BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 60, No. 3, 2012
DOI: 10.2478/v10175-012-0050-5

DATA MINING IN BIOENGINEERING

Analysis of complex-valued functional magnetic resonance imaging data: are we just going through a "phase"?

V.D. CALHOUN[1,2]* and T. ADALI[3]

[1] The Mind Research Network, Albuquerque, New Mexico 87106, USA
[2] Department of ECE, University of New Mexico, Albuquerque, New Mexico 87131, USA
[3] Department of CSEE, University of Maryland Baltimore County, Baltimore, MD 21250, USA

Abstract. Functional magnetic resonance imaging (fMRI) data are acquired as a natively complex data set, however for various reasons the phase data is typically discarded. Over the past few years, interest in incorporating the phase information into the analyses has been growing and new methods for modeling and processing the data have been developed. In this paper, we provide an overview of approaches to understand the complex nature of fMRI data and to work with the utilizing the full information, both the magnitude and the phase. We discuss the challenges inherent in trying to utilize the phase data, and provide a selective review with emphasis on work in our group for developing biophysical models, preprocessing methods, and statistical analysis of the fully-complex data. Of special emphasis are the use of data-driven approaches, which are particularly useful as they enable us to identify interesting patterns in the complex-valued data without making strong assumptions about how these changes evolve over time, something which is challenging for magnitude data and even more so for the complex data. Finally, we provide our view of the current state of the art in this area and make suggestions for what is needed to make efficient use of the fully-complex fMRI data.

Key words: fMRI, independent component analysis, ICA, phase, complex-valued, brain.

1. Introduction

Functional magnetic resonance imaging (fMRI) is a noninvasive, powerful tool that has been utilized in both research and clinical arenas since the early 1990s [1] and has provided valuable insights to the understanding of the human brain function. FMRI has enabled researchers to directly study the temporal and spatial changes in the brain as a function of various stimuli. Because it relies on the detection of small intensity changes over time, fMRI poses significant challenges for data analysis techniques. FMRI data is natively complex, and thus there is both magnitude and phase information available. To date, most fMRI analysis techniques to date have discarded the phase of the fMRI data. However, the phase information may be quite valuable for the analysis of the natively complex fMRI data. A number of direct benefits have been noted in the use of phase fMRI data as (i). In assessing functional connectivity, phase is more informative than magnitude data [2]; (ii) phase enables better detection of artifacts in both the magnitude and phase data which can then be effectively excluded from further analysis [3, 4]; (iii) the sensitivity and specificity of estimations increase as we have shown with a number of examples as part of our NSF-funded work [5–10]. That is our goal in this paper, to provide a review of various approaches for using the phase information along with the magnitude in fMRI data and to demonstrate the advantages. We first present some preliminaries on the fMRI signal and statistical properties. Next we discuss approaches to understand the underlying biophysics of the phase signal as well as approaches to preprocess and de-noise the data. The remainder of this article is devoted to model-based and data-driven analysis approaches to analyze the complex-valued fMRI data. Traditional model-based analysis approaches – such as linear regression – are robust, yet often too rigid to capture the richness of the human brain activation, in addition their limitation is more evident for incorporating phase into the analyses as still little is known about the nature of fMRI phase data. Independent component analysis (ICA), on the other hand, is a data-driven approach that provides a more flexible framework for the analysis of fMRI data. ICA facilitates the analysis of fMRI data in its complex form by eliminating the need to explicitly model the phase behavior. In addition, ICA is able to separate artifacts from signal more readily as well as capture signal from the phase which may be less predictable that that of the magnitude data. There are also a number of issues that require special consideration in the preprocessing and visualization of the complex fMRI data and we address those issues as well before discussing the two main analysis approaches.

1.1. FMRI data acquisition and the complex nature of FMRI data.

Most fMRI studies involve a neurobehavioral paradigm in which a participant is exposed to sensory stimuli and asked to perform a set of mental and/or motor tasks. A given volume is then collected through slices within a given repetition time, which is usually on the order of a few seconds. The acquired data set includes a brain volume movie with a temporal resolution specified by the time of repetition. The MRI signal is acquired as a quadrature signal using two

---

*e-mail: vcalhoun@unm.edu

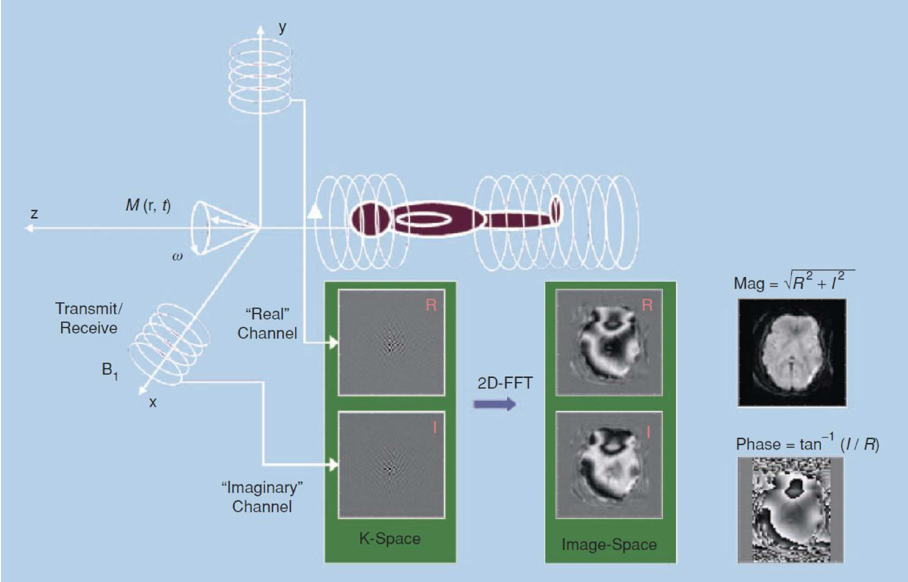

Fig. 1. Diagram of MRI acquisition where M(r, t) represents the magnetization at spatial position r and time t, and ω represents the net phase

or more orthogonal detectors as shown in Fig. 1. The signal that is acquired in the complex frequency space (k-space) is inverse Fourier transformed into the complex image space. This complex-valued fMRI signal change has been shown to contain physiologic in formation [11]. In spite of the presence of useful information in phase, it is usually discarded. Previous studies have reported task-related phase changes [11–14]. Several approaches for modeling the phase have been proposed [15–17]. Processing complex-valued fMRI data using independent component analysis was also proposed in [8]. Previous work has focused on filtering voxels with large phase changes [13, 18–20] based upon models that show that phase changes arise only from large non-randomly-oriented blood vessels. More recent studies from our group and others provide evidence that the randomly oriented microvasculature can also produce a non-zero BOLD-related phase change [20, 21] and we and others have also showed empirical evidence of changes in the phase which correspond to regions expected to be involved in the task [6]. These and much other work providing compelling evidence that the phase information contains useful physiologic information.

2. Preliminaries

Besides medical domain such as magnetic resonance imaging, complex-valued data are an integral part of many science and engineering problems, including those in communications, radar, geophysics, oceanography, electromagnetics, and optics, among others. The complex domain provides both a convenient representation for these signals and a natural way to capture the physical characteristics of these signals. Hence, working completely in the complex domain leads to the most efficient processing of these signals. The complex domain, however, also presents a number of challenges in derivation and analysis of the methods, and as a result, traditionally, the vast majority of algorithms developed for their processing have taken "engineering" shortcuts, thus failing to fully exploit the potential of complex-domain processing. The most common one among those shortcuts has been assuming the circularity of the signal, an assumption that discards the information conveyed by the relationship of real and imaginary parts of the signal, or equivalently by the phase, which is, one of the main reasons one would want to work in the complex domain, *i.e.*, take into account such information in a compact and effective way while using the power of complex calculus.

There have been important advances in this area within the last decade that clearly demonstrate that noncircularity is an intrinsic characteristic of many signals of practical interest, and when taken into account, the methods developed for their processing may provide significant performance gains [22–26]. The two key fundamental advances in this context have first been the development of methods that allow the use of complete statistical information without assuming circularity. What has greatly helped in this development is the development of a complete framework for optimization [22, 27] that is based on Wirtinger calculus [28].

2.1. Statistics. In almost all methods developed prior to 2000 for the complex domain, the circularity assumption has been invoked either explicitly, or implicitly, by simply using only partial statistical information in the development. For example when using second-order statistics, only the correlation information, $E\{\boldsymbol{x}\boldsymbol{x}^H\}$, is used ignoring the pseudo (complementary) correlation given by $E\{\boldsymbol{x}\boldsymbol{x}^T\}$, written for a random vector $\boldsymbol{x}$. A second-order circular (or improper) random vector is one for which $E\{\boldsymbol{x}\boldsymbol{x}^T\} = \mathbf{0}$. When taking full statistical information into account, circularity is defined in terms of the probability density function (pdf) such that a random variable x is circular if x and $xe^{j\theta}$ have the same pdf, *i.e.*, the pdf is rotation invariant [29]. In this case, the phase is noninformative and the pdf is a function of only the magnitude, $p(x) = g(|x|)$ where $g : R \mapsto R$. Data in many applications such as biomedical data analysis, array processing, and communications are however noncircular in nature, see e.g., [22, 26, 30, 31], hence taking their potential noncircularity into account is important for achieving the best performance.

In Fig. 2(a), we show the scatter plot of a motor component estimated using ICA of functional MRI data [3]. The paradigm used in the collection of the data is a simple motor task with a box-car type time-course, *i.e.*, the stimulus has periodic on and off periods. As can be observed in the figure, the distribution of the given fMRI motor component has a highly noncircular distribution. In Fig. 2(b) and (c), we show the spatial map for the same component using a Mahalanobis Z-score threshold, which we define in Section 0. This distribution is typical as most often the signal power is optimized to be mainly in one channel [32, 33] but as we noted, the signal component of the BOLD measurement appears in both the real and imaginary channels resulting in complex-valued fMRI data.

Hence, in the processing and analysis of fMRI data in its native complex form, it is important to account for noncircularity of the distribution. Wirtinger calculus, which we explain next allows derivation of algorithms that can fully take this property into account by making optimization much easier so that the common simplifying assumptions of circularity does not need to be invoked.

2.2. Optimization. The most important step in the derivation of algorithms, one has to compute gradient and Hessians of cost functions, such as a quadratic form or a likelihood function. Since cost functions are real valued, i.e., are scalar quantities in the complex vector space, they are not analytic, and hence not differentiable in a given open set. To overcome this basic limitation, a number of approaches have been traditionally adopted in the signal processing literature the most common of which is the evaluation of separate derivatives with respect to the real and complex parts of a given function.

