aNy-way Independent Component Analysis

Kuaikuai Duan*¶, Rogers F. Silva*¶, Vince D. Calhoun, Fellow, IEEE, and Jingyu Liu

***Abstract*— Multimodal data fusion is a topic of great interest. Several fusion methods have been proposed to investigate coherent patterns and corresponding linkages across modalities, such as joint independent component analysis (jICA), multiset canonical correlation analysis (mCCA), mCCA+jICA, and parallel ICA. JICA exploits source independence but assumes shared loading parameters. MCCA maximizes correlation linkage across modalities directly but is limited to orthogonal features. While there is no theoretical limit to the number of modalities analyzed together by jICA, mCCA, or the two-step approach mCCA+jICA, these approaches require the *same* number of sources/components for all modalities. Parallel ICA, on the other hand, simultaneously maximizes correlation between modalities and independence of sources, while allowing different number of sources for each modality. However, only a very limited number of modalities and linkage pairs can be optimized. To overcome these limitations, we propose aNy-way ICA, a new model to simultaneously maximize the independence of sources and correlations across modalities. ANy-way ICA combines infomax ICA and Gaussian independent vector analysis (IVA-G) via a shared weight matrix model without orthogonality constraints. Simulation results demonstrate that aNy-way ICA can accurately recover sources and loadings, as well as the true covariance patterns, whether different modalities have the same or different number of sources. Moreover, aNy-way ICA outperforms mCCA and mCCA+jICA in terms of source and loading recovery accuracy, especially under noisy conditions.**

***Index Terms*—Multimodal data fusion, Independence, Correlation, aNy-way ICA.**

***Clinical Relevance*—This establishes a model for N-way data fusion of any number of modalities and linkage pairs, allowing non-orthogonal mixings and different number of sources for different modalities.**

I. Introduction

With the fast advancement of technology, growing medical data from different imaging modalities and omics (e.g. genomics, transcriptomics, epigenomics, microbiomics) can be collected from participants for a variety of diseases. Data from different modalities/omics provide us unique and complementary features for understanding diseases. Thus, integrative/fusion methods are essential to identify unique and coherent biomarkers across modalities.

Several fusion approaches have been proposed to explore the latent shared structures in multimodal datasets in order to discover and establish links among modalities. To that end, joint independent component analysis (jICA) [1] has focused on identifying shared loadings across modalities, while mCCA+jICA [2] proposed a multiset canonical correlation analysis (mCCA) preprocessing in order to improve the correspondence over modalities and more closely satisfy the assumption of shared identical loadings prior to jICA, enabling N-way multimodal fusion. Parallel ICA [3], on the other hand, proposed to run separate (parallel) ICAs, one per modality, while simultaneously maximizing the correlation between specific multimodal loading pairs. This is attractive because both independence and linkage are optimized together rather than in two steps. However, Parallel ICA (and 3-way Parallel ICA) [3, 4] is limited in the number of modalities and linked components it can robustly detect since it optimizes individual correlation pairs rather than the entire underlying linkage structure.

Here, we propose a new model called aNy-way ICA, which overcomes these limitations, optimizing the entire loadings correlation structure of linked components via IVA-G [5], and simultaneously optimizing independence via separate ICAs. Therefore, aNy-way ICA is capable of detecting multiple linked sources over any number of modalities without requiring orthogonality constraints on sources and permitting different number of sources for different modalities. The remainder of the paper is organized as follows. Section II introduces the aNy-way ICA model. Section III describes simulation scenarios. Results and conclusions are shown in sections IV and V, respectively.

II. Method

ANy-way ICA aims to simultaneously maximize the independence of sources $\boldsymbol{s}$ while minimizing the mutual information between subspace component vectors (SCV) $\mathbf{a}_k$. As shown in Fig. 1, without loss of generality, we consider three modalities: structural MRI (sMRI), functional MRI (fMRI) and electroencephalography (EEG), each with 5, 4, and 6 sources, respectively. While the total number of sources C_m is different per modality, at most $K = 4$ SCVs (maximum number of one-to-one correspondences across all modalities) can be extracted. For each modality, ICA decomposes the $N \times L_m$ data $\mathbf{X}_m$ ($m = 1,2,3$ modalities, N is the number of subjects, L_m is the feature dimensionality) into a $C_m \times L_m$ source matrix $\mathbf{S}_\mathrm{m}$ with corresponding $N \times C_m$ loading matrices $\mathbf{A}_m$ (i.e., $\mathbf{X}_m = \mathbf{A}_m \mathbf{S}_m$).

