RESEARCH ARTICLE

Efficient Blind Source Separation Method for fMRI Using Autoencoder and Spatiotemporal Sparsity Constraints

MUHAMMAD USMAN KHALID[1], BILAL A. KHAWAJA[2], (Senior Member, IEEE), AND MALIK MUHAMMAD NAUMAN[3]

[1]College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University, Riyadh 11564, Saudi Arabia
[2]Faculty of Engineering, Department of Electrical Engineering, Islamic University of Madinah, Madinah 42351, Saudi Arabia
[3]Faculty of Integrated Technologies, Universiti Brunei Darussalam, Bandar Seri Begawan BE1410, Brunei

Corresponding author: Malik Muhammad Nauman (malik.nauman@ubd.edu.bn)

ABSTRACT Diversity measures exploited by blind source separation (BSS) methods are usually based on either statistical attributes/geometrical structures or sparse/overcomplete (underdetermined) representations of the signals. This leads to some inefficient BSS methods that are derived from either a mixing matrix (mm), sparse weight vectors (sw), or sparse code (sc). In contrast, the proposed efficient method, sparse spatiotemporal BSS (ssBSS), avoids computational complications associated with lag sets, deflation strategy, and repeated error matrix computation using the whole dataset. It solves the spatiotemporal data reconstruction model (STEM) with l_1-norm penalization and l_0-norm constraints using Neumann's alternating projection lemma and block coordinate descent approach to yield the desired bases. Its specific solution allows incorporating a three-step autoencoder and univariate soft thresholding for a block update of the source/mixing matrices. Due to the utilization of both spatial and temporal information, it can better distinguish between the sources and yield interpretable results. These steps also make ssBSS unique because, to the best of my knowledge, no mixing matrix based BSS method incorporates sparsity of both features and a multilayer network structure. The proposed method is validated using synthetic and various functional magnetic resonance imaging (fMRI) datasets. Results reveal the superior performance of the proposed ssBSS method compared to the existing methods based on mmBSS and swBSS. Specifically, overall, a 14% increase in the mean correlation value and 91% reduction in computation time over the ssICA algorithm was discovered.

INDEX TERMS Sparse representation, autoencoder, BSS, fMRI, activation maps, PCA, l_0-norm, l_1-norm.

I. INTRODUCTION

Over the last two decades, the statistical parametric mapping (SPM) toolbox's [1] general linear model (GLM) has been extensively used for fMRI analysis [2]. This hypothesis-driven model requires prior knowledge about the experimental paradigm, such as hemodynamic response function (HRF). However, the dynamics of the experiment are sometimes hard to model due to HRF variability across subjects [3] and the absence of stimulus for resting-state experiments. Instead, the data-driven approaches offer a promising alternative due to their adaptability to both task-based activation detection and resting-state functional connectivity analysis [4], [5]. In this regard, some mmBSS methods owing to their computational efficiency, have been very consequential for fMRI studies [6], [7], [8], [9], [10], [11], [12].

The associate editor coordinating the review of this manuscript and approving it for publication was Hasan S. Mir.

Blind signal extraction is a fundamental problem in many signal and image processing applications [13], [14], [15], [16], [17], [18] that aims at revealing hidden structures in the multivariate data through matrix decomposition [19]. While utterly blind to the mixing process, it makes some assumptions about the source signals by imposing constraints on the data matrix. In this regard, principal component analysis (PCA) assumes underlying signals are uncorrelated [20], independent component analysis (ICA) hypothesizes statistical independence [21], canonical correlation analysis (CCA) expects that signals are autocorrelated [22], sparse dictionary learning (SDL) considers that signals are overcomplete [23], partial least square (PLS) infers that sources have high autocovariance [24], and generalized morphological components analysis (GMCA) [25] encourages morphological diversity of signals.

Specifically, mmBSS methods are based on either deterministic or stochastic approaches. Deterministic methods impose geometrical assumptions on signals such as non-negative matrix factorization (NMF) [26] and spatial/temporal correlation [27]. In contrast, statistical methods are concerned with the probabilistic distribution of the signals such as Bayes' theorem assigns probability density function as priors to sources and mixing coefficients [28], and ICA explores the statistical properties of the source signals through higher order statistics [21]. Lower interpretability offered by statistical/geometrical assumptions is usually resolved by imposing sparsity as an additional constraint on the weight vectors, which results in sparse variants of BSS methods (swBSS), as discussed in the next section.

Alternatively, consider scBSS methods due to an overcomplete representation of the sparse signals such as SDL, GMCA, and sparse component analysis (SCA) [29] that naturally circumvent the interpretability problem but their computational cost is enormous. On the other hand, spatial ICA (sICA) has been very popular due to its numerical simplicity, better performance, and fMRI data's lower spatial variations [30]. Previously, it has been argued that independence is not adaptive for fMRI [31], and the sparsity of components rather than independence is a more fruitful assumption. Nevertheless, it was recently concluded that both sparsity and independence are reasonable assumptions for fMRI analysis [32]. More recently, a framework was developed that jointly exploited both source diversities to increase separation accuracy for fMRI [33].

Analyzing temporal and spatial information separately had been the convention, whereas exploiting both features concurrently may lead to better performance. Some mmBBS methods have attempted this strategy in their analysis. For instance, ECG specific BSS technique used spatial ICA to remove non-Gaussian interference followed by second-order blind identification (SOBI) [34] step to suppress temporal Gaussian noise [35], a low-dimensional CCA approach for fMRI that incorporated

straight forward image projection into temporal BSS model to exploit autocorrelations of both temporal and spatial features concurrently [36], and an SDL based method for multi-subject fMRI analysis that attempted to utilize information from both dimensions in terms of principal components and loadings [37].

More recent application-specific BSS methods are also worth mentioning. For instance, a hybrid approach was proposed that combined a variational autoencoder and bandpass filtering for speech processing [38], an efficient BSS method was proposed that utilized multi-channel Wiener filtering and joint diagonalization [39] for audio signal processing [40], two smooth variants of NMF namely smooth successive projection algorithm and smooth vertex component analysis have been presented for hyperspectral unmixing [41], and a novel BSS approach named semi-blind GMCA (sGMCA) was introduced that combined alternating least square algorithm with learning based regularization for hyperspectral imaging [42].

Some computationally efficient mmBSS methods, such as regular PCA and CCA, perform poorly for BSS. On the other hand, sparse variants of BSS strive to improve separation accuracy at the cost of the increased numerical burden. This is due to their reliance on either deflation strategy, iterative thresholding, repeated eigen/singular value decomposition (SVD), or error matrix computation of huge covariance/data matrices. This paper reached a compromise between the two scenarios by separating all source signals simultaneously in a transformed space through a sparse approach. Moreover, motivated by this discussion and the effectiveness of concurrent feature extraction [37], a new method named ssBSS has been proposed for fMRI data. The main contributions are

1) A novel approach that attempts to solve the proposed model in a computationally efficient manner through an autoencoder, BSS theory exploitation, and continuous but implicit training from the data.
2) Utilizing Neumann's successive projection lemma [43], each of the two sub-problems from the penalized and constrained STEM model is solved in an alternating least square manner.
3) A comprehensive validation of the proposed method and its comparison with existing mmBSS and swBSS methods using synthetic and experimental data.
4) In-depth analysis using performance metrics such as similarity measure, convergence analysis, and numerical cost.

The rest of the paper is organized into five sections where section II highlights the related work, section III provides some background on sparse BSS methods, section IV presents some direction on mathematical notations followed by the discussion on the proposed method and its related algorithm, section V produces experimental study, and section VI contains the concluding remarks.
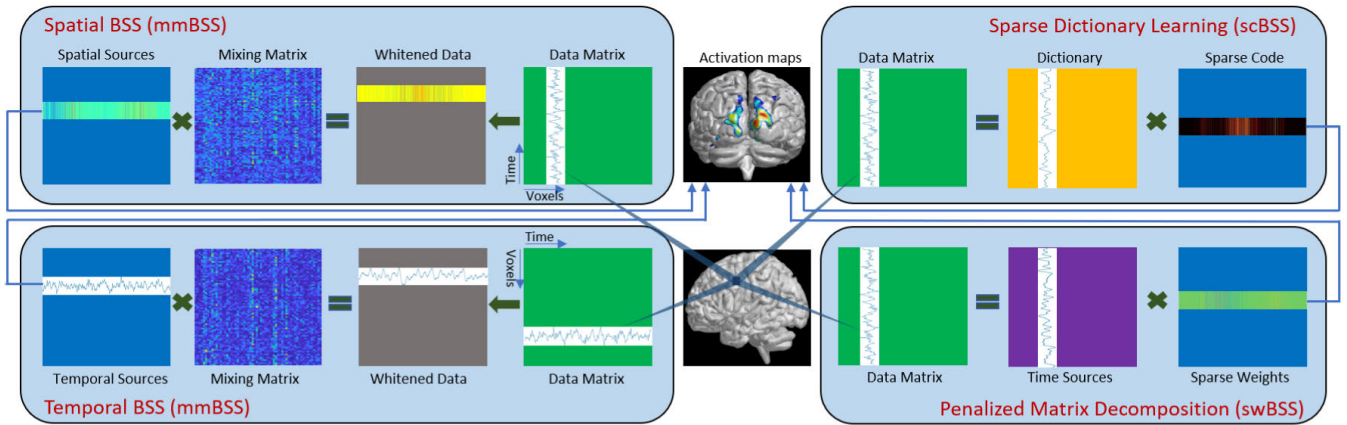
FIGURE 1. A pictorial description of different blind source separation strategies. The rendered brain is obtained using the BrainNet Viewer [44].

II. RELATED WORK

As shown in Fig. 1, BSS methods generally employ three strategies. These are

1) **mmBSS:** The whitening operation followed by a non-sparse mixing/rotation matrix estimation.
2) **swBSS:** Utilization of full data matrix to estimate sparse weight matrix through deflation.
3) **scBSS:** Computation of the error matrix using full data matrix to estimate dictionary and sparse code through alternating minimization.

State-of-the-art swBSS methods that have been quite successful for neuroimaging data include sparse PCA (sPCA) [45], sparse CCA (sCCA) [46], [47], sparse PLS (sPLS) [48], [49], multi-task sparse model (MTS) [50], and sparse two dimensional CCA (s2DCCA) [51]. On the other hand, improved fast ICA (ifICA) utilizes a fast ICA algorithm and l_0-norm [52] and sparse spatial ICA employs an entropy bound minimization algorithm (ssICA) and l_1-norm [33]. They both consist of the step-wise optimization approach for implementing independence and sparsity.

Sparse PCA improved the interpretation of original principal components by reformulating eigen-vector decomposition as a penalized regression problem to obtain sparse loadings [45]. Sparse CCA extends the sPCA strategy and adapts the elastic net to CCA to get sparse canonical weights [46]. Sparse PLS discovered an improved relationship between two datasets by a penalized matrix decomposition applied to cross matrix product to obtain sparse weight vectors [49]. Based on the l_0 norm, the improved fast ICA method was proposed by incorporating the sparsity constraint via Gaussian kernel to the fast ICA algorithm [52].

A regularized and a sparse CCA employed penalized rank-1 matrix approximation on the product of the orthogonal projectors of two sets of data to improve interpretability and performance of the canonical variates [47]. To simultaneously handle the sparse decomposition of mixed signals, the BSS problem was reformulated as a multi-task sparse problem that exploited the connection between tasks to achieve higher

retrieval accuracy [50]. While preserving the correlation structure of fMR images, two variants of s2DCCA, regularized and sparse, were proposed to improve the interpretability of the projected variables [51].

III. BACKGROUND

According to the uniquely/overdetermined case of BSS theory, where more or an equal number of mixtures than sources are observed, let an observed vector $y = [y_1, y_2, \ldots, y_N]^\top$ be a linear mixture of the unknown source vector $s = [s_1, s_2, \ldots, s_k]^\top$, i.e., $y = T^\top x$, where $T \in \mathbb{R}^{K \times N}$ is a transpose of the mixing matrix, and its size is $N \geq K$. This can be more formally expressed for multiple sources as

$$Y = T^\top S + E \qquad (1)$$

where E is the model error. When implementing (1) using ICA or CCA, prewhitening the data is the first step. This is carried out to reduce the correlation between features and dimensionality of the data. Only $D < N$ principal components are kept, followed by the estimation of mixing matrix $T^\top \in \mathbb{R}^{D \times K}$ as the second step as shown in Fig. 1.

From a general BSS perspective, there are two possible outcomes i) spatial BSS would require mixture matrix $Y \in \mathbb{R}^{N \times V}$ to be obtained by assembling time courses of length N across V voxels resulting in a K number of retrieved maps in the source matrix $S = TY \in \mathbb{R}^{K \times V}$, where T is assumed to be a semi-orthogonal demixing matrix whose pseudo-inverse is equal to its transpose. As an alternative, temporal BSS considers collecting mixture matrix as $Y^\top \in \mathbb{R}^{V \times N}$ to obtain K number of source signals as $T = SY^\top \in \mathbb{R}^{K \times N}$, where S is a semi-orthogonal demixing matrix.