The framework based on Wirtinger calculus [22, 28] – also called the CR calculus [27] – provides a simple and straightforward approach to performing derivatives in the complex plane, in particular for the important case we mention above, for non-analytic functions. More importantly, it allows one to perform all the derivations and the analyses in the complex domain without having to consider the real and imaginary parts separately. Hence, all computations can be carried out in a manner very similar to the real-valued case, and hence the derivations that use Wirtinger calculus can be directly adapted to the real case.

The main idea behind Wirtinger calculus is based on the definition of a more relaxed condition of differentiability for the complex domain as opposed to the classical definition whose main objective is to make sure that the derivative calculations parallel those in the real domain. Wirtinger calculus [28] relaxes the traditional definition of differentiability and only requires that $f(z)$ be differentiable when expressed as a function $f : R^2 \rightarrow R^2$. Such a function is called *real-differentiable*. Hence, if $u(z_r, z_i)$ and $v(z_r, z_i)$ have continuous partial derivatives with respect z_r to and z_i, f is real-differentiable. For such a function, we can write

$$\frac{\partial f}{\partial z} = \frac{1}{2}\left(\frac{\partial f}{\partial z_r} - j\frac{\partial f}{\partial z_i}\right) \quad \text{and} \quad \frac{\partial f}{\partial z^*} = \frac{1}{2}\left(\frac{\partial f}{\partial z_r} + j\frac{\partial f}{\partial z_i}\right) \tag{1}$$

which can be formally implemented by regarding f as a bivariate function $f(z, z^*)$ and treating z and z^* as *independent variables*. That is, when applying $\partial f/\partial z$, we take the derivative with respect to z, while formally treating z^* as a constant. Similarly, $\partial f/\partial z^*$ yields the derivative with respect to z^*, formally regarding z as a constant. Thus, there is no need to develop new differentiation rules. This was shown in [34] in 1983 without a specific reference to Wirtinger's earlier work [28]. If f is analytic, then the usual complex derivative $\partial f/\partial z$ and in (2) coincide. Hence, Wirtinger calculus contains standard complex calculus as a special case.

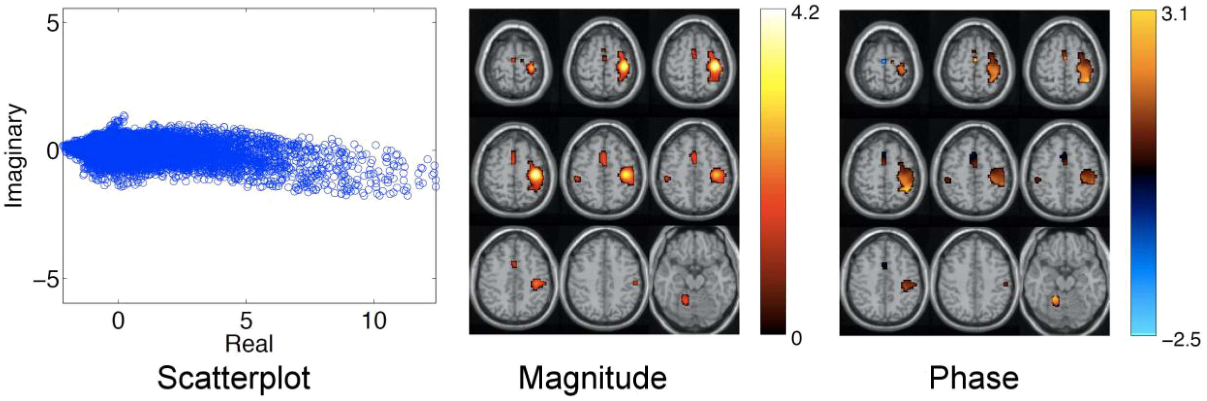

Fig. 2. a) Scatter plot of the average voxel values of the motor component estimated using ICA from 16 subjects. (b) Magnitude and (c) phase spatial maps using Mahalanobis Z-score thresholding; only voxels with are shown

The same approach, i.e., treating the variable and its complex conjugate as independent variables, can be used when taking derivatives of functions of matrix variables as well so that expressions given for real-valued matrix derivatives can be directly used. A good reference for real-valued matrix derivatives is [35] and a number of complex-valued matrix derivatives are discussed in detail in [11]. In [22, 27, 36], the complete framework for vector and matrix optimization using Wirtinger calculus is presented by using the gradient and Hessian relationships given in [37] but by keeping the whole development in the complex plane and in the original problem dimension C^N rather than doubling the dimension as in [37].

Hence, using Wirtinger calculus, all calculations can be carried out in a manner similar to real-valued calculus while keeping all the computations in the complex domain. It is also shown that the simplifying and unrealistic assumption of noncircularity can be avoided in both the algorithm development and in the analyses of the algorithms when one uses Wirtinger calculus, see e.g., [5, 36, 38–44].

3. Biophysical models

One powerful approach for understanding the underlying fMRI signal is to use biophysical modeling. To highlight the scope of what is possible, we present two models for calculating the complex BOLD signal. The first is a microscopic model that calculates the BOLD signal based on the detailed geometry of the micro-vessels, spatial distribution of susceptibility, and diffusion. The second is a macroscopic model that is defined at a resolution of the fMRI experiment and can be used in the inverse problem of combining magnitude/phase BOLD data to improve activation localization and detection.

3.1. Microscopic model. The microscopic model, we present now, follows the description presented in works [45–47]. A two-compartment model, consisting of the extravascular and the intravascular contribution to the BOLD signal is typically considered. Briefly it consists of a) defining a spatial susceptibility distribution $\chi(r) = \chi_m(r)V(r)$, as a product of the assumed network geometry $V(r)$ and the macroscopically varying susceptibility $\chi_m(r)$, b) calculating the magnetic field distribution $B_z(r)$ from $\chi(r)$ using a 3D fast Fourier transform (3DFFT) method, and c) calculating the BOLD signal from $B_z(r)$ by appropriate averaging over the voxel and taking into account diffusion. $B_z(k) = B_0(1/3 - k_z^2/k^2)\chi(k)$ is the 3DFFT of the magnetic field.

Figure 3a shows the geometry which consists of a random distribution of microvessels (infinite cylinders) with a radius = 2.5 um with blood volume fraction = 0.04. Figure 3b is a slice from the corresponding $\Delta B_0(r)$ calculated from the algorithm proposed in [48]) for a Hct = 0.4, Y = 0.5, $\Delta\chi_{do} = 0.27$ ppm $x4\pi$ (units in milliTesla). Assuming a diffusion coefficient $= 1 \times 10^{-9}$ m^2/s we compute the signal attenuation profile shown in Fig. 3c.

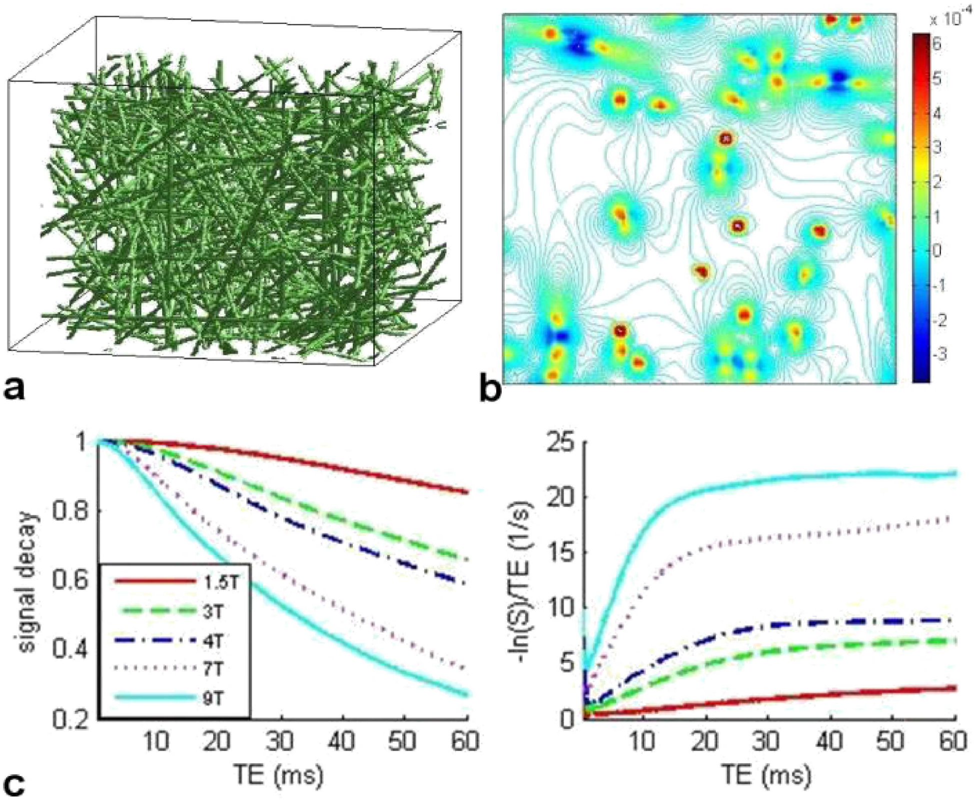

Fig. 3. Example simulation including diffusion: (a) geometry, (b) slice through simulation showing field changes, (c/d) signal decay at different field strengths

3.2. Macroscopic model. We can also use a macroscopic model which summarizes the results of the microscopic model at the level of fMRI voxel resolution. It reflects the experimentally observed patterns of magnitude and phase BOLD signals. We follow the methods developed by Yablonskiy and Haacke [49] and Marques and Bowtell [46], and a model for predicting phase change previously discussed by us [21].

Macroscopic models can be defined for a) random distribution of microvessels (capillaries), b) Oriented distribution of larger vessels (venules), and c) a single large vessel. As discussed before, the extravascular and the intravascular signal can be considered as separate compartments. Although both magnitude and phase effects depend on the underlying vascular geometry and the susceptibility change, they primarily depend on different magnetic field characteristics. To first order, the magnitude attenuation depends on the intra-voxel magnetic field inhomogeneity and the phase depends on the mean magnetic field at the voxel. The magnitude and phase changes have different models as described below.

Extravascular signal. $S_E = S_0 e^{-R_{2,E}^* T_E} e^{-i\phi}$, where $R_{2,E}^* = R_{2,GM} + aV^b \chi_m^c$. V is the voxel volume and $\chi_m(r)$ is the macroscopic susceptibility distribution. The parameters a, b, c are parameters obtained from numerical simulations with the microscopic model for different vessel dimensions, blood volume fraction, and vessel radius. The simulation work of Marques and Bowtell [46] has shown that a realistic model and the infinite cylinder model give similar values for a, b, c. The values for the realistic model were $a = 2.73$, $b = 1.13$, and $c = 1.29$. The phase $\phi = \gamma B_m T_E$ is calculated from macroscopic magnetic field B_m, where B_m is calculated from the macroscopic susceptibility distribution $\chi_m(r)$ by the 3DFFT method described earlier. This method was described earlier in Feng [21] and Fig. 4 shows an example.