¶These authors have contributed equally to this work.

This study was supported by the National Institutes of Health through the grant 1R01MH106655.

*Corresponding authors: phone: 404-413-4953; fax: 404-413-3393; e-mails: kuaikuaiduan@gmail.com, rsilva@gsu.edu.

Kuaikuai Duan, Rogers F. Silva, Vince D. Calhoun, and Jingyu Liu are with the Tri-institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS) [Georgia State University, Georgia Institute of Technology, Emory University], Atlanta, GA, 30303 USA. (vcalhoun@gsu.edu, jliu75@gsu.edu).

Kuaikuai Duan and Vince D. Calhoun are also with the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA.

Vince D. Calhoun is also with the Department of Psychology, Georgia State University, Atlanta, GA, 30303 USA.

Jingyu Liu is also with the Department of Computer Science, Georgia State University, Atlanta, GA, 30303 USA.

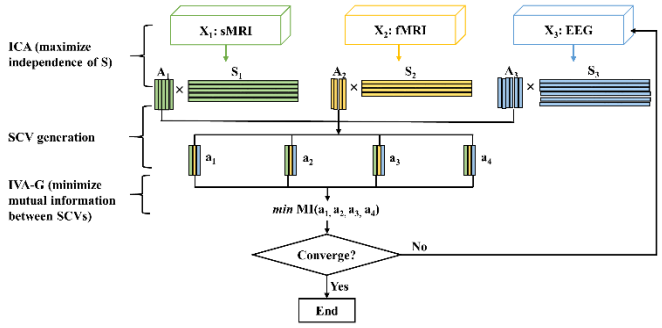

Figure 1 Diagram of aNy-way ICA. Sources $\mathbf{S}_m$ are estimated by maximization of independence in each modality separately. Corresponding loadings $\mathbf{A}_m$ are then organized into SCVs $\mathbf{a}_k$, followed by minimization of their mutual information with IVA-G, which amounts to maximization of the correlation structure within the SCV without orthogonality constraints [6]. The model is optimized with stochastic gradient descent until convergence.

The source matrix is estimated as the product of a weight matrix $\mathbf{W}_m$ and data as $\mathbf{S}_m = \mathbf{W}_m\mathbf{X}_m$. The loading matrix is then reconstructed as the product of data and pseudo-inverse of the source matrix ($\mathbf{A}_m = \mathbf{X}_m\mathbf{S}_m^-$), which is a key strategy of this work. SCVs are then formed by concatenation of linked loadings across the three modalities, where the total number of SCVs K is the minimum component number among the three modalities. Assuming each SCV forms a multivariate Gaussian distribution, mutual information among SCVs is minimized utilizing IVA-G since that has been shown to correspond to mCCA's generalized variance (GENVAR) approach without orthogonality constraints [6].

The cost function of aNy-way ICA is thus defined as:

$$max_{\mathbf{W}} \ \sum_{m=1}^{M} \mathrm{H}(\mathbf{y}_m) - \lambda \cdot \mathrm{I}(\mathbf{a}_1, \cdots, \mathbf{a}_K). \tag{1}$$

where M is the total number of modalities, $\mathrm{H}(\mathbf{y}_m)$ represents the entropy of $\mathbf{y}_m$ (the same as Infomax ICA cost [7]), $\mathbf{y}_m$ is a nonlinear transformation of the bias-adjusted source $\mathbf{u}_m$ ($\mathbf{y}_m = \frac{1}{1+e^{-\mathbf{u}_m}}, \mathbf{u}_m = \mathbf{s}_m + \mathbf{b}_m$), $\mathrm{I}(\mathbf{a}_1, \cdots, \mathbf{a}_K)$ is the mutual information among SCVs $\mathbf{a}_k$, λ is a regularizer to balance between independence (ICA) and correlation (IVA-G) maximization. The entropy of $\mathbf{y}_m$ [7] and mutual information among SCVs $\mathbf{a}_k$ [5] are defined in (2) and (3), respectively:

$$\mathrm{H}(\mathbf{y}_m) = -\mathbb{E}\big[\ln p_{\mathbf{y}_m}(\mathbf{y}_m)\big]. \tag{2}$$

$$\mathrm{I}(\mathbf{a}_1, \cdots, \mathbf{a}_K) = \sum_{k=1}^{K} \mathrm{H}(\mathbf{a}_k) - \mathrm{H}(\mathbf{a}_1, \cdots, \mathbf{a}_K). \tag{3}$$

