



## NIRS-ICA: A MATLAB Toolbox for Independent Component Analysis Applied in fNIRS Studies

Yang Zhao1t, Pei-Pei Sun1t, Fu-Lun Tan1, Xin Hou1 and Chao-Zhe Zhu1,2\*

<sup>1</sup> State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China, <sup>2</sup> IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China

Independent component analysis (ICA) is a multivariate approach that has been widely used in analyzing brain imaging data. In the field of functional near-infrared spectroscopy (fNIRS), its promising effectiveness has been shown in both removing noise and extracting neuronal activity-related sources. The application of ICA remains challenging due to its complexity in usage, and an easy-to-use toolbox dedicated to ICA processing is still lacking in the fNIRS community. In this study, we propose NIRS-ICA, an open-source MATLAB toolbox to ease the difficulty of ICA application for fNIRS studies. NIRS-ICA incorporates commonly used ICA algorithms for source separation, user-friendly GUI, and quantitative evaluation metrics assisting source selection, which facilitate both removing noise and extracting neuronal activity-related sources. The options used in the processing can also be reported easily, which promotes using ICA in a more reproducible way. The proposed toolbox is validated and demonstrated based on both simulative and real fNIRS datasets. We expect the release of the toolbox will extent the application for ICA in the fNIRS community.

#### OPEN ACCESS

#### Edited by:

Pedro Antonio Valdes-Sosa, University of Electronic Science and Technology of China, China

#### Reviewed by:

Ahmad Chaddad, Guilin University of Electronic Technology, China Hendrik Santosa, University of Pittsburgh, United States

#### \*Correspondence:

Chao-Zhe Zhu czzhu@bnu.edu.cn <sup>†</sup>These authors have contributed equally to this work

> Received: 22 March 2021 Accepted: 08 June 2021 Published: 14 July 2021

#### Citation:

Zhao Y, Sun P-P, Tan F-L, Hou X and Zhu C-Z (2021) NIRS-ICA: A MATLAB Toolbox for Independent Component Analysis Applied in fNIRS Studies. Front. Neuroinform. 15:683735. doi: 10.3389/fninf.2021.683735 Keywords: functional near-infrared spectroscopy, data processing, multivariate analysis, independent component analysis, blind source separation, MATLAB toolbox, software

#### INTRODUCTION

Functional near-infrared spectroscopy (fNIRS) is a non-invasive optical imaging technology that has been widely used in studying human brain activity (Boas et al., 2014). It has been utilized to investigate the task-evoked neuronal response of various cognitive functions as well as spontaneous neural activity, which is reflected by resting-state functional connectivity (White et al., 2009). Compared with fMRI, fNIRS is portable, less sensitive to head motion, and has high ecological validity, which make it very suitable to study human social interactions by hyper-scanning multiple brains in naturalistic environment, and special populations such as infants, patients, and the elderly population (Ferrari and Quaresima, 2012; Scholkmann et al., 2013). It also has higher temporal resolution and a more comprehensive measurement for hemoglobin changes in the cerebral cortex, which potentially provides a deeper understanding of the neurovascular coupling process (Steinbrink et al., 2006).

However, besides the sources of neural activity, various physiological and non-physiological noise components have been found in fNIRS signals (Scholkmann et al., 2014). Since fNIRS detects brain activity based on the banana-shaped light path traveled through both extra- and intra-cerebral compartments, the neuronal activity-unrelated physiological changes,

such as heart rate, respiration, Mayer waves, and blood pressure, in these compartments can induce changes in optical intensity, and further included as noise components in fNIRS signals (Liebert et al., 2004). Moreover, although fNIRS is less sensitive to head motion than fMRI, motion artifacts, generated from optode decoupling with the scalp, are still often presented in many fNIRS studies (Robertson and Douglas, 2010; Brigadoi et al., 2014; Cui et al., 2015). These latent noise processes are inherent in fNIRS signals and strongly reduce the quality of the recorded data. Therefore, the reduction of noises and extraction of neuronal activity-related components are essential for the fNIRS studies.

The methodology for processing fNIRS data is divided into two classes, i.e., univariate and multivariate methods by Scholkmann et al. (2014). Univariate methods independently process the data of each channel. These approaches include preprocessing methods such as bandpass or wavelet filtering and analysis methods like general linear model (GLM) regression, which have been implemented in public-available software such as NIRS-SPM and Homer2 (Huppert et al., 2009; Ye et al., 2009). Multivariate methods, another type of analysis method, have also been proposed and applied for fNIRS. Instead of processing data of each fNIRS channel separately, multivariate methods additionally exploit the correlation between the data of different channels and have shown better performance in removing noise or extracting neuronal activity-related sources compared to univariate methods (Scholkmann et al., 2014).

Independent component analysis (ICA) is a multivariate approach for processing fNIRS data. Two ways of using ICA had been presented in fNIRS literature. First, as a preprocessing method, it has been used to remove noises including both motion artifacts and physiological noises (Kohno et al., 2007; Robertson and Douglas, 2010; Virtanen et al., 2011; Chaddad, 2014). Second, ICA has been applied as a data-driven analysis method for exploring the neuronal activity-related sources, or sources of interest (SOI), including event-related fast optical signal, task-evoked hemodynamic response, and resting-state functional connectivity in typical functional systems (Morren et al., 2004; Katura et al., 2008; Medvedev et al., 2008; Markham et al., 2009; Zhang et al., 2010). For task data analysis, it can extract task-evoked hemodynamic responses without using a prior-defined hemodynamic response function (HRF), as in the conventional regression-based method, i.e., GLM. For analyzing resting-state functional connectivity, compared with the seedbased correlation analysis, ICA does not have the bias that originated from the selection of seed channels. Since the noises have been separated while extracting the SOI, the ICA approach usually outperforms these traditional methods (Kohno et al., 2007; Markham et al., 2009; Zhang et al., 2010).