Intravascular signal (case 1). Randomly distributed cylinders with blood volume fraction $= f_1$. $S_I = S_1 I(t) e^{-R_{2,I}^* T_E}$, where $R_{2,I}^* = a_1 V^{b_1} {\chi_m}^{c_1} + a_2 V^{b_2} {\chi_m}^{c_2}$, $I(t) = (1/2) \int\limits_0^\pi \sin\theta e^{ik(2\cos^2\theta - 1)t} d\theta$, and $k = (2/3)\pi\gamma B_)\chi$. Marques and Bowtell [46] estimate $a_1 = 3.5$, $b_1 = 1.0$, $c_1 = 1.2$, $a_2 = 40.0$, $b_2 = -0.4$, and $c_2 = 1.2$.

Intravascular signal (case 2). A single cylinder with blood volume fraction $= f_2$. The field inside the cylinder is orientation dependent but is spatially constant. $S_L = S_0 e^{-R_{2,B} T_E} e^{-i\phi_L}$, where $\phi_L = \gamma B_0 (\chi/6)(3\cos^2\theta - 1) T_E$. The total signal is given by $S_T = (1 - f_1 - f_2) S_E + f_1 S_I + f_2 S_L$. The magnitude signal in each compartment has the form $S = S_0 e^{-R_2^* T_E}$, and for small changes in R_2^*, $\Delta S/S = -T_E \Delta R_2^*$. ΔR_2^* can be related to changes in χ_m by models presented earlier. The magnitude of the blood/tissue susceptibility difference is modeled by $\chi = Hct(1-Y)\chi_{dHb}$, where $Hct = 0.4$, $\chi_{dHb} = 2.2$ ppm (MKS units). This gives $\chi = 0.36$ ppm in the resting state with an oxygenation fraction $Y = 0.6$ and $\chi = 0.18$ ppm in the active state with $Y = 0.8$. These numbers imply that the change in χ between the resting state and the activated state is estimated to be $\Delta\chi = 0.18$ ppm.

3.3. Intravascular effect. Here we estimate the maximum phase change expected in a single large vessel to the phase change calculated in a capillary bed under some reasonable parameter assumptions. If the cylinder is parallel to the main field $\phi_L = \gamma B_0 \chi T_E / 3 = 15.4\pi T_E$, for $B_0 = 3T$ and $\chi = 0.18$ ppm. The phase in the total signal will depend of the blood volume fraction f, but as can be seen it can become as large as π. Menon et al. have suggested that large phase changes can be used to detect and exclude large vessel artifacts [13]. When an analysis is done in combination with a physical model we can use the phase information to either enhance an activated region or suppress a false activation.

3.4. Phase changes also occur in parenchymal regions containing only small vessels. We assume that the macroscopic susceptibility change is 3D Gaussian.

$$\Delta\chi_m(r) = C_k \exp\left(-\frac{1}{2}\left[\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} + \frac{z^2}{\sigma_z^2}\right]\right), \tag{2}$$

where C_k is a scaling constant. We choose the value of C_k based on parameter values from the literature. We define $\Delta\chi'$ as the susceptibility difference between completely deoxygenated and completely oxygenated red blood cells ($0.264 \times 4\pi$ ppm in MKS units [50] with a hematocrit level of 0.4 [51], and oxygenation level Y is the fractional oxygenation in the red cells with $\Delta Y_{cap} = 0.08$ [11, 12]. Then for a blood volume fraction f of 0.05 and ignoring the cerebral blood volume change, $C_k = -f \cdot \Delta Y_{cap} \cdot 4\pi \cdot \Delta\chi' \cdot Hct$, C_k is approximately $-5.3e^{-9}$.

Figure 4 shows the simulation results of magnetic field/phase change corresponding to 3D Gaussian volume-averaged susceptibility/magnetization change for the cases of $\sigma_x : \sigma_y : \sigma_z = 1 : 1 : 1$, 1:1:2, 2:2:1 and 2:2:1, Eq. (1), rotated counter-clockwise around the x-axis by $\pi/3$, respectively. For the value of C_k selected above, the resulting maximum simulated phase change for all of these configurations is in the order of 1°. Depending on the spatial distribution of the susceptibility changes and the angle of the cut plane of the magnetic field change, the resulting phase shows patterns of dominantly positive, dominantly negative, or combinations of positive and negative phase changes due to the volume-averaged magnetization and demagnetization effects.

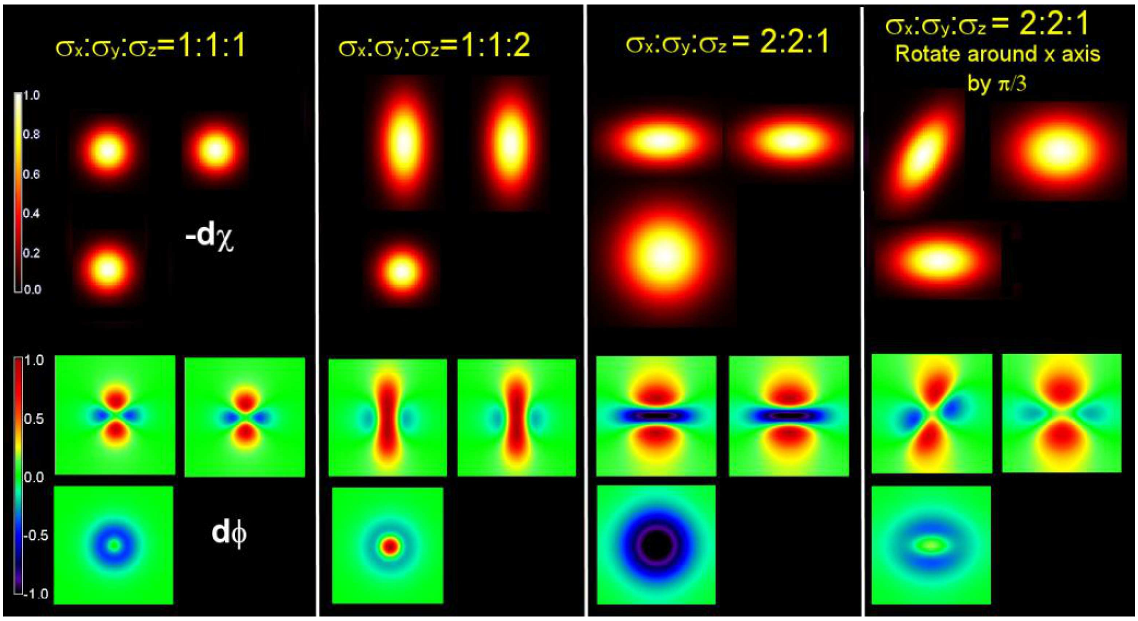

Fig. 4. Simulation of phase change corresponding to 3D Gaussian volume-averaged susceptibility change for the cases of $\sigma x : \sigma y : \sigma z = 1 : 1 : 1$, 1:1:2, 2:2:1 and 2:2:1 rotated counter-clockwise around the x-axis by $\pi/3$

3.5. Robustness to noise by using phase. A second simulated example is shown on how to fit the magnitude signal change or determine how much to smooth. The problem can be approached by finding an optimal smoothing of the magnitude response constrained by phase. The optimization cannot be performed without the phase information. Given a magnitude change we can calculate a signal proportional to phase change. Let $f_m(r)$ and $f_\phi(r)$ be the magnitude and phase change images. The error functions for magnitude and phase are written as $E1 = ||\chi(r) - f_m(r)||^2$, and $E2 = S^2||\alpha F(\chi(r) - f_\phi(r)||^2$. We now seek a function $\chi(r)$ and α, such that $E1$is approximately equal to $E2$ while $E1 + E2$ is a minimum. The additional phase information enables us to find a smooth solution for $\chi(r)$. The results of the proposed fitting are shown in Fig. 5 where we transform a noisy image magnitude/phase image pair (middle) to a pair which has error relative to the ground truth (left) reduced by a factor of 20 (right). The joint constraint allows us to determine the optimal smoothing and fit the magnitude data without ground truth knowledge.

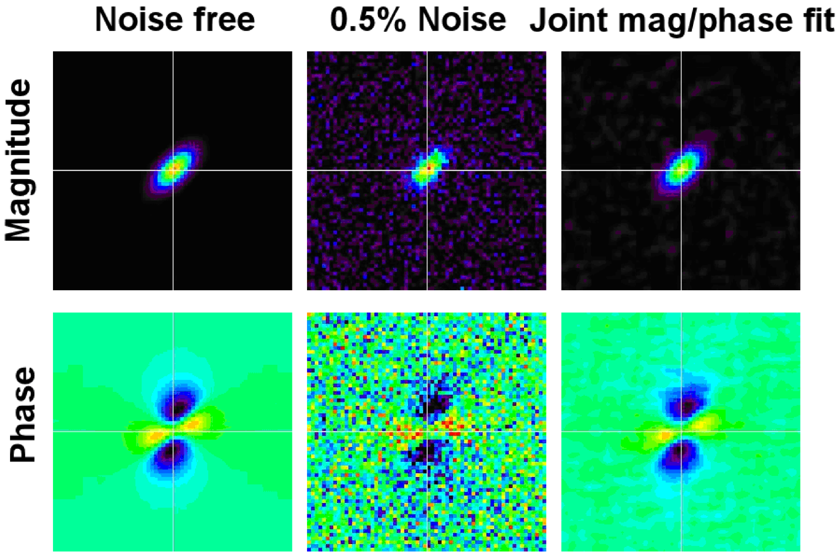

Fig. 5. Joint magnitude/phase fit to reduce noise

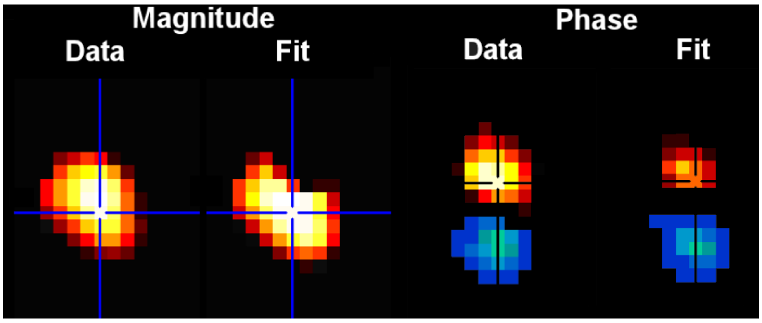

Fig. 6. Simultaneous fit for one subject

3.6. Fitting the model to real fMRI data. The macroscopic model suggests that both magnitude and phase of the BOLD signal depend on the macroscopic susceptibility distribution $\chi_m(r)$. The magnitude change depends $\chi_m(r)$ on through models for R_2^* and the phase through the 3DFFT model or the separate case of large vessels. We propose to estimate $\chi_m(r)$ by minimizing the goodness-of-fit function χ^2 used for fitting the magnitude and phase change models as defined by:

$$\chi^2 = \int\limits_r \left(\left(\frac{\Delta S(r)}{S(r)} - \frac{\Delta \widehat{S}(r)}{\widehat{S}(r)} \right) / \sigma_{ds}(r) \right)^2 dr + \int\limits_r \left(\frac{\Delta\phi(r) - \Delta\widehat{\phi}(r)}{\sigma_{d\phi}(r)} \right)^2 dr,$$

where $\Delta\widehat{S}(r)/\widehat{S}(r)$, $\Delta\widehat{\phi}(r)$ are theoretical scaled magnitude change and phase change, respectively, and $\Delta S(r)/S(r)$, $\Delta\phi(r)$ are the observed magnitude change data and phase change data; $\sigma_{ds}(r)$ and $\sigma_{d\phi}(r)$ are the standard deviation of magnitude and phase change at each voxel. The integration is over the volume of activation (VOA). Figure 6 shows the simultaneous fitting results for one subject showing data as well as the fits for magnitude and phase.