The BSS model given by equation (1) can be extended to a sparse model, which is formulated as [53]

$$||y_i - T^\top s_i||_2^2 + \lambda f(s_i), \quad i = 1, \ldots, V \qquad (2)$$

where V is the number of voxels, $||y_i - T^\top s_i||_2^2$ is a smooth and convex data fidelity term and $f(.)$ is a convex regularization function that promotes sparsity, and λ controls the trade-off
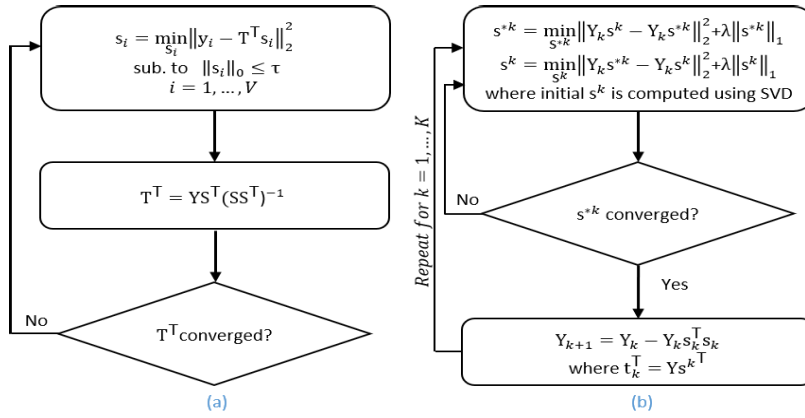
FIGURE 2. An algorithmic description of different sparse blind source separation approaches.

between sparsity and reconstruction error. This leads to the extraction of source signals from the observed signals in an overcomplete manner; however, this model during sparse coding decomposes each mixture signal in Y independently, resulting in an increased learning complexity as shown in Fig. 2a. One can refer to [54] for a block update of $T^\top$ and [23] for a sequential update.

Consider on the other hand deflation based sparse methods [46], [47], [48], [49], [55] that exploit a different fidelity term given by $||Y - \sum_{k=1}^{K} t_k^\top s^k||_F^2$, where K is the number of sources as shown in Fig. 2b. This results in reliance on estimating one source signal at a time and applying deflation to the data matrix to remove the effect described by the current sparse mixing vector. This creates learning inefficiency, just like SDL algorithms given by equation (2), and lowers the retrieval accuracy because this sort of penalized matrix decomposition ignores the important relationship between the source signals.

The proposed methodology described in the next section eliminates the shortcomings mentioned in the last two paragraphs. While performing source separation, the relationship among source signals was taken into account by continuously updating all sources until convergence. This resulted in better interpretability of retrieved sources compared to the deflation strategy. On the other hand, sparse assumption and reduced-dimension space allowed us to use equation (1) indirectly for the source separation problem, which resulted in a significant computational gain.

IV. METHODS

Throughout this paper, small italic or capital italic letters characterize the scalar values, small boldface letters signify vectors, and capital boldface letters are utilized to depict matrices. The specific column and row of the matrix are indicated by vectors attached with subscripts and superscripts, respectively. The error matrix for the k-th temporal source and its respective coefficient row is represented by the error matrix attached with a subscript k.

For the proposed method, equation (1) is extended and refined to produce a STEM model. Specifically, consider an example of fMRI blood-oxygen-level-dependent (BOLD) time series, which consists of N scanned volumes and V number of voxels arranged along the column direction in $Y = [y_1, y_2, \ldots, y_V] \in \mathbb{R}^{N \times V}$. It is assumed that due to sparseness along the row direction, each signal in Y can be represented as a linear combination of temporal sources from T according to the sparse coefficient strength in each column of the spatial source matrix S. Suppose that the BOLD signal at any voxel can be assumed smooth due to a neural response that is smoothed by the HRF [56]; then this can be accounted for as $T^\top = T_p A$. To achieve this decomposition, a sparsity constraint is imposed on coefficient rows and corresponding columns of the sparse representation matrix $A \in \mathbb{R}^{K_p \times K}$ [57] as

$$\min_{A,S} \left\| Y - T_p AS \right\|_F^2 + \lambda \left\| S \right\|_1,$$
$$\text{sub.to.} \left\| T_p a_k \right\|_2 = 1, \quad \left\| a_k \right\|_0 \leq \zeta \quad (3)$$

where $T_p \in \mathbb{R}^{N \times K_p}$ is the matrix containing DCT basis, and a_k is the k-th column of A, $K < K_p < N$, $\|.\|_0$ is the l_0 norm that induces sparsity by counting the number of non-zero elements, $\|S\|_1$ is the l_1 norm on **S** and is given as $\sum_{k=1}^{K} \sum_{j=1}^{V} |s_j^k|$, and λ is the sparsity parameter that controls coefficient values in the k-th row. Associated l_0 norm constrained least square is not part of the regularized problem and is solved separately [58]. One way to solve this problem is by computing error matrix of full data [56], [59], [60] to update the unknowns in a sequential manner as

$$\{a_k, x^k\} = \arg\min_{a_k, x^k} \left\| E_k - T_p a_k x^k \right\|_F^2$$

where $E_k = Y - \sum_{i=1, i \neq k}^{K} t^{i^\top} x^i$. This is another example of inefficient learning because it calls dataset L times for each of the K dictionary atoms/sparse codes, creating an analytic complexity of $4LK$ and computational complexity of approximately $LK(KNV + 3NV + \zeta^3 + K_p N)$ [56]. Because the proposed algorithm operates in a block manner, makes lesser

TABLE 1. Algorithm for spatiotemporal dimensionality reduction.

Given:
Training data ($Y \in \mathbb{R}^{N \times V}$)
Activation function (G)
Constants (K, L)

1. **Initialization**:
   For a given number of hidden neurons, generate the initial orthonormal weights/biases $^f A /^f b$ and temporal/spatial feature data $^f H /^f F$ according to equation 4
2. **Repeat** for $i = 1$ to L:
   a) Hidden nodes $^j A$ and $^j b$ are obtained in the decoding layer using least squares and invertible activation function given by equation 5
   b) Spatiotemporal feature data is generated using hidden nodes $^f A$ and $^f b$ from the previous layer and equation 6

Output: $^f H$ and $^f F$

TABLE 2. Algorithm for realizing l_0 constraint.

Given:
Matrix ($T \in \mathbb{R}^{P \times N}$)
Parameters (K_p, ζ)

Repeat for $p = 1$ to P:
   a) Compute correlation vector as
   $$c \leftarrow sort\left(\left|T_p^\top \left(X_t^\top u_p\right)\right|, descend\right)$$
   b) After sorting correlation vector in descending order keep ζ-largest values as
   $$i_s \leftarrow \{c_r\}_{r=1}^{\zeta}$$
   c) Compute representation vector as
   $$a_{p,i_s} \leftarrow (T_{p,i_s}^\top T_{p,i_s})^{-1} T_{p,i_s}^\top X_t^\top u_p$$
   d) Normalize as
   $$a_p \leftarrow a_p / \|T_p a_p\|_2$$

Output: A

calls to data, and performs lesser computations, the total complexity is brought down to approximately $2LK(K^2 + NV)$.

A. SPATIOTEMPORAL AUTOENCODER

A spatiotemporal PCA can be achieved by performing SVD on the data matrix Y to decompose it as

$$Y = \Omega \Delta \Gamma^\top$$

where $\Delta \in \mathbb{R}^{M \times M}$ consists of singular values on the diagonal, and matrices $\Omega \in \mathbb{R}^{N \times M}$ and $\Gamma \in \mathbb{R}^{V \times M}$ consists of M number of left and right singular vectors, respectively. Using these singular matrices that contain the eigenvectors of the temporal and spatial covariance matrices, one can obtain the reduced spatial and temporal feature matrices $X_t \in \mathbb{R}^{K \times N}$ and $X_s \in \mathbb{R}^{K \times V}$ by keeping only the first $K < M$ singular vectors with maximum variance as

$$X_t = \Delta \Gamma^\top Y^\top, \quad X_s = \Omega^\top Y$$

Mathematically, this is same as $X_t = \Omega^\top$, and $X_s = \Delta \Gamma^\top$. The objective of PCA is to encapsulate most information from the data into a few components. This may result in severe noise contamination of significant components because the hemodynamic response signal of fMRI is usually weak compared to the noise signals [8]. One way to resolve this is to order components by their autocorrelation profile rather than maximum variance, but producing corresponding spatial components is challenging. Due to this limitation and computational inefficiency associated with SVD operation, an autoencoder was introduced as an alternative to PCA.

A three-step temporal autoencoder [61] has been deployed for the proposed method. It is an extension of an extreme learning machine equipped with a feedback mechanism to pull the output data back to hidden nodes to enrich and accelerate the learning process. It is adapted for the proposed ssBSS case to learn the respective spatial features along with temporal ones. It is used as a dimensionality reduction step, and preferred over PCA-based whitening operation due to its learning efficiency and overcomplete nature. It atleast can preserve the quantity of the input information, and subsequent steps of the proposed method can improve quality.

It is a two-layer network where input is transferred to low-dimensional random feature space followed by the computation of hidden nodes using invertible functions and, finally, re-estimating low-dimensional feature space by pulling the weights of the decoding layer back to the hidden nodes. For a given data matrix $Y \in \mathbb{R}^{N \times V}$, orthogonal random weight matrix $^f A \in \mathbb{R}^{K \times V}$ and bias vector $^f b \in \mathbb{R}^K$ of the encoding layer, the first step constructs the first layer of the network through the orthogonal random process for (N, V) number of input neurons and (K, K) number of hidden neurons given as

$$^f H = G\left(^f A, ^f b, Y^\top\right),$$
$$^f A {}^f A^\top = I, \quad ^f b^\top {}^f b = 1,$$
$$^f F = G\left(^f H, Y\right). \tag{4}$$

where $\left(^f H, Y\right) = {}^f H({}^f H^\top {}^f H + \beta_N I_N)^{-1} Y$, and G is any general hidden neuron function such as sigmoid, wavelet, radial basis, and sinusoid. In the second step, invertibility of the activation function is exploited to obtain weights $^j A \in \mathbb{R}^{K \times N}$, $^j C \in \mathbb{R}^{K \times V}$ and biases $^j b \in \mathbb{R}^V$, $^j d \in \mathbb{R}^N$ of the decoding layer as

$$^j A = {}^f H^\dagger \left(\sin^{-1}(Y^\top)\right)^\top \quad if \quad G(.) = \sin(.),$$
$$^j C = -{}^f F^\dagger \left(\log\left(\frac{1}{Y} - 1\right)\right)^\top \quad if \quad G(.) = \frac{1}{1 + e^{-(.)}},$$
$$^j B = \sqrt{mse\left({}^f H^\top {}^j A - \left(\sin^{-1}(Y^\top)\right)^\top\right)},$$
$$^j D = \sqrt{mse\left({}^f F^\top {}^j C + \left(\log\left(\frac{1}{Y} - 1\right)\right)^\top\right)}. \tag{5}$$

where the sigmoid transfer function is just a proposition, and sinusoid can be used for the spatial case as well and vice-versa, $^f H^\dagger = ({}^f H^f H^\top + \beta_K I_K)^{-1} {}^f H$ and $^f F^\dagger = ({}^f F^f F^\top + \beta_K I_K)^{-1} {}^f F$. In the third step, feature data is updated using weights $^f A = {}^j A, {}^f C = {}^j C$ and biases $^f b = {}^j b_1 \in \mathbb{R}, {}^f d =$

TABLE 3. Algorithm for solving the minimization problem (3).

Given:
Training data ($Y \in \mathbb{R}^{N \times V}$)
Tuning parameters ($\lambda_1, \lambda_2, \lambda_3, \zeta, K_p$)
Constants ($K, P, L, l = 0, \tau = 0.05$)

1. **Initialization:**
   a) Initialize T using normally distributed random numbers and normalize its rows
   b) Update S using equation (1)
   c) Initialize U and A with zeros
   d) Initialize T_p using DCT basis
2. **Spatiotemporal Component Extraction:**
   Decompose the data matrix using SVD or algorithm given in Table 1
3. **Repeat while** $\left\| T^\top - T_l^\top \right\|_F / \left\| T_l^\top \right\|_F > \tau$:
   $l \leftarrow l + 1$
   $T_l^\top \leftarrow T^\top$
   a) Using S update T through equation (1)
   b) Normalize each row of T
   if update = 'block'
      c) Fix A, and update U by solving
      $\min_{U} \left\| X_t^\top - T_p A U^\top \right\|_F^2 + \lambda_1 \left\| U \right\|_1$, using equation (12)
      d) Fix U, and update A using algorithm given in Table 2 and compute $T^\top \leftarrow T_p A$
   else
      for $p \leftarrow 1$ to P
         c) $u_p^\top \leftarrow \text{sgn}\left(a_p^\top T_p^\top E_p^\top \right) \circ \left(|a_p^\top T_p^\top E_p^\top| - \frac{\lambda_{1p}}{2} \right)_+$
         d) $a_p \leftarrow \arg\min_{a_p} \left\| E_p^\top u_p - T_p a_p \right\|_2^2 \quad \text{sub.to. } \left\| a_p \right\|_0 \leq \zeta$
      end
   end if
   e) Using T update S through equation (1)
   f) Fix S, and update W by solving
   $\min_{W} \left\| X_s - WS \right\|_F^2 + \lambda_2 \left\| W \right\|_1$, using equation (16)
   g) Fix W, and update S by solving
   $\min_{S} \left\| X_s - WS \right\|_F^2 + \lambda_3 \left\| S \right\|_1$, using equation (17)

Output: T and S

$^j d_1 \in \mathbb{R}$ from the decoding layer as

$$^f H = G\left({}^f A, {}^f b, Y^\top \right),$$
$$^f F = G\left({}^f C, {}^f d, Y \right). \tag{6}$$

Steps 2 and 3 are repeated several times until convergence; however, we found that the computation of feature data $^f H$ just twice is sufficient to transform the data from high to low dimensional space while retaining the meaningful information from the observed data. The detailed description of these steps is given in Table 1 in the form of an algorithm.

