dynamic MRI to achieve high spatial and temporal resolution. In the multichannel case, considering a Q-channel receiver array, sub-Nyquist (accelerated) data are often modeled as samples of the continuous spatial Fourier transform as:

$$d_q\left(\vec{k},t\right) = \int c_q\left(\vec{x}\right)\rho\left(\vec{x},t\right)e^{-i2\pi\vec{k}\cdot\vec{x}}\,\mathrm{d}\vec{x} + \eta\left(\vec{k},t\right), \tag{1}$$

where $d_q(\vec{k},t)$ denotes a complex-valued sample in the $(\vec{k},t)$-space from the qth coil receiver, $q \in \{1,\ldots Q\}$; $c_q(\vec{x})$ is a complex-valued sensitivity map for the qth coil receiver; each $\vec{k} \in \mathbb{R}^D$ and $t \in \mathbb{R}$ denotes a k-space location and time point, respectively, at which the complex-valued underlying dynamic signal $\rho(\vec{x},t)$ is sampled in the $(\vec{k},t)$-space; $D \in \{2,3\}$ is the dimension of the image; $\vec{x} \in \mathbb{R}^D$ denotes a spatial location; and $\eta(\vec{k},t)$ denotes the white complex Gaussian noise observed in the sample. In accelerated scans, data are collected during a finite number of time frames T, and M k-space samples are acquired in each time frame out of an ideal total of N samples needed for a reconstruction according to the Nyquist theorem. After acquiring the accelerated data, the goal is to reconstruct the image sequence $\rho(\vec{x},t)$ from the QMT $(\vec{k},t)$-space measurements.

Many reconstruction methods have been proposed to solve this inverse problem [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], including methods that use partially separable function (PSF) models [4]. PSF models approximate the spatiotemporal signal $\rho(\vec{x},t)$ as a sum of products, each one consisting of a temporal function $\varphi_l(t)$ and a spatial function $\rho_l(\vec{x})$:

$$\rho\left(\vec{x},t\right) \approx \sum_{l=1}^{L_{\text{PSF}}} \varphi_l(t)\rho_l\left(\vec{x}\right). \tag{2}$$

The PSF model provides a parsimonious decomposition for $\rho(\vec{x},t)$ in cases where temporal/spectral characteristics are shared across different spatial locations. Notably, in these cases it can be theoretically shown that a Casorati matrix constructed from the data has low-rank characteristics [4], [7] that have been leveraged by many dynamic MRI reconstruction methods based on low-rank models [7], [8], [9], [10], [12], [16], [20]. On the other hand, the number of components in the PSF decomposition must be large in cases where a high degree of variability is observed for the temporal/spectral characteristics across different spatial locations, decreasing the effectiveness of low-rank methods. One way to overcome this limitation is to divide the spatial locations into multiple patches and assume that the Casorati matrix for each patch has low-rank characteristics, which is the principle behind locally low-rank (LLR) methods [10], [20], [21], [24], [25], [26], [27], [28], [29], [30], [31], [32],

Received 17 July 2025; revised 12 December 2025; accepted 10 January 2026. Date of publication 26 January 2026; date of current version 6 February 2026. This work was supported by the National Institutes of Health through under Grant R21EB034344, Grant R01EB035618 and Grant R01AT011665. The associate editor coordinating the review of this article and approving it for publication was Prof. Jon Tamir. *(Corresponding author: Rodrigo A. Lobos.)*

Rodrigo A. Lobos, Xiaokai Wang, Rex T. L. Fung, David Frey, Dinank Gupta, Zhongming Liu, and Douglas C. Noll are with the Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: rlobos@umich.edu; xiaokaiw@umich.edu; rexfung@umich.edu; djfrey@umich.edu; dinankg@umich.edu; zmliu@umich.edu; dnoll@umich.edu).

Yongli He is with the Department of Applied Physics, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: yonglihe@umich.edu).

Jeffrey A. Fessler is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: fessler@umich.edu).

This article has supplementary downloadable material available at https://doi.org/10.1109/TCI.2026.3657289, provided by the authors.

Digital Object Identifier 10.1109/TCI.2026.3657289

[33]. However, implementing LLR methods involves further challenges: patches should overlap to avoid blocky artifacts, and the patch size should be chosen assuming that the underlying PSF decomposition for each patch is parsimonious [21], [25], which might not hold always. This work examines if it is possible to provide a parsimonious decomposition in a smaller scale than patches. More specifically, if different parsimonious decompositions can be provided for different spatial locations, such that temporal/spectral characteristics at different spatial locations could be captured separately.

We propose a signal model of the form

$$\rho\left(\vec{x}, t\right) \approx \sum_{l=1}^{L(\vec{x})} s_l(\vec{x}, t)\, \rho_l(\vec{x}), \tag{3}$$

where we call the temporal functions $s_l(\vec{x}, t), l = 1, \ldots, L(\vec{x})$ *spatiotemporal maps* (STMs) and the number of components can depend on the spatial location. We call the decomposition in (3) the STM model. The STM model has the potential to provide a more parsimonious representation than PSF models, as dynamic behaviors that substantially vary between spatial locations are modeled separately. This is equivalent to STM needing fewer components than PSF, which can be leveraged in reconstruction settings.

The first question that we theoretically explore in this work regards the scenarios in which the STM model, as presented in (3), is applicable. We show that the applicability of spatiotemporal maps is based on establishing the existence of shift-invariant linear-predictability (SILP) relationships in $(\vec{k}, t)$-space [34]. Moreover, we show that a sufficient condition for the existence of these SILP relationships, is that the time series for each spatial location should have a spectral support in the $(\vec{x}, f)$-space mostly limited to a finite set of frequency bands, i.e., such that $\rho(\vec{x}, t)$ could be modeled as a multiband signal [35], where the set of frequency bands is not restricted to be the same across spatial locations.

The second question that we address is how to efficiently calculate spatiotemporal maps from the acquired accelerated data. Recent work in subspace-based sensitivity map estimation in multichannel MRI [36], has shown that the existence of SILP relationships across the k-space of different channels is closely related to the estimation of sensitivity maps. By considering the different time frames as virtual channels in a parallel imaging setting, and showing that SILP relationships exist across the k-space of different time frames, we show how to calculate spatiotemporal maps using analogous principles to the ones used in subspace-based sensitivity map estimation. Notably, this allows us to use recently developed computational methods for the estimation of sensitivity maps to efficiently compute spatiotemporal maps from autocalibration data. Specifically, we draw inspiration from PISCO [36], [37], [38], a set of computational methods that has enabled fast sensitivity map estimation.

In many dynamic MRI applications the number of time frames is much larger than the number of coils in typical receiver arrays, so some of the original PISCO techniques are slower than in sensitivity map estimation. This issue is exacerbated in 3D data. One of the most demanding tasks in PISCO is calculating a singular value decomposition (SVD) of a big Hankel/Toeplitz convolutional matrix whose dimensions increase with the number of channels, which corresponds to the number of time frames in the dynamic MRI setting. Here we extend the PISCO method by adding a randomized linear algebra technique that calculates a proxy for this SVD using sketching [39]. This technique significantly reduces computation time, enabling STM calculation for long time series while negligibly affecting representation quality. One PISCO technique uses power iterations to estimate a single nullspace vector for a family of Hermitian matrices. Here we extend this technique using orthogonal iteration [40, p. 454] instead of a power iteration, because calculating spatiotemporal maps requires a nullspace basis for each of the aforementioned Hermitian matrices, rather than just one nullspace vector.

Finally, we show that the STM model in (3) can be easily incorporated into a reconstruction framework for accelerated dynamic MRI. Inspired by reconstruction methods based on the PSF model [9], our strategy consists of computing the spatial functions in the STM model, i.e., $\{\rho_l(\vec{x})\}_{l=1}^{L(\vec{x})}$, from undersampled $(\vec{k}, t)$-space, assuming that the number of components $L(\vec{x})$ per voxel is much smaller than the number of frames. Using spatiotemporal maps that we precompute from autocalibration data, the final reconstructed time series uses the model (3). This approach can be complemented by adding regularizers for the spatial functions in cases where the data is heavily undersampled. We explore the combination of the STM model with several regularizers, and we evaluate this approach in two challenging dynamic MRI datasets: 2D single-channel animal gastrointestinal MRI data and 3D multichannel human functional MRI (fMRI) data. A preliminary version of this work was presented as a short abstract [41].

The formulation of the STM model in (3) resembles the one used in the blind compressed sensing (BCS) method proposed in [11]. However, the principles guiding the learning of the voxel-dependent temporal functions are distinct. In BCS, a dictionary of temporal functions is learned from the undersampled $(\vec{k}, t)$-space data using an energy constraint based on the Frobenius norm. Then, $\rho(\vec{x}, t)$ is modeled as a linear combination of the dictionary temporal functions, with voxel-dependent coefficients chosen via sparsity constraints. In the STM model, the temporal functions (i.e., the spatiotemporal maps) are learned without constructing a dictionary. Instead, the STMs are learned independently for each voxel by applying a subspace-based method to autocalibration data, that rely on SILP relationships in $(\vec{k}, t)$-space.

This paper is organized as follows. Section II presents the theory behind the applicability of spatiotemporal maps for dynamic MRI reconstruction. Section III reviews the computational methods used to compute spatiotemporal maps, and introduces novel techniques that extend PISCO. Section IV provides a reconstruction framework for accelerated dynamic MRI using spatiotemporal maps. Section V describes the experiments that evaluate the proposed STM model and reconstruction framework, and Section VI provides the corresponding results. Lastly, Section VII provides discussion and final remarks.

II. THEORY

This section establishes a theoretical framework for the applicability of spatiotemporal maps. We observe that calculating spatiotemporal maps is closely related to estimating sensitivity maps in parallel imaging. In fact, by setting $L(\vec{x}) = 1$ in (3) and considering time frames as virtual channels, the STM model recovers the commonly used parallel imaging model.[1] In this case, spatiotemporal maps become sensitivity maps. We extend the theoretical framework proposed in [36] for subspace-based sensitivity map estimation to calculating spatiotemporal maps. This theoretical framework starts by assuming that SILP relationships exist across channels in a parallel imaging scenario [34], [36], [37]. Therefore, here we show that SILP relationships can also exist across time frames (for both single coil and multichannel MRI) in the dynamic MRI setting. For this purpose, we first formally introduce the concept of SILP relationships in $(\vec{k}, t)$-space and study sufficient conditions for their existence.

A. *Shift-Invariant Linear Predictability Relationships in $(\vec{k}, t)$-Space*

We assume that the spatial support of $\rho(\vec{x}, t)$ is finite for all time, i.e., that the set of spatial locations $\Omega \triangleq \{\vec{x} \in \mathbb{R}^D : \exists t \in \mathbb{R},\, |\rho(\vec{x}, t)| > 0\}$ is completely contained within a hypercube Γ, corresponding to the field-of-view (FOV). We normalize the coordinates so that each FOV width is unity. In other words, the object being imaged remains within the FOV over time. Analogously, after normalizing the temporal frequency range, we assume that $\rho(\vec{x}, t)$ is approximately bandlimited to the set $\mathcal{F} = [-\frac{1}{2}, \frac{1}{2}]$ for all $\vec{x} \in \Gamma$. In other words, $\check{\rho}(\vec{x}, f) \approx 0, \forall f \notin \mathcal{F}$, where $\check{\rho}(\vec{x}, f)$ denotes the 1D temporal Fourier transform [2] of $\rho(\vec{x}, t)$:

$$\check{\rho}\left(\vec{x}, f\right) \triangleq \int \rho\left(\vec{x}, t\right)\, e^{-i2\pi f t}\, \mathrm{d}t\,. \tag{4}$$

These assumptions imply the following Fourier series representations:

$$\check{\rho}(\vec{x}, f) \approx \mathbb{1}_{\mathcal{F}}(f) \sum_{t \in \mathbb{Z}} \rho(\vec{x}, t)\, e^{-i2\pi f t}, \tag{5}$$

$$\rho(\vec{x}, t) = \mathbb{1}_{\Gamma}(\vec{x}) \sum_{\vec{k} \in \mathbb{Z}^D} \tilde{\rho}\left(\vec{k}, t\right)\, e^{i2\pi \vec{k} \cdot \vec{x}}, \tag{6}$$

where $\rho(\vec{x}, t)$ in (5) denotes temporal samples of the dynamic MRI signal according to the Nyquist rate based on $\mathcal{F}$; $\tilde{\rho}(\vec{k}, t)$ denotes samples of the spatial Fourier transform of $\rho(\vec{x}, t)$, located on a rectilinear Nyquist grid with k-space sample spacing given by a sampling period $\Delta k = 1$ (in each dimension), which follows from our FOV assumption; and $\mathbb{1}_{\mathcal{U}}(u)$ is the indicator function of the set $\mathcal{U}$ which is equal to 1 if $u \in \mathcal{U}$ or equal to 0 otherwise.