where $p_{\mathbf{y}_m}(\mathbf{y}_m)$ is the probability density function of the vector $\mathbf{y}_m$. The k-th SCV $\mathbf{a}_k$ can be formed as $\mathbf{a}_k = \mathbf{A}\mathbf{P}_k^\top$, where $\mathbf{P}_k^\top$ is akin to a permutation matrix that selects the loadings from the k-th component of each modality, and $\mathbf{A} = [\mathbf{A}_1, \cdots, \mathbf{A}_M]$ is a concatenation of loadings from all modalities. $\mathrm{H}(\mathbf{a}_k)$ represents the marginal entropy of the k-th SCV and $\mathrm{H}(\mathbf{a}_1, \cdots, \mathbf{a}_K) \approx \sum_m^M \log|\det(\mathbf{S}_m^-\mathbf{P}_m^\top)|$ is the joint entropy of all SCVs, which simplifies to a determinant function on the sources after noting the entropy of the data is constant. The proposed aNy-way ICA is then solved by stochastic gradient descent (SGD), utilizing the relative gradients [8] of (2) and (3) with respect to $\mathbf{W}$, which are given, respectively, as:

$$\nabla\mathbf{W}_m^H = (\mathbf{I} + (\mathbf{1} - 2\mathbf{y}_m)\mathbf{u}_m^\top)\mathbf{W}_m. \tag{4}$$

$$\nabla\mathbf{W}_m^I = -\mathbf{W}_m^{-\top}\mathbf{K}_m\mathbf{V}_{1m}^-\dot{\mathbf{S}}_m^-\mathbf{W}_m \tag{5}$$
$$-\mathbf{W}_m^{-\top}\mathbf{F}_m^\top\mathbf{T}_m^{-\top}\mathbf{Q}_m^\top\mathbf{W}_m.$$

where $\mathbf{I}$ is an identity matrix, $\mathbf{1}$ is a column vector of ones, $\top$ is the matrix transpose operation, $\mathbf{K}_m = \mathrm{diag}\big(\mathrm{std}(\mathbf{V}_{1m}^\top)\big)$ is a diagonal matrix, $\mathrm{std}(\mathbf{V}_{1m}^\top)$ is a vector with the standard deviations from each row of $\mathbf{V}_{1m}^\top$, $\mathbf{V}_{1m}^\top = \mathbf{U}^\top\mathbf{X}_m$ is based on the eigenvalue decomposition (EVD) of $\mathbf{\Sigma} \overset{\text{EVD}}{=} \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top$, with $\mathbf{\Sigma} = \frac{1}{M}\sum_{m=1}^{M}\frac{\mathbf{\Sigma}_m}{\mathrm{trace}(\mathbf{\Sigma}_m)}$ being the average of normalized sample covariance matrices from each modality $\mathbf{\Sigma}_\mathrm{m} = \frac{\mathbf{X}_m\mathbf{X}_m^\top}{L_\mathrm{m}-1}$ ($\mathbf{X}_m$ has zero mean rows and columns), and $\dot{\mathbf{S}}_m^- = [\dot{\mathbf{S}}_{m1}^- \cdots \dot{\mathbf{S}}_{mK}^-]\mathbf{P}_m$ is an intermediate gradient for the pseudo-inverse of the source matrix of the m-th modality, and $\dot{\mathbf{S}}_{mk}^- = -\frac{1}{N-1}\mathbf{X}_m^\top\mathbf{z}_{mk}$ with $\mathbf{z}_{mk} = \mathbf{a}_k(\mathbf{a}_k^\top\mathbf{a}_k)^{-1}\mathbf{P}_{1m}^\top$. Note that $\mathbf{P}_{1m}^\top$ selects the m-th column from the matrix to its left, $\mathbf{P}_m$ expands the concatenation of $\dot{\mathbf{S}}_{mk}^-$ vectors from $L_m \times K$ to $V_m \times C_m$ (effectively padding zero columns to it as needed). Finally, approximating the determinants in the SCV joint entropy term as the product of singular values of $\mathbf{S}_m^-\mathbf{P}_m^\top \overset{\text{SVD}}{=} \mathbf{U}_{2m}\mathbf{D}_m\mathbf{V}_{2m}^\top$ (via singular value decomposition (SVD)) and collecting constant terms into $\mathbf{F}_m = \mathbf{U}_{2m}^\top\mathbf{V}_{1m}^{-\top}\mathbf{K}_m$ and $\mathbf{Q}_m = \mathbf{P}_m^\top\mathbf{V}_{2m}$, yields the final form in (5), where $\mathbf{T}_m = \mathbf{F}_m\mathbf{W}_m^{-1}\mathbf{Q}_m$.