B. ssBSS

Using blind source separation theory, one can obtain the following pair of mixture models given as

$$X_t^\top = T^\top U^\top, \quad X_s = WS \tag{7}$$

where the unknowns U, $W \in \mathbb{R}^{K \times P}$ are the mixing matrices whereas $T \in \mathbb{R}^{P \times N}$ and $S \in \mathbb{R}^{P \times V}$ are the source matrices, and $K \geq P$. For a non-orthogonal demixing matrices, equation (7) and (1) can be manipulated together to obtain

the unknowns W and U as

$$W = X_s (Y^\top Y + \beta_V I_V)^{-1} Y^\top X_t^\top (UU^\top + \beta_K I_K)^{-1} U,$$
$$U^\top = (W^\top W + \beta_P I_P)^{-1} W^\top X_s (Y^\top Y + \beta_V I_V)^{-1} Y^\top X_t^\top$$

where $\beta_K I_K$ and $\beta_P I_P$ are the Tikhonov regularization terms to avoid overfitting and singularity of $UU^\top$ and $W^\top W$, respectively [55]. From the above equation, one can infer that taking the inverse of the data's covariance matrix is cumbersome. Therefore, it would be more convenient to indirectly exploit the information from the observed mixture matrix Y to estimate the source signals. Specifically, the source estimates S and T obtained directly from the entire data matrix can be used to estimate mixing matrices and then re-estimate the source matrices. This approach will produce an optimal solution by properly unmixing the PC/autoencoder components. For instance, by imposing sparsity on the matrices W, U, S, and T, the orthogonality condition of the mixing matrices, e.g., $W^\dagger = W^\top$ can be violated and results far superior to PCA/autoencoder can be obtained. This is because dimensionality reduction methods' spatial and temporal overfitting can be retracted.

Therefore, we extend this idea to efficiently solve the non-smooth constrained optimization problem given in equation (3). For this purpose, consider solving equation (3) using (7), which breaks (3) into following pair of problems

$$\min_{A,U} \left\| X_t^\top - T_p A U^\top \right\|_F^2 + \lambda_1 \left\| U \right\|_1,$$
$$\text{sub.to. } \left\| T_p a_p \right\|_2 = 1, \quad \left\| a_p \right\|_0 \leq \zeta \tag{8}$$
$$\min_{S,W} \left\| X_s - WS \right\|_F^2 + \lambda_2 \left\| W \right\|_1 + \lambda_3 \left\| S \right\|_1 \tag{9}$$

According to Neumann's alternating projection strategy [43], these two subproblems can be solved iteratively and alternately until the convergence of the unknowns, and they will ultimately result in an optimal solution of equation (3) while mutually converging to the subspace formed by them. This can be accomplished by letting T estimated from equation (8) to play its role in optimizing (9) and using S estimated from equation (9) in solving (8) through their mutual connection given by equation (1). Therefore, temporal sources obtained from the above problem can then be plugged into equation (1) to obtain the spatial sources and vice versa.

Due to l_1 norm in (3), proximal splitting methods [62] are deployed to minimize functions that are not differentiable by splitting the cost function into a sum of convex functions that are minimized in an alternating manner. For instance, one way to approach this problem is through gradient descent, giving rise to the thresholded Landweber problem given as

$$s_k^{n+1} = M_\lambda \left(s_k^n - T(T^\top s_k^n - y_k) \right),$$
$$M_\lambda(z) = \text{sgn}(z) \circ \left(|z| - \frac{1\lambda}{2} \right)_+$$

where $(z)_+$, sgn(.), and $\circ$ define the component-wise max between (0, z), the component-wise sign, and the Hadamard product, respectively [63]. For the proposed ssBSS method,
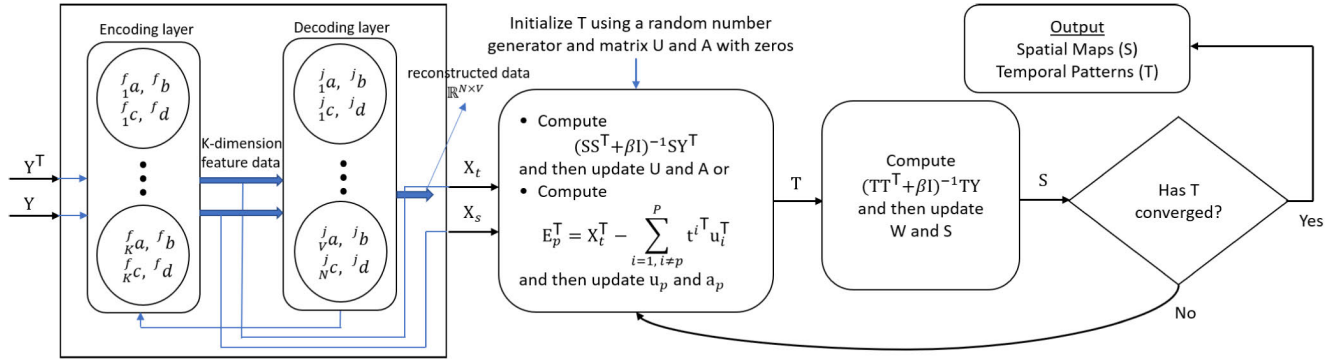
FIGURE 3. A block diagram of the proposed method where the autoencoder performs dimensionality reduction and the rest of the blocks attempt to recover underlying spatial and temporal sources.
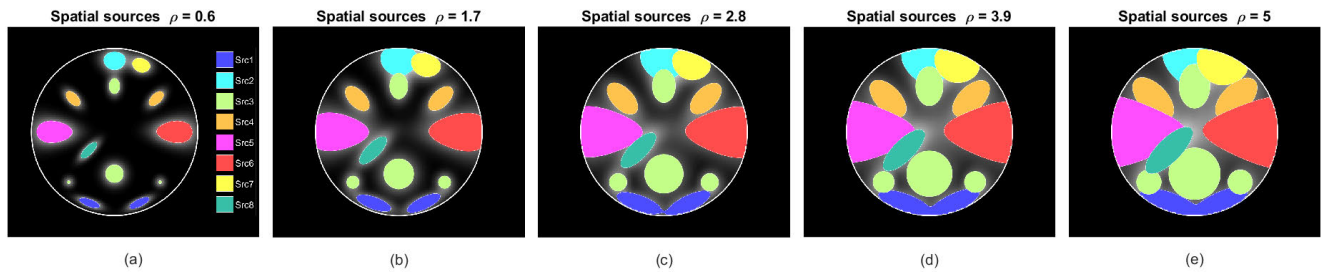

FIGURE 4. The size of 8 different activation blobs controlled using spread parameter ρ to create none to significant spatial overlap.

the sparsity pattern, however, has been imposed row-wise, then $s^{k,n+1} = M_\lambda\left(s^{k,n} - t_k(t_k^\top s^{k,n} - Y)\right)$. It is speculated that similar conditions, that is, when $T^\top$ holds the finite basis injective property and $s^{\dagger k}$ possesses a strict sparsity pattern, linear convergence rate would be achieved as $s^{k,n} \to s^{\dagger k}$ such that $\left\|s^{k,n} \to s^{\dagger k}\right\| \leq C\alpha$ where $C > 0$ and $0 < \alpha < 1$ [64].

Similarly, it is also postulated that both mixing matrices must hold i) strict sparsity pattern because none of their columns can be completely zero for any case, and ii) both source matrices S/T hold unique sources/basis that cannot be constructed as a linear combination of the other two. When a scenario such as a bad source occurs (whose spatial source consists of all zeros), the respective temporal source is replaced with the least represented normalized data element from the mixture matrix. Before attempting an efficient solution, equation (8) can be re-written for a sequential update as

$$\min_{a_p, u_p^\top} \left\|E_p^\top - T_p a_p u_p^\top\right\|_F^2 + \sum_{k=1}^K \lambda_{1p}^k |u_p^k|,$$

$$\text{sub.to. } \left\|T_p a_p\right\|_2 = 1, \quad \left\|a_p\right\|_0 \leq \zeta$$

where $E_p^\top = X_t^\top - \sum_{i=1, i\neq p}^P t^{i^\top} u_i^\top \in \mathbb{R}^{N \times K}$. Then the following closed-form solution can be reached by solving this

equation with respect to the unknowns as shown in [56]

$$u_p^\top = \text{sgn}\left(a_p^\top T_p^\top E_p^\top\right) \circ \left(|a_p^\top T_p^\top E_p^\top| - \frac{\lambda_{1p}}{2}\right)_+,$$

$$a_p = \arg\min_{a_p} \left\|E_p^\top u_p - T_p a_p\right\|_2^2 \quad \text{sub.to. } \left\|a_p\right\|_0 \leq \zeta \quad (10)$$

where $\lambda_{1p} = [\lambda_{1p}^1, \ldots, \lambda_{1p}^k]$. Although this error matrix based solution in combination with equation (9) is not based on a full data matrix, it is still only efficient for the synthetic data. As the size of the components (K) increase, it becomes an expensive update, and therefore we shift our focus toward the block update for all experimental data. Taking the Lagrangian of equation (8) provides

$$L(A, U) = X_t X_t^\top - 2X_t^\top UT + UTT^\top U^\top + \lambda_1 \|U\|_1 \quad (11)$$

Ignoring penalization on the mixing matrix and after some manipulations, a block update for U can be obtained as

$$U = X_t T^\top (TT^\top + \beta_P I_P)^{-1} \quad (12)$$

Now taking penalization of U into account, univariate soft thresholding can be performed on it as $u_p = \text{sgn}(u_p) \circ (|u_p| - \frac{1\lambda_1}{2})_+$, and block update for **T** is obtained as

$$T^\top = X_t^\top U(U^\top U + \beta_P I_P)^{-1} \quad (13)$$

where 1 is a vector of ones of length K. Using (13) an update for p-th column of A can be obtained as $a_p = (T_p^\top T_p)^{-1} T_p^\top X_t^\top u_p$. Using this solution of a_p and imposing
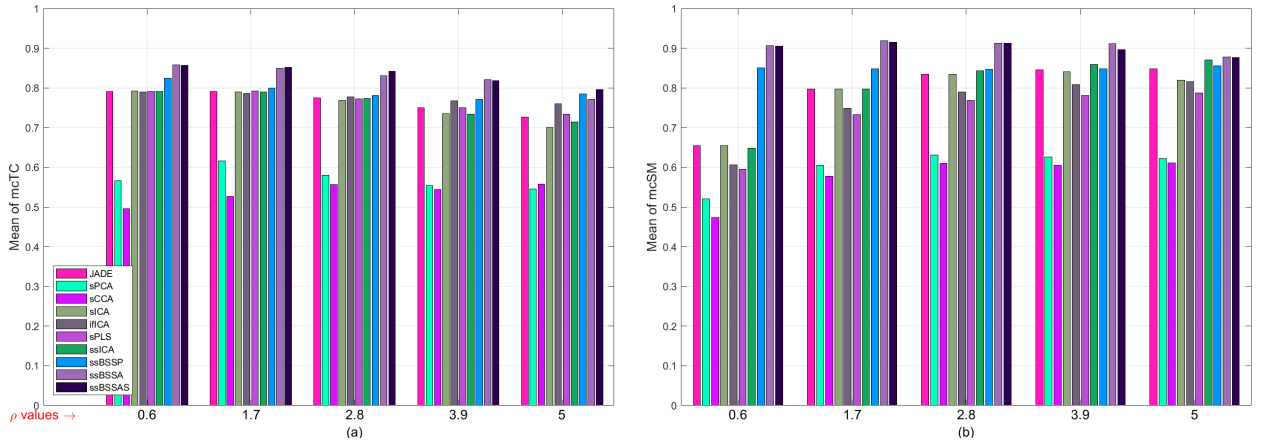
FIGURE 5. Mean values of a) mcTC and b) mcSM over 100 trails where sources are recovered with respect to GT-SMs.