In real applications, we only have access to samples in the $(\vec{k}, t)$-space for a finite number of time frames T. We index the time samples using the set $\{1, \ldots, T\}$.

An SILP relationship across the $(\vec{k}, t)$-space of the T available time frames is equivalent to the existence of a nonzero multi-frame finite impulse response (FIR) filter $\tilde{h}(\vec{k}, t)$ that satisfies:

$$\sum_{t=1}^{T} \sum_{\vec{l} \in \Lambda} \tilde{h}\left(\vec{l}, t\right)\, \tilde{\rho}\left(\vec{k} - \vec{l}, t\right) \approx 0, \quad \forall \vec{k} \in \mathbb{Z}^D, \tag{7}$$

where $\Lambda \subset \mathbb{Z}^D$ is the finite support set of each $\tilde{h}(\vec{k}, t)$, $t \in \{1, \ldots, T\}$, in k-space. If (7) holds, then there is an annihilation relationship among the k-space samples of different time frames that is shift-invariant [34]. The following subsection provides sufficient conditions for the existence of filters satisfying (7). These conditions are based on the assumptions that $\rho(\vec{x}, t)$ has a limited support in the $(\vec{x}, f)$-space located in a small set of disjoint bands, and that its spectral characteristics smoothly vary across the FOV. These assumptions provide sufficient conditions for the existence of SILP relationships as in (7); however, we are not providing necessary conditions. There might be other cases where SILP relationships could be established.

B. *Sufficient Conditions for the Existence of SILP Relationships in the $(\vec{k}, t)$-Space*

So far we have assumed that the frequency support of $\check{\rho}(\vec{x}, f)$ is restricted to the same band $\mathcal{F} = [-\frac{1}{2}, \frac{1}{2}]$ for every spatial location in the FOV. However, it is possible to provide more refined modeling assumptions when each spatial location is considered separately, as this enables more parsimonious models where $\rho(\vec{x}, t)$ exhibits distinct behavior across spatial locations in the FOV. Hereafter we assume that for every $\vec{x} \in \Gamma$, $\check{\rho}(\vec{x}, f)$ is further restricted to be (approximately) supported on no more than $J(\vec{x}) \in \mathbb{N}$ disjoint intervals (bands) in $\mathcal{F}$, which is equivalent to assuming that $\rho(\vec{x}, t)$ is a multiband signal [35]. We denote the location-dependent union of these bands by $\mathcal{B}(\vec{x}) \subset \mathcal{F}$. We further assume that the set of bands smoothly varies across spatial locations in the FOV, akin to locally low-rank models that assume $\rho(\vec{x}, t)$ exhibits similar behaviors for nearby spatial locations.

Under these support conditions, one can define a nonzero function $\check{b}(\vec{x}, f)$ that is spatially smooth, bandlimited to $\mathcal{F}$, and that for each $\vec{x} \in \Gamma$ its frequency support is mostly defined on $\mathcal{F} \setminus \mathcal{B}(\vec{x})$, where "$\setminus$" denotes the set difference. Such a function satisfies:

$$\check{b}\left(\vec{x}, f\right)\, \check{\rho}\left(\vec{x}, f\right) \approx 0, \quad \forall f \in \mathcal{F}, \forall \vec{x} \in \Gamma. \tag{8}$$

Under our on-going assumptions there are many functions that satisfy (8). This approximation could be arbitrarily close to zero if we allow $\check{b}(\vec{x}, f)$ to have arbitrary rapid variations in frequency in cases where the number of bands $J(\vec{x})$ is relatively large; however, such variations would complicate computing $\check{b}(\vec{x}, f)$ in practice. Thus, we restrict our analysis to annihilator functions (i.e., functions that satisfy (8)) that are smooth in f, i.e., whose temporal support is concentrated in a small time interval.

---

[1] By setting $L(\vec{x}) > 1$ it is possible to recover the case when more than one set of sensitivity maps is needed, for example, when aliasing exists due to a small field-of-view [36], [42].

[2] We use $\check{\gamma}(\vec{x}, f)$ and $\tilde{\gamma}(\vec{k}, t)$ to denote the temporal Fourier transform and the spatial Fourier transform of the function $\gamma(\vec{x}, t)$, respectively. The variables $\vec{k}$ and t can be continuous or discrete, as should be clear from the context.

Specifically, we require that the 1D inverse discrete time Fourier transform (IDTFT) of $\check{b}(\vec{x},f)$ given by

$$b\left(\vec{x},t\right) \triangleq \frac{1}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} \check{b}\left(\vec{x},f\right)\, e^{i2\pi f t}\,\mathrm{d}f, \quad t\in\mathbb{Z}, \tag{9}$$

is approximately zero when $|t-\lfloor T/2\rfloor-1| > \lfloor T_b(\vec{x})/2\rfloor$, where $T_b(\vec{x})\in\mathbb{N}$ can depend on the spatial location and $\lfloor\cdot\rfloor$ denotes the flooring operation. For simplicity, hereafter we assume that $b(\vec{x},t)\approx 0$ when $|t-\lfloor T/2\rfloor-1| > \lfloor T_b/2\rfloor$, where $T_b \triangleq \max_{\vec{x}\in\Gamma} T_b(\vec{x})$. It follows from (8) and the DTFT convolutional theorem that

$$\sum_{t\in\mathbb{Z}} \rho(\vec{x},t)\, b(\vec{x},\tau-t) \approx 0, \quad \forall\tau\in\mathbb{Z},\ \forall\vec{x}\in\Gamma. \tag{10}$$

If T is large enough such that

$$\check{\rho}(\vec{x},f) \approx \mathbb{1}_{\mathcal{F}}(f)\sum_{t=1}^{T}\rho(\vec{x},t)\, e^{-i2\pi f t}, \tag{11}$$

then (10) implies that

$$\sum_{t=1}^{T} h_\tau(\vec{x},t)\,\rho(\vec{x},t)\approx 0,\ \forall\vec{x}\in\Gamma, \tag{12}$$

for $\tau\in\{-\lfloor T_b/2\rfloor+1,\ldots,T+\lfloor T_b/2\rfloor\}$, where we define $h_\tau(\vec{x},t)\triangleq b(\vec{x},\tau-t)$ and for simplicity we assume T and T_b are both odd. Next, we apply the Fourier transform convolution theorem on the spatial domain to obtain that

$$\sum_{t=1}^{T}\sum_{\vec{l}\in\Lambda}\tilde{h}_\tau\left(\vec{l},t\right)\,\tilde{\rho}\left(\vec{k}-\vec{l},t\right)\approx 0, \quad \forall\vec{k}\in\mathbb{Z}^D, \tag{13}$$

$\tau\in\{-\lfloor T_b/2\rfloor+1,\ldots,T+\lfloor T_b/2\rfloor\}$, where $\tilde{h}_\tau(\vec{k},t)$, the spatial discrete Fourier transform of $h_\tau(\vec{x},t)$, has its support Λ restricted by our assumption on the spatial smoothness of $\check{b}(\vec{x},f)$. Finally, each function $\tilde{h}_\tau(\vec{k},t)$ corresponds to a multiframe FIR filter satisfying (7).

The previous analysis shows that many FIR filters satisfying (7) can be found for one function $\check{b}(\vec{x},f)$; however, infinitely many other functions can be defined with the same characteristics of $\check{b}(\vec{x},f)$. This property suggests that multiple SILP relationships should exist in $(\vec{k},t)$-space across time frames when $\rho(\vec{x},t)$ follows the multiband model at each spatial location. Our theoretical analysis is inspired by similar observations in previous work [34], [37], [43].

Our previous analysis assumes that the set of bands where $\check{\rho}(\vec{x},f)$ is nonzero varies smoothly across the FOV. However, this condition may not hold in real applications. If $\check{\rho}(\vec{x},f)$ is not spatially smooth, then $\check{b}(\vec{x},f)$ should be defined accordingly to ensure that the approximation in (8) remains valid. Allowing $\check{b}(\vec{x},f)$ to vary rapidly across the FOV can improve the approximation in (8), and enable multiframe FIR filters to achieve a better approximation in (13). However, these filters would require a larger support Λ, making their computation more challenging in practice (*cf.*, Section III-A). The next subsection shows that choosing the filter support size involves a trade-off between computational complexity and representation accuracy of the STM model.

The following subsection first shows how SILP relationships connect with the applicability of spatiotemporal maps.

C. SILP Relationships in $(\vec{k},t)$-Space and the Applicability of Spatiotemporal Maps

Recent work in multichannel MRI has shown that the existence of SILP relationships across channels in a parallel imaging setting is closely related to the estimation of sensitivity maps using subspaces [36]. Given that SILP relationships can also exist across time frames in $(\vec{k},t)$-space, this section shows how to extend the theoretical framework in [36] to STM calculation.

Section II-B showed that many multiframe FIR filters exist that satisfy (7) by assuming a multiband behavior for $\rho(\vec{x},t)$ at each spatial location. Hereafter we assume that P of these filters are available, denoted by $\tilde{h}_p(\vec{k},t), p\in\{1,\ldots,P\}$. Therefore, we have P equations following the same structure as (7). Using analogous derivations to the ones in [36], which involve the Fourier transform convolutional theorem, we equivalently express these P equations in the spatial domain using matrix-vector multiplications. By using the spatial representation

$$h_p\left(\vec{x},t\right)\triangleq\sum_{\vec{k}\in\Lambda}\tilde{h}_p\left(\vec{k},t\right)\, e^{i2\pi\vec{k}\cdot\vec{x}}, \quad t\in\{1,\ldots,T\}, \tag{14}$$

it follows that

$$\boldsymbol{H}(\vec{x})\,\boldsymbol{\rho}(\vec{x})\approx\boldsymbol{0}, \quad \forall\vec{x}\in\Gamma, \tag{15}$$

where $\boldsymbol{H}(\vec{x})\in\mathbb{C}^{P\times T}$ has its (p,t)th entry defined as $[\boldsymbol{H}(\vec{x})]_{pt}\triangleq h_p(\vec{x},t)$, and $\boldsymbol{\rho}(\vec{x})\in\mathbb{C}^T$ has elements $[\boldsymbol{\rho}(\vec{x})]_t\triangleq\rho(\vec{x},t)$. In words, (15) says that $\boldsymbol{\rho}(\vec{x})$ is an approximate nullspace vector of $\boldsymbol{H}(\vec{x})$. Because the number of filters P tends to be much larger than the number of frames T, it is more convenient to use the nullspace relationship

$$\boldsymbol{G}(\vec{x})\,\boldsymbol{\rho}(\vec{x})\approx\boldsymbol{0}, \quad \forall\vec{x}\in\Gamma, \tag{16}$$

where $\boldsymbol{G}(\vec{x})\triangleq\boldsymbol{H}^H(\vec{x})\boldsymbol{H}(\vec{x})\in\mathbb{C}^{T\times T}$, whose (t',t)th entry is given by

$$[\boldsymbol{G}(\vec{x})]_{t't}=\sum_{p=1}^{P}h_p^*(\vec{x},t')\,h_p(\vec{x},t), \tag{17}$$

where $(\cdot)^*$ denotes the complex conjugate operation. In the k-space domain, (16) is equivalent to reducing the initial P SILP equations of the form (7) to T SILP equations of the form

$$\sum_{t=1}^{T}\sum_{\vec{l}\in\Theta}\tilde{g}_{t'}\left(\vec{l},t\right)\,\tilde{\rho}\left(\vec{k}-\vec{l},t\right)\approx 0, \quad \forall\vec{k}\in\mathbb{Z}^D, \tag{18}$$

$t'\in\{1,\ldots,T\}$, where

$$\tilde{g}_{t'}\left(\vec{k},t\right)\triangleq\sum_{p=1}^{P}\tilde{h}_p^*\left(-\vec{k},t'\right)*\tilde{h}_p\left(\vec{k},t\right); \tag{19}$$

here $*$ denotes k-space domain convolution; and Θ is the finite support of the multiframe filters $\{\tilde{g}_{t'}(\vec{k},t)\}_{t'=1}^{T}$.