Some of the fNIRS software packages have already included ICA as a module of their analysis pipeline (Strangman et al., 2009; Raggam, 2020). However, the usages of ICA in these toolboxes are limited to certain types of fNIRS devices, or processing steps (e.g., preprocessing). Therefore, a toolbox that can fully unveil the potential of ICA analysis in the fNIRS field is still lacking. In this study, we propose NIRS-ICA, a user-friendly, systemic integrated, public-available MATLAB toolbox to ease the ICA application for the fNIRS dataset. NIRS-ICA

integrates commonly used ICA algorithms for source separation and provides user-friendly GUIs which can be used for both noise reduction and extracting SOI. In addition, NIRS-ICA also incorporates quantitative evaluation metrics to evaluate the separated sources. Therefore, users can speed up the procedure of source selection by ranking the sources based on the value of the metrics. The options used in the processing can also be output as a report easily, which facilitates reproducing the ICA method in another study.

This manuscript is structured as follows: In Section "Materials and Methods," the mathematical model and applications of ICA used in the NIRS-ICA are first illustrated (sections "Mathematical Model of ICA" and "Applications"). Then we present the simulative and real fNIRS dataset used for validating the functionality of noise reduction and extracting SOI (section "Validations"). In Section "Implementations and Results," the developed GUIs and usages of NIRS-ICA for noise reduction and extracting SOI are demonstrated based on the simulative and real fNIRS datasets, respectively. The results of the simulative and real fNIRS experiments are also compared with conventional analysis methods. Finally, the proposed approach and toolbox are discussed in Section "Discussion."

### MATERIALS AND METHODS

#### Mathematical Model of ICA

The mathematical model of ICA applied in fNIRS data analysis is briefly introduced as follows: In general, ICA is a datadriven method that decomposes the data of fNIRS into multiple source components and their corresponding weights, which is mathematically expressed as:

$$X = A \cdot S, \tag{1}$$

where *X* is the matrix containing the observed fNIRS data. *S* and *A* are the matrices of source signals and their associated weights, respectively. Since *A* and *S* are both underdetermined, constraints must be imposed for the decomposition. In the ICA framework, it assumes the source signals in *S* are statistically independent of each other. This constraint can be imposed either on the temporal or spatial dimension of the fNIRS data of a subject, namely, temporal ICA (TICA) or spatial ICA (SICA) (Calhoun et al., 2001b). The bidimensional decomposition model is schematically described in **Figure 1**. In TICA, each channel of the recorded fNIRS data is considered as a random variable and its values in different time points are regarded as samples. Therefore, each time course of the fNIRS data is modeled as a weighted linear combination of temporally independent source time courses, i.e.,

$$\begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \cdot \begin{bmatrix} s_1(t) \\ \vdots \\ s_m(t) \end{bmatrix}, \quad (2)$$

where  $x_i(t)$ :  $(i \quad 1, 2, ..., n, t = 1, 2, ..., T)$  denotes the observed fNIRS time course of the *i*th channel.  $s_i(t)$ : (i = 1, 2, ..., m, t = 1, 2, ..., T) is the *i*th source



time course. The *i*th column of the mixing matrix, i.e.,  $a_i = [a_{1i}, a_{2i}, \ldots, a_{ni}]^T$ , contains the contributions of  $s_i(t)$  to the observed signals of *n* fNIRS channels, which represent the spatial map of the source. Note that  $m \leq n$  otherwise Eq. 2 is underdetermined. In SICA, the fNIRS data of different time points are regarded as different spatial maps, which are weighted linear combinations of spatially independent source spatial maps, i.e.,

$$\begin{bmatrix} x_1(s) \\ \vdots \\ x_T(s) \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{T1} & \cdots & a_{Tm} \end{bmatrix} \cdot \begin{bmatrix} s_1(s) \\ \vdots \\ s_m(s) \end{bmatrix}, \quad (3)$$

where  $x_i(s)$  : (*i* =  $1, 2, \ldots, n, s$  $1, 2, \ldots, n$ =denotes the spatial map at the *i*th time point.  $s_i(t)$ :  $(i = 1, 2, \dots, m, t = 1, 2, \dots, T)$  is the *i*th source spatial map and  $m \leq T$ .  $a_i = [a_{1i}, a_{2i}, \dots, a_{Ti}]^T$  is the time course of the ith spatial map. Since the TICA and SICA decompose fNIRS data in different ways, their results should be interpreted correctly. TICA is better for finding latent processes such as motion artifacts and neuronal activity-related hemodynamic responses which are temporally independent which each other, but their spatial distributions can be related. In contrast, SICA identifies sources that happened in independent spatial locations but may have correlated temporal dynamics. Theoretically, both TICA and SICA can be applied to fNIRS data. However, in practice, TICA is more often used than SICA since fNIRS often has few samples in the spatial dimension, i.e., has a small number of channels. Performing ICA with a small number of samples may lead to overfitting problems and result in low reproducibility (Weihong et al., 2012). Therefore, SICA should be used with caution especially when the number of channels is small.

## Applications

#### **Decomposition Algorithms**

Before ICA, a prewhitening step is performed using principal component analysis (PCA), which means applying a transformation matrix to the data, i.e.,

$$X' = V \cdot X, \tag{4}$$

where V is an orthogonal matrix obtained by singular value decomposition of the covariance matrix of X (Huppert et al., 2009). The prewhitening step limits the transformation matrix searched in the ICA step as an orthogonal matrix, which can reduce the computational cost (Hyvarinen et al., 2004). It also reduces the data dimension so that it prevents the over-fitting problem of ICA. The number of sources (m) is determined in this step. By default, NIRS-ICA determines the number of sources based on retaining 99% of the data variance using PCA (Zhang et al., 2010). For increasing flexibility, determining the number of sources based on information-theoretic criteria, i.e., Akaike's information criterion (AIC), Bayesian information criterion (BIC), or users' input is also implemented. Other preprocessing steps such as detrend and bandpass filtering may also be helpful to achieve an ideal decomposition (Vergara et al., 2017). Therefore, we incorporate commonly used pre-processing methods in the NIRS-ICA package.