Experiments were performed on a 3T Siemens TRIO TIM system using a standard Siemens gradient-echo EPI sequence. We used a Field-of-View (FOV) = 240 mm, Slice thickness = 3.5 mm, Slice Gap = 1 mm, 32 slices, Matrix size = 64×64, TE = 29 ms, and TR = 2 s. The fMRI experiment used a block design with alternating 30s finger tapping. The total experiment time was 5.5 minutes.

Data were preprocessed using the SPM software (http://www.fil.ion.ucl.ac.uk/spm/software/spm/5/). Complex images were corrected by dividing each time point by the first time point, and then recalculating the phase images. Further phase unwrapping was not required. Data were motion corrected [51], spatially smoothed with a 10 mm^3 full width at half-maximum Gaussian kernel, and spatially normalized into the Montreal Neurological Institute space. Activation maps were computed using the multiple regression framework within SPM5 in which regressors are created from the stimulus onset times and convolved with a standard hemodynamic response function. A contrast was created for each individual subject for finger tapping versus rest. A group analysis was performed using the activation maps from each individual subjects and entering them into voxelwise one-sample t-tests.

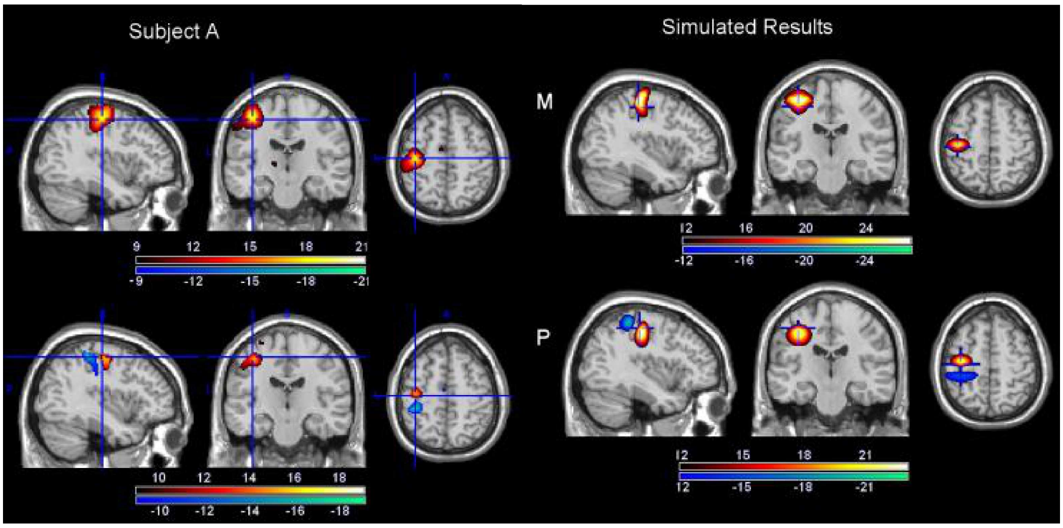

Fig. 7. Magnitude (M) and phase (P) changes (t-values) for representative subject (left) and simulated results (right). The color bars for the subjects show the t-value ranges. The colorbar for the simulated results indicate the relative strength of susceptibility and phase change

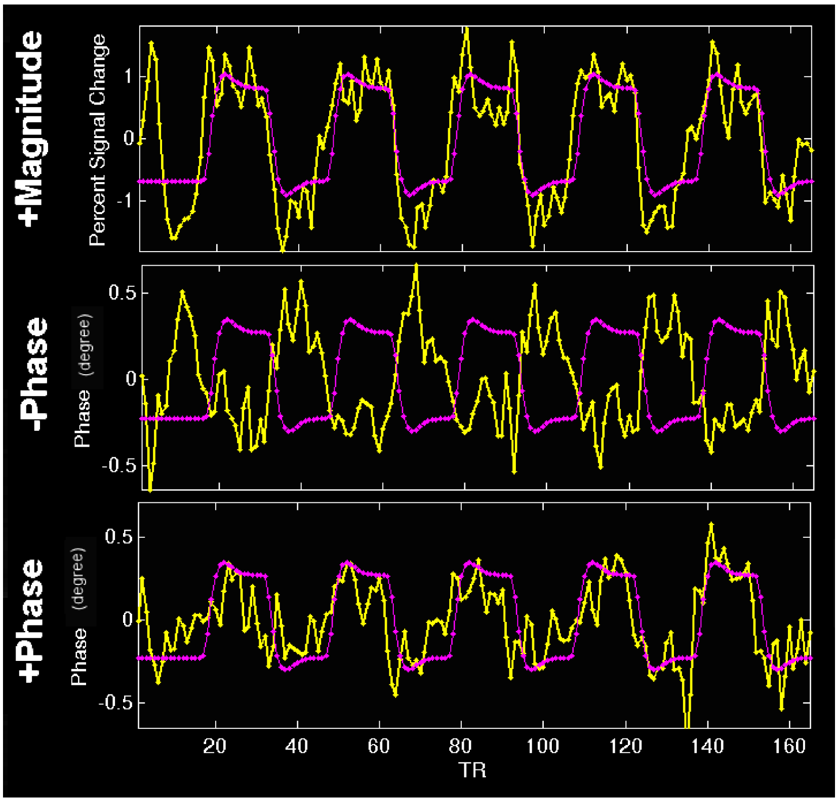

Fig. 8. Phase and magnitude change time courses

In Fig. 7, the left panel show the magnitude and phase change of the results (thresholded at $t = 8$ ($p < 1 \times 10^{-13}$)) for subject A (for a complete reporting of results see [21]. The panel on the right shows the susceptibility change and phase change of simulated results. Here, we assume the magnitude change is approximately linear to the volume-averaged susceptibility change. The highest magnitude change was observed in the motor cortex. Further observation of Fig. 7 indicates the peak of the magnitude change is not located where the phase change peaks; instead, it is closer to the sign change of the phase change (the minimum absolute phase value). The simulations to match the observed fMRI phase/magnitude changes began by approximating the observed phase change pattern to that obtained by Gaussian distributions, and calculating the susceptibility distribution by an inverse calculation [53]. Then a matched phase distribution was calculated by a forward model calculation. The results show that our model can closely match patterns observed experimentally.

Figure 8 shows the phase and magnitude change time courses from a single voxel (the one showing maximal phase change) for a representative subject. The time evolutions of the phase and magnitude change are similar to each other, suggesting that both changes originate from the same source, the deoxyhemoglobin-induced susceptibility change. The measured voxel phase change (unsmoothed) is around 1°, on the same order of the simulation results as shown in Fig. 4. They are both on the same order of a measured voxel phase change (no large vessel present) 0.028 radians or 1.6° in [13].

4. Preprocessing and visualization

4.1. Motion correction. Existing methods for motion correction and spatial normalization can be modified to work for an analysis including the phase. A straightforward approach is to simply use the magnitude image to compute the parameters for both motion correction and for spatial normalization and then to apply the computed transformation to the phase images. This appears to work quite well although a potential down side is it does not consider the possibility that the phase data may contain useful information about movement or structure in general. As a simple extension one can try to incorporate a cost function which uses both the magnitude and the phase. A typical least-squares cost function for motion correction is given below, CF $= \sum_{i=2:N}[M_i - M_1]^2$ where M_i is the i-th time point image. It is possible to implement a complex motion correction approach by realigning the images using the information present in the complex image (obtained by combining the magnitude and phase data). A proposed cost function for the complex algorithm is given below,

$$\text{Cost} = \sum_{(} i = 2 : N) \left[\frac{(R_i - R_1)^2}{1/(\text{Norm}(R_1))} + \frac{(I_i - I_1)^2}{1/(\text{Norm}(I_1))} \right],$$

where is the real part of the time point and is the imaginary part of the time point. We applied this approach to fMRI data from several subjects performing a motor tapping task. The results using the magnitude only cost function and the complex-valued cost function are similar but not identical. The resulting T-maps for both magnitude and phase show less activity at the edges of the brain (40% smaller T-values on average) when performing motion correction using the complex-valued data suggesting that the complex motion correction algorithm does a better job in handling the motion. Overall the patterns observed pre and post motion correction are highly similar with pre/post motion corrected maps spatially correlated at 0.97 for phase and 0.98 for magnitude. The significance is slightly higher for the unified approach compared to an approach in which we used only the magnitude data and applied the parameter estimates to the phase data. The phase for the unified approach shows increases from T = 18.8 to T = 18.9 and magnitude changed from T = 21.3 to T = 21.4. The T-value improvements are encouraging, but relatively modest.

4.2. Spatial normalization. Just as in motion correction we can use two approaches, one using the magnitude information and pulling the phase information along, and the other using both magnitude and phase to computed the normalization parameters. For the latter the images can be spatially normalized to the Montreal Neurological Institute (MNI) template and we use the cost function in [2] to derive the nonlinear parameter estimates. A study specific template (including both real and imaginary images) was created using an initial registration based upon the magnitude data and then applying the parameters to the real and imaginary data and averaging across subjects to create a template that has real and imaginary images. The data can then be renormalized to this group template using the cost function in [2] to provide a more accurate estimate. We tested this approach on a group of 17 subjects. We converted the data to magnitude and phase and a statistical analysis was performed on the magnitude and phase data independently. Results were overall quite similar to the magnitude-only approach, but again the T-maps were slightly higher for the unified approach. Results are thus encouraging, but more work is needed.