The annihilation property (16) suggests that $\boldsymbol{\rho}(\vec{x})$ could be approximated by a scaled version of the eigenvector of $\boldsymbol{G}(\vec{x})$ with the smallest eigenvalue, or by some linear combination of

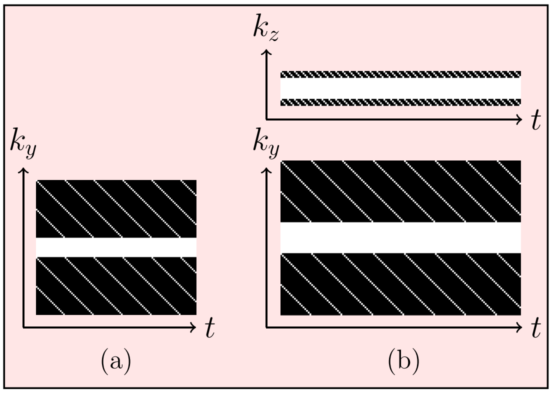
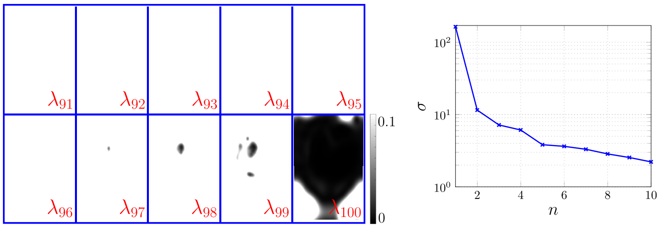


Fig. 1. (a) k-space mask used to retrospectively undersample 2D+T dataset (A.1). White indicates the PE lines considered in each time frame. (b) Analogous k-space mask used to retrospectively undersample 3D+T dataset (B). In this case two PE directions are considered.



(a) (b)

Fig. 2. (a) Eigenvalue maps representing the last 10 eigenvalues (after normalizing and sorting in decreasing order) of the matrices $\boldsymbol{G}(\vec{x})$ for each spatial location in the FOV using dataset (A.1). (b) First 10 singular values of the Casorati matrix for dataset (A.1).

a set of eigenvectors whose corresponding eigenvalues are all zero or nearly zero. We call such eigenvectors the "nullspace" of $\boldsymbol{G}(\vec{x})$, though in practice the smallest eigenvalues might not be exactly zero so "very small space" might be a more apt name. In other words, (16) suggests that we can represent $\boldsymbol{\rho}(\vec{x})$ (approximately) using a basis for the nullspace of $\boldsymbol{G}(\vec{x})$. Therefore, as detailed in Section III, we compute spatiotemporal maps by first computing an orthonormal basis for the nullspace of $\boldsymbol{G}(\vec{x})$.

For each $\vec{x}$, let $\boldsymbol{s}_l(\vec{x}) \in \mathbb{C}^T, l \in \{1, \ldots, L(\vec{x})\}$, denote orthonormal vectors composing a basis for the nullspace of $\boldsymbol{G}(\vec{x})$, where $L(\vec{x})$ is the nullspace dimension. Then we propose the following signal model:

$$\boldsymbol{\rho}(\vec{x}) \approx \sum_{l=1}^{L(\vec{x})} \boldsymbol{s}_l(\vec{x}) \rho_l(\vec{x}), \quad \forall \vec{x} \in \Gamma, \tag{20}$$

where $\{\rho_l(\vec{x})\}_{l=1}^{L(\vec{x})} \subset \mathbb{C}$ are scalar coefficients that must be computed by an image reconstruction algorithm. We define spatiotemporal maps (STMs) as follows:

$$s_l(\vec{x}, t) \triangleq [\boldsymbol{s}_l(\vec{x})]_t, \quad \forall \vec{x} \in \Gamma, \tag{21}$$

for $l \in \{1, \ldots, L(\vec{x})\}$ and $t \in \{1, \ldots, T\}$.

The signal decomposition provided in (20) has all its elements depending on the spatial location, including the number of components $L(\vec{x})$. We selected this parameter by studying the eigenvalues of the $\boldsymbol{G}(\vec{x})$ matrices (*cf.*, Fig. 2). For each spatial location $L(\vec{x})$ can be the number of eigenvalues below a user-selected threshold. However, for simplicity, for the empirical results in Section V we used the same number of components for each spatial location. Specifically, we set the number of components for each spatial location to $L \triangleq \max_{\vec{x} \in \Gamma} L(\vec{x})$, $\forall \vec{x} \in \Gamma$, and we relied on the scalar coefficients $\rho_l(\vec{x})$ to be zero in cases where the number of components was overestimated. Having the number of components depend on the spatial location could help further reduce degrees of freedom in future work.

The accuracy of the approximation in (20) depends primarily on how well the basis for the (approximate) nullspace of $\boldsymbol{H}(\vec{x})$ can be computed. Additionally, the validity of the nullspace relationship in (15) plays a significant role. This latter relationship, in turn, heavily depends on how closely the multiframe FIR filters satisfy the SILP relationships in (7). If these conditions are not met–meaning the approximation in (7) deviates significantly from zero–then the approximation in (20) will also be less accurate, ultimately decreasing the representation capabilities of the STM model. Therefore, accurately computing multiframe FIR filters is an important step in calculating STMs.

Our derivation assumes that P multiframe FIR filters exist that satisfy (7). Our theoretical analysis in Section II-B indicates that the accuracy of the approximation in (7) may depend on the size of the support Λ. Specifically, increasing the size of the support Λ can lead to a better approximation in (7), which in turn improves the nullspace approximation in (15) throughout the FOV. However, a larger support Λ can also allow for more rapid variations in the spatial domain of the multiframe FIR filters. As a consequence, the matrix $\boldsymbol{H}(\vec{x})$ may change more rapidly from voxel to voxel. This leads to the nullspace of each matrix $\boldsymbol{H}(\vec{x})$ varying more quickly across the FOV, resulting in more rapid changes in their bases. Consequently, the resulting STMs are less smooth. Additionally, as we show in the following section, computing multiframe FIR filters with a larger support increases computational complexity of calculating STMs (*cf.*, Section III-A).

III. COMPUTING SPATIOTEMPORAL MAPS

This section summarizes the practical steps for computing spatiotemporal maps efficiently from autocalibration k-space data. Each of these steps relies on PISCO [36], a set of computational methods originally proposed for efficient subspace-based sensitivity map estimation in multichannel MRI. We show that PISCO can be adapted for computing spatiotemporal maps due to the theoretical connections with sensitivity map estimation shown in previous sections. For simplicity, we first present the steps for computing STMs for the single-channel case, leaving the extension to the multichannel case to the end of this section.

A. Step 1: Calculating Multiframe FIR Filters

Assuming that SILP relationships exist in $(\vec{k}, t)$-space, the first step in computing spatiotemporal maps corresponds to calculating P multiframe FIR filters satisfying (7). Let $\tilde{\boldsymbol{h}}_p \in \mathbb{C}^{|\Lambda| T}$

denote the vectorized version of one of these multiframe filters. One can rewrite (7) in matrix-vector form to show that each $\tilde{\boldsymbol{h}}_p$ is an approximate nullspace vector of the matrix [44]

$$\boldsymbol{C} \triangleq [\boldsymbol{C}_1 \quad \boldsymbol{C}_2 \quad \cdots \quad \boldsymbol{C}_T] \in \mathbb{C}^{I \times |\Lambda| T}, \tag{22}$$

where $\boldsymbol{C}_t \in \mathbb{C}^{I \times |\Lambda|}$ is a matrix with a Hankel/Toeplitz convolutional structure constructed from $\tilde{\rho}(\vec{k}, t)$, i.e., from k-space samples from the tth time frame, with each of its I rows corresponding to a vectorized neighborhood of k-samples of size $|\Lambda|$ where (7) holds [34], [36], [43], [44]. The type of linear relationship described in (7) is shift-invariant in k-space; consequently, any neighborhood with size Λ available in k-space could be used to form a row of $\boldsymbol{C}$. In practice, we construct $\boldsymbol{C}$ from autocalibration data that is Nyquist-sampled and has sufficiently large enough dimensions to allow the extraction of k-space neighborhoods of size Λ from it. We consider autocalibration data from the center of k-space as these samples commonly exhibit higher signal-to-noise ratio than samples from outer regions. Then, we use an SVD to compute an orthonormal basis for its (approximate) nullspace. This basis, denoted $\boldsymbol{N} \in \mathbb{C}^{|\Lambda| T \times P}$, is expected to generate the subspace of multiframe filters with support Λ that satisfy (7). Two options for the FIR filter support Λ were studied in [36] when computing $\boldsymbol{C}$: an ellipsoidal shape i.e., $\Lambda = \{\vec{k} \in \mathbb{Z}^D : \|\vec{k}\|_2 \leq R\}$, and a rectangular shape i.e., $\Lambda = \{\vec{k} \in \mathbb{Z}^D : \|\vec{k}\|_\infty \leq R\}$, where $\|\cdot\|_2$ and $\|\cdot\|_\infty$ denote the ℓ_2 and infinity norms, respectively. The experiments in Section V used FIR filters with an ellipsoidal shape that saves computation without sacrificing representation quality compared to using a rectangular shape [36], [45]. This choice is particularly relevant when $D = 3$.

As discussed in previous sections, using multiframe FIR filters with a larger support Λ can improve the representation accuracy of the STM model; particularly, when the data characteristics in the $(\vec{x}, f)$-space vary non-smoothly across the FOV. This suggests that choosing a larger value for R may be beneficial. However, increasing R also enlarges the dimensions of $\boldsymbol{C}$, which raises memory usage and computation time, especially when using SVD to compute an orthonormal basis for its nullspace. Therefore, the selection of R poses a trade-off between representation capabilities of the STM model and computational complexity. Furthermore, $\boldsymbol{C}$ is computed using k-space neighborhoods of size $|\Lambda|$ extracted from autocalibration data. If R is set too large relative to the size of the available autocalibration data, only a limited number of k-space neighborhoods could be extracted. This restriction reduces the number of rows in $\boldsymbol{C}$, which may compromise the accuracy with which its nullspace can be computed.

The dimensions of $\boldsymbol{C}$ tend to be undesirably large, which affects memory usage and computation time when computing a basis for its nullspace using SVD. In many applications $I > |\Lambda| T$, so it is preferable to work with $\boldsymbol{C}^H \boldsymbol{C}$ instead of $\boldsymbol{C}$, as they share the same nullspace. PISCO provides FFT-based computational methods to efficiently calculate an approximation of $\boldsymbol{C}^H \boldsymbol{C}$ without calculating $\boldsymbol{C}$ first, by leveraging the convolutional structure of $\boldsymbol{C}$; for details see [36, Section IV.A].

B. Step 2: Calculating $\boldsymbol{G}(\vec{x})$ for Each Spatial Location

After computing P multiframe FIR filters from $\boldsymbol{C}$, the next step is to calculate $\boldsymbol{G}(\vec{x})$ for each $\vec{x} \in \Gamma$. Evaluating (14) for each $\vec{x} \in \Gamma$ and then computing $\boldsymbol{G}(\vec{x}) = \boldsymbol{H}(\vec{x})^H \boldsymbol{H}(\vec{x})$ would be computationally expensive when the FOV is big. Fortunately, one of the computational methods provided in PISCO can directly calculate $\boldsymbol{G}(\vec{x})$ using an FFT-based approach without forming $\boldsymbol{H}(\vec{x})$ first; see [36, Section IV.C] for details. In general terms, this method expresses the entries of $\boldsymbol{G}(\vec{x})$ using the k-space representation of the P FIR multiframe filters. Specifically, this equivalent expression uses the entries of the matrix $\boldsymbol{W} \triangleq \sum_{p=1}^{P} \tilde{\boldsymbol{h}}_p \tilde{\boldsymbol{h}}_p^H$. Because step 1) computes the filters $\{\tilde{\boldsymbol{h}}_p\}_{p=1}^P$ as nullspace vectors of the $\boldsymbol{C}$ matrix and correspond to the columns of the matrix $\boldsymbol{N}$, it follows that $\boldsymbol{W} = \boldsymbol{N}\boldsymbol{N}^H$. Therefore, we solely need $\boldsymbol{N}$ from step 1) to calculate $\boldsymbol{G}(\vec{x})$ for each $\vec{x} \in \Gamma$.