Principal component analysis expresses the raw data using orthogonal vectors, which may not be able to fully characterize the true mixing model of sources (McKeown et al., 1998; Calhoun et al., 2009). After the PCA-based prewhitening step, the whitened data are further projected to the sources space for estimation of independent sources, i.e.,

$$\hat{S} = W \cdot X', \tag{5}$$

where W is the unmixing matrix and  $\hat{S}$  is the estimated sources time courses. This procedure can be considered as an optimization problem whose objective function is the statistical independence between sources in  $\hat{S}$ . Then the mixing matrix Ais estimated as  $\hat{A} = (W \cdot V)^{-1}$ . There are mainly two types of ICA algorithms to achieve independence in previous fNIRS studies (Adali et al., 2014). The first type of ICA algorithm involves using high-order statistics (HOS) such as FastICA and Infomax, which decompose X by maximizing non-Gaussianity of  $s_i(t)$  or  $s_i(s)$  to achieve independence between source time courses or spatial maps (Hyvarinen, 1999). The second type exploits the sample dependence (SOS). For example, secondorder blind identification (SOBI) involves joint diagonalization of the sample-delayed correlation matrix of X (Belouchrani et al., 1997). Since FastICA and SOBI are the most commonly used algorithms in fNIRS literature, we incorporate them in the current version of NIRS-ICA. The parameters of the above algorithms are listed in the Supplementary Material. The ambiguity of order, signs, and scales always remain in the sources decomposed by ICA (Hyvarinen et al., 2004). However, only the ambiguity of signs may influence the selection of sources. Therefore, a button to change the sign of a source is implemented in the interface of detailed displaying sources.

#### Source Selection

In the source selection step, the separated sources are manually selected based on their spatial and temporal features. Taking the sources of motion artifacts as an example, the spike-shaped waveforms are often found in their time course, which are generated by the decoupling of the optode and scalp when participants move their heads (Cui et al., 2015). Meanwhile, its corresponding spatial map may present a global pattern since the head motion causes displacement of the optode holder which influences the contact of almost all of the optodes. Other types of noises include heart rate, breath, and Mayer wave, which can generate sources whose time course usually has a narrow frequency spectrum, and spatial map depicts global pattern (Zhang et al., 2010). In NIRS-ICA, we also propose several quantitative metrics to rank the separated sources, which can facilitate the manual source selection procedure. Three metrics are implemented to facilitate identifying noise-related sources. The spike-shaped metric evaluates sources' time course for identifying motion artifacts based on Scholkmann et al. (2010). Specifically, the moving standard deviation (MSD) is first calculated using the equation:

MSD(t)

$$= \frac{1}{2k+1} \left[ \sum_{j=-k}^{k} \hat{s}^2(t+j) - \frac{1}{2k+1} \left( \sum_{j=-k}^{k} \hat{s}(t+j) \right)^2 \right]^{\frac{1}{2}}, \quad (6)$$

where  $t \in \{k + 1, k + 2, ..., T - k\}$  and T is the length of the time course. W = 2k + 1 is the length of the window for calculating the MSD(t), which can be specified by users. In general, it means that when the time course depicts a high standard deviation, i.e., contains many large spikes, in a time window W around time point t, it will have high MSD(t).

Therefore, the *W* should be similar to the period of spikes that users aimed to identify. Second, the time points whose *MSD* are outliers, i.e., greater than the threshold:  $mean(MSD) \pm n \cdot std(MSD)$ , are determined as motion artifacts. The default value of *W* and *n* are 2s and 3, respectively. Finally, the value of the spike-shaped metric of one source is derived using the number of outliers identified in its time course.

Correlation with external input (*CEI*) metric evaluates sources by correlating their time courses with the time course recorded by external devices such as accelerometer, physiological instruments, or short channel of fNIRS. Pearson's correlation coefficient is used for calculating the correlation.

The spatial homogenous metric quantifies the spatial map of sources, which is evaluated based on the coefficient of spatial uniformity (CSU) (Kohno et al., 2007):

$$CSU(i) = \left| \frac{\overline{c}_i}{\sigma(c_i)} \right|, \tag{7}$$

where  $\bar{c}_i$  and  $\sigma(c_i)$  represent the mean and standard deviation of the spatial map of the *i*th source. Therefore, spatial maps with high mean and low standard deviation will have high CSU values and can be regarded as global noises.

Two metrics are proposed to explore neuronal activity-related sources in an fNIRS dataset. To investigate neuronal activityrelated source evoked by task stimuli, it often involves selecting SOI using a reference time course, which can be created by convolving the HRF and a square curve of task design of an experiment (McKeown et al., 1998). The Pearson correlation coefficient is used to compare the time course of sources and the generated reference time course. Task onsets and durations can be input to generate the reference time course when the user selects the metric of the reference time course. For RSFC detection, sources whose spatial map follows the hypothetic RSFC pattern and time course has prominent low-frequency (around the band of 0.01–0.1 Hz) spectrum are often selected (Cordes et al., 2001). The goodness of fit (GOF) metric is implemented to measure the similarity between the spatial map of the hypothesis and those of the separated sources:

$$GOF(i) = \left| \frac{\mathbf{y} \cdot \mathbf{c}_i}{|\mathbf{y} \cdot \mathbf{c}_i|} \right|,$$
 (8)

in which  $y = [y_1, y_2, ..., y_n]$  denotes a vector of the Boolean variable representing the mask of the spatial template and ! is the logical operator of negation. In other words, users can let  $y_i = 1$  if the *i*th channel belongs to the region of interest (ROI), otherwise  $y_i = 0$ . Spatial maps with high value inside ROI and low value outside ROI will have a high value of GOF.

#### Postprocessing

When using ICA to remove noises, the indices of noise sources are recorded, and the cleaned fNIRS data are reconstructed by discarding these noise sources using the equation:

$$X_c = \sum_{j \in H} \hat{a}_j \otimes \hat{s}_j, \qquad (9)$$



where  $H \subseteq \{1, \ldots, m\}$  is the set including the indices of nonnoise sources and  $X_c$  is the noise-cleaned fNIRS data. The cleaned fNIRS data can then be analyzed by traditional methods such as GLM regression using other fNIRS software packages.