4.3. Preprocessing and visualization.

Spatial smoothing. Data are typically spatially smoothed with e.g., a Gaussian kernel to improve the contrast-to-noise ratio [6, 8]. Smoothing is useful for group data as it both reduces the amount of high-frequency spatial noise as well as desensitizes the images to variability of functional activation and anatomy among subjects. We analyze both unsmoothed and smoothed data with different kernels to evaluate the impact of smoothing upon the analysis. There is likely room for improvement in smoothing schemes to move beyond Gaussian, e.g. wavelet smoothing [54]. In the complex-domain there is also the possibility of performing natively complex smoothing, for example one could use a modified complex anisotropic diffusion filtering [55]. However, this is an application area which has not yet been fully evaluated.

Phase denoising. A physiologically motivated denoising method is given in [4] and uses the phase to identify noisy voxels and eliminates them or introduces a weighting scheme depending on their noise level. The quality map phase denoising (QMPD) uses gradient information to determine the noisy voxels and eliminates them from further analysis [4]. It is shown that the voxels identified as noisy are in areas that are known to suffer from susceptibility artifacts, such as the area from the orbitofrontal cortex due to air in the sinuses. The final component of this method includes an important smoothing step that if done before eliminating noisy voxels can spread their detrimental effects to their surroundings. It is also noted that the ICA results obtained with the QMPD method provides higher Z-scores (11.87 versus 10.83) and a larger number of active voxels (1589 versus 1238) as compared to those obtained with the MTEE method [4].

Phase correction for ICA group studies. We note the importance of phase correction for the analysis results for groups of subjects as a simple rotation of the estimated distribution of the activation map can have serious detrimental effects in the group results when we compute averages. There have been two approaches proposed to address this problem. If information on the distribution of the original fMRI data is available, then this prior information can be used for selecting the appropriate nonlinearity in the ICA algorithm as shown in Fig. 9 (right) [38, 40, 43]. As shown in [38], a number of trigonometric functions and their hyperbolic counterparts can be effectively used for achieving ICA, and for the fMRI data we use where the signal is maximized in the real part, the function shown in Fig. 9 (right) provides a perfect match and eliminates the phase rotation in the estimated components as we show in [56].

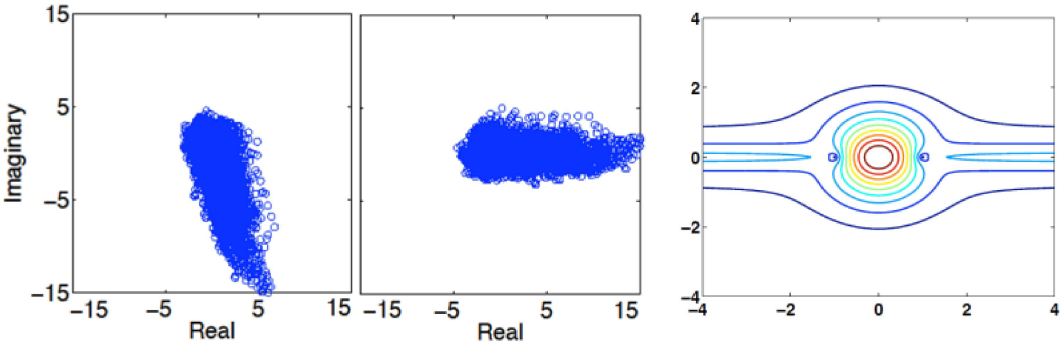

Fig. 9. Complex scatter plot of an estimated motor task related source before (left) and after (middle) applying the PCA-based phase correction scheme. (right) form of pdf implied by the score function

A second approach uses direct manipulation of the data to correct for phase rotation in the estimated fMRI components: 1) Find rotation angle θ that maximizes: argmaz$_\theta E[(\mathbf{M}\widehat{\mathbf{s}})^2]$, 2) Resolve 180° phase ambiguity: argmax$_\theta E[(\mathbf{M}\widehat{\mathbf{s}})^3]$ where and $\mathbf{M} = [\cos\theta \ \ \sin\theta; \ - \in \theta \ \ \cos\theta]$ and $\widehat{\mathbf{s}}_k = [\widehat{s}_{k.re}, \widehat{s}_{k,im}]^T$. Figure 9 shows the scatter plot of the real and imaginary data of an estimated source before and after the PCA based phase correction. In Fig. 13, we show group ICA results for 20 subjects performing a motor tapping task.

5. Model-based analysis of complex fMRI data

5.1. fMRI data analysis. We now discuss statistical analysis using a model-based approach. Typically, the acquired fMRI data are first preprocessed, for example by the 1) correction of slices for the slight time shift within each volume, 2) registration to correct for subject motion during the scan, and 3) spatial normalization to enable comparisons among subjects and neuro-anatomical labeling, and 4) smoothing. Following the preprocessing step, the data are analyzed to determine the voxels with significant temporal signal change, which are then super-thresholded and overlaid on an anatomical image. The volume data is then organized into a matrix X such that each row is formed by concatenation of the slices at a given time instant resulting in the $T \times V$ matrix shown in Fig. 10.

5.2. General linear model. The most widely used method for the analysis of fMRI data is linear regression using the general linear model (GLM) [83]. As shown in Fig. 10 along with an example time course (regressor) for a simple on-off paradigm, X is the matrix of input data, R is a design matrix, and M is the matrix of activation maps. The time course is correlated with the fMRI data to determine the voxels that show activity related to the chosen time course. The goal is find the matrix M, i.e., compute the regression coefficients (entries of matrix M) that are deemed to be active. An important limitation of this method is that the regressors rk that form the matrix R need to be specified a priori.

Several methods for approaching the analysis of the phase information within the GLM framework were proposed by Rowe et al. [14–17, 57]. The general complex fMRI model proposed by Rowe is given in [15]:

$$\begin{gathered} y_t = (\rho_t \cos\theta_t + \eta_{Rt}) + i(\rho_t \sin\theta_t + \eta_{It}), \\ \rho_t = x_t'\beta = \beta_0 + \beta_1 x_{1t} + \cdots + \beta_{q1} x_{q1t}, \\ \theta_t = u'\gamma = \gamma_0 + \gamma_1 u_{1t} + \cdots + \gamma_{q2t} u_{q2t}, \\ t = 1, \ldots, n \end{gathered}$$

where $(\eta_{Rt}, \eta_{It})' \sim N(0, \sigma^2 I)$, x_t' is the t-th row of an $n \times (q_1 + 1)$ design matrix X for the magnitude, u_t' is the t-th row of an $n \times (q_1 + 1)$ design matrix U for the phase, while β and γ are the regression coefficients for the magnitude and phase.

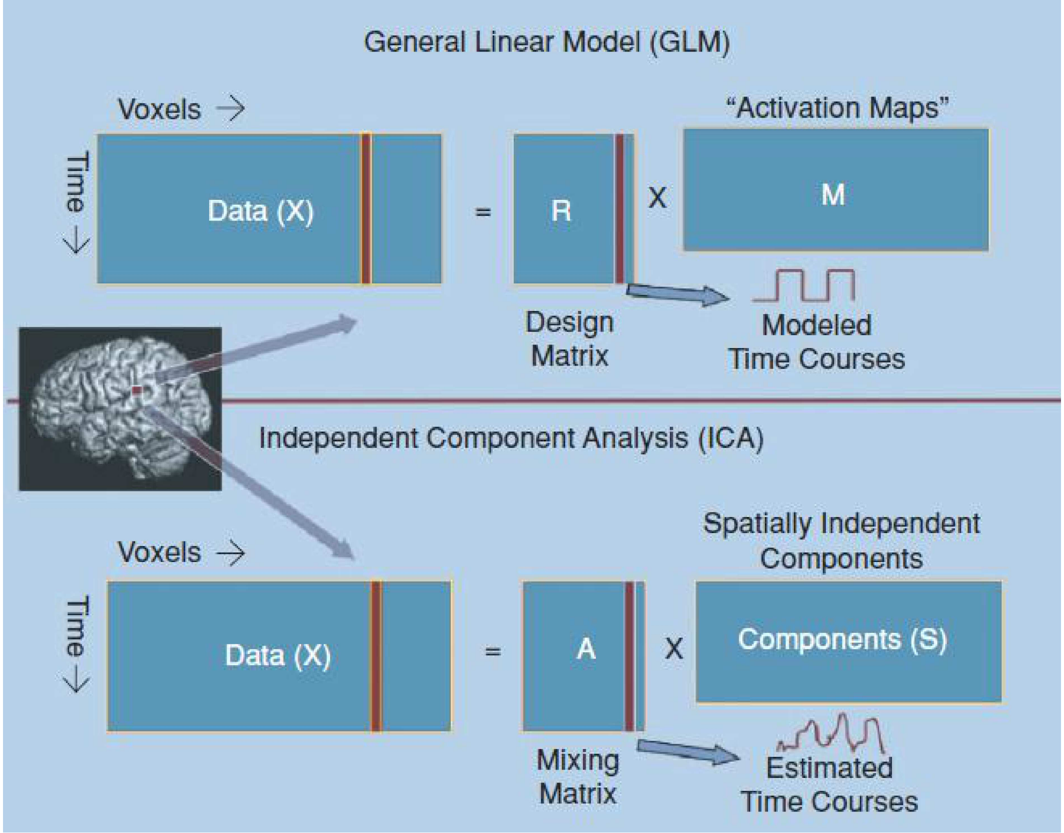

Fig. 10. Application of GLM and ICA to fMRI data

Rowe considers a number of cases for modeling the phase, in [15] and [14], a task-related phase is considered, in [17] the phase is modeled as an arbitrary value (and under this model, i.e., uninformative phase, the model is shown to be equivalent to magnitude only case), and finally in [16], a constant value (identical at all time-points) has been used for modeling the phase. With the given linear model, the hypotheses regarding task related magnitude and phase changes can be evaluated on an individual complex-valued voxel-wise basis, using maximum likelihood estimators.

5.3. GLM group analysis. Few studies have examined phase data in a larger group of subjects for multiple types of fMRI tasks, nor have studies examined phase changes due to event-related stimuli. In recent work from our group, we evaluate the correspondence between the magnitude and phase changes at a group level in a block-design motor tapping task and in an event-related auditory oddball task [6]. The results for both block-design and event-related tasks indicate the presence of task related information in the phase data with phase-only and magnitude-only approaches showing signal changes in the expected brain regions. Although there is more overall activity detected with magnitude data, the phase-only analysis also reveals activity in regions expected to be involved in the task, some of which were not significantly activated in the magnitude-only analysis, suggesting that the phase might provide some unique information. In addition, the phase can potentially increase sensitivity within regions also showing magnitude changes. The identification of regions which 1) show signal changes for magnitude data only, 2) show signal changes for phase data only, or 3) show signal changes for both magnitude and phase data were of particular interest.