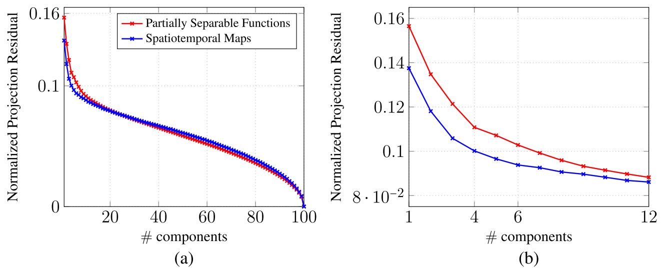
C. Step 3: Construction of Spatiotemporal Maps

The last step is to construct the spatiotemporal maps using (21) by calculating a basis for the nullspace of $\boldsymbol{G}(\vec{x})$ for each $\vec{x} \in \Gamma$. A natural approach would be to use an SVD for each matrix $\boldsymbol{G}(\vec{x})$, which would be computationally expensive. When the nullspace dimension of $\boldsymbol{G}(\vec{x})$ is equal to one – which is usually the case found in subspace-based sensitivity map estimation – PISCO provides a method based on power iteration to compute the nullspace vectors for all the spatial locations simultaneously [36, Section IV.E]. For dynamic imaging, we expect to need $L(\vec{x}) > 1$, i.e., STMs require nullspace bases with more than one vector. Thus, here we used the orthogonal iteration [40, p. 454] [42] instead of the power iteration. This approach computes nullspace bases for all spatial locations simultaneously for a general number of basis vectors (i.e., the number of components L).

The aforementioned steps and computational techniques enable efficient STM computation in many cases. However, computational efficiency can be substantially improved by assuming that STMs are spatially smooth. Under this assumption, we can compute spatiotemporal maps on a grid with coarser resolution than the original and then interpolate to the desired resolution. PISCO also provides an FFT-based method for this procedure that we directly use to compute STMs; see [36, Section IV.D] for details. Nevertheless, the degree of interpolation should be properly selected, as smoothness of STMs can be reduced when choosing a large value for R, as previously discussed.

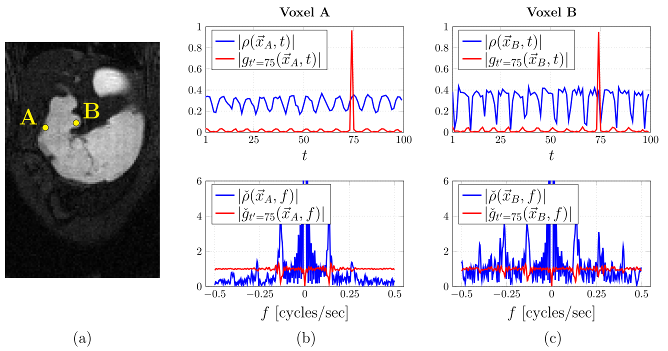
D. Sketched SVD for the Projection Onto the Nullspace of $\boldsymbol{C}$

Computing STMs is particularly challenging when $D = 3$ and T is large, because finding $\boldsymbol{N}$ in step 1) involves an SVD of $\boldsymbol{C}^H \boldsymbol{C} \in \mathbb{C}^{|\Lambda| T \times |\Lambda| T}$, whose dimensions become inconveniently large in such cases. When $D = 3$, using multiframe FIR filters with an ellipsoidal shape support reduces $|\Lambda|$ considerably compared to using a rectangular shape support [45]. However, these savings are still insufficient when T is large, i.e., for long time series (*e.g.*, fMRI). Given that $\boldsymbol{N}$ is used solely to calculate $\boldsymbol{N}\boldsymbol{N}^H$ in step 2), here we propose a novel technique based



(a) (b)

Fig. 3. Normalized projection residual versus the number of components of the PSF and STM models. (b) is a zoom in of (a).



(a) (b) (c)

Fig. 4. Visualization of the dynamic behavior of two voxels using dataset (A.1). (a) Approximate locations of the two voxels overlaid with an image corresponding to the first time frame. (b) Time evolutions of $\rho(\vec{x}, t)$ and $g_{t'}(\vec{x}, t)$ for voxel A (top), and their respective frequency spectra (bottom). The frequency range is shown after normalization in cycles/sec. (c) Analogous results for voxel B.

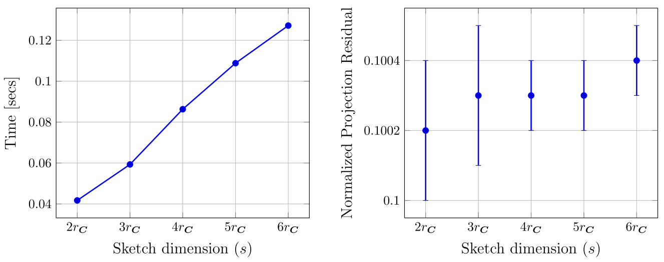


Fig. 5. (a) Computation time of the sketched SVD versus sketch dimension for dataset (A.1). The median over 50 realizations is reported in each case. (b) Mean and standard deviations of the NPR obtained in each case.

on randomized linear algebra to efficiently approximate $\boldsymbol{N}\boldsymbol{N}^H$ without calculating the SVD of $\boldsymbol{C}^H\boldsymbol{C}$.

Our approach uses a sketched SVD [39], a randomized linear algebra method, to efficiently calculate approximations for the singular values and singular vectors of a large matrix. Applied to our case, we propose to use a sketched SVD of the matrix $\boldsymbol{C}^H\boldsymbol{C}$. Sketched SVD uses a random matrix that satisfies the Johnson-Lindenstrauss property [46, p. 110], called the sketch matrix. We denote this matrix by $\boldsymbol{\Phi} \in \mathbb{C}^{s\times|\Lambda|T}$, where $1 \leq s \ll |\Lambda|T$ denotes the sketch dimension. Using the theoretical results from [39], it is possible to choose s greater than the rank of $\boldsymbol{C}$, such that the singular values and singular vectors of the matrix $\boldsymbol{Y} \triangleq \boldsymbol{\Phi}\boldsymbol{C}^H\boldsymbol{C}$ are similar to those of $\boldsymbol{C}^H\boldsymbol{C}$. If $r_{\boldsymbol{C}}$ denotes the rank of $\boldsymbol{C}$, it is commonly observed that $r_{\boldsymbol{C}} \ll |\Lambda|T$, because the matrix $\boldsymbol{C}$ has low-rank characteristics [34], [43], [44]. By choosing the sketch dimension such that $r_{\boldsymbol{C}} < s \ll |\Lambda|T$, calculating an SVD of $\boldsymbol{Y}$ requires less computation than calculating an SVD of $\boldsymbol{C}^H\boldsymbol{C}$. If the sketch dimension is appropriately chosen [39], the right singular vectors of $\boldsymbol{Y}$ approximate the ones of $\boldsymbol{C}^H\boldsymbol{C}$. However, this approximation would hold only for the *first* $r_{\boldsymbol{C}}$ right singular vectors [39]. Because $\boldsymbol{N}$ is a *nullspace* basis, we need approximations for the *last* $|\Lambda|T - r_{\boldsymbol{C}}$ right singular vectors, which are not provided by the traditional sketched SVD.

Fortunately it is not necessary to find $\boldsymbol{N}$ to compute the projection matrix $\boldsymbol{N}\boldsymbol{N}^H$. Instead we construct an approximation for $\boldsymbol{N}\boldsymbol{N}^H$ from the matrix whose columns are the *first* $r_{\boldsymbol{C}}$ right singular vectors of $\boldsymbol{Y}$, that we denote by $\tilde{\boldsymbol{V}} \in \mathbb{C}^{|\Lambda|T\times r_{\boldsymbol{C}}}$. Using a sketched SVD to compute $\tilde{\boldsymbol{V}} \approx \boldsymbol{V}$, where $\boldsymbol{V} \in \mathbb{C}^{|\Lambda|T\times r_{\boldsymbol{C}}}$ is the matrix whose columns are the first $r_{\boldsymbol{C}}$ right singular vectors of $\boldsymbol{C}^H\boldsymbol{C}$, it follows from the fundamental theorem of linear algebra [47], [48, p. 181] that

$$\boldsymbol{N}\boldsymbol{N}^H = \boldsymbol{I}_{|\Lambda|T} - \boldsymbol{V}\boldsymbol{V}^H \tag{23}$$

$$\approx \boldsymbol{I}_{|\Lambda|T} - \tilde{\boldsymbol{V}}\tilde{\boldsymbol{V}}^H, \tag{24}$$

where $\boldsymbol{I}_{|\Lambda|T}$ is the $|\Lambda|T \times |\Lambda|T$ identity matrix.

The authors in [39] provide sufficient conditions to select the sketch dimension s such that an accurate singular vector approximation holds with high probability. However, we have observed in our experiments that this selection rule can provide a sketch dimension that is too large to significantly reduce computation time. In view of this, we provide a heuristic approach to select s. We start by calculating a regular SVD of $\boldsymbol{C}^H\boldsymbol{C}$ considering a small number of time frames, such that the SVD computation is less computationally demanding. Then, we make a first estimation of $r_{\boldsymbol{C}}$ by selecting the point where the singular value curve (sorted in decreasing order) starts to flatten out. Using this estimation of $r_{\boldsymbol{C}}$, we select the sketch dimension $s \in [2r_{\boldsymbol{C}}, 6r_{\boldsymbol{C}}]$, and proceed to estimate the projector matrix using (24). A subsequent refined estimation of $r_{\boldsymbol{C}}$ can be made using the singular values of $\boldsymbol{Y}$. We found that this heuristic approach significantly reduces computation time while negligibly affecting STM representation quality (*cf.*, Fig. 5).

E. Spatiotemporal Map Computation for Multichannel Data

The previous computation process assumes that an autocalibration region sampled at the Nyquist rate (i.e., ACS data) is acquired in each time frame. In the single-channel case, we immediately proceed with the STM computation using the ACS data to calculate $\boldsymbol{C}^H\boldsymbol{C}$ in the first step given in Section III-A. However, in the multichannel case, we apply a preliminary step to construct a virtual single-channel dataset from the ACS data. In our experiments using multichannel data we performed a SENSE-based coil combination in each time frame using sensitivity maps [49], after zero-padding everything outside the ACS data. This creates a single-channel low-resolution dataset that we used subsequently for spatiotemporal map computation.

IV. RECONSTRUCTION OF UNDERSAMPLED DYNAMIC MRI DATA USING SPATIOTEMPORAL MAPS.

This section shows how to use (3) to reconstruct accelerated dynamic MRI data. Assuming a Cartesian grid with N points, we rewrite (3) in a vectorized form as $\boldsymbol{\rho} \approx \sum_{l=1}^{L} \boldsymbol{S}_l \boldsymbol{\rho}_l$, where $\boldsymbol{\rho} \in \mathbb{C}^{NT}$ is the vectorized version of $\rho(\vec{x}, t)$ when considering all the samples and time frames; $\boldsymbol{S}_l \in \mathbb{C}^{NT \times N}$ is the matrix version of $s_l(\vec{x}, t)$, which consists of a stack of T diagonal matrices where each one contains the spatiotemporal map entries of one time frame; and $\boldsymbol{\rho}_l \in \mathbb{C}^{N}$ is the vectorized version of $\rho_l(\vec{x})$. Then, we rewrite the signal model in (1) as:

$$\boldsymbol{d} \approx \boldsymbol{A} \left(\sum_{l=1}^{L} \boldsymbol{S}_l \boldsymbol{\rho}_l \right) + \boldsymbol{\eta}, \tag{25}$$

where $\boldsymbol{d} \in \mathbb{C}^{QMT}$ is the vectorized version of the available samples in the $(\vec{k}, t)$-space considering all the coils; $\boldsymbol{\eta} \in \mathbb{C}^{QMT}$ is the vectorized version of the noise present in each sample; and $\boldsymbol{A} \in \mathbb{C}^{QMT \times NT}$ is a system matrix that includes multiplication with sensitivity maps, spatial Fourier transform of each time frame, and an undersampling operation that is time-frame dependent. The reconstruction problem now becomes estimating $\boldsymbol{\rho} = \sum_{l=1}^{L} \boldsymbol{S}_l \boldsymbol{\rho}_l$ from $\boldsymbol{d}$, given $\boldsymbol{A}$ and given the STMs $\{\boldsymbol{S}_l\}_{l=1}^{L}$ computed in Section III. One could apply any of numerous reconstruction methods from the literature to the measurement model (25).