Using NIRS-ICA to explore SOI for a group of participants, the SOI is first selected for each participant based on methods mentioned in Section "Source Selection." Then the grouplevel ICA results are derived using the selected source of each individual. Specifically, to view the group spatial map of the SOI, the individual spatial map is first transformed to z-map (Calhoun et al., 2001a). Then the group z-map is derived by averaging individual z-map, and the t-map is calculated by performing the two-tailed one-sample *t*-test (against zero) in a channel-wise manner (Zhang et al., 2010). The group-level time course is visualized based on the design of the experiment. For experiments that involve external task stimuli, the group-level time course of SOI is viewed by averaging individual time courses of SOI based on the task stimuli (Katura et al., 2008). Note that since the length of blocks may be different both within and across participants, the time courses are adjusted to have the minimum length of the blocks by removing the time points in the middle of the block. For the RSFC detection, since the sources of RSFC often have low-frequency spectrums, the group-level frequency spectrum of sources is shown instead of the time course (White et al., 2009).

#### Validations

#### Simulative fNIRS Dataset for Noise Reduction

A simulative dataset is generated for validating the noise reduction mode of NIRS-ICA. Channels arranged as an  $8 \times 8$  matrix are used for the detection of neuronal activity in an

experiment with block-design tasks. Six sources, including one neuronal activity-related source and five noise sources, are generated for reconstructing the recorded fNIRS data using Eq. 1. Specifically, for the neuronal activity-related source, its temporal mode is simulated by convoluting a square wave of the task design (block-design) with HRF and its spatial mode is set to be locally distributed. For physiological noises, including heart rate, breath, Mayer wave, and low-frequency noise, their temporal modes are generated using sinusoidal waves with noise-specific frequencies (Supplementary Table 1). Their spatial modes are designed as global distribution. In addition, we simulate task-related motion artifacts on the edge of the channel matrix, which can lead to the detection of falsepositive activation. Finally, channel-specific Gaussian noise is added to the data of each channel. Representative temporal and spatial modes of the simulated sources are shown in Supplementary Figure 1. The above simulations are repeated 10 times to produce a dataset acquired from a group of subjects. The detailed procedure of generating neuronal and noise sources can be found in the Supplementary Material. NIRS-ICA is used to remove noises for the generated dataset. The decomposition parameters of ICA are TICA and SOBI with 100 time-delayed correlation matrices. Ten sources are retained using PCA. To demonstrate noise reduction performance, both ICA-processed data and raw data are analyzed using the conventional GLM method implemented in NIRS-SPM (Ye et al., 2009). Preprocessing procedures, including a DCT-based detrending and HRF-based high-pass filtering implemented in NIRS-SPM, are used before GLM regression. Group-level βmap is calculated by averaging individuals' β-maps and t-map

	o noo halana da bib in an actuar a fara a biblio da anti-anti-anti-anti-anti-anti- noo hacubbalan list a biblio da biblio da biblio anti-anti-anti-anti-anti-anti-anti-anti-			
		2500 3000 3500 40000	<u>i i i i i i i i</u> 500 1000 1500 2000 254	20 2000 2500 4000
 O Noise Reduction 🔽 Spil	kelike Spatial Homogene	Similar to External Input	1 2 3 5 6	Res

is derived by performing a two-tailed one-sample *t*-test in a channel-wise manner.

noise-related source.

#### Real fNIRS Dataset for Extracting Sources of Interest

We use previously conducted finger-tapping experiments to demonstrate the usages of extracting SOI (Zhao et al., 2020). Nine participants were recruited in the experiment and informed consent was obtained from all of them. The experiment consists of eight blocks of rest and task periods, each last 15 s. In the task period, participants performed a finger-tapping task, in which they alternatively pressed their index and ring finger on a keyboard. Eight sources and seven detectors were automatically arranged according to the transcranial brain atlas (TBA) to cover the pre- and post-central gyrus (see Supplementary Figure 2) (Zhao et al., 2020). LABNIRS (Shimadzu) system was used to record the brain activity of participants during the experiment. The recorded dataset is first preprocessed using DCT-based detrending and HRF-based high-pass filtering with NIRS-SPM. Then the preprocessed data are analyzed using conventional GLM regression and ICA implemented in NIRS-SPM and NIRS-ICA, respectively. The decomposition parameters of ICA are TICA and SOBI with 100 time-delayed correlation matrices. The number of sources is determined by retaining 99% of data variance using PCA (Zhang et al., 2010).

## IMPLEMENTATIONS AND RESULTS

## General Requirements and Main Interfaces

NIRS-ICA is developed using MATLAB 2012a (The MathWorks Inc., Natick, MA, United States), and had been tested in multiple platforms including Windows, MacOS, and Linux. To use NIRS-ICA, users need first install the MATLAB and add the NIRS-ICA folder to the search path, then tap "NIRS\_ICA" in the MATLAB command line to open the main interface.

The data input interface is implemented for inputting the fNIRS dataset and configuring the process of decomposition (Figure 2A). In the Data input and configuration panel, one can specify the input data, hemoglobin type to be processed, and the number of sources to be retained. The input data format supported by NIRS-ICA currently is the format converted using NIRS-KIT, which is a MATLAB toolbox for conducting conventional data analysis for fNIRS data developed by our group (Hou et al., 2021). NIRS-KIT provides a data conversion module for converting data recorded from various fNIRS devices. It also enables input probe montage information with the TopoMaker module, which is useful to visualize the spatial map of separated sources. The generation of probe montage used in the real fNIRS dataset (section "Real fNIRS Dataset for Extracting Sources of



Interest") is demonstrated in **Supplementary Figure 3**. If NIRS-ICA detects there is probe montage information in the input data structure, the spatial maps of the separated sources are displayed according to the probe montage. Otherwise, the spatial maps of the sources will be displayed as a matrix ordered by channel number.