l_0 norm constraint on it gives rise to the following constrained problem given as

$$a_p = \arg\min_{a_p} \left\| X_t^\top u_p - T_p a_p \right\|_2^2 \quad \text{sub.to.} \quad \left\| a_p \right\|_0 \leq \zeta \quad (14)$$

Solution for (14) is $a_{p,i_s} = (T_{p,i_s}^\top T_{p,i_s})^{-1} T_{p,i_s}^\top X_t^\top u_p$, where finding the indices set i_s using correlation based thresholded values [56], [65], [66] is the main task as described in Table 2 in the shape of an algorithm. Equation (9) can be written as

$$L(S, W) = X_s^\top X_s - 2X_s^\top WS + S^\top W^\top WS + \lambda_2 \|W\|_1 + \lambda_3 \|S\|_1 \quad (15)$$

While disregarding penalization on mixing matrix, from (15) an update for W is given as

$$W = X_s S^\top (SS^\top + \beta_P I_P)^{-1} \quad (16)$$

To accommodate penalization on W, take into account the soft thresholding given by $w_p = \text{sgn}(w_p) \circ (|w_p| - \frac{1\lambda_2}{2})_+$. Similarly, S, while ignoring penalization on its entries, can be updated as

$$S = (W^\top W + \beta_P I_P)^{-1} W^\top X_s \quad (17)$$

and its soft thresholding is implemented as $s^p = \text{sgn}(s^p) \circ (|s^p| - \frac{1\lambda_3}{2})_+$. Using equation (12, (14), (16), and (17)) a block update for four unknowns U, A, W, and S can be obtained by an alternating minimization approach where one variable is kept fixed while other one is updated. The related optimization steps are given as an algorithm in Table 3 and in terms of a flow chart in Fig. 3. The soft thresholding for these matrices, which is applied row/column-wise, can also be realized in a block manner for a threshold $\upsilon > 0$ and the matrix M as $(\mathcal{T}_\upsilon(M))_i^j = \text{sgn}(m_i^j) \circ (|m_i^j| - \frac{\upsilon}{2})_+$.

V. EXPERIMENTS
To validate the proposed ssBSS method, its evaluation is carried out in this section. It consists of several comparative studies implemented using a synthetic dataset, three

experimental fMRI datasets, and existing BSS algorithms based on mixing matrices and sparse weights strategies. The participating algorithms are introduced in Table 4 along with their code availability. These are JADE, sICA, sPCA, ifICA, ssICA, sPLS, sCCA, and the variants of the proposed method i) ssBSSP, ii) ssBSSA, and iii) ssBSSAS.

The synthetic dataset was prepared using the Simtb toolbox [67], the block design fMRI dataset of a single subject was acquired from the quarter 3 release of the Human Connectome Project (HCP) [68], [69], the resting-state dataset of a single subject was obtained from the first set of the S900 release of the HCP [68], [69], and the event-related dataset of a single subject was provided by the Korea Advanced Institute of Science and Technology (KAIST) and has been previously used in other papers [70], [71], [72]. Using these datasets, all participating BSS algorithms are compared to assess their ability to recover the ground truth.

Before commencing with the analysis, it is imperative to mention the following specifics, which are common to all datasets:

1) For the variants of the proposed method, the matrix T was generated by drawing random numbers from the standard normal distribution when using both synthetic and experimental datasets.
2) Each algorithm's tuning parameter settings were managed by trying their different combinations. The ones that produced the best results in terms of the sum of the correlation strength between the recovered and the ground truth sources were considered.
3) It was aimed to preserve most of the information in the data when performing dimensionality reduction by PCA or autoencoder. For this purpose, trying different values and selecting the best one was preferred over the model selection approach such as Akaike's Information criterion [73].
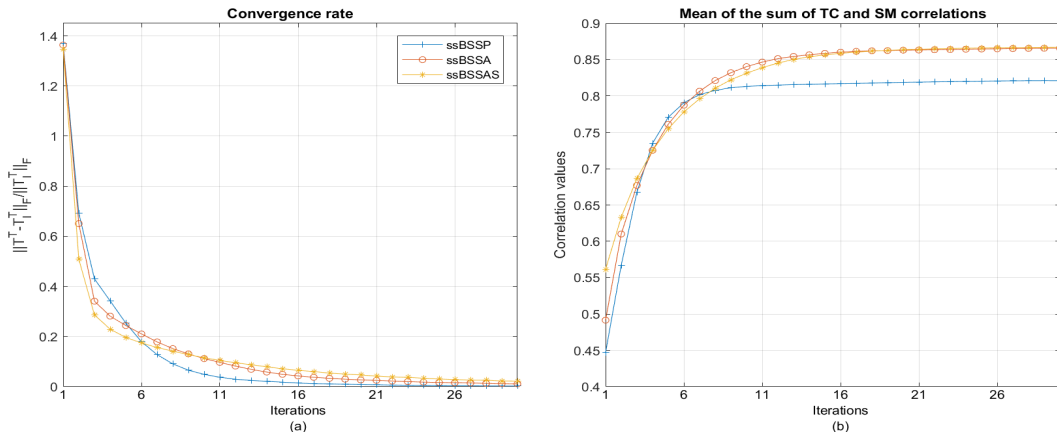4) While selecting the total number of DCT basis and their associated sparsity parameter for the proposed method,

FIGURE 6. For all variants of the proposed method, a) convergence rate, and b) the mean of the sum of the highest correlations between ground truth SMs/TCs and recovered sources shown as a function of algorithm iterations.

TABLE 4. A summary of the BSS algorithms that have been included in the comparative study.

Algorithms	Approach	Matlab code availability
JADE [39]	Data whitening followed by the computation of fourth-order cumulants and rotation matrix	http://www2.iap.fr/users/cardoso/compsep_classic.html
sICA [21]	Data whitening followed by the gradient descent method to estimate the demixing matrix	https://research.ics.aalto.fi/ica/fastica/
sPCA [45]	Singular value decomposition followed by the estimation of sparse weights	http://www2.imm.dtu.dk/projects/spasm/
ifICA [52]	Two-step optimization strategy i) maximize independence, ii) refine the results in the previous step using the sparsity constraint	https://bit.ly/3nGxxUl
ssICA [33]	Three-step optimization approach i) initial estimate of the demixing matrix, ii) refinement through orthogonal ICA, iii) further refinement through non-orthogonal ICA	https://zoisboukouvalas.github.io/Code.html
sPLS [49]	Utilize rank-1 decomposition of the covariance matrix to estimate sparse weights	
sCCA [46]	Utilize iterative partial least square and soft thresholding to estimate sparse weights	https://pypi.org/project/cca-zoo/
Proposed (ssBSSP)	Undercomplete basis obtained through dimensionality reduction using SVD/PCA followed by the block update of sparse mixing/source matrices	
Proposed (ssBSSA)	Overcomplete basis obtained through dimensionality reduction using autoencoder followed by the block update of sparse mixing/source matrices	https://github.com/usmankhalid06/ssBSS_2023
Proposed (ssBSSAS)	Overcomplete basis obtained through dimensionality reduction using autoencoder followed by the sequential update of temporal sparse mixing/source matrix and the block update of spatial sparse mixing/source matrix	

our approach relied on a bias-variance tradeoff to avoid preserving too much information and ignoring essential features in the temporal source. For this purpose, several bases were tried between 60 and N before reaching an optimal number.

5) For sPLS and sCCA, the second multivariate data set was generated using a lagged version of SVD's left singular vectors.

6) The number of underlying sources is unknown for experimental fMRI data, so their selection was based on trial and error. Several numbers ranging between 30 and 60 were tested, and the ones that gave the best results were used for each algorithm.

7) Because sICA and ssICA gave different results and computation times when applied repeatedly on the same dataset, they were applied ten times on each

experimental data. The run with best results was acknowledged, and the mean computation time over ten runs was considered.

8) The MRIcron toolbox [74] is used for all brain image plots of experimental fMRI dataset.

A. SYNTHETIC DATASET
1) SYNTHETIC DATASET GENERATION

In order to produce a dataset that mimics the experimental fMRI signals, the Simtb toolbox was employed. It allowed constructing eight distinct temporal and spatial sources for synthetic dataset generation. Each time source consisted of 240 points with a repetition time (TR) of 1 sec, and each spatial source comprised a square image of 150×150 voxels. The spatial components were generated by these source IDs {3, 6, 8, 10, 22, 23, 26, 30}. Activations within the image

slices were created by combining Gaussian distributions parameterized by location, orientation, and spread. The spatial extent of these activations was controlled using the spread parameter (ρ), which is the mean of the Gaussian distribution $\sim \mathcal{N}(\rho, 0.01^2)$. The spatial maps (SMs) with none to significant spatial overlaps are shown in Fig. 4a-e, and their respective time courses (TCs) are shown in Fig. 7, which are also considered as the ground truth (GT) TCs and SMs when retrieving the underlying sources.

The reason for including 4a is to create a scenario with no spatial dependence but considerable sparseness of the spatial sources. These sources were then utilized to generate a synthetic dataset using a linear mixture model given as $Y = \sum_{i=1}^{8}(\text{tc}_i + \psi_i)(\text{sm}^i + \phi^i)$, where the noise generating matrices $\Psi \in \mathbb{R}^{240 \times 8}$ and $\Phi \in \mathbb{R}^{8 \times 22500}$ were produced using Gaussian distribution $\sim \mathcal{N}(0, 0.6)$ and $\sim \mathcal{N}(0, 0.01)$, respectively, TC $\in \mathbb{R}^{240 \times 8}$ contains the time courses, and SM $\in \mathbb{R}^{8 \times 22500}$ contains the spatial maps obtained after reshaping the image slices. Depending on the value of ρ, the resulting dataset Y was then generated and used for source retrieval by all algorithms.

2) SYNTHETIC DATASET SOURCE SEPARATION
To keep the convergence rate comparable, the total number of iterations was set to 30 for all algorithms. Unlike real fMRI data, ground truth about the number of generating sources is known for the synthetic dataset. Therefore, the same number of components were trained for each algorithm. The total number of components to be extracted was set to 8 for JADE, sPCA, sPLS, and sCCA. Using SVD dimensionality of the data was reduced to 16, from which eight components were extracted using sICA, ssICA, ifICA, and ssBSSP. Whereas 250 components were kept from the autoencoder, and then eight sources were retrieved for ssBSSA and ssBSSAS. The ability of the autoencoder to extract a set of overcomplete basis is exhibited by reducing the data dimensionality to 250, which is greater than the total number of time points. As the selection of tuning parameters was based on getting the best results in terms of correlation values, many values were tried before setting the sparsity parameter to 300 for sPCA, 30 for sCCA, and 5 for sPLS, and default parameter settings were used for ifICA. The sparsity and smoothing parameter for ssICA was set to 0.7 and 10^5, respectively. For ssBSSP, the sparsity parameters were set as $\lambda_1 = \lambda_2 = 0.01$ and $\zeta = 90$, whereas $\lambda_1 = 2$, $\lambda_2 = 3$, and $\zeta = 60$ for ssBSSA, and $\lambda_1 = \lambda_2 = 2$ and $\zeta = 60$ for ssBSSAS, and the total number of DCT basis was set to $K_p = 150$ and $\lambda_3 = 12$ for all three variants.

3) SYNTHETIC DATASET RESULTS
This section discusses simulation results to demonstrate the robustness and consistency of the proposed method to spatial dependence. In this regard, the spread parameter ρ was varied between 0.6 and 5 to gradually create activation overlaps as shown in Fig. 4a-e and Fig. 7. Initially, five different activation instances were created using five different values
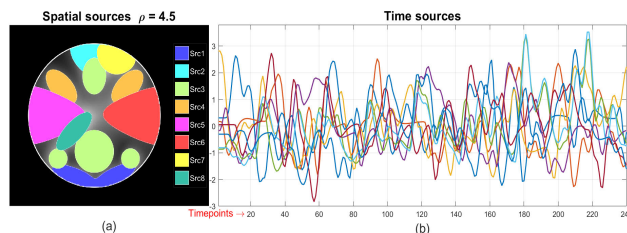

FIGURE 7. Activation maps with a different spatial dependence compared to Fig. 4 and the corresponding time courses.

TABLE 5. Mean correlation values over 10 trials for $\rho = 4.5$.