For the experimental results in Section V, we simply focused on model-based image reconstruction approaches where one estimates $\boldsymbol{\rho}$ as follows:

$$\hat{\boldsymbol{\rho}} = \sum_{l=1}^{L} \boldsymbol{S}_l \hat{\boldsymbol{\rho}}_l. \tag{26}$$

Here

$$\{\hat{\boldsymbol{\rho}}_1, \ldots, \hat{\boldsymbol{\rho}}_L\} = \underset{\{\boldsymbol{\rho}_l\}_{l=1}^{L} \subset \mathbb{C}^N}{\operatorname{argmin}} \frac{1}{2} \left\| \boldsymbol{A} \left(\sum_{l=1}^{L} \boldsymbol{S}_l \boldsymbol{\rho}_l \right) - \boldsymbol{d} \right\|_2^2 + \beta\, \mathcal{R}(\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_L), \tag{27}$$

β denotes a regularization term, and $\mathcal{R}$ corresponds to a regularizer that imposes prior assumptions on the spatial functions of the STM decomposition. The STM formulation reduces the reconstruction of a dynamic time series to computing a few "static" images, where the temporal dynamics are encoded in the spatiotemporal maps. Therefore, we can use standard regularizers developed for non-dynamic MRI such as Tikhonov, total variation + ℓ_1, or regularizers based on structured low-rank models [43], [50], among numerous possible options. Moreover, because this formulation includes the spatiotemporal maps in a data consistency term, it could potentially be combined with machine learning reconstruction methods developed for non-dynamic MRI.

The following section explores the representation capabilities of the STM decomposition and tests the illustrative reconstruction methods on realistic dynamic MRI data.

V. EXPERIMENTS

We start describing the MRI datasets used in our experiments. All animal procedures followed a protocol approved by the Unit for Laboratory Animal Medicine and the Institutional Animal Care & Use Committee at the University of Michigan. For human subjects, approval for all ethical and experimental procedures and protocols was granted by the University of Michigan Medical School Institutional Review Board.

A. Data Description

(A.1) *Fully sampled 2D single-channel animal data:* Gastrointestinal in vivo MRI data of a rat was acquired on a 7 T single-coil small-animal scanner using a T1-weighted contrast and a gradient echo sequence. Gating was used to reduce respiratory motion. The slice orientation allowed visualizing movement of the rat's stomach during digestion. The data was acquired using a 128×84 Cartesian grid where the dimensions corresponded to the readout (RO) and phase encoding (PE) directions, respectively; 100 time frames were acquired with a temporal resolution of ~ 1.6 secs. Fully sampled data were acquired in each time frame. Further details of the acquisition can be found in [51]. Supplementary video S1 visualizes the time frames.

(A.2) *Prospectively accelerated 2D single-channel animal data:* Dataset with the same acquisition characteristics as dataset (A.1). In this case the Cartesian grid had dimensions 128×48 and the data was prospectively accelerated using an undersampling mask that in each time frame contained a 128×12 ACS region in the center of the k-space, and 12 evenly spaced PE lines outside the ACS region. These PE lines were shifted at different time frames such that each k-space location was sampled at least once every 3 time frames. Therefore, the undersampling mask produced $\times 2$ accelerated data. For illustrative purposes our results using this dataset were zero-padded in the k-space domain to match the dimensions of dataset (A.1) (i.e., 128×84). Supplementary video S2 visualizes the time frames.

Fully sampled 3D multichannel human BOLD fMRI data: BOLD fMRI data was acquired on a 3 T MRI scanner using a 32-coil receiver array and a 3D gradient-echo EPI sequence. A right-hand finger tapping task was performed with alternating 20 secs blocks of tap/rest for 5 cycles. The data was acquired using a Cartesian grid with dimensions $90 \times 90 \times 20$ corresponding to the RO direction, the PE direction in k_y, and the PE direction in k_z, respectively, with a $2.4\,\text{mm}^3$ isotropic resolution; 140 time frames were acquired with a temporal resolution of ~ 1.6 secs. Fully sampled data were acquired in each time frame. The data were coil-compressed to 10 virtual coils using a standard SVD approach [52].

All datasets presented their own challenges for accelerated reconstruction. On the one hand, only a single channel was available for datasets (A.1) and (A.2); therefore, no parallel imaging models could be used. In addition, they exhibited a high level of noise (*cf.*, supplementary videos S1 and S2), and the temporal/spectral characteristics varied considerably at different spatial locations. On the other hand, dataset (B) was 3D, which increased computation time for both spatiotemporal map computation and reconstruction. In addition, the temporal variations of interest in fMRI are inherently small in magnitude. Finally, there were many time frames for both datasets, so the proposed sketched SVD approach in Section III-D was crucial for practical spatiotemporal map computation.

All datasets exhibited a dynamic behavior that was quasi-periodic for most spatial locations, that also varied not too rapidly across the FOV. Therefore, according to our theoretical analysis in Section II-B, we expected to have a good approximation using the STM model.

B. Undersampling K-Space Masks for Retrospective Data Acceleration

For our reconstruction experiments we simulated accelerated acquisitions by retrospectively undersampling the previously described datasets (A.1) and (B). For dataset (A.1) we used an undersampling mask that in each time frame contained a 128×12 ACS region in the center of the k-space, and 4 evenly spaced PE lines outside the ACS region. These PE lines were shifted at different time frames such that each k-space location was sampled at least once every 18 time frames. Therefore, the undersampling mask produced $\times 5.25$ accelerated data. Fig. 1(a) illustrates this undersampling mask.

For dataset (B) the undersampling mask was designed analogously. It contained a $90 \times 18 \times 12$ ACS region in the center of the k-space of each time frame, and 4 PE lines in the k_y direction outside the ACS region. In addition, this mask included PE lines in the k_z direction outside the ACS region. Specifically, 2 lines were considered. Therefore, this mask produced $\times 5.84$ accelerated data. Fig. 1(b) illustrates this undersampling mask. Though this sampling mask produced heavily undersampled data, the acquisition time would not be considerably decreased in a prospectively accelerated scenario, as the EPI sequence rapidly traverses the full (k_x, k_y)-space. Our goal for investigating this undersampling mask was to assess, as a proof-of-concept for fMRI, the STM model and the proposed reconstruction framework in a heavily undersampled scenario. Our goal in future work is to use the STM model for other undersampled 3D k-space patterns for fMRI, for example, [53].

C. Implementation Details for Computing Spatiotemporal Maps

For all experiments we computed spatiotemporal maps using the steps in Section III. For datasets (A.1) and (A.2), in step 1) we used the ACS region of the undersampling mask described in Section V-B, and we calculated $\boldsymbol{C}^H\boldsymbol{C}$ using the FFT-based approach provided in PISCO with an ellipsoidal neighborhood of radius $R = 3$ for Λ (*cf.*, Section III-A). Step 2) computed a projection matrix for the nullspace of $\boldsymbol{C}$ using the sketched SVD approach proposed in Section III-D with a complex-valued Gaussian matrix for the sketch matrix $\boldsymbol{\Phi}$. We selected the sketch dimension $s = 2r_{\boldsymbol{C}}$ using the heuristic approach described in Section III-D. Finally, step 3) computed spatiotemporal maps by calculating an approximate nullspace basis for each $\boldsymbol{G}(\vec{x})$ matrix using an SVD performed for each spatial location in the FOV. If not specified otherwise, we used $L = 4$ components. To further reduce computation time, we computed spatiotemporal maps on a grid with coarser resolution than the original and interpolated using the FFT-based PISCO approach. The 12.5% of each spatiotemporal map was interpolated in k-space if not specified otherwise. For dataset (B) the procedure was analogous. The main differences with respect to datasets (A.1) and (A.2) were: the ACS region corresponded to the one considered in the undersampling mask for dataset (B) described in Section V-B; we selected $R = 4$ if not specified otherwise; and we used the orthogonal iteration to find the nullspace basis of each matrix $\boldsymbol{G}(\vec{x})$.

D. Representation Capabilities: Comparison With the PSF Model

We compared the representation capabilities of the STM model against the PSF model using dataset (A.1). We computed spatiotemporal maps and the temporal functions of the PSF decomposition using the same ACS data, and we calculated the normalized projection residual (NPR) for both models while varying the number of components. NPR corresponds to the normalized error after projecting the fully sampled data onto the space spanned either by the spatiotemporal maps or the PSF temporal functions. For the STM model we define

$$\mathrm{NPR}(L) \triangleq \|\boldsymbol{\rho} - \hat{\boldsymbol{\rho}}(L)\|_2 \,/\, \|\boldsymbol{\rho}\|_2 \,, \tag{28}$$

where $\boldsymbol{\rho}$ denotes the reference image reconstructed from fully sampled data and $\hat{\boldsymbol{\rho}}(L)$ is calculated using (26) with no regularization, assuming an STM model with L components, and no undersampling operation in the system matrix $\boldsymbol{A}$. Due to the high level of background noise in dataset (A.1), we calculated the NPR over a region of interest (ROI) around the stomach as illustrated in Fig. S2 of the supplementary material.

E. Sketched SVD for a Projector Matrix Calculation: Evaluation of Computation Time and Representation Quality

We tested the method proposed in Section III-D using datasets (A.1) and (B). Given that a random matrix is used, it is relevant to study whether different realizations can induce a high variability in the final STM computation. We varied the sketch dimension s in the range $[2r_{\boldsymbol{C}}, 6r_{\boldsymbol{C}}]$, and calculated the NPR for different realizations of the sketch matrix; we reported the mean and standard deviation using 50 realizations in each case. We also report the median time to compute an SVD with and without sketching.

F. Reconstruction Experiments

We tested the reconstruction framework in Section IV using datasets (A.1), (A.2) and (B). For datasets (A.1) and (B) we simulated accelerated acquisitions by retrospectively undersampling the data using the undersampling k-space masks described in Section V-B. We reconstructed each accelerated dataset using the proposed reconstruction framework where, as a proof-of-principle, we explored Tikhonov regularization and a P-LORAKS regularizer [44], where the latter considered each STM spatial function as a virtual channel. For simplicity, we refer to these two methods as STM & Tikhonov and STM & P-LORAKS hereWe studied how the parametersafter. Our motivation for exploring a LORAKS-type regularizer was based on our empirical observation that SILP relationships can be present among the spatial components of the STM model, and therefore structured low-rank models could be used [34]. One theoretical explanation for the existence of these relationships could be based on the shared spatial support of the spatial components [37]. The two regularizers that we used were selected for simplicity, as both involved only one regularization parameter. Alternatively, one could use different regularizers for each spatial function in the STM decomposition, with individual regularization parameters.

When using Tikhonov regularization, we minimized (27) using conjugate gradient. When using P-LORAKS regularization, we solved the inverse problem using the majorize-minimize algorithm in [50]. We compared the reconstruction results with five baseline methods. The first method used data sharing to substitute each missing k-space sample with a sample available in a neighboring time frame [1]. We refer to this method as Data Sharing. The second method used a locally low-rank (LLR) model that was implemented using nonlinear conjugate gradient [21]. The third method corresponded to blind compressed sensing (BCS) [11], using the software provided by the authors. The fourth method used a low-rank plus sparsity model [12] that was implemented using POGM [16]. We refer to this method as L + S. The fifth method corresponded to an analogous version of the reconstruction framework in Section IV, where instead of using the STM model in (26) and computing the STM spatial functions in (27), we used the PSF model and computed the PSF spatial functions, respectively. The PSF temporal functions were computed from the same ACS data used to compute spatiotemporal maps, and we used Tikhonov regularization when computing the PSF spatial functions. We refer to this method as PSF & Tikhonov hereafter.

We studied how the parameters related to the spatial resolution of spatiotemporal maps affected reconstruction performance. Using dataset (B) we studied how reconstruction performance varied when choosing the radius parameter $R \in \{2, 3, 4, 5\}$. In addition, we reconstructed the retrospectively undersampled data using spatiotemporal maps computed after interpolating 12.5% , 25% , and 40% of their k-space using the FFT-based PISCO approach.

For datasets (A.1) and (B) we report the normalized root mean square error (NRMSE) in each case; for dataset (A.1) the NRMSE was calculated over the same ROI used when calculating the NPR. For dataset (B) t-score maps were calculated to assess functional activity. For the multichannel experiments we calculated sensitivity maps using PISCO [36] and the ACS data of the first time frame. Each method was implemented in-house using MATLAB R2023a on a local server with an Intel Zeon Silver 4216 2.10 GHz CPU and 263 GB RAM. Open-source code for computing STMs is available at: https://github.com/ralobos/STM_MRI.git. This code extends the PISCO software [38], by including the sketched SVD and the orthogonal iteration computational methods.