Users can click the ICA settings button to specify the parameters of ICA processing (Figure 2B). NIRS-ICA provides commonly used preprocessing steps, including data crop, detrend, band-passed filtering, and subsampling, which can be selected before ICA. By clicking the aimed ICA algorithm in the Algorithm panel, one can select an ICA algorithm and set its associated parameters (Figure 2C). An example of parameters of preprocessing and the ICA algorithm is shown in Figure 2B. The Source evaluation metrics panel includes metrics for evaluating the separated sources based on the feature of their spatial maps or time courses. As discussed in Section "Source Selection," three metrics for noise reduction and two metrics for selecting neuronal activityrelated sources are implemented to evaluate the separated sources after ICA. Note that multiple evaluation metrics can be simultaneously selected.

After input data and configuration for ICA processing, NIRS-ICA processes the input data using the selected preprocessing methods, decomposes the filtered data using the selected ICA algorithm and calculates the evaluation metrics of the separated sources, and opens the interface of source selection (**Figure 3**). User-friendly interfaces for selecting sources of noise or SOI are provided by NIRS-ICA. Users can switch between the mode of noise reduction and source extraction by clicking the radio button of Feature in the Display control panel (**Figure 3B**). The selection of the different modes of using NIRS-ICA will trigger different interfaces of detailed displaying the sources and results output, which will be demonstrated in later sections.

The time courses of the separated sources are overviewed in a grid manner as shown in Figure 3. Users can choose to view source spatial maps or frequency spectrums of the time course by clicking the pop-up menu of the Features in the Displaying control panel (Figure 4). If evaluation metrics have been selected, separated sources are displayed according to the value of metrics in descending order; otherwise, they are shown by the output order of the selected ICA algorithm. If multiple metrics are chosen, the order is determined by averaging the normalized value of the selected metrics. Users can also choose whether to incorporate a metric by clicking its corresponding checkbox in the Display control panel. For each source, if the value of selected metrics surpasses a certain threshold, the top bar will be highlighted with a different color (instead of gray). Users can mark target sources by clicking the checkbox on the left of the top bar of each source. Sources can be displayed in detail by pressing the small magnifier on the right corner of the top bar. After the selection of sources, users can press the save icon in the menu bar to output the processed ICA results.

# Noise Reduction for the Simulative Dataset

A representative separation result of the simulative dataset is shown in **Figure 3**. It can be seen from the separation result that one source of spike-shaped and four sources of global noise are automatically labeled (top bars are highlighted in red and purple). Spike-shaped motion artifacts can be visually identified from the grid displaying the time courses (**Figure 3C**). From the spatial maps of sources, noises showing a global pattern (the sources with purple top bar) can be easily identified (**Figure 4A**). The time courses of the global noises also have narrow frequency bands as the generated true sources (**Figure 4B**).



In the noise reduction mode, users can press the small magnifier on the top bar of the source to enter the interface of detailed displaying of a source (Figure 5). The detailed information of a separated source, including its time course, the frequency spectrum of the time course, and spatial map, is displayed in the left panel. To facilitate inspecting the relationship between the time course of the current source and the original signal of each fNIRS channel, the fNIRS time course is displayed in the right panel. Users can examine the contribution of the current source to the time course of each fNIRS channel by pressing the channel buttons on the right panel. To enable users to preview the performance of removing the current source, the cleaned fNIRS time course, i.e., reconstructed without the current source using Eq. 9, can be overlaid with the raw fNIRS time course by selecting the checkbox at the bottom of the left panel. The fNIRS time course or baseline time course displayed can also be the time course without the already-selected noise sources by selecting User-defined Baseline NIRS Data at the bottom of the left panel. The periods of motion artifact, i.e., MSD is larger than a threshold, in the time course of the source can be highlighted by pressing the button in the bottom left corner of the source time course. A representative source of motion artifact is shown in Figure 5. Its time course showed a spike-shaped waveform, and the spatial map has high values on the edge of the channel matrix, which are the same as the simulated true source. It can be seen that the spike-shaped waveform is greatly reduced after this source is removed in fNIRS data as shown in the right panel. After the selection of noise sources, users can press the preview button in the main interface to examine the data quality after the noise sources are removed (Figure 3B). Users are able to press the save button on the top menu of the main interface to save the cleaned fNIRS data, which is derived using Eq. 9. The saved fNIRS

data has the same format as the input file, i.e., NIRS-KIT format, which facilitates further processing using other methods such as conventional GLM regression.

The group-level GLM results of both with and without ICA preprocessing are depicted in **Figure 6**. False-positive value can be found in  $\beta$ -map, T-map, and significance map (p < 0.05) due to the noise contamination using only conventional GLM analysis. With ICA preprocessing, the false-positive values are reduced and the resultant spatial maps of GLM are localized to the position of the simulated true sources.

## Extracting Neuronal Activity-Related Sources in the Real fNIRS Dataset

To facilitate extracting neuronal activity-related sources (SOI), NIRS-ICA provides interfaces to input reference time courses or spatial templates. Users can press the corresponding checkboxes in the Source evaluation metrics of Data input and configuration interfaces (Figure 2A), which enable creating reference time courses and spatial templates, respectively (Figure 7). By inputting or loading the onset times and durations of the task stimuli, the reference time course is generated as described in Section "Source Selection" (Figure 7A). The spatial template is created by selecting the channels within ROI in the user interface (Figure 7B). Users can generate multiple reference time courses or spatial templates since several task conditions or ROIs may be of interest in a study. The created reference time courses or spatial templates can also be saved as a MATLAB file and reloaded by pressing the load button on their respective interfaces.

A representative source of the task-evoked hemodynamic response of real fNIRS data is shown in the interface of the



FIGURE 6 | Group-level activation results of the simulated dataset. (A) The results of activation maps using conventional GLM analysis. (B) The results of activation maps using ICA preprocessing and conventional GLM analysis.



detailed displaying of SOI (**Figure 8**). By clicking the buttons on the bottom panel, the reference time course and spatial template are overlaid with the time course and spatial map of the current source, which enables users to examine the correlations between them in detail. It can be seen the time course of the representative source follows the reference time course consistently with a high correlation coefficient (r = 0.81). Its spatial map also depicts high value in channels within ROI (blue circle). After the selection of SOI for an individual participant, the selected source can be saved by pressing the save button on the top menu in the main interface (**Figure 3**).