Srcs	JADE	sCCA	sPLS	sICA	ssICA	ssBSSP	ssBSSA
TC1	0.762	0.506	0.777	0.745	0.731	0.819	**0.821**
TC2	0.703	0.513	0.638	0.687	0.646	**0.820**	0.796
TC3	0.687	0.634	0.732	0.603	0.750	0.787	**0.851**
TC4	0.671	0.485	0.776	0.700	0.728	0.825	**0.827**
TC5	0.807	0.738	0.771	0.804	0.809	0.815	**0.855**
TC6	0.816	0.654	0.806	0.700	0.814	0.793	**0.863**
TC7	0.730	0.493	0.697	0.742	0.761	0.737	**0.826**
TC8	0.716	0.487	0.704	0.702	0.681	**0.765**	0.744
SM1	0.866	0.576	0.850	0.865	0.880	0.923	**0.956**
SM2	0.761	0.631	0.815	0.780	0.862	0.839	**0.928**
SM3	0.842	0.784	0.877	0.681	0.877	0.754	**0.909**
SM4	0.884	0.520	0.811	0.888	0.898	0.884	**0.932**
SM5	0.868	0.593	0.760	0.862	**0.893**	0.856	0.859
SM6	0.895	0.557	0.716	0.830	**0.922**	0.829	0.921
SM7	0.827	0.659	0.814	0.815	0.774	0.896	**0.885**
SM8	0.835	0.632	0.641	0.831	0.843	**0.914**	0.912
Mean	0.792	0.591	0.761	0.765	0.804	0.829	**0.868**

of ρ. For each activation instance, all algorithms attempted to recover the underlying sources 100 times, where every time the dataset Y was generated using Gaussian distribution $\sim \mathcal{N}(0, 0.6)$ and $\sim \mathcal{N}(0, 0.01)$, for spatial and temporal noise respectively.

Underlying source TCs/SMs were recovered by correlating every algorithm's extracted sources with the ground truth sources and keeping the indices of the highest correlation values. These values were computed with respect to GT-SMs and saved as cTC/cSM. Except for JADE and ifICA, where temporal sources were recovered first, all other algorithms recovered spatial maps and then obtained corresponding temporal sources using least squares. For each of the 100 trials and five spatial overlap instances, the mean of the highest correlation values for eight spatiotemporal sources is saved as mcTC/mcSM, and their mean is plotted in Fig. 5. The convergence rate of all variants of the proposed method as a function of algorithm iterations is plotted in Fig. 6a. The progression of the mean of the sum of correlation values between the recovered spatiotemporal sources and ground truth sources over the algorithm iterations is shown in Fig. 6b.

It can be deduced from Fig. 5 that the proposed method's variants achieved the highest source recovery performance. Although this performance decreased with increasing spatial overlaps, it remained higher than all other competing algorithms. Among all the variants, the ssBSSAS algorithm seemed to have performed slightly better than ssBSSA for
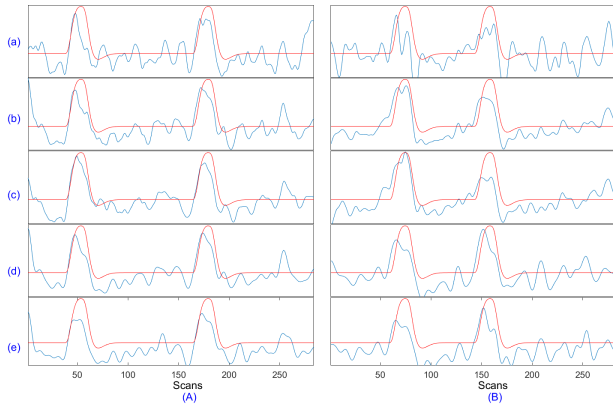
FIGURE 8. For the block design dataset's left toe (A) and tongue (B) movement tasks, the most correlated temporal components with MHR obtained using a) JADE, b) sICA, c) ssICA, d) ssBSSP, and e) ssBSSA. The corresponding correlation values are given in Table 6.

TABLE 6. Correlation values of most correlated temporal components with MHRs for six different block design tasks obtained using five BSS algorithms and two variants of the proposed method.

Algos	VC	LT	LF	RT	RF	T	Mean
JADE	0.679	0.622	0.393	0.521	0.407	0.451	0.512
sICA	0.643	0.699	0.700	0.734	0.405	**0.790**	0.662
ssICA	0.716	**0.791**	0.593	**0.738**	0.501	0.744	0.681
sPLS	0.585	0.410	0.654	0.689	0.579	0.737	0.609
sCCA	0.554	0.618	0.517	0.526	0.356	0.764	0.556
ssBSSP	0.733	0.742	0.705	0.555	0.571	0.734	0.673
ssBSSA	**0.754**	0.756	**0.743**	0.668	**0.659**	0.725	**0.717**

some cases; however, this performance gain is at the cost of increased computational complexity. Moreover, compared to other algorithms, the proposed algorithms had performed exceptionally well in unveiling spatial maps when the sources were very sparse, that is, $\rho = 0.6$. It is worth mentioning that ssICA turned out to be a runner-up as the spatial dependence was increased, and JADE also accomplished impressive results over all cases. Fig. 6 illustrates that all variants of the proposed method have converged around the fifteenth iteration, after which they also stopped accumulating the correlation strength. After the fifteenth iteration, this strength was maintained as the algorithms settled down after convergence to a local minimum.

Later on, in order to obtain component-wise analysis and keep computational analysis on equal footing to the real fMRI datasets, the experiment was repeated for $\rho = 4.5$ as shown in Fig. 7. For this experiment and the real fMRI dataset in subsequent sections, ssBSSAS, due to its higher computational cost, ifICA and sPCA, due to their inferior results in terms of spatial maps, have been dropped from further analysis. This experiment was repeated ten times with a random spatial and temporal noise of variance 0.6 and 0.01, respectively. The mean of correlation values over ten trials is given in Table 5 where one can see that the block variant of the proposed method with components from autoencoder (ssBSSA) has outperformed all other algorithms by recovering both spatial and temporal sources with the highest correlation values except the few occurrences. In this case, the total computation time in seconds consumed by all algorithms over all trials is given in the top row of Table 9 with the lowest values highlighted in bold.

B. EXPERIMENTAL fMRI DATASET
1) EXPERIMENTAL fMRI DATASET PREPROCESSING
Excluding the resting-state dataset, which the HCP had already processed using their preprocessing pipelines, the block design, and the event-related dataset were prepro-

cessed using the SPM-12 toolbox [1]. The preprocessing steps comprised realignment, normalization, spatial smoothing, and masking. More details about these steps can be found in [56]. To correct for motion artifacts, functional images were realigned to the first image, then all images were spatially normalized to a standard Tailarach template, resampled to $2 \times 2 \times 2$ mm^3 voxels, and spatially smoothed using a $6 \times 6 \times 6$ mm^3 full-width at half-maximum (FWHM) Gaussian kernel followed by the masking step to remove any data outside the scalp. After the masking step, the four-dimensional datasets were reshaped and stored in a 2-dimensional matrix Y to be considered as a whole brain dataset. This resulted in the size of Y to be 284×236115 for the block design dataset, 311×166390 for the event-related dataset, and 400×230367 for the resting-state dataset. As a second last step, temporal filtering was performed on all three datasets, which consisted of a high-pass DCT basis set filter to remove scanner-induced trends and a low-pass FWHM filter to remove physiological noise. The cutoff for a DCT-based filter was set to 1/150 Hz for block design, whereas 1/128 Hz for event-related and resting-state datasets. FWHM of 1 sec was used for both block design resting-state dataset, and 1.5 secs for event-related. These cutoffs were assigned according to the previous studies [70], [72]. As a final step, all column entries of Y were normalized to have zero mean and unit variance for all datasets.

2) BLOCK DESIGN DATASET
The motor task 3T MRI unprocessed block design dataset used in this section was acquired from the quarter 3 release of the HCP [68], [69]. To map the primary motor cortex of the brain experiment was run for 204 secs; the visual cues were presented to the subjects to tap their right or left fingers, squeeze their right or left toes, or move their tongue. Subjects were shown a three-second visual cue followed by a specific movement task lasting 12 seconds. There were ten movement tasks, including left/right finger, left/right toe, and two tongue movements. Thus, there were a total of 13 blocks, including three fixation blocks that lasted for 15 secs. To obtain GT-TCs, six modeled HRF (MHRs) were produced by using the convolution operation between the canonical HRF and task stimuli related to 5 movement types {left toe (LT), left finger (LF), right toe (RT), right finger (RF), tongue (T)} and visual type cue (VC). A Siemens 3 Tesla (3T) scanner was used to
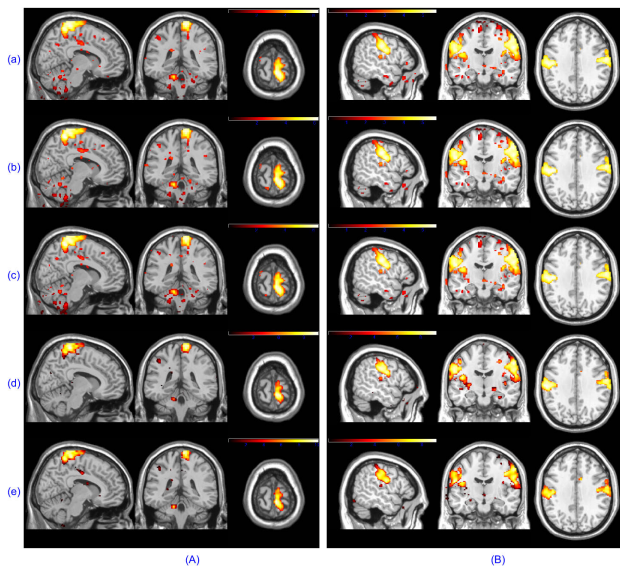
FIGURE 9. Thresholded activation maps for left toe (A) and tongue (B) movement tasks at a random field correction $p < 0.001$ obtained using a) JADE, b) sICA, c) ssICA, d) ssBSSP, and e) ssBSSA.

FIGURE 10. For the even-related dataset, thresholded activation maps for the right finger tapping task at a random field correction $p < 0.001$ and corresponding most correlated temporal source with MHR function obtained using a) JADE, b) sICA, c) ssICA, d) ssBSSP, and e) ssBSSA. The respective correlation values are given in Table 7.

acquire fMRI images for each subject. The acquisition parameters were TR = 0.72 secs, echo time (TE) = 33.1 ms, field of view (FOV) = 208 × 180 mm, flip angle (FA) = 52°, matrix size = 104 × 90, slice thickness = 2 mm with 72 contiguous slices, and 2 mm isotropic voxels, echo spacing = 0.58 ms, BW = 2290 Hz/Px, and 284 EPI volumes were collected. The dataset of a single subject with ID 163129, aged between 26 and 30 years, was used in our analysis.

3) BLOCK DESIGN DATASET SOURCE SEPARATION

Many values of the tuning parameters were tested, and their best combination in terms of the sum of correlation strength with the ground truth was considered. The total number of components to be estimated was set to 40 for sPLS, and sCCA. While performing dimensionality reduction 40/60/50 components were kept from PCA, and then 40/60/50 sources were uncovered using JADE/sICA/ssICA. Whereas 60/105 components were kept from PCA/autoencoder, and then 40/35 were extracted using ssBSSP/ssBSSA. The total number of iterations was set to 30 for all algorithms. The sparsity parameter was set to 15 for sCCA and sPLS. The sparsity and smoothing parameter for ssICA was set to 50 and 10^5. For ssBSSP, the sparsity parameters were set as $\lambda_1 = \lambda_2 = 0.01$, $\lambda_3 = 16$, $\zeta = 50$, and $K_p = 60$, whereas $\lambda_1 = 0.3$, $\lambda_2 = 0.11$, $\lambda_3 = 16$, $\zeta = 50$, and $K_p = 60$ were used for ssBSSA.