Figure 11 shows the magnitude change and phase change of the results for motor tapping and auditory oddball. As expected the highest magnitude change for motor tapping was observed in the left motor cortex and for the auditory oddball highest change was in bilateral temporal lobe. Similarly, maximal phase changes were also observed in motor cortex for MT and in temporal lobe for AOD. The images in the top-right and bottom-right panels of Fig. 1 are the RGB (R-red, G-greed, B-blue) color maps for MT and AOD similar to the display provided in [15]. The areas in red are where only significant magnitude signal changes were observed, the ones in green are for significant phase-only signal changes and the areas in blue are where both significant magnitude and significant phase signal changes were observed. The resulting signal change changes for phase and magnitude data for both motor tapping and AOD were cluster thresholded to correct for multiple comparisons at family wise error (FWE).

The color activation maps in Fig. 11 show the correspondence between the magnitude and phase responses. The regions of interest in each of these maps are labeled such that red shows magnitude only areas, green shows phase only, and the areas for magnitude and phase are shown by blue. Tapping movement mainly activates regions in the motor cortex, hence for the motor tapping paradigm it is expected to see peaks in precentral gyrus. The results are encouraging and corroborate with patterns observed in the ROI analysis. The presence of these areas in phase only activation maps (without any magnitude signal change) suggests that using the phase date in fMRI may provide useful information beyond the magnitude data.

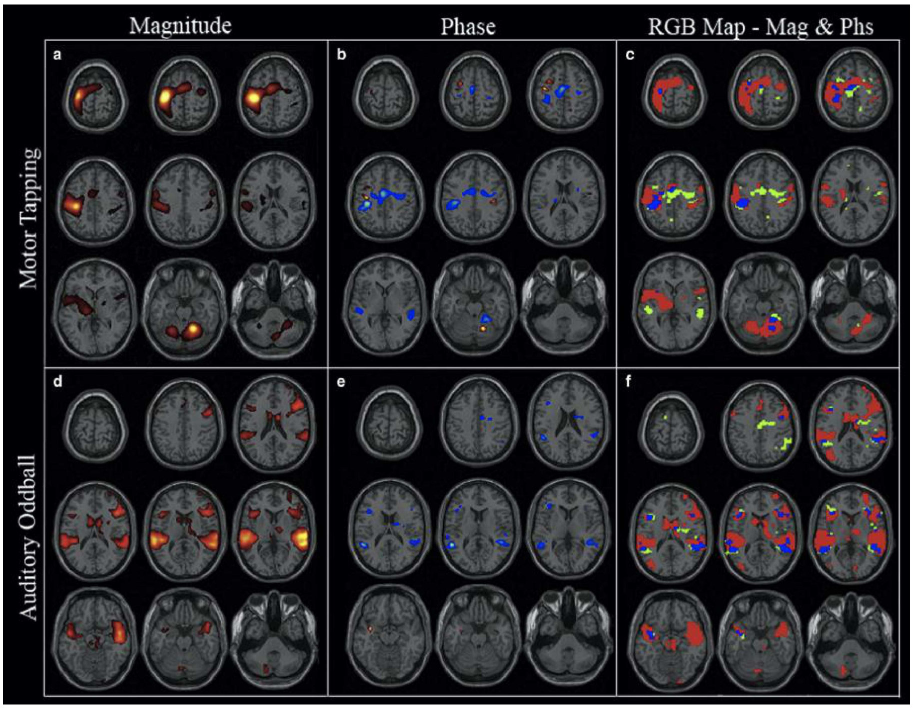

Fig. 11. Whole brain analysis results

Separate analyses of phase and magnitude fMRI data at a group level for two different paradigms were analyzed. The group statistical results show significant phase changes in both block design and event related design. The presence of phase activation in the regions expected to be activated by the task, suggests that the information in the phase might help increase the ability to isolate the task-related functional changes.

6. Data-driven analysis of complex fMRi data

6.1. Blind Source Separation (BSS) and Independent Component Analysis (ICA). Data-driven methods are based on a simple generative model and hence can minimize the assumptions on the nature of the data. They have emerged as promising alternatives to the traditional model-based approaches in many applications where the underlying dynamics are hard to characterize. Blind source separation, in particular, has been a popular data-driven approach and an active area of research. Most BSS formulations start with the linear mixing model $\boldsymbol{x} = \boldsymbol{As}$ with the possibility of an additive noise term), where $\boldsymbol{x}$ is the mixture that is factorized into latent variables through two matrices (a mixing matrix $\boldsymbol{A}$ and a component (source) vector $\boldsymbol{s}$ where each entry correspond to a source in random variable notation) and replaced by the observation matrix $\boldsymbol{X}$ and component matrix $\boldsymbol{S}$in an implementation. For uniqueness of the decomposition (subject to certain ambiguities), constraints are applied to the two matrices such as sparsity, uncorrelatedness, or independence of the components. ICA is a popular blind source separation technique that imposes the constraint of statistical independence on the components, i.e., source distributions and hence can recover the original sources by estimating a demixing matrix $\boldsymbol{W}$ such that $\boldsymbol{u} = \boldsymbol{Wx}$ subject to only a scaling and permutation ambiguity. It has been successfully applied to numerous data analysis problems in areas as diverse as biomedicine, communications, finance, geophysics, and remote sensing [22, 61].

To solve the source separation problem, different properties of source signals have been exploited including non-Gaussianity, non-stationarity, and sample correlation – see e.g. [58–63]. The most commonly used property among those has been non-Gaussianity. The natural cost in this context that leads to ICA is the mutual information among separated components, which can be shown to be equivalent to maximum likelihood estimation, and to negentropy maximization [40, 59, 61, 64] when we constrain the demixing matrix to be orthogonal. In these approaches, one either estimates a parametric density model [61, 63, 65, 85] along with the demixing matrix, or maximizes the information transferred in a network of non-linear units [58, 67], or estimates the entropy using a parametric or nonparametric approach [58, 63, 68, 69]. A recent semi-parametric approach uses the maximum entropy bound to estimate the entropy given the observations, and uses a numerical procedure thus resulting in accurate estimates for the entropy [42]. We have showed that the method can successfully approximate a wide class of source distributions by selecting few measuring functions, and when incorporated into ICA, the flexible density matching in this approach, ICA by entropy bound minimization (ICA-EBM) provides a very attractive trade-off between performance and computational cost [42, 70]. As presented in [71, 72], one way to incorporate prior information to an ICA algorithm is by working in a constrained optimization framework and directly adding the constraints through Lagrange multipliers. ICA-EBM, on the other hand, besides the use of such a direct constrained approach allows easy incorporation of prior information in a number of ways, in particular by selecting the best nonlinearities to model the underlying source densities.

We have made considerable progress in the development of data-driven algorithms for processing complex-valued fMRI data. Many of the approaches discussed in this paper are available in two Matlab software tools, the Group ICA of fMRI Toolbox (GIFT; http://mialab.mrn.org/software) and the LibrarY of Complex Independent component analysis Algorithms (LYCIA; http://mlsp.umbc.edu/lycis/lycia.html).

6.2. Complex ICA. When performing ICA in the complex domain, all quantities are assumed to be complex and an important result in the complex case is that one can make use of *noncircularity* to achieve source separation. Specifically, when all the sources in the mixture are improper with distinct circularity coefficients, we can achieve ICA through joint diagonalization of covariance and complementary covariance matrices to achieve source separation as in the strongly uncorrelating transform (SUT) [73, 74]. For the real-valued case, separation using second-order statistics can be achieved only when the sources have sample-to-sample correlations.

Algorithms such as joint approximate diagonalization of eigenmatrices (JADE) [75] explicitly calculate the higher-order statistics, the cumulants in the case of JADE, and can be directly used for ICA of complex-valued data. A recent extension for these algorithms [76] enables joint diagonalization of matrices that can be Hermitian and/or complex symmetric and hence can be used for more efficient ICA solutions using both the commonly used statistics and the complementary statistics that have been traditionally neglected. The algorithms that rely on joint diagonalization of cumulant matrices are robust. However, their performance suffers as the number of sources increases, and the cost of computing and diagonalizing cumulant matrices might become prohibitive for separating a large number of sources. On the other hand, ICA techniques that exploit non-Gaussianity are the more attractive solutions for the complex case as well. As in other areas for complex-valued processing, circularity assumption was a common one for the extension of popular ICA algorithms to the complex case as in complex Infomax [77, 78] and complex FastICA [78]. As expected, in the presence of noncircular sources, the performance of those algorithms suffer. There are now a number of powerful solutions available for complex ICA for the general case where sources can be either circular or noncircular, e.g., [36, 38, 41, 43, 44, 80] as well as those that adapt to different source distributions using general models such as complex generalized Gaussian distributions [43, 81], or more flexible models through efficient entropy estimation techniques as in ICA by entropy bound minimization (ICA-EBM) [70].

6.3. An application example: importance of accounting for noncircularity. Two most commonly used ICA techniques are based on maximum likelihood (ML) or maximization of negentropy (MN), and the two are equivalent when the demixing matrix is constrained to be unitary [40]. Infomax [58] has been the most widely used algorithm for analysis of fMRI data following its first application to the problem [82] and can be shown to be equivalent to maximum likelihood when the score function is matched to the source density, and in the case of orginal infomax [58] the super-Gaussian source that corresponds to the sigmoid nonlinearity (score function). Its first extension to the complex domain, circular infomax [77], assumes again a fixed nonlinearity, again of sigmoid form that matches most fMRI sources well as they tend to be mostly super-Gaussian. However, it also assumes a circular distribution, which might be quite limiting as discussed in Subsec. 2.2. Circular infomax uses the nonlinear score function that is a good match to typical super-Gaussian pdf given as

$$\varphi(u) = -\log\frac{\partial p(u)}{\partial u} = \mathrm{sign}(u)\frac{1-\exp(-|u|)}{1+\exp(-|u|)},$$

where only the magnitude of the data has been considered and all the phase information has been discarded. In this example, we compare its performance with that of the complex version of ICA-EBM [69] as discussed in detail in [10], which is a powerful ICA approach using an adaptive density model.