After the source of interest has been chosen for each participant, group-level SOI can be derived using the method described in Section "Postprocessing." The averaged spatial map of SOI in brain space is visualized using a module inherited from NIRS-KIT (**Figure 9**). Note that users need to additionally input fiducial makers and channel locations recorded on a head model (or representative subject) using a 3D-digitizer, which is



necessary to estimate brain locations measured by the channels (Hou et al., 2021). It can be seen the spatial map derived by ICA (z-map) is similar to the  $\beta$ -map of GLM, and both methods derive peak value at hand knob area, which is consistent with literature studies (**Figures 9A,B**). The block-averaged time course of ICA also depicts the hemodynamic response curve, which indicates that the current SOI is a task-evoked neuronal source (**Figure 9C**).

## Outputs

For both removing noise and extracting SOI, NIRS-ICA saves the parameters of the ICA process, the value of evaluation metrics, as well as the separation results in the output data structure. This output file can be reloaded to NIRS-ICA, which enables users to reproduce the ICA process and correct the source selection. The saved file can also be used to generate a document that includes the options used in the ICA processing. A representative document of the ICA processing used in analyzing the real fNIRS data is shown in **Table 1**. This document can be further reported in an fNIRS study that involves the ICA processing.

## DISCUSSION

In this study, we propose NIRS-ICA, a MATLAB toolbox for the applications of ICA for fNIRS studies. NIRS-ICA incorporates



commonly used ICA algorithms for source separation, userfriendly GUIs, and quantitative evaluation metrics for source selection. It also supports two applications: noise reduction and exploration of neuronal activity-related sources in fNIRS signals. These functionalities are validated based on simulative and real fNIRS datasets. The GUIs and usages are also demonstrated based on the used datasets.

It should be noted that although ICA has been successfully applied in fNIRS studies, the interpretation of ICA results should be made with caution. ICA decomposes fNIRS data relying on the assumption of either temporal or spatial independence between sources of noise and sources of neuronal activity. Therefore, whether ideal decomposition can be achieved depends on how well these assumptions are satisfied in the fNIRS data to be processed. For example, fNIRS studies have shown that there is task-evoked physiological noise, which can temporally correlate with task-evoked hemodynamic response (Kirilina et al., 2012). In that case, TICA may fail to separate noise and neuronal activity-related sources. For SICA, since fNIRS often has a limited number of channels in the spatial dimension, how many channels are enough to obtain a reliable and reproducible decomposition needs to be investigated in future studies. The decomposition can

TABLE 1	Options of ICA	processing	generated by	V NIRS-ICA.

General information	Usage:	Extract sources of interest
	Hb type:	OXY
	Task or RSFC:	Task
	Number of participants:	9
ICA parameters	Number of sources:	PCA, 99%
	Algorithm:	SOBI
	Number of correlation matrices to be diagonalized:	100
Evaluation metrics	Similarity to the reference time course (R)	$0.60 \pm 0.16$
	R for each participant	1(0.60), 2(0.40), 3(0.41), 4(0.66), 5(0.82), 6(0.70), 7(0.74), 8(0.67), 9(0.41)

be also influenced by the preprocessing parameters, the number of sources retained, and the separation algorithm used, which is still an open problem in the fNIRS field. Since the above factors will influence the ICA results, it is important to report these options for the reproducibility of an fNIRS study. With the assistant of NIRS-ICA, researchers can report the parameters for the decomposition, as well as the criterion for source selection, i.e., source evaluation metrics, thus improving the reproducibility of ICA applications for an fNIRS dataset, and further promoting the standardization of using ICA by the fNIRS community.

Several fNIRS toolboxes have already included ICA as a processing step of their analysis pipeline. For example, for data recorded by the NIRx system, one can use NICA and NAVI to denoise the fNIRS data or extracting features from the fNIRS data (Raggam, 2020). FastICA is also implemented in NyPin as a preprocessing step for noise reduction (Strangman et al., 2009). However, since these toolboxes are not dedicated to ICA processing, they have limitations such as availability for only certain types of devices, lack of decomposition algorithms, and methods facilitating source selection. Using the data conversion module provided by NIRS-KTI, NIRS-ICA supports various fNIRS devices such as ETG-4000/7000 (Hitachi Medical Company), LABNIRS (Shimadzu), or NIRx (Hou et al., 2021). It also includes two commonly used ICA algorithms and metrics facilitating source selection, which makes it more flexible and user-friendly than previously proposed toolboxes.

The future updates of NIRS-ICA will be focused on the following aspects: First, in the decomposition step, the current version of NIRS-ICA included two algorithms for source separation since they are commonly used and validated in fNIRS literature. Since the source separation performance may depend on the data quality, other algorithms such as infomax and JADE, which are widely used in other imaging modalities, may still be useful for fNIRS data (Correa et al., 2007). The decomposition performance can also be increased by adding external information, which can be acquired by additional devices. For example, multi-distance channels have been used to discriminate deep and shallow components in the ICA process (Funane et al., 2014). von Lühmann et al. (2019) also combined a

3D accelerometer with ICA to achieve a better noise reduction result. More preprocessing methods, such as excluding bad channels and data segmentation, will also be added. Second, in the source selection step, more evaluation metrics can be developed and included in the source selection step such as mean inter-block cross-correlation (MITC), which is used to select task-related sources (Katura et al., 2008). These metrics may potentially lead to automatic source selection, which will further reduce the subjectiveness in the source selection step. Third, since the current version of NIRS-ICA adopts manual source selection, the sources are manually matched between different subjects in the postprocessing step. Other automatic approaches of matching sources such as clustering and group-ICA will be included in the toolbox (Calhoun et al., 2001a; Spadone et al., 2012). Fourth, since ICA has been utilized in the domain of fNIRS hyper-scanning, it is hopeful to make NIRS-ICA supporting the analysis of multi-brain dataset so that it can be used for fNIRS hyper-scanning studies (Zhao et al., 2017). Finally, NIRS-ICA will soon be integrated into NIRS-KIT<sup>1</sup>, which is a toolbox for analyzing both task-related and resting-state fNIRS data using conventional methods developed by our group (Hou et al., 2021).