4) BLOCK DESIGN DATASET RESULTS

Due to the absence of ground truth for spatial maps, we decided to base our analysis on the BOLD time series via six constructed MHRs. This was accomplished by correlating the MHRs with temporal components retrieved by all algorithms and saving the highest correlation values. These values
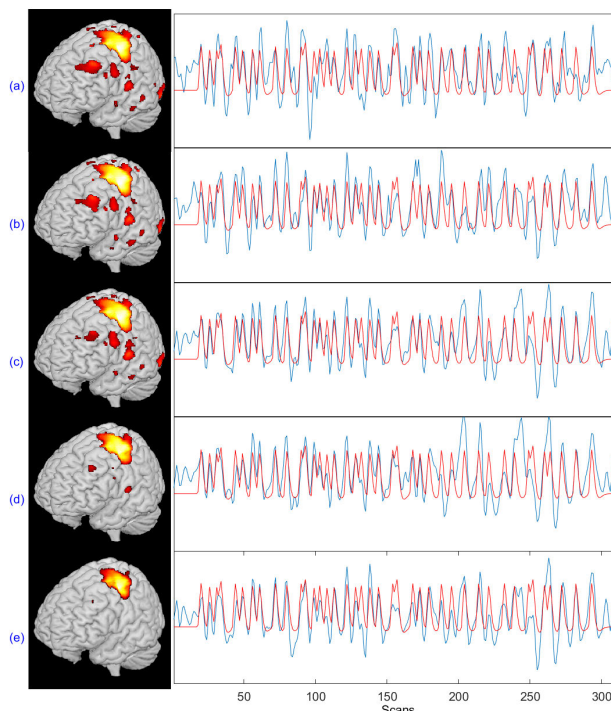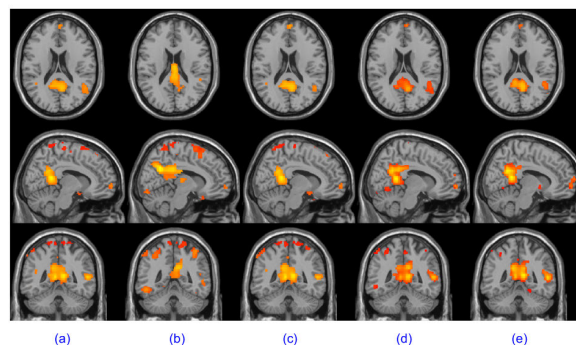

FIGURE 11. For the event-related dataset, default mode network retrieved by keeping the most correlated spatial map with R4 RSN template obtained from a) JADE, b) sICA, c) ssICA, d) ssBSSP, and e) ssBSSA.

are mentioned in Table 6 with the highest values highlighted in bold, and a couple of recovered temporal components (left toe and tongue along with the MHRs in red) and their corresponding spatial maps are given in Fig. 8 and 9, respectively. Only two activation maps have been shown to avoid unnecessarily increasing the paper length. The table shows that the proposed ssBSSA algorithm has the highest similarity with the ground truth overall. Fig. 9 shows that the variants of the proposed (ssBSS) method, specifically ssBSSA, retrieved activation maps that are very specific to the motor area compared to other algorithms. Although, ssBSSA has
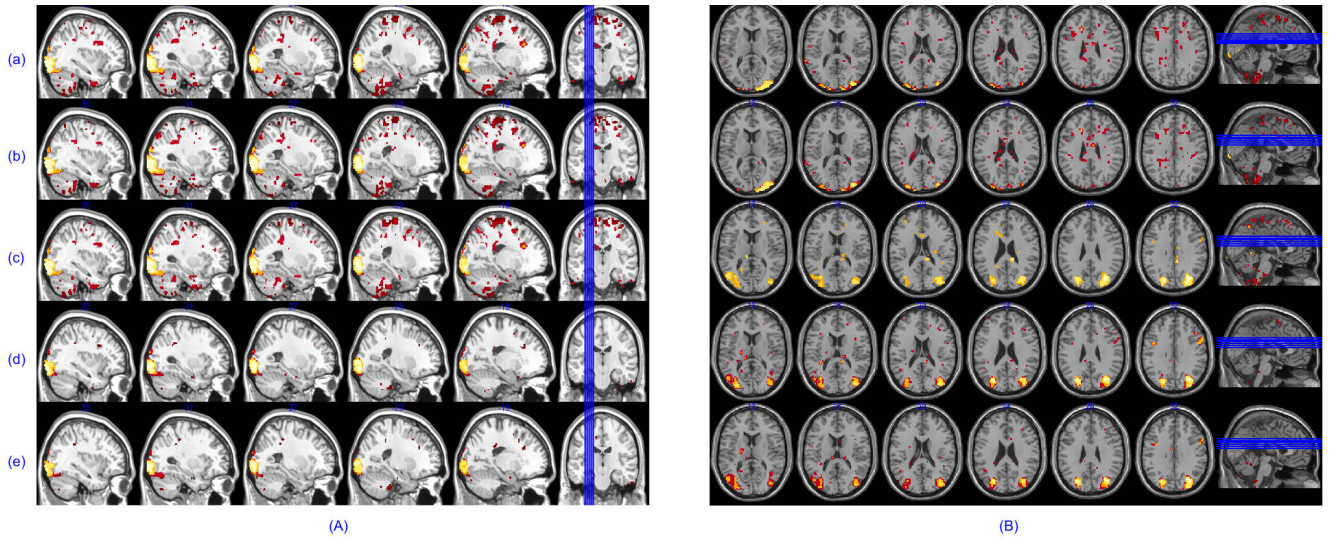
FIGURE 12. For the resting-state dataset, occipital pole visual (A) and lateral visual (B) network retrieved by keeping the most correlated spatial maps with R2 and R3 RSN templates, respectively, obtained from a) JADE, b) sICA, c) ssICA, d) ssBSSP, and e) ssBSSA.

not produced the highest correlation values compared to other algorithms for these plotted tasks, the recovered temporal dynamics are comparable and spatial maps are distinct but analogous. These results for the motor task dataset are very similar to what has been reported in other papers [56], [72], [75].

5) EVENT RELATED DATASET

The motor task 3T MRI unprocessed event-related dataset used in this section was provided by KAIST [70]. During the experiment, which was recorded for 650 secs, subjects were asked to perform a right finger tapping task to map the left motor area of the brain. The session began with an initial preparation period of 30 secs and a resting period of 30 secs, followed by 40 repeated trials of the task and resting period. Here each trial's duration is 28 secs, always starting with a task period and ending with a rest. This was followed by an additional 30 secs resting period. The interstimulus interval varied between 4 and 20 secs, with a mean of 12 secs within each repeated trial. The brain scans were acquired using the 3T ISOL Technologies MRI system Korea. The acquisition parameters were matrix size $= 64 \times 64$, TR/TE $= 2000/35$ms, FA $= 80^o$, slice thickness $= 4$ mm with 24 contiguous slices, voxel size $= 3.44 \times 3.44 \times 4$ mm^3. A total of 325 EPI volumes were collected, from which the first 15 preparation scans were discarded. The dataset of a single subject with ID s3sh aged between 23 and 28 years approximately was used in our analysis.

6) EVENT RELATED DATASET SOURCE SEPARATION

The same strategy used for the block design dataset was adapted for event-related dataset analysis. We commenced by selecting the optimal number of sources and tuning parameters by experimentation. To keep the comparison unbiased, all

TABLE 7. Correlation values of most correlated temporal components with event-related finger tapping MHR and most correlated spatial map with few RSN templates using five BSS algorithms and two variants of the proposed method. The best values are highlighted in bold.

Algos	FT	R1	R2	R4	R9	R10	Mean
JADE	0.620	0.416	0.518	0.462	0.245	0.217	0.413
sICA	**0.651**	0.533	0.515	0.357	0.324	0.253	0.439
ssICA	0.637	0.489	0.468	0.451	0.314	0.258	0.436
sPLS	0.481	0.425	0.409	0.379	0.345	0.281	0.387
sCCA	0.427	0.382	0.220	0.482	0.362	0.202	0.346
ssBSSP	0.558	**0.544**	0.426	0.572	**0.527**	0.378	0.501
ssBSSA	0.645	0.523	**0.565**	**0.583**	0.450	**0.426**	**0.532**

TABLE 8. For resting-state dataset, correlation values of the most correlated spatial component with few RSN templates using seven different algorithms, including the variants of the proposed method.

Algos	R1	R2	R3	R4	R9	R10	Mean
JADE	0.420	0.617	0.227	0.331	0.316	0.344	0.376
sICA	0.496	0.629	0.235	0.333	0.347	0.337	0.396
ssICA	0.486	0.611	0.327	0.276	0.275	0.298	0.379
sPLS	0.455	0.551	0.394	0.329	0.269	0.254	0.375
sCCA	0.273	0.284	0.231	0.316	0.141	0.209	0.242
ssBSSP	**0.678**	0.704	0.437	0.424	0.351	0.372	0.494
ssBSSA	0.640	**0.712**	**0.448**	**0.485**	**0.380**	**0.374**	**0.507**

algorithms were run for 30 iterations. For sPLS, and sCCA, 40 components were estimated. For JADE, sICA and ssICA, 40 components were estimated after retaining the same number of components from PCA. Whereas for ssBSSP/ssBSSA, the dimensionality of the data was reduced to 60/105 using PCA/autoencoder to extract 40 components in both cases. For sCCA and sPLS, the sparsity parameter was set to 25 and 8, respectively. For ssICA, the smoothing/sparsity parameters were set to $10^5/5$. The sparsity parameters were set to $\lambda_1 = \lambda_2 = 0.025$, $\lambda_3 = 11.5$, $\zeta = 150$, and $K_p = 300$ for ssBSSP.

On the other hand, $\lambda_1 = 0.5$, $\lambda_2 = 0.01$, $\lambda_3 = 11.5$, $\zeta = 150$, and $K_p = 300$ were used for ssBSSA.

7) EVENT RELATED DATASET RESULTS

This section deployed task-related MHR and some resting-state networks (RSNs: R1-R10) described in [76] as the ground truth. Using recovered components from each algorithm, both temporal and spatial sources correlated with the ground truth, and the highest correlation values were saved. These values are reported in Table 7 with highest values highlighted in bold, task-related (right finger tappig) recovered time course and its associated spatial maps are plotted in Fig. 10, and recovered default mode network (R4) is plotted in Fig. 11. Overall, the ssBSSA algorithm has produced the highest correlation values. Like the block design dataset, task-related activation maps from all algorithms seem to have recovered activations very specific to the motor area, and all algorithms have successfully recovered the default mode network. In both cases, the maps obtained using ssBSSA seems to be more specific compared to other algorithms. The recovered activation maps are similar to earlier reported results for this dataset [37], [70], [71], [72], [77].

8) RESTING STATE DATASET

The resting-state dataset was obtained from the first set of 3T MRI preprocessed S900 release of the HCP [68], [69]. The same acquisition parameters as for the block design dataset given in Section V-B2 were used. The experiment was run for 873 secs, resulting in a total of 1200 scans. This data was acquired two times in a single session for two different phase encoding directions. Subjects were asked to keep their eyes fixated on a bright crosshair presented in a darkened room, and the second run that consisted of left to right phase encoding was considered for our analysis. The first 20 scans were discarded, the following 400 scans were kept for analysis, and the remaining 780 scans were discarded. Before applying temporal filtering, the preprocessed data downloaded from the HCP website also underwent spatial smoothing using $6 \times 6 \times 6$ mm^3 FWHM Gaussian kernel. The dataset of a single subject with ID 100206, aged between 26 and 30 years, was used in our analysis.

9) RESTING STATE DATASET SOURCE SEPARATION

Similar to the last three datasets, all algorithms were iterated 30 times for an unbiased comparison. The number of sources to be uncovered by each algorithm was decided by trying numbers between 30 and 60. Different tuning parameter combinations were tested, and only the best were considered. The number of components estimated using sPLS and sCCA was 40. Whereas the dimensionality of the data was reduced to 40/60/40, and the number of components estimated was set to 40/40/30 for JADE/sICA/ssICA. On the other hand, the dimensionality reduction was set to 60/105 for ssBSSP/ssBSSA, while only 40 components for both algorithms were eventually retained. The best sparsity settings

TABLE 9. Total and mean computation time in seconds and correlation values by each algorithm for four different datasets.

Data	sCCA	ssICA	JADE	sPLS	sICA	ssBSSP	ssBSSA
SD	107	70	**2**	15	7	25	33
BD	450	531	415	211	134	33	**21**
ER	370	232	71	145	**19**	40	25
RS	510	541	135	195	45	54	**42**
Mean	359	344	156	142	51	38	**30**
SD	0.591	0.804	0.792	0.761	0.765	0.829	**0.868**
BD	0.556	0.681	0.512	0.609	0.662	0.673	**0.717**
ER	0.346	0.436	0.413	0.387	0.439	0.501	**0.532**
RS	0.242	0.379	0.376	0.375	0.396	0.494	**0.507**
Mean	0.434	0.575	0.523	0.533	0.566	0.624	**0.656**

for sCCA/sPLS were found to be 60/12. The best smoothing/sparsity parameters for ssICA were noted as $10^5/10$. The sparsity parameters were set to $\lambda_1 = \lambda_2 = 0.01$, $\lambda_3 = 13$, $\zeta = 150$, and $K_p = 300$ for ssBSSP. In contrast, $\lambda_1 = 0.2$, $\lambda_2 = 0.2$, $\lambda_3 = 13$, $\zeta = 150$, and $K_p = 300$ were used for ssBSSA.

10) RESTING STATE DATASET RESULTS

The resting-state networks (R1-R10), as reported in [76], were factored in as the ground truth. Six of these networks were found to be recovered by all algorithms. The recovered components were correlated with six of these resting-state templates, and the highest correlation values were kept, which are reported in Table 8 with best values highlighted in bold. The recovered activation maps that were most correlated with R2 and R3 are plotted in Fig. 12. Overall, the ssBSSA algorithm has produced the highest correlation values. All algorithms successfully recovered the activation network for the occipital pole region; however, only ssICA, ssBSSA, and ssBSSP could correctly recover the ground truth for the lateral visual network. Maps from both ssBSSA and ssBSSP were found to be more specific, with lower false positives compared to other algorithms.