Since the source/loading ordering from separate ICAs per modality would be arbitrary (not linked), it is important *not* to over-emphasize independence early during the optimization. Rather, we propose to project the multimodal data into a *shared* variance-adjusted space, i.e., along the shared basis $\mathbf{U}$ defined above, which provides a good (linked) starting point for the optimization. Also, we observed improved loading (SCV) linkage and source alignment by initially optimizing only IVA-G for a few steps/epochs (here, 5) before optimizing (1). Moreover, the regularizer parameter λ is initialized with 1 and adaptively adjusted by annealing it according to the entropy trend during the optimization: whenever entropy decreases with a slope larger than a threshold (here, 10^{-4}), this triggers λ to be reduced by a factor (here, 0.8).

III. SIMULATION

Simulated sMRI, fMRI and EEG data were employed to validate the proposed aNy-way IVA, where sMRI and fMRI sources were simulated with the SimTB toolbox [9]. For EEG sources, we utilized a subset of the simulated independent Laplacian timeseries in [10], which have an autocorrelation between timepoint t and $t-1$ in the order of 0.85, and between t and $t-10$ in the order of 0.2. Each SCV was sampled from a zero mean m-dimensional multivariate Gaussian distribution, each SCV with a unique covariance structure with correlation values in the range [0.2,0.8]. All SCVs were then organized into loading matrices $\mathbf{A}_m$. If applicable, extra loadings for non-linked sources were sampled from a standard Normal distribution and padded to $\mathbf{A}_m$. Finally, the data for each modality was generated as the product of the loading matrix and corresponding source matrix as $\mathbf{X}_m = \mathbf{A}_m\mathbf{S}_m$. The following scenarios were then examined:

Scenario 1: All 3 modalities had the same number of sources $C_m = 7$. The number of subjects was fixed at $N = 500$. Feature dimensions (L_m) for sMRI, fMRI and EEG data were 31064 (voxels), 17420 (voxels), 4444 (time points), respectively (the same for scenario 2). Comparisons were performed against both mCCA and mCCA+jICA.

Scenario 2: sMRI, fMRI, and EEG data had different number of sources. In part (a), we assessed the effect of varying the number of linked SCVs K from 3 to 11 with a step size of 2 while keeping $N = 500$ fixed. Therefore, $C_m = K + 1, K, K + 2$, respectively, such that the fMRI data always had the minimum C_m (for SCV number from 3 to 9), and sMRI and EEG data had one and two extra components, respectively, with exception of the last case where $C_m = 11,11,12$. In part (b) the number of sources C_m for sMRI, fMRI and EEG data were fixed at 8, 7, 9, respectively, while the subject number N varied from 100 to 1000. Finally, in part (c) we assessed the effect of additive Gaussian white noise with C_m fixed at 8, 7, 9 for sMRI, fMRI and EEG, respectively, N fixed at 500, and the signal-to-noise ratio (SNR) varying from 0db to 30dB with a step size of 5dB. For each SNR level, the noise samples $\mathbf{E}_{im}$ were generated 5 times yielding noisy data $\mathbf{X}_{im}^{Noisy} = \mathbf{X}_m + \mathbf{E}_{im}, i = 1, \ldots, 5$ for modality m. For clarity, note that scenarios 1, 2(a), and 2(b) were free from noise.

Several component-wise performance measures for each modality were computed for aNy-way ICA, mCCA and mCCA+jICA in each of the scenarios above, including correlations between recovered *sources* and ground-truth, as well as correlations between recovered *loadings* and ground-truth. For mCCA and mCCA+jICA, we estimated both $K = \min_m C_m$ and $K = \max_m C_m$ for all modalities (two separate runs). To examine SCV recovery, we also calculated the MSE value between the cross-correlation of the recovered SCVs and that of the ground-truth, considering the whole matrix, as well as the diagonal blocks and off-diagonal portions separately. For scenarios 2(a) and 2(b), correlations were averaged across matched components per modality, while for scenario 2(c), correlations were averaged across matched components and runs. MSE values were averaged across runs.

IV. RESULT

Fig. 2 illustrates the seven sources for (a) sMRI, (b) fMRI, and (c) EEG data in scenario 1. Their recovery accuracies are plotted in Fig. 3: (a)-(d), under noise-free conditions, aNy-way ICA recovered all designed SCV structures very well (the MSE value was 2.8E-4), while both mCCA and mCCA+jICA failed to recover, as indicated by the large values on off-diagonal positions likely due to overfitting of correlations and inability to estimate non-orthogonal sources in mCCA, plus potential residual source misalignments from mCCA leaking into jICA for mCCA+jICA. Fig. 3 (e)-(f) demonstrates that aNy-way ICA accurately recovered both sources and loading coefficients for in all three modalities, contrary to mCCA and mCCA+jICA.