## DATA AVAILABILITY STATEMENT

The data analyzed in this study are subject to the following licenses/restrictions: The data that support the findings of this study are available from the corresponding author upon reasonable request. Requests to access these datasets should be directed to C-ZZ, czzhu@bnu.edu.cn.

<sup>1</sup>https://www.nitrc.org/projects/nirskit/

#### REFERENCES

- Adali, T., Anderson, M., and Fu, G.-S. (2014). Diversity in independent component and vector analyses: identifiability, algorithms, and applications in medical imaging. *IEEE Signal Process. Mag.* 31, 18–33. doi: 10.1109/MSP.2014.2300511
- Belouchrani, A., Abed-Meraim, K., Cardoso, J. F., and Moulines, E. (1997). A blind source separation technique using second-order statistics. *IEEE Trans. Signal Process.* 45, 434–444. doi: 10.1371/journal.pone.0024893
- Boas, D. A., Elwell, C. E., Ferrari, M., and Taga, G. (2014). Twenty years of functional near-infrared spectroscopy: introduction for the special issue. *NeuroImage* 85, 1–5. doi: 10.1016/j.neuroimage.2013.11.033
- Brigadoi, S., Ceccherini, L., Cutini, S., Scarpa, F., Scatturin, P., Selb, J., et al. (2014). Motion artifacts in functional near-infrared spectroscopy: a comparison of motion correction techniques applied to real cognitive data. *NeuroImage* 85, 181–191. doi: 10.1016/j.neuroimage.2013.04.082
- Calhoun, V. D., Adali, T., Pearlson, G. D., and Pekar, J. J. (2001a). A method for making group inferences from functional MRI data using independent component analysis. *Hum. Brain Mapp.* 14, 140–151. doi: 10.1002/hbm
- Calhoun, V. D., Adali, T., Pearlson, G. D., and Pekar, J. J. (2001b). Spatial and temporal independent component analysis of functional MRI data containing a pair of task–related waveforms. *Hum. Brain Mapp.* 13, 43–53. doi: 10.1002/ hbm.1024
- Calhoun, V. D., Liu, J., and Adali, T. (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *NeuroImage* 45, S163–S172. doi: 10.1016/j.neuroimage.2008.10.057
- Chaddad, A. (2014). Brain Function Diagnosis Enhanced Using Denoised fNIRS Raw Signals. J. Biomed. Sci. Eng. 7, 218–227. doi: 10.4236/jbise.2014.74025

#### ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University. The patients/participants provided their written informed consent to participate in this study.

#### **AUTHOR CONTRIBUTIONS**

YZ: conceptualization, methodology, data curation, formal analysis, and writing—original draft. P-PS: conceptualization, methodology, software, and resources. F-LT: methodology and resources. XH: software and resources. C-ZZ: conceptualization, methodology, supervision, and writing—review and editing. All authors contributed to the article and approved the submitted version.

#### FUNDING

This work was supported by the National Natural Science Foundation of China (Grant Nos. 61431002, 82071999, 31521063, and 61273287) and the National 973 Program (Grant No. 2014CB846100).

### SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fninf. 2021.683735/full#supplementary-material

- Cordes, D., Haughton, V. M., Arfanakis, K., Carew, J. D., Turski, P. A., Moritz, C. H., et al. (2001). Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. *Am. J. Neuroradiol.* 22, 1326–1333.
- Correa, N., Adali, T., and Calhoun, V. D. (2007). Performance of blind source separation algorithms for fMRI analysis using a group ICA method. *Magn. Reson. Imaging* 25, 684–694. doi: 10.1016/j.mri.2006.10.017
- Cui, X., Baker, J. M., Liu, N., and Reiss, A. L. (2015). Sensitivity of fNIRS measurement to head motion: an applied use of smartphones in the lab. *J. Neurosci. Methods* 245, 37–43. doi: 10.1016/j.jneumeth.2015.02.006
- Ferrari, M., and Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *NeuroImage* 63, 921–935. doi: 10.1016/j.neuroimage.2012. 03.049
- Funane, T., Atsumori, H., Katura, T., Obata, A. N., and Sato, H. (2014). Quantitative evaluation of deep and shallow tissue layers' contribution to fNIRS signal using multi-distance optodes and independent component analysis. *NeuroImage* 85, 150–165. doi: 10.1016/j.neuroimage.2013.02.026
- Hou, X., Zhang, Z., Zhao, C., Duan, L., Gong, Y., Li, Z., et al. (2021). NIRS-KIT: a MATLAB toolbox for both resting-state and task fNIRS data analysis. *Neurophoton* 8:010802. doi: 10.1117/1.NPh.8.1.010802
- Huppert, T. J., Diamond, S. G., Franceschini, M. A., and Boas, D. A. (2009). HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain. *Appl. Opt.* 48, D280–D298.
- Hyvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Netw.* 10, 626–634. doi: 10.1109/72. 761722