C. DISCUSSION

Some statistics about the proposed method and its overall comparison with other algorithms are illustrated in Fig. 13 and 14. The convergence rate of the block variants of the proposed method for block design (BD), event-related (ER), and resting-state (RS) datasets as a function of algorithm iterations is plotted in Fig. 13a. The growth of the mean of the sum of correlation coefficients between the retrieved spatiotemporal and ground truth sources over the algorithm iterations is shown in Fig. 13b. To give an idea about the sparseness of the mixing matrix for the ssBSSA algorithm, its image with scaled colors for the event-related dataset is drawn in Fig. 14a. The total computation time and correlation values by all participating algorithms for all datasets, including the synthetic dataset (SD), are given in Table 9. The pictorial version of this performance comparison is provided in Fig. 14b, where the mean computation time values
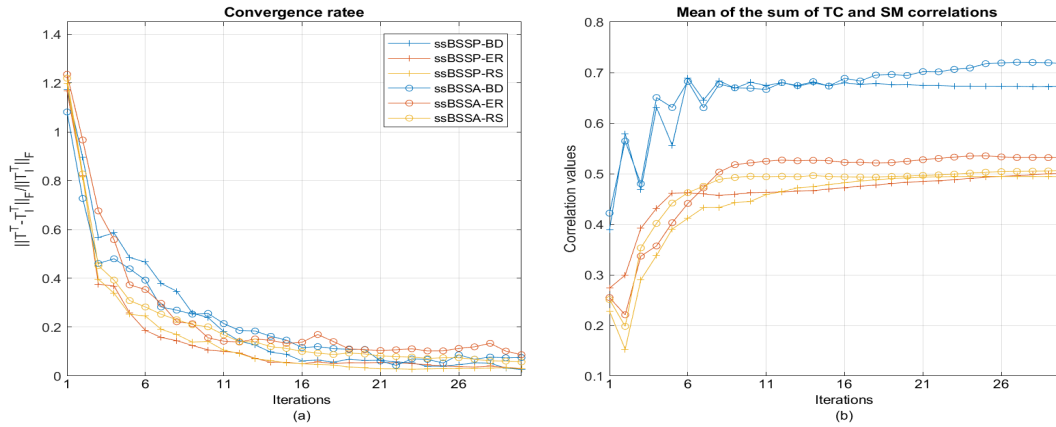
FIGURE 13. For two variants of the proposed method, a) convergence rate, and b) the mean of the sum of highest correlations between ground truth spatiotemporal sources and retrieved sources shown as a function of algorithm iterations.
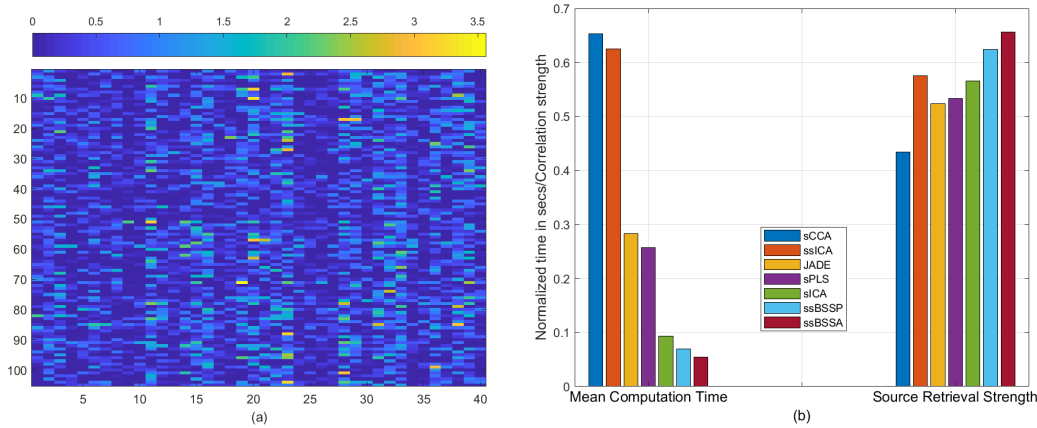

FIGURE 14. a) Temporal mixing matrix (U) of the event-related dataset revealing sparseness of its entries, b) comparison of all participating algorithms in terms of computational time and source retrieval accuracy.

have been normalized to accommodate correlation values and computation time in the same graph.

Fig. 13a shows that apart from some fluctuations, both variants of the proposed method converged for all experimental datasets consistently; however, this consistency was more evident for ssBSSP. Fig. 13b shows the gradual development of source retrieval where the ssBSSA has been shown to have performed better than ssBSSP. After observing both subfigures, one can conclude that both algorithms for all three datasets started to settle down after the fifteenth iteration except for ssBSSA in case of the BD dataset. Fig. 14b displays the overall performance by all algorithms where the variants of the proposed method have exhibited superior results in terms of computation time and source recovery strength, with ssICA being the runner-up in source recovery category and sICA in computation time.

VI. CONCLUSION

A new source separation method is proposed for fMRI data, and all possible settings for its solution are presented, including thresholding, PCA, autoencoder, and discussion on convergence. The effectiveness of the proposed ssBSS

method is demonstrated using synthetic and experimental fMRI datasets. Its performance was consistent across experiments, and its computational simplicity makes it favorable over swBSS and scBSS variants such as SOBI, JADE, sPCA, sCCA, sPLS, sICA, ssICA, and ifICA. Its reduced computational complexity is in terms of analytical complexity associated with lesser calls to data and arithmetic complexity associated with lesser operations required to solve the method [78]. It can be considered a promising alternative to ICA, sparse ICA, and other swBSS-based neuroimaging data analysis methods. Investigating the fluctuations exhibited in the convergence of the ssBSSA algorithm and using model selection criteria for the optimal number of PCA/autoencoder components can be pursued as part of future work.

Unlike conventional time-consuming strategies, such as joint diagonalization of covariance matrices, deflation, independent decomposition of each mixture signal during sparse coding, and repeated error matrix computation of full data matrix, the proposed method concurrently exploited spatial and temporal information from reduced-dimension space in a sparse manner, leading to better source recovery at a reduced computational cost. The STEM model that takes into account the temporal smoothness and the spatial sparseness

along the fMRI data matrix columns and rows, respectively, is computationally intensive to solve; however, by manipulating the BSS theory, it was disintegrated into two sub-optimization problems. Based on Neumann's alternating projection lemma, its solution is based on performing the block/sequential update of the source/mixing matrices from these two sub-problems in an iterative manner. This strategy allowed us to solve the STEM model in a computationally efficient manner. It was speculated and then found empirically that convergence of the proposed model is guaranteed if finite basis injective property and a strict sparsity pattern are sustained. It produced results that were specifically superior to strategies based on the combination of sparsity and independence.

ACKNOWLEDGMENT

REFERENCES

[1] W. Penny, K. Friston, J. Ashburner, S. Kiebel, and T. Nichols, *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. New York, NY, USA: Academic, 2007.

[2] K. J. Friston, P. Jezzard, and R. Turner, "Analysis of functional MRI time-series," *Hum. Brain Mapping*, vol. 1, no. 2, pp. 153–171, 1994.

[3] G. K. Aguirre, E. Zarahn, and M. D'Esposito, "The variability of human, BOLD hemodynamic responses," *NeuroImage*, vol. 8, no. 4, pp. 360–369, Nov. 1998.

[4] M. U. Khalid and A.-K. Seghouane, "Constrained maximum likelihood based efficient dictionary learning for fMRI analysis," in *Proc. IEEE 11th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2014, pp. 45–48.

[5] M. U. Khalid and A.-K. Seghouane, "A single SVD sparse dictionary learning algorithm for FMRI data analysis," in *Proc. IEEE Workshop Stat. Signal Process. (SSP)*, Jun. 2014, pp. 65–68.

[6] M. J. Mckeown, S. Makeig, G. G. Brown, T.-P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski, "Analysis of fMRI data by blind separation into independent spatial components," *Hum. Brain Mapping*, vol. 6, no. 3, pp. 160–188, 1998.

[7] A. H. Andersen, D. M. Gash, and M. J. Avison, "Principal component analysis of the dynamic response measured by fMRI: A generalized linear systems framework," *Magn. Reson. Imag.*, vol. 17, no. 6, pp. 795–815, Jul. 1999.

[8] O. Friman, M. Borga, P. Lundberg, and H. Knutsson, "Exploratory fMRI analysis by autocorrelation maximization," *NeuroImage*, vol. 16, no. 2, pp. 454–464, Jun. 2002.

[9] M. U. Khalid, A. Shah, and A.-K. Seghouane, "Adaptive 2DCCA based approach for improving spatial specificity of activation detection in functional MRI," in *Proc. Int. Conf. Digit. Image Comput. Techn. Appl. (DICTA)*, Dec. 2012, pp. 1–6.

[10] A. Shah, M. U. Khalid, and A.-K. Seghouane, "Comparing causality measures of fMRI data using PCA, CCA and vector autoregressive modelling," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 6184–6187.

[11] W. Lin, H. Wu, Y. Liu, D. Lv, and L. Yang, "A CCA and ICA-based mixture model for identifying major depression disorder," *IEEE Trans. Med. Imag.*, vol. 36, no. 3, pp. 745–756, Mar. 2017.

[12] M. A. Qadar and A. Seghouane, "A projection CCA method for effective fMRI data analysis," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 11, pp. 3247–3256, Nov. 2019.

[13] M. U. Khalid and A.-K. Seghouane, "Improving functional connectivity detection in fMRI by combining sparse dictionary learning and canonical correlation analysis," in *Proc. IEEE 10th Int. Symp. Biomed. Imag.*, Apr. 2013, pp. 286–289.

[14] M. Pal, R. Roy, J. Basu, and M. S. Bepari, "Blind source separation: A review and analysis," in *Proc. IEEE Int. Conf. Oriental COCOSDA Held Jointly Conf. Asian Spoken Language Res. Eval. (O-COCOSDA/CASLRE)*, Mar. 2013, pp. 1–5.

[15] C. Ruan, D. Zhao, W. Jia, C. Chen, Y. Chen, and X. Liu, "A new image denoising method by combining WT with ICA," *Math. Problems Eng.*, vol. 2015, pp. 1–10, Jan. 2015.

[16] A.-K. Seghouane and M. U. Khalid, "Learning dictionaries from correlated data: Application to fMRI data analysis," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 2340–2344.

[17] S. Hinderer, "Blind source separation of radar signals in time domain using deep learning," in *Proc. 23rd Int. Radar Symp. (IRS)*, Sep. 2022, pp. 486–491.

[18] L. Y. Taha and E. Abdel-Raheem, "Blind source separation: A performance review approach," in *Proc. 5th Int. Conf. Signal Process. Inf. Secur. (ICSPIS)*, Dec. 2022, pp. 148–153.

[19] S. Choi, A. Cichocki, H.-M. Park, and S.-Y. Lee, "Blind source separation and independent component analysis: A review," *Neural Inf. Process. Lett. Rev*, vol. 6, no. 1, pp. 1–57, 2004.

[20] I. T. Jolliffe, *Principal Component Analysis for Special Types of Data*. Springer, 2002.

[21] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Netw.*, vol. 13, nos. 4–5, pp. 411–430, Jun. 2000.

[22] O. Friman, J. Cedefamn, P. Lundberg, M. Borga, and H. Knutsson, "Detection of neural activity in functional MRI using canonical correlation analysis," *Magn. Reson. Med.*, vol. 45, no. 2, pp. 323–330, Feb. 2001.

[23] M. Aharon, M. Elad, and A. Bruckstein, "*rmK*-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[24] A. Krishnan, L. J. Williams, A. R. McIntosh, and H. Abdi, "Partial least squares (PLS) methods for neuroimaging: A tutorial and review," *NeuroImage*, vol. 56, no. 2, pp. 455–475, May 2011.

[25] J. Rapin, J. Bobin, A. Larue, and J.-L. Starck, "Sparse and non-negative BSS for noisy data," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5620–5632, Nov. 2013.

[26] P. Paatero and U. Tapper, "Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics*, vol. 5, no. 2, pp. 111–126, Jun. 1994.

[27] X. Zhuang, Z. Yang, and D. Cordes, "A technical review of canonical correlation analysis for neuroscience applications," *Human Brain Mapping*, vol. 41, no. 13, pp. 3807–3833, Sep. 2020.

[28] S. Moussaoui, D. Brie, O. Caspary, and A. Mohammad-Djafari, "A Bayesian method for positive source separation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, 2004, pp. V–485.

[29] P. Georgiev, F. Theis, A. Cichocki, and H. Bakardjian, "Sparse component analysis: A new tool for data mining," in *Data Mining in Biomedicine*, vol. 8. 2007, pp. 91–116.

[30] V. D. Calhoun and T. Adali, "Multisubject independent component analysis of fMRI: A decade of intrinsic networks, default mode, and neurodiagnostic discovery," *IEEE Rev. Biomed. Eng.*, vol. 5, pp. 60–73, 2012.

[31] I. Daubechies, E. Roussos, S. Takerkart, M. Benharrosh, C. Golden, K. D'Ardenne, W. Richter, J. D. Cohen, and J. Haxby, "Independent component analysis for brain fMRI does not select for independence," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 26, pp. 10415–10422, Jun. 2009.