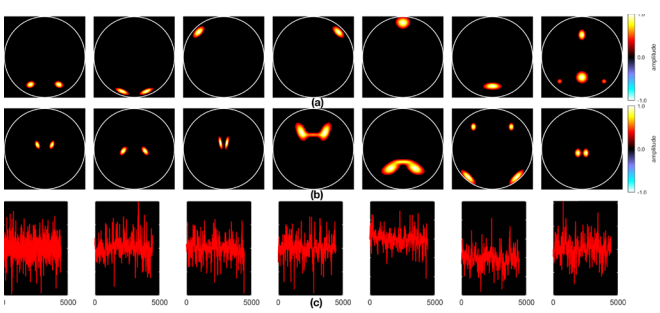

Figure 2 *Scenario 1* sources for (a) sMRI, (b) fMRI, and (c) EEG data.

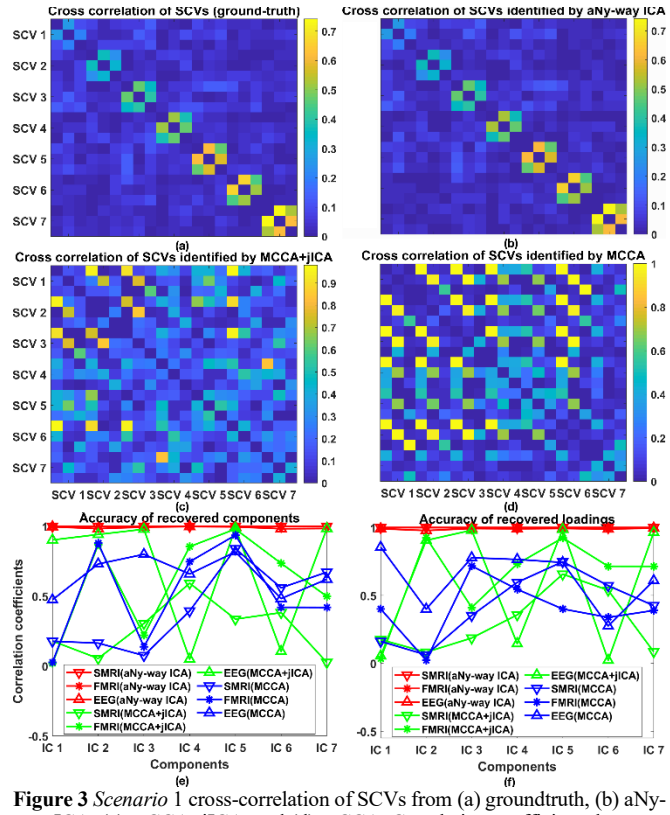

Figure 3 *Scenario* 1 cross-correlation of SCVs from (a) groundtruth, (b) aNy-way ICA, (c) mCCA+jICA, and (d) mCCA. Correlation coefficients between (e) recovered sources and ground-truth, and between (f) recovered loadings and ground-truth for each independent component (IC): sMRI (downward triangle), fMRI (star), and EEG data (upward triangle) from aNy-way ICA (red), mCCA+jICA(green) and mCCA(blue). Note shapes and colors are consistent throughout the paper.

Performance measures for scenario 2(a) are shown in Fig. 4. ANy-way ICA accurately recovered components and loadings for sMRI and fMRI data, with slightly reduced recovery accuracy for EEG data (especially for 9 and 11 SCV). The inter-symbol interference [11] for the recovered loading matrix and minimum distance indices [12] for the recovered source matrix corroborate these observations and, thus, are omitted here. Lower accuracies for EEG data are likely due to its relatively small feature dimensionality ($L_3 = 4444$) and designed weaker correlation linkage with other modalities within each SCV. Overall, aNy-way ICA outperformed mCCA+jICA and mCCA in terms of source and loading recovery. mCCA+jICA and mCCA decompositions with the minimum component number of the three modalities had higher accuracies compared to those obtained with the maximum component number (similarly for scenarios 2(b) and 2(c), which have been omitted). As the SCV number increased (e.g., 9 or 11 SCVs), source and loading accuracies from both mCCA and mCCA+jICA dropped. Table I lists MSE values for aNy-way ICA's recovered covariance structure. ANy-way ICA fully recovered the designed covariance structure when SCV number was 3, 5, and 7, and largely recovered it when SCV number was 9 and 11 (actually, sMRI and fMRI were aligned well but not sMRI-EEG and fMRI-EEG).