TABLE I
COMPUTE TIME AND REPRESENTATION QUALITY WHEN USING SKETCHED SVD AND REGULAR SVD FOR THE COMPUTATION OF SPATIOTEMPORAL MAPS USING THE 3D DATASET (B)

	Time [min]	NPR
Regular SVD	115.14	0.0565 ± 0.0001
Sketched SVD	0.18	0.0575 ± 0.0003

VI. RESULTS

Fig. 2(a) shows the last 10 eigenvalues (after sorting them in decreasing order) of the $\boldsymbol{G}(\vec{x})$ matrices found when computing the STMs using dataset (A.1). The number of eigenvalues near zero for each matrix $\boldsymbol{G}(\vec{x})$ determines the number of components needed in the STM model, as this is an estimation of the nullspace dimensionality. The number of components varies across the FOV. For example, areas that move slowly over time or that are motionless and only present contrast variations, need only one component. However, areas with more complex dynamics like the antrum, which is in the distal part of the stomach, need more components. The eigenvalue maps suggest that four components could capture the dynamic behavior of these areas. On the other hand, the PSF model needs more components to represent the dynamic behaviors across the FOV.

Fig. 2(b) shows the singular values (sorted in decreasing order) of the Casorati matrix used to find the PSF temporal functions. The number of relatively large singular values can be used to approximate the number of components needed in the PSF model (i.e., L_{PSF}); Fig. 2(b) suggests that 4-6 components should capture most of the dynamic behavior of the data.

Fig. 3 quantifies the representation error of both models by measuring the NPR while varying the number of components (i.e., L in the STM model and L_{PSF} in the PSF model). The NPR curve decreased faster for the STM model than for the PSF model when considering a small number of components. The NPR for the STM model using $L = 4$ components (i.e., the number of components used in our experiments), was achieved by the PSF model when using at least $L_{\text{PSF}} = 6$ components. Figure S1 shows qualitative results for the approximations given by the STM and PSF models when $L = L_{\text{PSF}} = 4$ for three representative time frames. Each approximation was obtained after calculating the spatial components of each model from the fully sampled data. Error images are also shown. Overall, the STM approximation obtained smaller differences than the PSF approximation, which is consistent with the NPR results in Fig. 3. The better representation capabilities of the STM model can be attributed to the construction of different temporal subspaces for each spatial location. The PSF model construct a single temporal subspace that is the same for all spatial locations. Given that only 4 components were considered, the single PSF temporal subspace was not able to properly represent the temporal behavior at every voxel.

Section II-B provided sufficient conditions for the existence of SILP relationships in the $(\vec{k}, t)$-space. Using dataset (A.1), we explored how these theoretical results can manifest in practice.

Specifically, we studied the SILP relationships given by the multiframe FIR filters $\{\breve{g}_{t'}(\vec{k},t)\}_{t'=1}^{T}$ defined in (19), which we calculated from the $\boldsymbol{G}(\vec{x})$ matrices. Fig. 4 shows the temporal behavior of $\rho(\vec{x},t)$ and $g_{t'}(\vec{x},t)$ for one specific $t' \in \{1,\ldots,T\}$, considering two different voxels located in two areas with complicated dynamic behaviors. It also shows $\breve{\rho}(\vec{x},f)$ and $\breve{g}_{t'}(\vec{x},f)$, from which we see that for both voxels $\breve{\rho}(\vec{x},f)$ has most of its energy in a finite set of bands that depends on the spatial location of the voxel. Furthermore, $\breve{g}_{t'}(\vec{x},f)$ has its support defined such that $\breve{g}_{t'}(\vec{x},f)\breve{\rho}(\vec{x},f) \approx 0$, $\forall f \in \mathcal{F}$, which is related to the sufficient conditions provided in Section II-B.

Table I compares the compute time of the regular SVD versus the sketched SVD when computing spatiotemporal maps using the 3D dataset (B). The NPR obtained in the final STM computation is also shown in each case. Both SVD approaches were applied on the matrix $\boldsymbol{C}^H\boldsymbol{C}$ whose dimensions were 35980×35980, which caused regular SVD to be computationally expensive. The sketched SVD was ~ 640-fold faster than the regular SVD, reducing the computation from hours to seconds. The spatiotemporal maps obtained for both cases were quite similar in terms of their representation capabilities, which is reflected on the small differences in the NPR. Both cases experience randomness from the random initialization used by the orthogonal iteration when calculating the nullspace bases for the matrices $\boldsymbol{G}(\vec{x})$. This explains why the NPR result for regular SVD is shown with a standard deviation.

Given our heuristic approach to selecting the sketch dimension s, we also explored the effects of varying its value. Using dataset (A.1), Fig. 5 shows the computation times of using sketched SVD for different values of the sketched dimension, and it also shows the NPR obtained in each case after computing the corresponding STMs. Increasing the sketch dimension increased the computation time as expected; however, STM representation quality was only slightly affected as the NPR displayed negligible fluctuations. In comparison to using regular SVD, which took 5.77 secs and obtained an NPR equal to 0.1005, a considerable acceleration was obtained for each case considered in Fig. 5.

Figs. 6 and 7 show the reconstruction results using dataset (A.1). Fig. 6 shows the NRMSE obtained by STM & Tikhonov and PSF & Tikhonov when varying the number of components. (For reference, the NRMSE obtained by Data Sharing is also shown.) The NRMSE for both methods decreased initially as the number of components increased, as expected because the representation capabilities of both models improve when increasing the number of components. The NRMSE achieved a minimum for both methods and then increased. This is also expected as the number of unknowns increases as more components are included. For all cases STM & Tikhonov obtained a lower NRMSE than PSF & Tikhonov, which can be attributed to the better representation capabilities of the STM model compared to the PSF model (*cf.*, Fig. 3).

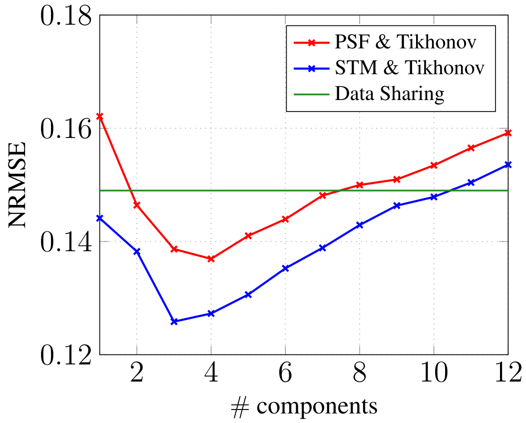


Fig. 6. NRMSE versus the number of components of the PSF & Tikhonov and STM & Tikhonov methods. The NRMSE obtained by Data Sharing is shown as a reference.

TABLE II
TIME-SERIES NRMSE FOR BOTH DATASETS

	2D dataset (A.1)	3D dataset (B)
Zero filled	0.200	0.346
Data Sharing	0.149	0.206
Locally Low-rank	0.133	n/a
Blind Compressed Sensing	0.119	n/a
L + S	0.109	0.116
PSF & Tikhonov ($L_{\text{PSF}} = 4$)	0.137	0.124
PSF & Tikhonov ($L_{\text{PSF}} = 6$)	0.144	0.127
STM & Tikhonov	0.127	0.097
STM & P-LORAKS	0.113	0.097

Fig. 7 shows the reconstruction results for the methods based on the STM model when setting the number of components to $L = 4$, PSF & Tikhonov when the number of components is set to $L_{\text{PSF}} = 4$ and $L_{\text{PSF}} = 6$, and the other considered reconstruction methods. The idea behind the considered values for L_{PSF} was to compare STM & Tikhonov and PSF & Tikhonov when the number of components was the same, and also when the latter had more components than the former. The first time frame is shown for each reconstruction method, as well as one line in the PE direction for all the time frames. This line covers several areas with a complex dynamic behavior, and its evolution over time shows how several voxels exhibit different temporal characteristics. For the first time frame, the regularized methods all obtained similar NRMSE values that were lower than the one obtained by Data Sharing (indicated in Fig. 7). However, when considering the whole time series, L + S and STM & P-LORAKS obtained better NRMSE than STM & Tikhonov and PSF & Tikhonov as reported in Table II. Although L + S obtained a slightly lower time-series NRMSE than STM & P-LORAKS, note that this value was calculated with respect to a fully sampled dataset that exhibited a high level of noise. A method with denoising characteristics can obtain a higher time-series NRMSE in comparison to methods that preserve noise. In our case, STM & P-LORAKS had better qualitative denoising characteristics as seen in supplementary video S1. However, higher quantitative errors can be expected for voxels where noise was decreased. Figure S3 shows the reconstructed time course for voxel B in Fig. 4(a). Even though STM & P-LORAKS obtained a higher NRMSE than BCS or L + S for that specific voxel (*cf.*, Table S1), it was able to better follow the intensity time profile of the fully sampled data.

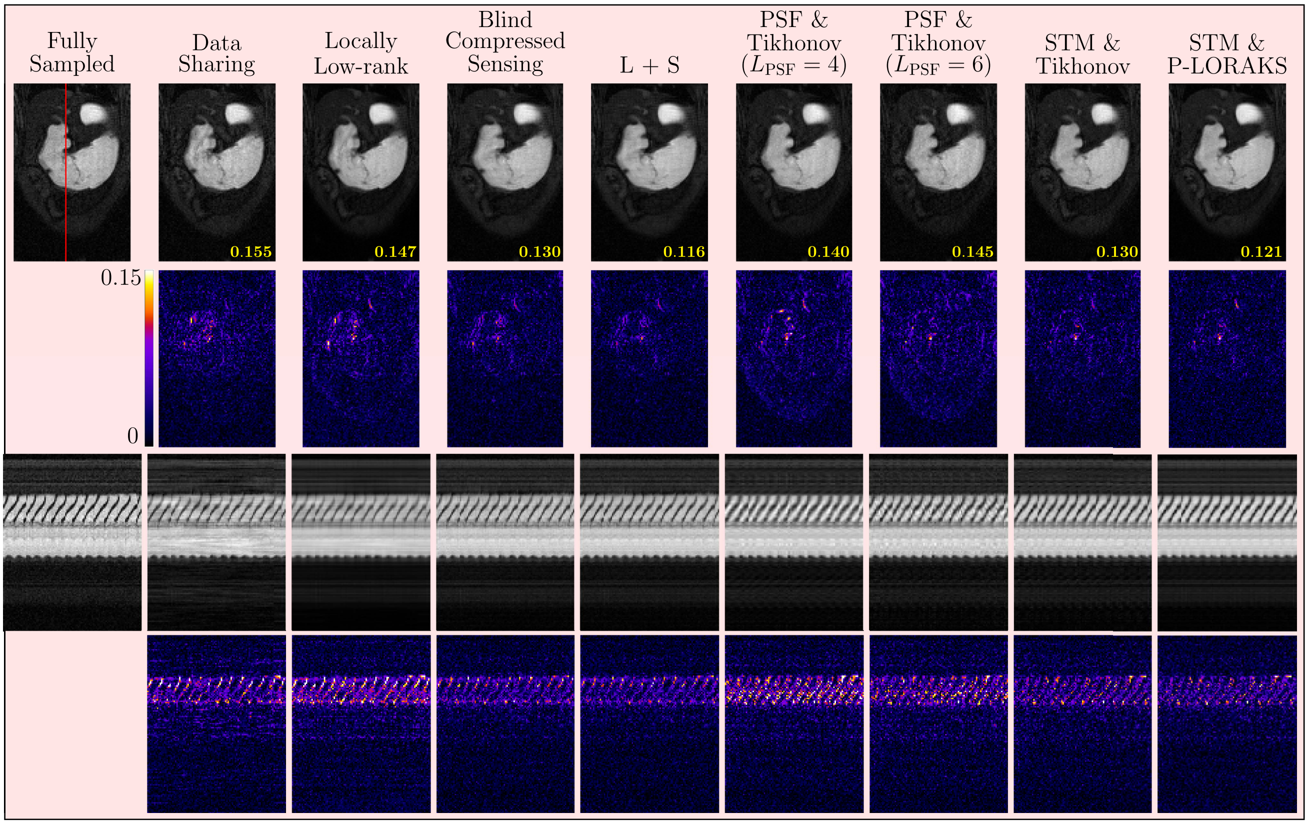


Fig. 7. Retrospective reconstruction results using dataset (A.1). The first time frame of the fully sampled data is shown on the left with a time profile of one line in the PE direction (indicated in red). Then, reconstruction results are shown for each method, where the NRMSE for the displayed time frame is indicated in yellow at the bottom right corner. Error magnitude images are provided for each method below the reconstructed images.