Complex ICA-EBM algorithm. The complex ICA-EBM algorithm uses the minimization of the mutual information principle, which is equivalent to ML, to perform source separation. The cost function can be written as [42] $I(y) = \sum_{k=1}^{n} H(u_k) - 2\log|\det(W)| - H(x)$ here $H(u_k)$ is the entropy of the k-th spatial map. Instead of estimating $H(u_k)$ directly, complex ICA-EBM estimates the tightest bound of the entropy by assuming that the density of the sources is either weighted linear combinations or elliptical distribution. Complex ICA-EBM can obtain a reliable estimate of the bound of entropy by solving for the maximum entropy distribution that maximizes the entropy under certain constraints. The associated maximum entropy distribution includes many bivariate distributions, such as Gaussian, uniform, (double) exponential, Student t, and GGD. Let us define two density forms as $p(u) = A\exp[-au^2 - bu - cG_m(u)]$, $q(|u|) = A\exp[-a|u|^2 - cG(|u|)]$ where G_m is one of a set of pre-determined measure functions and the parameters k, a, b, and c are solved by using a normalization constraint, the associated maximum entropy distribution of s could be $kp(s_r)p(s_i)$ or $kq(|u|)$ for two different entropy bounds. Four function forms of G, including the unbounded fourth order symmetric and bounded second-order asymmetric, are considered for the first entropy bound where the density is weighted linear combinations. Two function forms of G, including fourth order symmetric and first order asymmetric, are considered for the second entropy bound where the sources are elliptical distributions. Among all the entropy estimates, only the minimum one is used as the final estimate of the entropy.

Dataset. The dataset used in the experiment is from 16 subjects performing a finger-tapping motor task while receiving auditory instructions [10]. The data are first preprocessed using the quality map phase de-noising (QPMD) [4] discussed in Subsec. 4.3. The fMRI data were multiplied by the mask generated using QMPD and smoothed for the real and imaginary parts separately. Then components are estimated using both the complex versions of Infomax and ICA-EBM.

6.4. Results of bivariate t-maps and difference t-maps. To study the statistics across subjects, we can define a bivariate t-map using the Hotelling $\boldsymbol{T}^2$-test defined as [81]

$$T^2 = N < \mathbf{s} >^T \mathbf{C}_{\overline{S}}^{-1} < \mathbf{s} >, \tag{3}$$

where N is the number of realizations of random vector s. For Hotelling T^2-test, $< s >$ in (3) represents the sample mean vector of a set of realizations from a multivariate Gaussian distributed s. For a group of subjects, we can calculate the mean image $< \mathbf{s} >$ of 16 subjects, where $< \mathbf{s} >$ represents the sample mean vector of 16 2×1 vectors for each voxel. The covariance matrix $C_{\overline{S}}$ is also defined with respect to the 16 2×1 vectors, where each 2×1 vector is treated as a realization of a 2×1 Gaussian random vector. Therefore we can construct an image of size v using (3) and the value of each pixel represents a Hotelling T^2 value, where we have assumed that the 16 2×1 vectors for each voxel are the realizations of a multivariate Gaussian distribution. Such an image might be called a group bivariate t-map. It should be noted that 16 may be too small as a sample size and the multivariate Gaussian distribution across different subjects is also assumed for simplicity.

Using the Hotelling T^2 statistic defined in (3), we construct a bivariate t-map for the two algorithms to check for voxels that are significant active across the 16 subjects. However, the fMRI images estimated by the ICA algorithms have phase rotation ambiguity and we cannot simply apply the T^2-test. The phase ambiguity, is due to the ICA model $\mathbf{x} = \mathbf{As}$. As we can observe, $\mathbf{x} = \mathbf{As} = \mathbf{A}'\mathbf{s}'$ where $\mathbf{A}' = \mathbf{AD}^{-1}$, $\mathbf{s}' = \mathbf{Ds}$ and $\mathbf{D}$ represents a diagonal matrix whose entries are complex, and hence include a magnitude and phase part. Therefore there are infinitely many solutions of A and s since D is an arbitrary diagonal phase rotation matrix. That is, for the general ICA problem, it is impossible to recover the original scale of the sources, which in the complex case includes a magnitude and a phase term.

Hence, we perform the phase rotation on each estimate to ensure that the largest magnitude of the estimate is on the real axis since that is how the data we are using is acquired. The value for each voxel in the bivariate t-map tells us how the voxel values are distributed across different subjects. High T^2-values in such figures might be regarded as an index indicating that the voxel values are probably high for all the subjects at that specific pixel and with probably lower variations across different subjects. We observed that the bivariate t-maps looked visually similar for all the four algorithms. However, in terms of hypothesis testing, the results were not the same. For hypothesis testing, we are testing $H_0 : \mu_s = 0$

versus $H_1 : \mu_s \neq 0$ where μ_s is the mean spatial map of 16 subjects. We accept H_0 if $T^2 < T^2_{\alpha,d,N-1}$ where α is the probability of Type I error (accept H_1 when H_0 is in fact true), $d-2$ is the dimension of random vector, and $N-1 = 15$ is the degrees of freedom. For instance, $T^2_{0.05,2,15} = 8.01$. We then evaluate the number of voxels for the right motor component when thresholded $T^2 > T^2_{0.05,2,15}$ at as 2761 for Infomax and as 2951 for ICA-EBM, hence resulting in significantly greater number of voxels for the flexible ICA-EBM that also accounts for noncircularity of the sources.

We can also compute difference t-maps such that given any two sets of estimated spatial maps with voxels X_{ijk} and Y_{ijk}, where X_{ijk} denotes the k-th voxel in the j-th component of the i-th subject, each voxel of 16 difference images are calculated as $D_{ijk} = X_{ijk} - Y_{ijk}$. The difference images of 16 subjects were calculated first, then a T^2-test was performed. Results of difference bivariate t-maps show that the adaptive ICA algorithms have significantly higher activation within the motor area for each class as shown in Fig. 12.

Thresholding of complex analysis results. Estimated sources for complex-valued data require a method which takes into account both phase and magnitude. The thresholding method introduced in [4] takes into account the phase by using a Mahalanobis distance metric in the real and imaginary data of the estimated sources given by $d_{k,i} = \sqrt{[\widehat{\mathbf{s}}_{k,i} - \boldsymbol{\mu}_k]^T \mathbf{C}_k^{-1} [\widehat{\mathbf{s}}_{k,i} - \boldsymbol{\mu}_k]}$ where $\widehat{\mathbf{s}}_{k,i} = [\widehat{s}_{k,i,r,e}, \widehat{s}_{k,i,im}]^T$; and μ_k and CZ_k are the corresponding mean and covariance of the estimated sources. Figure 13 shows results with this new Z-score that takes the complex nature of data fully into account. An additional challenge for fMRI group studies using complex-valued ICA is the well-known inherent scaling ambiguity of ICA algorithms, which in the complex case includes a phase term. Hence, for complex-valued ICA, the phase term in the scaling ambiguity presents an additional problem, since the estimated distribution of matching components across subjects then will have different unknown rotations (in the complex plane), without phase ambiguity correction, they can add destructively, hence creating group average component images with lower magnitude and noisy phase images. In [4], two effective approaches are introduced to correct for the phase ambiguity such that successful group results can be obtained and presented using complex fMRI data. In Fig. 12, we show the estimation results for a motor component where a total of 30 components were estimated using circular Infomax [76] using data from 16 subjects performing a finger-tapping task, and using the Mahalanobis z-score.

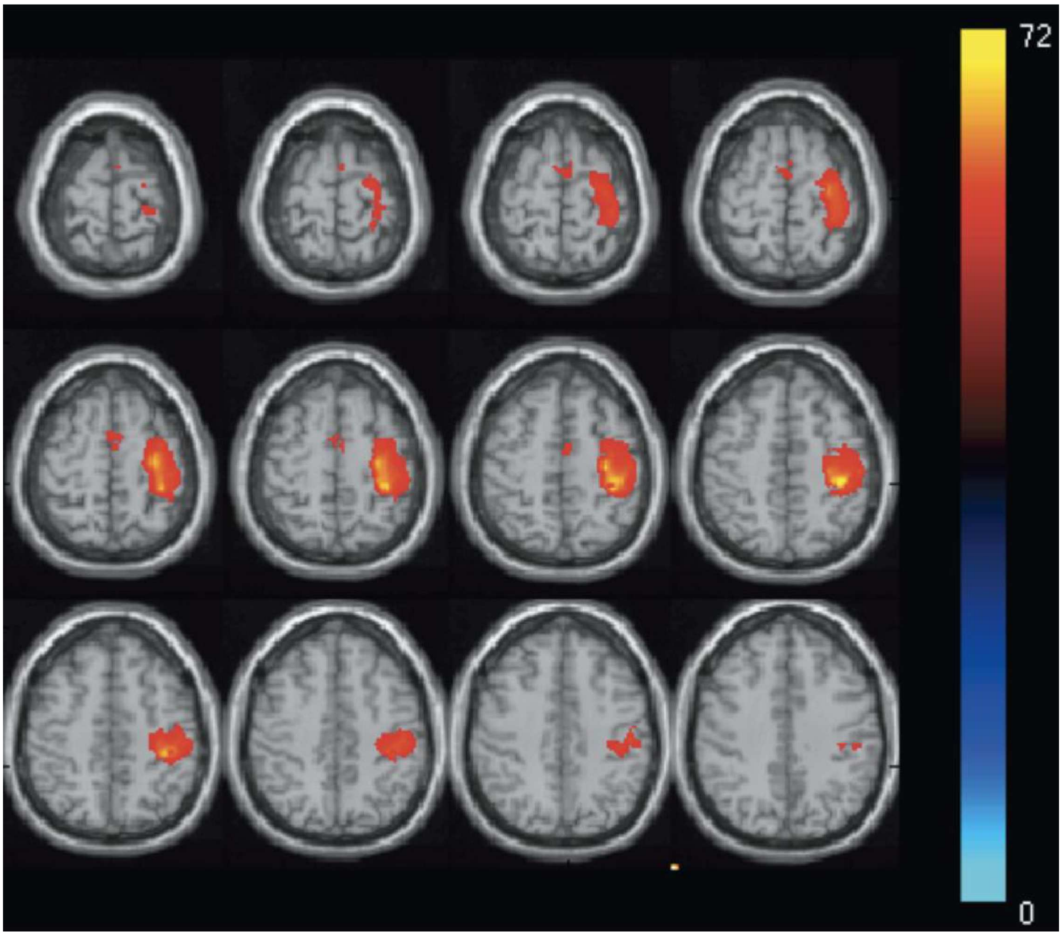

Fig. 12. Difference bivariate t-map for the motor component with ICA-EBM and circular Infomax

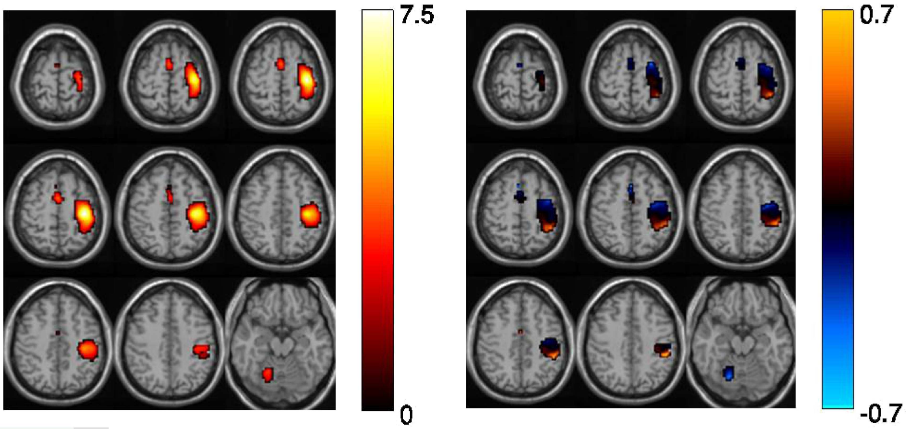

Fig. 13. Mean magnitude (left) and phase (right) maps for the motor component (Mahalanobis z-score of 4)

7. Summary

In summary, we believe there is great potential in using the phase information in an fMRI analysis. The convergence of biophysical models and simulation approaches, high-field and high-resolution data acquisition, preprocessing and denoising approaches, and statistical modeling approaches which utilize the fully complex data are pushing the field forward. However significant challenges still remain and need to be addressed before complex-valued fMRI data will become a mainstream approach.