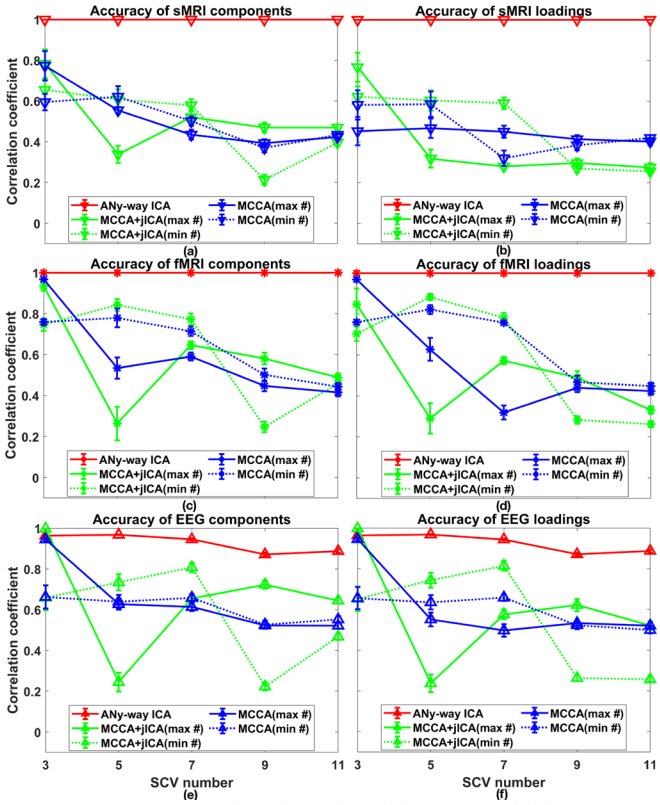

Figure 4 *Scenario 2(a)*, where the number of the SCVs varied from 3 to 11. Recovered *source* accuracies for (a) sMRI, (c) fMRI, and (e) EEG data, and *loading* accuracies for (b) sMRI, (d) fMRI, and (f) EEG data using aNy-way ICA (red), mCCA+jICA with maximum (solid green) or minimum component number (dashed green), mCCA with maximum (solid blue) or minimum component number (dashed blue).

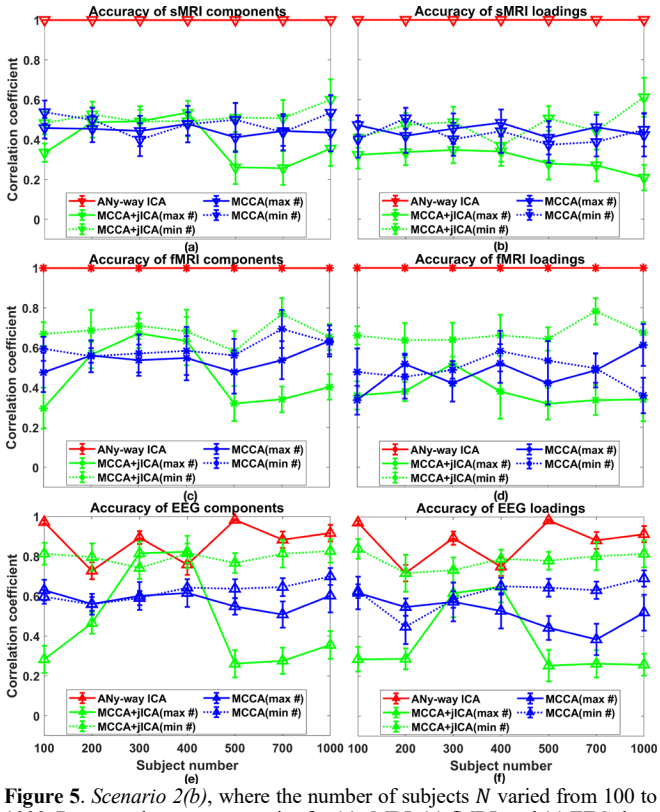

Figure 5. *Scenario 2(b)*, where the number of subjects N varied from 100 to 1000. Recovered *source* accuracies for (a) sMRI, (c) fMRI, and (e) EEG data, and *loading* accuracies for (b) sMRI, (d) fMRI, and (f) EEG data.