Fig. 8 shows the reconstruction results using the 3D dataset (B) for the same methods shown in Fig. 7, except for BCS and LLR.[3] In this case STM & P-LORAKS and STM & Tikhonov obtained a similar time-series NRMSE, which was better than those obtained by the other methods (*cf.*, Table II), and all methods produced quite similar qualitative results. In addition to image quality, it was also important to capture functional activity. Fig. 8 shows the t-score maps obtained for each reconstruction method overlaid on the corresponding reconstructed images. Even though the reconstructed images for each method exhibited good quality, only STM & P-LORAKS and STM & Tikhonov were able to obtain t-score maps reflecting functional activity similarly to the fully sampled data. Fig. 8 also shows a plot with the intensity over time for one voxel with a high t-score considering the fully sampled data and the reconstructed data for each method, in addition to a "task" curve that shows the tap/rest blocks. STM &P-LORAKS and STM & Tikhonov followed the task behavior better than L + S and Data Sharing. The voxel time-series for PSF & Tikhonov($L_{\text{PSF}} = 4$) and PSF & Tikhonov($L_{\text{PSF}} = 6$) were quite different than the task behavior. Table S1 reports the NRMSE for the time course obtained by each reconstruction method. For this particular voxel, methods using STMs obtained better values than L + S, and similar values to methods based on PSF.

We observed (not shown) that the t-score maps for L + S were improved when using low acceleration factors; however, for high accelerations a high degree of global low-rank regularization was needed to obtain good reconstruction quality, causing the final reconstruction to have reduced temporal variability. On the other hand, the STM model encodes dynamics at a voxel level, which better captures phenomena that occur in a voxel-wise fashion, as happens in fMRI.

Figure S4 shows reconstruction results for dataset (B) using STM & P-LORAKS while varying the radius parameter R when computing STMs. Given that R and the degree of interpolation affect the spatial resolution of STMs, we used no interpolation to study the isolated effect of the former in the reconstruction. Even though the time-series NRMSE was similar in each case as shown in Table S2, high values of R better captured functional activity than low values, as shown by the corresponding t-score maps and the time plots of one specific voxel with a high t-score. This is expected as higher values of R allows the computation of STMs with a higher spatial resolution that better capture abrupt changes in dynamic behavior across spatial locations. Figure S5 shows representative STMs corresponding to various values of R for a specific time frame and the slice shown in Fig. 8. As

---

[3] The implementation of BCS for the 3D case was not available, and extending the provided 2D implementation required a nontrivial selection of parameters that was beyond the scope of this paper; LLR for the 3D case required a prohibitive computation time, mainly due to the high data dimensions and the number of shifts needed such that overlapping patches could be processed to avoid blocky artifacts [21].

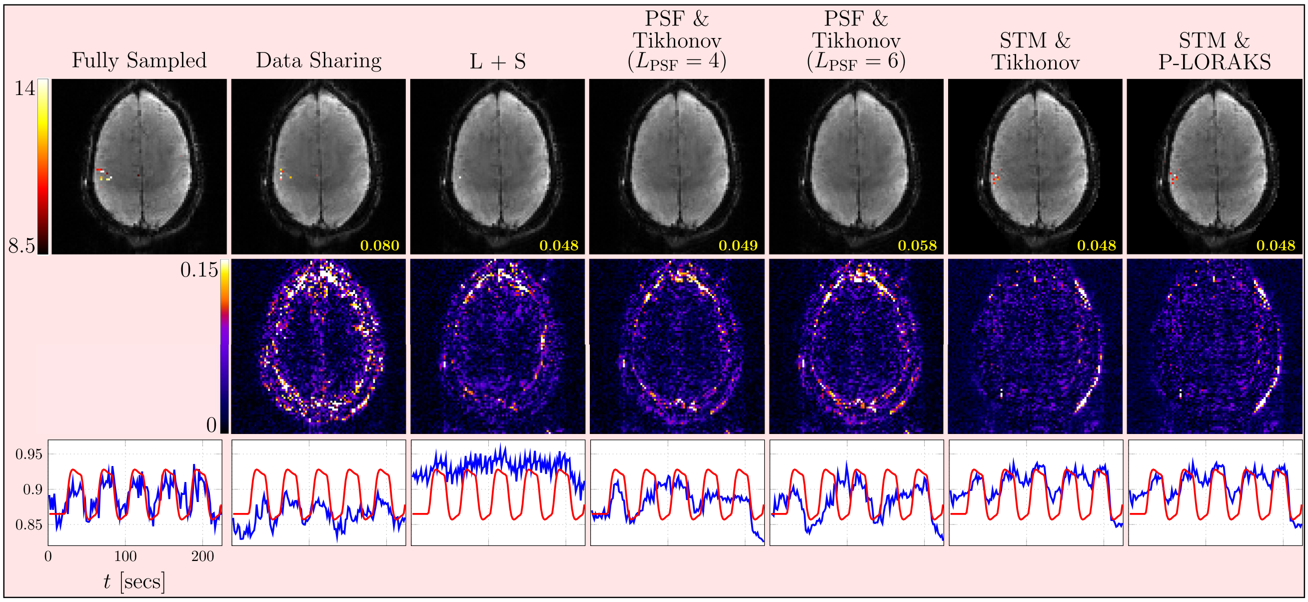


Fig. 8. Retrospective reconstruction results using dataset (B). Row 1: The first time frame of a representative slice of the fully sampled data is shown on the top left overlaid with a t-score map calculated using the task information, followed by reconstruction results for each method; the NRMSE for the displayed time frame is indicated in yellow at the bottom right corner. The t-score maps obtained for each reconstruction method are overlaid on the reconstructed images. Row 2: Error magnitude images are provided for each method. Row 3: Signal intensity plots; the first plot corresponds to the signal evolution for one specific voxel with a high t-score using the fully sampled data (in blue), and a task curve (in red) showing the tap/rest blocks. Analogous plots are also shown for all the reconstructed datasets for the same voxel location.

illustrated, increasing the value of R resulted in less smooth STMs. However, larger values of R increase computational complexity and memory usage and require larger autocalibration regions (*cf.*, Section III-A).

Table S3 shows the time-series NRMSE when reconstructing dataset (B) using STM & P-LORAKS and different degrees of interpolation when computing STMs. Similar quantitative results were obtained in each case. Even though similar image quality and time-series NRMSE were obtained for each case (*cf.* Table S3), variations in functional activity were observed as shown in Fig. S6. The reconstructed time courses had higher variations with respect to the task curve for higher interpolation degrees. The dynamic behavior in dataset (B) is expected to be localized in brain regions related to the task, and non-smooth variations in the dynamic behavior might be expected when transitioning to voxels outside these regions. This can require STMs that are not necessarily smooth in those regions, which decreases the degree of interpolation to accurately capture the dynamic behavior of voxels activated during the task.

Supplementary video S2 shows the reconstruction results for the prospectively undersampled dataset (A.2). Qualitatively the reconstructions were similar to the ones obtained when reconstructing restrospectively undersampled data using dataset (A.1). BCS, L + S, STM & Tikhonov, and STM & P-LORAKS exhibited less aliasing artifacts than the other methods. Methods based on STM exhibited less noise than BCS and L + S; although, sharper edges were observed for BCS and L + S.

VII. DISCUSSION AND CONCLUSIONS

The STM model proposed in this work for dynamic MRI reconstruction has been presented as an extension of the widely used PSF model. Unlike the PSF model, the STM model represents the dynamic MRI signal using temporal functions (i.e., spatiotemporal maps) that depend on the spatial location. We have shown that this spatial dependency allows dynamic MRI reconstruction methods based on the STM model to obtain better reconstruction than methods based on the PSF model (*cf.*, Figs. 6, 7, 8 and Table II).

Another feature of the STM model is that the number of components can depend on the spatial location. We have not exploited this flexibility in the experiments shown in this work; however, this spatial dependency could be used, for example, to focus the reconstruction on regions of interest. In particular, we hypothesize that spatiotemporal maps could have a synergistic effect when combined with Region-Optimized Virtual (ROVir) coils [54].

The number of components in the STM model vary depending on the dynamics present in the data. In our experiments we observed that some quasi-periodic behaviors related to small deformations or contrast changes were well-approximated using a few components. However, we have observed (not shown) that more complex dynamics (*e.g.*, respiratory motion) require several components. Our future work includes extending the STM framework to more challenging contexts, and theoretically studying how the number of needed components varies depending on the data dynamics.

We also provided a theoretical framework that connects the properties of spatiotemporal maps with autoregressive properties of the $(\vec{k}, t)$-space. Specifically, we have shown that establishing shift-invariant linear predictability relationships in the k-space of different time frames, leads to a subspace-based computation framework for spatiotemporal maps. Notably, our theory allows a direct connection with subspace-based estimation of sensitivity maps in multichannel MRI. Using this analogue, we provided a spatiotemporal map construction procedure that relies on recent advanced signal processing computational methods, originally proposed for sensitivity map estimation. Moreover, we have extended these computational methods using randomized linear algebra techniques to make spatiotemporal map computation efficient for long time series. As an example, using 3D fMRI data, the computational method proposed in this work based on sketched SVD was $\sim 640\times$ faster than using the analogous computational method based on the previous SVD approach originally proposed for sensitivity map estimation. Even though our proposed computational methods were used in the context of STM computation, they could also be used for subspace-based sensitivity map estimation [36], [42], where compute time could be greatly reduced for receiver arrays having many coils.

We have shown that spatiotemporal maps can be incorporated in the signal model, supporting numerous reconstruction frameworks for accelerated dynamic MRI. As proof-of-principle illustrations, we investigated a model-based reconstruction framework that includes spatiotemporal maps as part of a data-consistency term, where they encode the information about the dynamics of the data. The resulting inverse problem considers the spatial functions of the STM model as the optimization variables. This formulation reduces the reconstruction of a long dynamic series of images to the estimation of a few non-dynamic component images. Therefore, regularizers developed for non-dynamic MRI can be used. We evaluated the proposed reconstruction framework using a simple Tikhonov regularizer and a LORAKS-type regularizer. Using highly accelerated data, both regularizers obtained similar NRMSE metrics as the L + S method based on low-rank models. However, the STM reconstruction framework showed additional denoising characteristics, and better captured small dynamic changes that are particularly relevant in fMRI. As an example, t-scores maps produced from fMRI images reconstructed using the STM-based approach clearly revealed functional activity in expected spatial areas, similarly to the gold standard; this was not the case when using a reconstruction based on a low-rank + sparsity model or other methods based on the PSF model. Interestingly, a simple data-sharing reconstruction exhibited a periodic pattern similar to that of the gold standard. A weighted version of this method could improve reconstruction performance and serve as initialization for other methods. We emphasize that the regularizers used in our experiments are not necessarily optimal, and they were just used as a proof-of-principle. Another option would be to use regularizers based on machine/deep learning models that were originally proposed in non-dynamic or dynamic settings. Moreover, given that spatiotemporal maps can be incorporated as part of the system forward operator, they could be used in other reconstruction frameworks that leverage on scan-specific machine learning methods, *e.g.*, [23]. Exploring these ideas is part of our future work.

The computation time for reconstruction frameworks based on the STM model can vary depending on how these are employed. In this work, we have presented a proof-of-principle, model-based reconstruction framework, where the total computation time includes both STM calculation and the time spent solving the optimization problem–which itself depends on the choice of regularizer. STM computation can be performed efficiently using PISCO and our proposed sketched SVD technique (*cf.*, Section III). For instance, STM computation required only ~ 5 secs for dataset (A.1). However, solving the optimization problem in (27) using iterative methods necessitates performing a number of FFTs equal to the number of time frames. For long time series, this process can be time-intensive. This can be considered as a disadvantage compared to reconstruction frameworks based on the PSF model. As shown in certain applications, the PSF model's structure can be exploited to reduce the number of required FFTs–from the total number of time frames down to the number of spatial components [55]. Additionally, the STM model requires more memory to store its parameters: a different set of temporal functions needs to be stored for each voxel, in contrast to the PSF model that requires only a single set of temporal functions for all voxels. In future work, we plan to investigate more efficient implementations for reconstruction frameworks based on the STM model.

Our future work also includes using the STM model in other applications. For example, given the existence of linear predictability in the $(\vec{k}, t)$-space of images with different contrast [34], [56], the STM model could be used to reconstruct accelerated quantitative MRI data (*cf.*, Section II-C).

REFERENCES

[1] R. Jones, O. Haraldseth, T. Müller, P. Rinck, and A. Øksendal, "K-space substitution: A novel dynamic imaging technique," *Magn. Reson. Med.*, vol. 29, no. 6, pp. 830–834, Jun. 1993.

[2] Z.-P. Liang, H. Jiang, C. P. Hess, and P. C. Lauterbur, "Dynamic imaging by model estimation," *Int. J. Imag. Syst. Tech.*, vol. 8, no. 6, pp. 551–557, 1997.