[32] V. D. Calhoun, V. K. Potluru, R. Phlypo, R. F. Silva, B. A. Pearlmutter, A. Caprihan, S. M. Plis, and T. Adalı, "Independent component analysis for brain fMRI does indeed select for maximal independence," *PLoS ONE*, vol. 8, no. 8, Aug. 2013, Art. no. e73309.

[33] Z. Boukouvalas, Y. Levin-Schwartz, V. D. Calhoun, and T. Adalı, "Sparsity and independence: Balancing two objectives in optimization for source separation with application to fMRI analysis," *J. Franklin Inst.*, vol. 355, no. 4, pp. 1873–1887, Mar. 2018.

[34] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 434–444, Feb. 1997.

[35] F. Castells, J. J. Rieta, J. Millet, and V. Zarzoso, "Spatiotemporal blind source separation approach to atrial activity estimation in atrial tachyarrhythmias," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 2, pp. 258–267, Feb. 2005.

[36] M. Li, Y. Liu, G. Feng, Z. Zhou, and D. Hu, "OI and fMRI signal separation using both temporal and spatial autocorrelations," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 8, pp. 1917–1926, Aug. 2010.

[37] M. U. Khalid and A. Seghouane, "Multi-subject fMRI connectivity analysis using sparse dictionary learning and multiset canonical correlation analysis," in *Proc. IEEE 12th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2015, pp. 683–686.

[38] H. D. Do, S. T. Tran, and D. T. Chau, "Speech source separation using variational autoencoder and bandpass filter," *IEEE Access*, vol. 8, pp. 156219–156231, 2020.

[39] J.-F. Cardoso, "High-order contrasts for independent component analysis," *Neural Comput.*, vol. 11, no. 1, pp. 157–192, Jan. 1999.

[40] N. Ito, R. Ikeshita, H. Sawada, and T. Nakatani, "A joint diagonalization based efficient approach to underdetermined blind audio source separation using the multichannel Wiener filter," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 29, pp. 1950–1965, 2021.

[41] N. Nadisic, N. Gillis, and C. Kervazo, "Smoothed separable nonnegative matrix factorization," 2021, *arXiv:2110.05528*.

[42] R. Carloni Gertosio, J. Bobin, and F. Acero, "Semi-blind source separation with learned constraints," *Signal Process.*, vol. 202, Jan. 2023, Art. no. 108776.

[43] J. Von Neumann, *Functional Operators Volume 2: The Geometry of Orthogonal Spaces*. Princeton, NJ, USA: Princeton Univ. Press, 1951.

[44] M. Xia, J. Wang, and Y. He, "BrainNet viewer: A network visualization tool for human brain connectomics," *PLoS ONE*, vol. 8, no. 7, Jul. 2013, Art. no. e68910.

[45] H. Zou, T. Hastie, and R. Tibshirani, "Sparse principal component analysis," *J. Comput. Graph. Statist.*, vol. 15, no. 2, pp. 265–286, Jan. 2006.

[46] S. Waaijenborg, P. C. Verselewel de Witt Hamer, and A. H. Zwinderman, "Quantifying the association between gene expressions and DNA-markers by penalized canonical correlation analysis," *Stat. Appl. Genet. Mol. Biol.*, vol. 7, no. 1, Jan. 2008.

[47] A. Aïssa-El-Bey and A.-K. Seghouane, "Sparse and smooth canonical correlation analysis through rank-1 matrix approximation," *EURASIP J. Adv. Signal Process.*, vol. 2017, no. 1, p. 25, Dec. 2017.

[48] D. M. Witten, R. Tibshirani, and T. Hastie, "A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis," *Biostatistics*, vol. 10, no. 3, pp. 515–534, Jul. 2009.

[49] J. M. Monteiro, A. Rao, J. Shawe-Taylor, and J. Mourão-Miranda, "A multiple hold-out framework for sparse partial least squares," *J. Neurosci. Methods*, vol. 271, pp. 182–194, Sep. 2016.

[50] B. Du, S. Wang, C. Xu, N. Wang, L. Zhang, and D. Tao, "Multi-task learning for blind source separation," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4219–4231, Sep. 2018.

[51] M. A. Qadar, A. Aïssa-El-Bey, and A.-K. Seghouane, "Two dimensional CCA via penalized matrix decomposition for structure preserved fMRI data analysis," *Digit. Signal Process.*, vol. 92, pp. 36–46, Sep. 2019.

[52] R. Ge, Y. Wang, J. Zhang, L. Yao, H. Zhang, and Z. Long, "Improved FastICA algorithm in fMRI data analysis using the sparsity property of the sources," *J. Neurosci. Methods*, vol. 263, pp. 103–114, Apr. 2016.

[53] A. Hyvärinen, J. Hurri, and P. O. Hoyer, *Natural Image Statistics a Probabilistic Approach to Early Computational Vision*. New York, NY, USA: Springer-Verlag, 2009.

[54] K. Engan, S. O. Aase, and J. H. Husoy, "Method of optimal directions for frame design," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, Mar. 1999, pp. 2443–2446.

[55] D. R. Hardoon and J. Shawe-Taylor, "Sparse canonical correlation analysis," *Mach. Learn.*, vol. 83, no. 3, pp. 331–353, Jun. 2011.

[56] M. U. Khalid, "Sparse group bases for multisubject fMRI data," *IEEE Access*, vol. 10, pp. 83379–83397, 2022.

[57] R. Rubinstein, M. Zibulevsky, and M. Elad, "Double sparsity: Learning sparse dictionaries for sparse signal approximation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1553–1564, Mar. 2010.

[58] M. Nikolova, "Relationship between the optimal solutions of least squares regularized with ℓ_0-norm and constrained by k-sparsity," *Appl. Comput. Harmon. Anal.*, vol. 41, no. 1, pp. 237–265, Jul. 2016.

[59] A. Seghouane and A. Iqbal, "Basis expansion approaches for regularized sequential dictionary learning algorithms with enforced sparsity for fMRI data analysis," *IEEE Trans. Med. Imag.*, vol. 36, no. 9, pp. 1796–1807, Sep. 2017.

[60] M. U. Khalid, A. Shah, and A. Seghouane, "Sparse dictionary learning for fMRI analysis using autocorrelation maximization," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 4286–4289.

[61] Y. Yang, Q. M. J. Wu, and Y. Wang, "Autoencoder with invertible functions for dimension reduction and image reconstruction," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 48, no. 7, pp. 1065–1079, Jul. 2018.

[62] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. Springer, 2011, pp. 185–212.

[63] W. J. Fu, "Penalized regressions: The bridge versus the lasso," *J. Comput. Graph. Statist.*, vol. 7, no. 3, pp. 397–416, Sep. 1998.

[64] K. Bredies and D. A. Lorenz, "Linear convergence of iterative soft-thresholding," *J. Fourier Anal. Appl.*, vol. 14, nos. 5–6, pp. 813–837, Dec. 2008.

[65] R. Gribonval, H. Rauhut, K. Schnass, and P. Vandergheynst, "Atoms of all channels, unite! Average case analysis of multi-channel sparse recovery using greedy algorithms," *J. Fourier Anal. Appl.*, vol. 14, nos. 5–6, pp. 655–687, Dec. 2008.

[66] A. Shah, M. U. Khalid, and A. Seghouane, "Recovering HRFs from overlapping Rois in fMRI data using thresholding correlations for sparse dictionary learning," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 5756–5759.

[67] E. Erhardt, E. Allen, Y. Wei, T. Eichele, and V. Calhoun, "SimTB, a simulation toolbox for fMRI data under a model of spatiotemporal separability," *NeuroImage*, vol. 59, pp. 4160–4167, Dec. 2011.

[68] D. C. Van Essen, K. Ugurbil, E. Auerbach, D. Barch, T. E. J. Behrens, and R. Bucholz, "The human connectome project: A data acquisition perspective," *NeuroImage*, vol. 62, no. 4, pp. 2222–2231, Oct. 2012.

[69] D. M. Barch, G. C. Burgess, M. P. Harms, S. E. Petersen, B. L. Schlaggar, M. Corbetta, M. F. Glasser, S. Curtiss, S. Dixit, C. Feldt, D. Nolan, E. Bryant, T. Hartley, O. Footer, J. M. Bjork, R. Poldrack, S. Smith, H. Johansen-Berg, A. Z. Snyder, and D. C. Van Essen, "Function in the human connectome: Task-fMRI and individual differences in behavior," *NeuroImage*, vol. 80, pp. 169–189, Oct. 2013.

[70] K. Lee, S. Tak, and J. Chul Ye, "A data-driven sparse GLM for fMRI analysis using sparse dictionary learning with MDL criterion," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1076–1089, May 2011.

[71] C. Ting, A. Seghouane, M. U. Khalid, and S. Salleh, "Is first-order vector autoregressive model optimal for fMRI data?" *Neural Comput.*, vol. 27, no. 9, pp. 1857–1871, Sep. 2015.

[72] A.-K. Seghouane and A. Iqbal, "Consistent adaptive sequential dictionary learning," *Signal Process.*, vol. 153, pp. 300–310, Dec. 2018.

[73] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 6, pp. 716–723, Dec. 1974.

[74] C. Rorden, H.-O. Karnath, and L. Bonilha, "Improving lesion-symptom mapping," *J. Cognit. Neurosci.*, vol. 19, no. 7, pp. 1081–1088, Jul. 2007.

[75] S. Zhao, J. Han, J. Lv, X. Jiang, X. Hu, Y. Zhao, B. Ge, L. Guo, and T. Liu, "Supervised dictionary learning for inferring concurrent brain networks," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 2036–2045, Oct. 2015.

[76] S. M. Smith, P. T. Fox, K. L. Miller, D. C. Glahn, P. M. Fox, C. E. Mackay, N. Filippini, K. E. Watkins, R. Toro, A. R. Laird, and C. F. Beckmann, "Correspondence of the brain's functional architecture during activation and rest," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 31, pp. 13040–13045, Aug. 2009.

[77] M. U. Khalid, "Dictionary learning algorithms for functional magnetic resonance imaging," Ph.D. dissertation, Austral. Nat. Univ., Canberra, ACT, Australia, 2015.

[78] Y. Nesterov, *Introductory Lectures on Convex Optimization*, vol. 87. Springer, 2004.

MUHAMMAD USMAN KHALID received the M.Sc. degree in electrical engineering from the Royal Institute of Technology, Stockholm, in 2010, and the Ph.D. degree in engineering and computer science from Australian National University, Canberra, in 2015. He was a Visitor Academics with the Department of Electrical and Electronic Engineering, The University of Melbourne, and a Research Fellow with the School of Mathematics and Statistics, The University of Melbourne. His research interests include sparse representation, blind source separation, numerical optimization, and medical imaging.

BILAL A. KHAWAJA (Senior Member, IEEE) received the B.S. degree in computer engineering from the Sir Syed University of Engineering and Technology, Karachi, Pakistan, in 2002, the M.Sc. degree in communication engineering and signal processing from the University of Plymouth, Plymouth, U.K., in 2005, and the Ph.D. degree in electrical engineering from the University of Bristol, Bristol, U.K., in 2010. From 2010 to 2016, he was an Assistant Professor with the Electronics and Power Engineering Department, PN Engineering College, National University of Science and Technology (NUST), Karachi. In 2015, he was a Visiting Postdoctoral Researcher with the Lightwave Systems Research Laboratory, Queens University, Kingston, Canada, involved in the Natural Sciences and Engineering Research Council (NSERC)-Canada CREATE Next Generation Optical Network (NGON) Project on the characterization and measurements of 25 GHz RF signal generation optical comb sources. Currently, he is an Associate Professor with the Department of Electrical Engineering, Faculty of Engineering, Islamic University of Madina (IUM), Madina, Saudi Arabia. He has authored and coauthored more than 60 peer-reviewed international journals and conference papers. His current research interests include antennas/antenna arrays design and characterization, meta-materials, 5G and next-generation wireless systems, RF/microwave and mm-wave systems, and the use of artificial intelligence techniques in these systems.

MALIK MUHAMMAD NAUMAN is currently an accomplished academic and a researcher with over 11 years of experience in manufacturing, energy engineering, and innovation and entrepreneurship. He has worked in various universities across the globe, including in Turkey and Saudi Arabia, before joining the Faculty of Integrated Technologies, Universiti Brunei Darussalam, as an Associate Professor, and the Deputy Dean. He has published numerous research papers and articles in reputed international journals and conferences in these fields, in addition to his contributions to 2D/3D printing of energy devices involving functional and smart materials, thin film solar cells, memristors, and other electronic devices, and the inverse estimation of heat and mass transfer. Through his active involvement in various academic and professional activities, including organizing and participating in several workshops, seminars, and conferences, both locally and internationally, he has demonstrated his unwavering commitment to academic excellence and the pursuit of knowledge. Overall, his exceptional achievements and vast experience in mechatronics engineering and related disciplines make him an invaluable member of the academic community. His research interests include energy devices and encompasses other areas, such as image processing and shape memory alloys.

● ● ●