TABLE I. MSE VALUES BETWEEN THE CROSS-CORRELATION OF THE RECOVERED SCVS AND THAT OF THE GROUND-TRUTH BY ANY-WAY ICA WHILE VARYING THE NUMBER OF SCVS

MSE	Number of SCVs (K)				
	3	*5*	*7*	*9*	*11*
All*	2×10^{-4}	5×10^{-4}	6×10^{-4}	2×10^{-2}	6×10^{-3}
Diag*	3×10^{-5}	2×10^{-5}	2×10^{-4}	8×10^{-2}	3×10^{-2}
Off*	3×10^{-4}	7×10^{-4}	6×10^{-4}	9×10^{-3}	3×10^{-3}

*Note: 'All' denotes MSE of whole matrix, 'Diag' denotes MSE of diagonal blocks only, 'Off' denotes MSE of off-diagonal blocks only. The same for all tables.

Fig. 5 demonstrates performance measures for scenario 2(b). As number of subjects increases, the recovery accuracies of sources and loadings, as well as the covariance structure of SCVs, increased, although there were some variations when the subject number was less than 400. Variation of accuracies of EEG data from aNy-way ICA were probably due to weak covariance structures generated from few samples. MSE values for the whole, diagonal blocks as well as off-diagonal blocks of the recovered cross-correlation matrix of SCV are listed in Table II (a). When number of subjects is larger than 400, the MSE values are smaller (except for N=100).

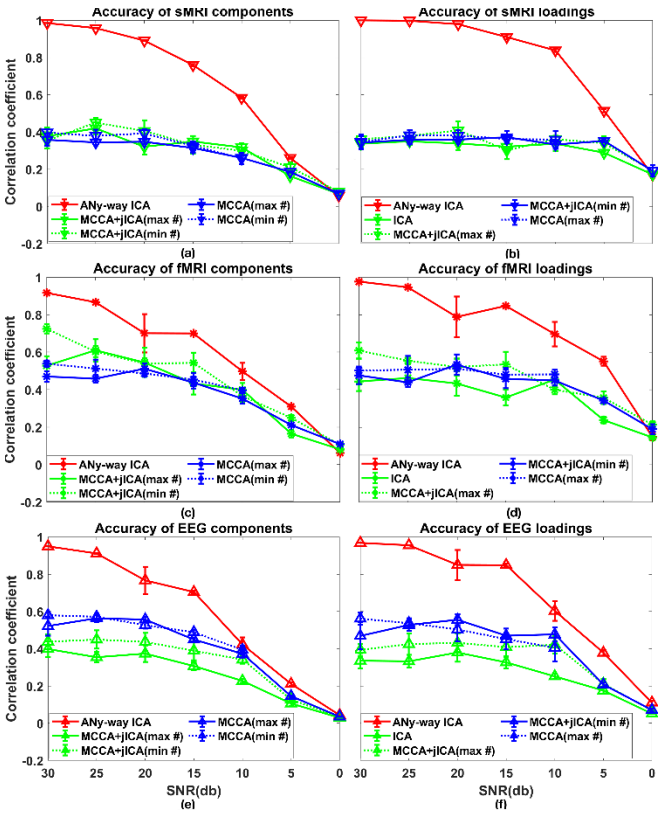

Figure 6. *Scenario 2(c)*, where SNR varied from 30db to 0db. Recovered *source* accuracies for (a) sMRI, (c) fMRI, and (e) EEG data, and *loading* accuracies for (b) sMRI, (d) fMRI, and (f) EEG data.

Performance measures for scenario 2(c) are plotted in Fig. 6. Increasing the strength of noise superimposed on data, recovery accuracies of sources and loadings dropped. ANy-way ICA outperformed mCCA and mCCA+jICA when SNR $\geq$ 5dB. When SNR = 0db, aNy-way ICA, mCCA, as well as mCCA+jICA failed to recover the sources and loadings. Table II (b) summarizes MSE values of the whole, diagonal blocks and off-diagonal blocks of the cross-correlation matrix of

recovered SCVs, reflecting larger MSE values for smaller SNR (i.e. heavier noise).

TABLE II. MSE VALUES BETWEEN THE CROSS-CORRELATION OF THE RECOVERED SCVS AND THAT OF THE GROUND-TRUTH BY ANY-WAY ICA WHILE VARYING THE NUMBER OF SUBJECTS OR SNR.