[3] B. Madore, G. H. Glover, and N. J. Pelc, "Unaliasing by fourier-encoding the overlaps using the temporal dimension (UNFOLD), applied to cardiac imaging and fMRI," *Magn. Reson. Med.*, vol. 42, no. 5, pp. 813–828, 1999.

[4] Z.-P. Liang, "Spatiotemporal imaging with partially separable functions," in *Proc. IEEE Intl. Symp. Biomed. Imag.*, 2007, pp. 988–991.

[5] J. Tsao, P. Boesiger, and K. P. Pruessmann, "K-T BLAST and.k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations," *Magn. Reson. Med.*, vol. 50, no. 5, pp. 1031–1042, 2003.

[6] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye, "k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI," *Magn. Reson. Med.*, vol. 61, no. 1, pp. 103–116, 2009.

[7] J. P. Haldar and Z.-P. Liang, "Spatiotemporal imaging with partially separable functions: A matrix recovery approach," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2010, pp. 716–719.

[8] S. G. Lingala, Y. Hu, E. DiBella, and M. Jacob, "Accelerated dynamic MRI exploiting sparsity and low-rank structure: K-t SLR," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1042–1054, May 2011.

[9] B. Zhao, J. P. Haldar, A. G. Christodoulou, and Z.-P. Liang, "Image reconstruction from highly undersampled (K, t)-space data with joint partial separability and sparsity constraints," *IEEE Trans. Med. Imag.*, vol. 31, no. 9, pp. 1809–20, Sep. 2012.

[10] J. Trzasko and A. Manduca, “Local versus global low-rank promotion in dynamic MRI series reconstruction,” in *Proc. Int. Soc. Magn. Reson. Med.*, 2011, p. 4371.[Online]. Available: http://archive.ismrm.org/2011/4371.html

[11] S. G. Lingala and M. Jacob, “Blind compressive sensing dynamic MRI,” *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 1132–1145, Jun. 2013.

[12] R. Otazo, E. Candes, and D. K. Sodickson, “Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components,” *Magn. Reson. Med.*, vol. 73, no. 3, pp. 1125–36, Mar. 2015.

[13] L. Feng, L. Axel, H. Chandarana, K. T. Block, D. K. Sodickson, and R. Otazo, “XD-GRASP: Golden-angle radial MRI with reconstruction of extra motion-state dimensions using compressed sensing,” *Magn. Reson. Med.*, vol. 75, no. 2, pp. 775–788, 2016.

[14] S. Poddar and M. Jacob, “Dynamic MRI using smoothness regularization on manifolds (SToRM),” *IEEE Trans. Med. Imag.*, vol. 35, no. 4, pp. 1106–1115, Apr. 2016.

[15] S. Ravishankar, B. E. Moore, R. R. Nadakuditi, and J. A. Fessler, “Low-rank and adaptive sparse signal (LASSI) models for highly accelerated dynamic imaging,” *IEEE Trans. Med. Imag.*, vol. 36, no. 5, pp. 1116–1128, May 2017.

[16] C. Y. Lin and J. A. Fessler, “Efficient dynamic parallel MRI reconstruction for the low-rank plus sparse model,” *IEEE Trans. Comput. Imag.*, vol. 5, no. 1, pp. 17–26, Mar. 2019.

[17] S. Biswas, H. K. Aggarwal, and M. Jacob, “Dynamic MRI using model-based deep learning and SToRM priors: MoDL-SToRM,” *Magn. Reson. Med.*, vol. 82, no. 1, pp. 485–494, 2019.

[18] C. Qin, J. Schlemper, J. Caballero, A. N. Price, J. V. Hajnal, and D. Rueckert, “Convolutional recurrent neural networks for dynamic MR image reconstruction,” *IEEE Trans. Med. Imag.*, vol. 38, no. 1, pp. 280–290, Jan. 2019.

[19] Z. Ke et al., “Learned low-rank priors in dynamic MR imaging,” *IEEE Trans. Med. Imag.*, vol. 40, no. 12, pp. 3698–3710, Dec. 2021.

[20] G. Cruz et al., “Low-rank motion correction for accelerated free-breathing first-pass myocardial perfusion imaging,” *Magn. Reson. Med.*, vol. 90, no. 1, pp. 64–78, 2023.

[21] R. A. Lobos, J. S. Cavazos, R. R. Nadakuditi, and J. A. Fessler, “Smooth optimization using global and local low-rank regularizers,” May 2025, *arXiv:2505.06073*.

[22] J. Feng et al., “Spatiotemporal implicit neural representation for unsupervised dynamic MRI reconstruction,” *IEEE Trans. Med. Imag.*, vol. 44, no. 5, pp. 2143–2156, May 2025.

[23] H. Yu, J. A. Fessler, and Y. Jiang, “Bilevel optimized implicit neural representation for scan-specific accelerated MRI reconstruction,” Feb. 2025, *arXiv:2502.21292*.

[24] J. V. Manjón, P. Coupé, L. Concha, A. Buades, D. L. Collins, and M. Robles, “Diffusion weighted image denoising using overcomplete local PCA,” *PLoS One*, vol. 8, no. 9, 2013, Art. no. e73021.

[25] A. Saucedo, S. Lefkimmiatis, N. Rangwala, and K. Sung, “Improved computational efficiency of locally low rank MRI reconstruction using iterative random patch adjustments,” *IEEE Trans. Med. Imag.*, vol. 36, no. 6, pp. 1209–1220, Jun. 2017.

[26] J. I. Tamir et al., “T2 shuffling: Sharp, multicontrast, volumetric fast spin-echo imaging,” *Magn. Reson. Med.*, vol. 77, no. 1, pp. 180–95, Jan. 2017.

[27] L. Cordero-Grande, D. Christiaens, J. Hutter, A. N. Price, and J. V. Hajnal, “Complex diffusion-weighted image estimation via matrix recovery under general noise models,” *NeuroImage*, vol. 200, pp. 391–404, Oct. 2019.

[28] S. Guo, J. A. Fessler, and D. C. Noll, “High-resolution oscillating steady-state fMRI using patch-tensor low-rank reconstruction,” *IEEE Trans. Med. Imag.*, vol. 39, no. 12, pp. 4357–68, Dec. 2020.

[29] L. Vizioli et al., “Lowering the thermal noise barrier in functional brain mapping with magnetic resonance imaging,” *Nature Commun.*, vol. 12, 2021, Art. no. 5181.

[30] Y. Zhao et al., “Joint denoising of diffusion-weighted images via structured low-rank patch matrix approximation,” *Magn. Reson. Med.*, vol. 88, no. 6, pp. 2461–2474, Dec. 2022.

[31] P.-A. Comby, Z. Amor, A. Vignaud, and P. Ciuciu, “Denoising of fMRI volumes using local low rank methods,” in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2023, pp. 1–5.

[32] N. K. Meyer et al., “Enhanced clinical task-based fMRI metrics through locally low-rank denoising of complex-valued data,” *Neuroradiol. J.*, vol. 36, no. 3, pp. 273–88, Jun. 2023.

[33] X. Chen et al., “Improved structured low-rank reconstruction for 3D multi-shot EPI with joint motion modelling,” in *Proc. ISMRM Workshop Data Sampling Image Reconstruction*, 2023.

[34] J. P. Haldar and K. Setsompop, “Linear predictability in MRI reconstruction: Leveraging shift-invariant fourier structure for faster and better imaging,” *IEEE Signals Process. Mag.*, vol. 37, no. 1, pp. 69–82, Jan. 2020.

[35] M. Mishali and Y. C. Eldar, “Blind multiband signal reconstruction: Compressed sensing for analog signals,” *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 993–1009, Mar. 2009.

[36] R. A. Lobos, C.-C. Chan, and J. P. Haldar, “New theory and faster computations for subspace-based sensitivity map estimation in multichannel MRI,” *IEEE Trans. Med. Imag.*, vol. 43, no. 1, pp. 286–296, Jan. 2024.

[37] R. A. Lobos, C.-C. Chan, and J. P. Haldar, “Extended version of ‘new theory and faster computations for subspace-based sensitivity map estimation in multichannel MRI,” 2023, *arXiv:2302.13431*.

[38] R. A. Lobos, C.-C. Chan, and J. P. Haldar, “PISCO software version 1.0,” University of Southern California, Los Angeles, CA, USA, Tech. Rep. USC-SIPI-458, Mar. 2023. [Online]. Available: https://sipi.usc.edu/reports/abstracts.php?rid=sipi-458

[39] A. C. Gilbert, J. Y. Park, and M. B. Wakin, “Sketched SVD: Recovering spectral features from compressive measurements,” 2012, *arXiv:1211.0361*.

[40] G. Golub and C. van Loan, *Matrix Computations*, 4th ed. Baltimore, MD, USA: The Johns Hopkins University Press, 2013.

[41] R. A. Lobos, X. Wang, Z. Liu, J. A. Fessler, and D. C. Noll, “Spatiotemporal maps for dynamic MRI reconstruction: A proof-of-principle demonstration on single-coil animal gastrointestinal data,” in *Proc. Int. Soc. Magn. Reson. Med.*, 2025, p. 2622.

[42] M. Uecker et al., “ESPIRiT–an eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA,” *Magn. Reson. Med.*, vol. 71, pp. 990–1001, Mar. 2014.

[43] J. P. Haldar, “Low-rank modeling of local k-space neighborhoods (LORAKS) for constrained MRI,” *IEEE Trans. Med. Imag.*, vol. 33, no. 3, pp. 668–681, Mar. 2014.

[44] J. P. Haldar and J. Zhuo, “P-LORAKS: Low-rank modeling of local K-space neighborhoods with parallel imaging data,” *Magn. Reson. Med.*, vol. 75, no. 4, pp. 1499–1514, Apr. 2016.

[45] R. A. Lobos and J. P. Haldar, “On the shape of convolution kernels in MRI reconstruction: Rectangles versus ellipsoids,” *Magn. Reson. Med.*, vol. 87, no. 6, pp. 2989–2996, Jun. 2022.

[46] R. Vershynin, *High-Dimensional Probability: An Introduction With Applications in Data Science*. Cambridge, U.K.: Cambridge Univ. Press, 2018.

[47] G. Strang, “The fundamental theorem of linear Algebra,” *Amer. Math. Monthly*, vol. 100, no. 9, pp. 848–55, 1993.

[48] J. A. Fessler and R. R. Nadakuditi, “Linear Algebra for Data Science,” in *Machine Learning, and Signal Processing*. Cambridge, U.K.: Cambridge Univ. Press, 2024.

[49] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, “SENSE: Sensitivity encoding for fast MRI,” *Magn. Reson. Med.*, vol. 42, no. 5, pp. 952–62, Nov. 1999.

[50] T. H. Kim, K. Setsompop, and J. P. Haldar, “LORAKS makes better SENSE: Phase-constrained partial fourier SENSE reconstruction without phase calibration,” *Magn. Reson. Med.*, vol. 77, no. 3, pp. 1021–1035, Mar. 2017.

[51] X. Wang et al., “Diffeomorphic surface modeling for MRi-based characterization of gastric anatomy and motility,” *IEEE Trans. Biomed. Eng.*, vol. 70, no. 7, pp. 2046–2057, Jul. 2023.

[52] M. Buehrer, K. P. Pruessmann, P. Boesiger, and S. Kozerke, “Array compression for MRI with large coil arrays,” *Magn. Reson. Med.*, vol. 57, no. 6, pp. 1131–9, Jun. 2007.

[53] H. Xiang, J. A. Fessler, and D. C. Noll, “Model-based reconstruction for looping-star MRI,” *Magn. Reson. Med.*, vol. 91, no. 5, pp. 2104–2113, May 2024.

[54] D. Kim, S. F. Cauley, K. S. Nayak, R. M. Leahy, and J. P. Haldar, “Region-optimized virtual (ROVir) coils: Localization and/or suppression of spatial regions using sensor-domain beamforming,” *Magn. Reson. Med.*, vol. 86, no. 1, pp. 197–212, Jul. 2021.

[55] M. Mani, M. Jacob, V. Magnotta, and J. Zhong, “Fast iterative algorithm for the reconstruction of multi-shot non-Cartesian diffusion data,” *Magn. Reson. Med.*, vol. 74, no. 4, pp. 1086–94, Oct. 2015.

[56] B. Bilgic et al., “Improving parallel imaging by jointly reconstructing multi-contrast data,” *Magn. Reson. Med.*, vol. 80, no. 2, pp. 619–632, Jan. 2018.