- Hyvarinen, A., Karhunen, J., and Oja, E. (2004). *Independent Component Analysis*. New York, NY: John Wiley & Sons, doi: 10.1002/0471221317
- Katura, T., Sato, H., Fuchino, Y., Yoshida, T., Atsumori, H., Kiguchi, M., et al. (2008). Extracting task-related activation components from optical topography measurement using independent components analysis. J. Biomed. Opt. 13:054008. doi: 10.1117/1.2981829
- Kirilina, E., Jelzow, A., Heine, A., Niessing, M., Wabnitz, H., Brühl, R., et al. (2012). The physiological origin of task-evoked systemic artefacts in functional near infrared spectroscopy. *NeuroImage* 61, 70–81. doi: 10.1016/j.neuroimage.2012. 02.074
- Kohno, S., Miyai, I., Seiyama, A., and Oda, I. (2007). Removal of the skin blood flow artifact in functional near-infrared spectroscopic imaging data through independent component analysis. J. Biomed. Opt. 12:062111. doi: 10.1117/1. 2814249
- Liebert, A., Villringer, A., Wabnitz, H., Obrig, H., Rinneberg, H., Steinbrink, J., et al. (2004). Time-resolved multidistance near-infrared spectroscopy of the adult head: intracerebral and extracerebral absorption changes from moments of distribution of times of flight of photons. *Appl. Opt.* 43, 3037–3047. doi: 10.1364/AO.43.003037
- Markham, J., White, B. R., and Zeff, B. W. (2009). Blind identification of evoked human brain activity with independent component analysis of optical data. *Hum. Brain Mapp.* 30, 2382–2392. doi: 10.1002/hbm.20678
- McKeown, M. J., Makeig, S., Brown, G. G., Jung, T. P., Kindermann, S. S., Bell, A. J., et al. (1998). Analysis of fMRI data by blind separation into independent spatial components. *Hum. Brain Mapp.* 6, 160–188.
- Medvedev, A. V., Kainerstorfer, J., Borisov, S. V., Barbour, R. L., and VanMeter, J. (2008). Event-related fast optical signal in a rapid object recognition task: improving detection by the independent component analysis. *Brain Res.* 1236, 145–158. doi: 10.1016/j.brainres.2008.07.122
- Morren, G., Wolf, M., Lemmerling, P., and Wolf, U. (2004). Detection of fast neuronal signals in the motor cortex from functional near infrared spectroscopy measurements using independent component analysis. *Med. Biol. Eng. Comput.* 42, 92–99.
- Raggam, P. (2020). NICA: a novel toolbox for near-infrared spectroscopy calculations and analyses. *Front. Neuroinform.* 14:26. doi: 10.3389/fninf.2020. 00026
- Robertson, F. C., and Douglas, T. S. (2010). Motion artifact removal for functional near infrared spectroscopy: a comparison of methods. *IEEE Trans. Biomed. Eng.* 57, 1377–1387. doi: 10.1109/TBME.2009.2038667
- Scholkmann, F., Holper, L., Wolf, U., and Wolf, M. (2013). A new methodical approach in neuroscience: assessing inter-personal brain coupling using functional near-infrared imaging (fNIRI) hyperscanning. *Front. Hum. Neurosci.* 7:813. doi: 10.3389/fnhum.2013.00813
- Scholkmann, F., Spichtig, S., Muehlemann, T., and Wolf, M. (2010). How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation. *Physiol. Meas.* 31, 649–662. doi: 10.1088/0967-3334/31/5/004
- Scholkmann, F., Wolf, U., Wolf, M., Kleiser, S., Metz, A. J., Jaakko, A., et al. (2014). A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *NeuroImage* 85, 6–27. doi: 10.1016/j.neuroimage.2013.05.004
- Spadone, S., de Pasquale, F., Mantini, D., and Penna Della, S. (2012). A Kmeans multivariate approach for clustering independent components from

magnetoencephalographic data. NeuroImage 62, 1912-1923. doi: 10.1016/j. neuroimage.2012.05.051

- Steinbrink, J., Obrig, H., Villringer, A., Kempf, F., Haux, D., and Boden, S. (2006). Illuminating the BOLD signal: combined fMRI-fNIRS studies. *Magn. Reson. Imaging* 24, 495–505. doi: 10.1016/j.mri.2005. 12.034
- Strangman, G. E., Zhang, Q., and Zeffiro, T. (2009). Near-infrared neuroimaging with NinPy. Front. Neuroinform. 3:12. doi: 10.3389/neuro.11.012.2009
- Vergara, V. M., Mayer, A. R., Damaraju, E., Hutchison, K., and Calhoun, V. D. (2017). The effect of preprocessing pipelines in subject classification and detection of abnormal resting state functional network connectivity using group ICA. *NeuroImage* 145, 365–376. doi: 10.1016/j.neuroimage.2016.03.038
- Virtanen, J., Noponen, T., Kotilahti, K., Virtanen, J., and Ilmoniemi, R. J. (2011). Accelerometer-based method for correcting signal baseline changes caused by motion artifacts in medical near-infrared spectroscopy. J. Biomed. Opt. 16:087005. doi: 10.1117/1.3606576
- von Lühmann, A., Boukouvalas, Z., Müller, K.-R., and Adali, T. (2019). A new blind source separation framework for signal analysis and artifact rejection in functional Near-Infrared Spectroscopy. *NeuroImage* 200, 1–41. doi: 10.1016/j. neuroimage.2019.06.021
- Weihong, D., Yebin, L., Jiani, H., and Jun, G. (2012). The small sample size problem of ICA: a comparative study and analysis. *Pattern Recognit.* 45, 4438–4450. doi: 10.1016/j.patcog.2012.06.010
- White, B. R., Culver, J. P., Snyder, A. Z., Cohen, A. L., Petersen, S. E., Raichle, M. E., et al. (2009). Resting-state functional connectivity in the human brain revealed with diffuse optical tomography. *NeuroImage* 47, 148–156. doi: 10. 1016/j.neuroimage.2009.03.058
- Ye, J. C., Tak, S., Jang, K. E., Jung, J., and Jang, J. (2009). NIRS-SPM: statistical parametric mapping for near-infrared spectroscopy. *NeuroImage* 44, 428–447. doi: 10.1016/j.neuroimage.2008.08.036
- Zhang, H., Zhang, Y. J., Lu, C. M., Ma, S. Y., Zang, Y. F., and Zhu, C. (2010). Functional connectivity as revealed by independent component analysis of resting-state fNIRS measurements. *Neuroimage* 51, 1150–1161. doi: 10.1016/j. neuroimage.2010.02.080
- Zhao, Y., Dai, R.-N., Xiao, X., Zhang, Z., Duan, L., Li, Z., et al. (2017). Independent component analysis-based source-level hyperlink analysis for twoperson neuroscience studies. J. Biomed. Opt. 22, 027004–027004. doi: 10.1117/ 1.JBO.22.2.027004
- Zhao, Y., Xiao, X., Jiang, Y.-H., Sun, P.-P., Zhang, Z., Gong, Y.-L., et al. (2020). Transcranial brain atlas-based optimization for functional near-infrared spectroscopy optode arrangement: theory, algorithm, and application. *Hum. Brain Mapp.* 2:020801. doi: 10.1002/hbm.25318

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Zhao, Sun, Tan, Hou and Zhu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.