MSE							
(a) Varying Number of Subjects (N)				***(b) Varying SNR***			
N	All*	Diag*	Off*	SNR	All*	Diag*	Off*
100	8×10^{-4}	3×10^{-4}	9×10^{-4}	0	4×10^{-2}	2×10^{-1}	1×10^{-2}
200	2×10^{-2}	6×10^{-2}	1×10^{-2}	5	3×10^{-2}	2×10^{-1}	4×10^{-3}
300	2×10^{-2}	6×10^{-2}	9×10^{-3}	10	3×10^{-2}	1×10^{-1}	1×10^{-2}
400	7×10^{-3}	2×10^{-2}	5×10^{-3}	15	1×10^{-2}	7×10^{-2}	2×10^{-3}
500	4×10^{-4}	2×10^{-5}	5×10^{-4}	20	8×10^{-3}	3×10^{-2}	4×10^{-3}
700	1×10^{-3}	4×10^{-3}	7×10^{-4}	25	2×10^{-3}	5×10^{-3}	1×10^{-3}
1000	6×10^{-3}	2×10^{-2}	3×10^{-3}	30	9×10^{-4}	4×10^{-4}	1×10^{-3}

V. CONCLUSION

The proposed aNy-way ICA simultaneously optimizes independence of sources and correlations across modalities without orthogonality constraints. It can be flexibly applied to N-way multimodal fusion with the same or different number of sources per modality. Simulation results support that aNy-way ICA can recover sources and loadings, as well as the true covariance patterns with improved recovery accuracies compared to mCCA and mCCA+jICA, especially under noisy cases.

REFERENCES

[1] M. Moosmann, T. Eichele, H. Nordby, K. Hugdahl, and V. D. Calhoun, "Joint independent component analysis for simultaneous EEG-fMRI: principle and simulation," *Int J Psychophysiol,* vol. 67, no. 3, pp. 212-21, Mar 2008.

[2] J. Sui *et al*., "Discriminating schizophrenia and bipolar disorder by fusing fMRI and DTI in a multimodal CCA+ joint ICA model," *Neuroimage,* vol. 57, no. 3, pp. 839-55, Aug 1 2011.

[3] J. Y. Liu, G. Pearlson, A. Windemuth, G. Ruano, N. I. Perrone-Bizzozero, and V. Calhoun, "Combining fMRI and SNP Data to Investigate Connections Between Brain Function and Genetics Using Parallel ICA," (in English), *Hum Brain Mapp,* vol. 30, no. 1, pp. 241-255, Jan 2009.

[4] V. M. Vergara, A. Ulloa, V. D. Calhoun, D. Boutte, J. Y. Chen, and J. Y. Liu, "A three-way parallel ICA approach to analyze links among genetics, brain structure and brain function," (in English), *Neuroimage,* vol. 98, pp. 386-394, Sep 2014.

[5] M. Anderson, T. Adali, and X. L. Li, "Joint Blind Source Separation With Multivariate Gaussian Model: Algorithms and Performance Analysis," (in English), *Ieee T Signal Proces,* vol. 60, no. 4, pp. 1672-1683, Apr 2012.

[6] M. Anderson, X. L. Li, and T. Adali, "Nonorthogonal Independent Vector Analysis Using Multivariate Gaussian Model," (in English), *Latent Variable Analysis and Signal Separation,* vol. 6365, pp. 354-361, 2010.

[7] A. J. Bell and T. J. Sejnowski, "An Information Maximization Approach to Blind Separation and Blind Deconvolution," (in English), *Neural Comput,* vol. 7, no. 6, pp. 1129-1159, Nov 1995.

[8] C. J. Pierre Comon, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, 1st ed. The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK: Academic Press. , 2010.

[9] E. B. Erhardt, E. A. Allen, Y. H. Wei, T. Eichele, and V. D. Calhoun, "SimTB, a simulation toolbox for fMRI data under a model of spatiotemporal separability," (in English), *Neuroimage,* vol. 59, no. 4, pp. 4160-4167, Feb 15 2012.

[10] Rogers F. Silva, Sergey M. Plis, Tulay Adali, Marios S. Pattichis, and V. D. Calhoun, "Multidataset Independent Subspace Analysis with Application to Multimodal Fusion," *arXiv preprint,* 2019.

[11] S. Amari, A. Cichocki, and H. H. Yang, "A new learning algorithm for blind signal separation," (in English), *Adv Neur In,* vol. 8, pp. 757-763, 1996.

[12] P. Ilmonen, K. Nordhausen, H. Oja, and E. Ollila, "A New Performance Index for ICA: Properties, Computation and Asymptotic Analysis," (in English), *Latent Variable Analysis and Signal Separation,* vol. 6365, pp. 229-+, 2010.