Acknowledgements. This work was supported by NSF grants 0715022, 0840895, 0635129 and 0612076.

REFERENCES

[1] S. Ogawa, D.W. Tank, R. Menon, J.M. Ellermann, S.G. Kim, H. Merkle , and K. Ugurbil, "Intrinsic signal changes accompanying sensory stimulation, functional brain mapping with magnetic resonance imaging", *Proc. Natl. Acad. Sci.* 89 (13), 5951–5955 (1992).

[2] N. Petridou, A. Schafer. P. Gowland, and R. Bowtell, "Phase vs. magnitude information in functional magnetic resonance imaging time series, toward understanding the noise", *Magn. Reson. Imaging* 27 (8), 1046–57 (2009).

[3] P. Rodriguez, V.D. Calhoun, and T. Adali, "Phase ambiguity correction and visualization techniques for complex-valued ICA of group fMRI data, *Pattern Recognition* 45 (6), 2050–2063 (2012).

[4] P. Rodriguez, N. Correa, T. Adali, T. Eichele, and V.D. Calhoun, "Quality map thresholding for de-noising of complex-valued fmri data and its application to ICA of fMRI", *J. Signal Processing Systems* 1, 1–16 (2009).

[5] T. Adali and V.D. Calhoun, "Complex ICA of brain imaging data", *IEEE Signal Proc. Magazine* 24 (5), 136–139 (2007).

[6] S. Arja, Z. Feng, Z. Chen, A. Caprihan, K.A. Kiehl, T. Adali, and V.D. Calhoun, "Changes in fMRI magnitude data and phase data observed in block-design and event-related tasks", *NeuroImage* 49 (4), 3149–3160 (2010).

[7] V.D. Calhoun and T. Adali, "Complex ICA for fMRI analysis, performance of several approaches", *Proc. ICASSP* 1, CD-ROM (2003).

[8] V.D. Calhoun, T. Adali, G.D. Pearlson, P.C. van Zijl and J.J. Pekar, "Independent component analysis of fMRI data in the complex domain", *Magn. Reson. Med.* 48 (1), 180–192 (2002).

[9] H. Li, T. Adali, N. Correa, P. Rodriguez, and V.D. Calhoun, "Flexible complex ICA of fMRI data", *Proc. ICASSP* 1, CD-ROM (2010).

[10] H. Li, N. Correa, V.D. Calhoun, and T. Adali, "Application of independent component analysis with adaptive density model to complex-valued fMRI data", *IEEE Trans. Biomed. Eng.* I 58 (10), 2794–2803 (2011).

[11] F. G. Hoogenraad, P.J. Pouwels, M.B. Hofman, J.R. Reichenbach, M. Sprenger, and E.M. Haacke, "Quantitative differentiation between BOLD models in fMRI", *Magn. Reson. Med.* 45 (2), 233–246 (2001).

[12] F.G Hoogenraad, J.R. Reichenbach, E.M. Haacke, S. Lai, K. Kuppusamy, and M. Sprenger, "In vivo measurement of changes in venous blood-oxygenation with high resolution functional MRI at 0.95 tesla by measuring changes in susceptibility and velocity", *Magn. Res. Med.* 39 (1), 97–107 (1998).

[13] R. Menon, "Postacquisition suppression of large-vessel BOLD signals in high-resolution fMRI", *Magn. Res. Med.* 47 (1), 1–9 (2002).

[14] D.B. Rowe, "Parameter estimation in the magnitude-only and complex-valued fMRI data models", *Neuroimage* 25 (4), 1124–32 (2005b).

[15] D.B. Rowe, "Modeling both the magnitude and phase of complex-valued fMRI data", *Neuroimage* 25(4), 1310–24 (2005a).

[16] D.B. Rowe and B.R. Logan, "A complex way to compute fMRI activation", *Neuroimage* 23 (3), 1078–92 (2004).

[17] D.B. Rowe and B.R. Logan, "Complex fMRI analysis with unrestricted phase is equivalent to a magnitude-only model", *Neuroimage* 24 (2), 603–6 (2005).

[18] A.S. Nencka and D.B. Rowe, "Reducing the unwanted draining vein BOLD contribution in fMRI with statistical post-processing methods", *Neuroimage* 37 (1), 177–88 (2007).

[19] D.G. Tomasi and E.C. Caparelli, "Macrovascular contribution in activation patterns of working memory", *J. Cereb. Blood Flow Metab.* 27 (1), 33–42 (2007).

[20] F. Zhao, T. Jin, P. Wang, X. Hu, and S.G. Kim, "Sources of phase changes in BOLD and CBV-weighted fMRI", *Magn. Reson. Med.* 57 (3), 520–7 (2007).

[21] Z. Feng, A. Caprihan, K. Blagoev and V.D. Calhoun, "Biophysical modeling of phase changes in BOLD fMRI", *NeuroImage* 47, 540–548 (2009).

[22] T. Adali and S. Haykin, *Adaptive Signal Processing Next Generation Solutions*, Wiley-IEEE Press, New York, 2010.

[23] T. Adali, P.J. Schreier, and L.L. Scharf, "Complex-valued signal processing, The proper way to deal with impropriety", *IEEE Trans. Signal Processing* 59 (11), 5101–5123 (2011).

[24] B. Picinbono and P. Chevalier, "Widely linear-estimation with complex data", *IEEE Trans. on Signal Processing* 43 (8), 2030–2033 (1995).

[25] P. Schreier and L.L. Scharf, "Second-order analysis of improper complex random vectors and processes", *IEEE Trans. on Signal Processing* 51 (3), 714–725 (2003).

[26] P. Schreier and L.L. Scharf, "Statistical signal processing of complex-valued data", *The Theory of Improper and Noncircular Signals* 1, CD-ROM (2010).

[27] K. Kreutz-Delgago, *The Complex Gradient Operator and the CR-Calculus*, University of California, San Diego, 2007.

[28] W. Wirtinger, Zur formalen theorie der funktionen von mehr complexen veräänderlichen, *Math. Ann.* 97, 357–375 (1927).

[29] B. Picinbono, "On circularity", *IEEE Trans. Signal Processing* 42, 3473–3482 (1994).

[30] P. Chevalier and F. Pipon, "New insights into optimal widely linear array receivers for the demodulation of BPSK, MSK, and GMSK signals corrupted by noncircular interferences – application to SAIC", *IEEE Trans. Signal Processing* 54 (3), 870–883 (2006).

[31] F. Roemer and M. Haardt, "Efficient 1-D and 2-D DOA estimation for non-circular sources with hexagonal shaped espar arrays", *Proc. IEEE Int. Conf. Acoust. Speech, Signal Processing (ICASSP)* 1, 881–884 (2007).

[32] E. Hardy, D. Hoferer, D. Mertens, and G. Kasper, "Automated phase correction via maximization of the real signal", *Mag. Res. Imag.* 27, 393–400 (2009).

[33] A. Macovski, "Noise in MRI", *Magn. Res. Med.* 36 (3), 494–497 (1996).

[34] D.H. Brandwood, "A complex gradient operator and its application in adaptive array theory", *Proc. Inst. Elect. Eng.* 1, 11–16 (1983).

[35] K.B Petersen and M.S. Pedersen, *The Matrix Cookbook*, Technical University of Denmark, Copenhagen, 2008.

[36] H. Li, T. Adali, "Complex-valued adaptive signal processing using nonlinear functions", *J. Advances in Signal Processing* B, 1–9 (2008).

[37] A. Van den Bos, "Complex gradient and Hessian", *IEE Proc. Vision, Image, and Signal Processing* 1, 380–382 (1994).

[38] T. Adali, T. Kim, and V.D. Calhoun, "Independent Component analysis by complex nonlinearities", *Proc. ICASSP* 1, 525–528 (2004).

[39] T. Adali and H. Li, *Complex-Valued Adaptive Signal Processing, Adaptive Signal Processing, Next Generation Solutions*, Wiley, New York, 2009.

[40] T. Adali, M. Novey, and J.F. Cardoso, "Complex ICA using nonlinear functions", *IEEE Trans. Signal Processing* 59 (9), 4356–4544 (2008).

[41] H. Li and T. Adali, "A class of complex ICA algorithms based on the kurtosis cost function", *IEEE Trans. Neural Netw.* 19 (3), 408–420 (2008).

[42] X. Li and T. Adali, "Independent component analysis by entropy bound minimization", *IEEE Trans. on Signal Processing* 58 (10), 5151–5164 (2010).

[43] M. Novey and T. Adali, "Complex ICA by negentropy maximization", *IEEE Trans. Neural Networks* 19 (4), 596–609 (2008).

[44] M. Novey and T. Adali, "On extending the complex FastICA algorithm to noncircular sources", *IEEE Trans. Signal Processing* 56 (5), 2148–2154 (2008).

[45] J. Martindale, A.J. Kennerley, D. Johnston, Y. Zheng, and J.E. Mayhew, "Theory and generalization of Monte Carlo models of the BOLD signal source", *Magnetic Resonance in Medicine* 59 (3), 607–618 (2008).

[46] J.P. Marques and R.W. Bowtell, "Using forward calculations of the magnetic field perturbation due to a realistic vascular model to explore the BOLD effect", *NMR Biomed.* 21 (6), 553–65 (2008).

[47] Z. Chen, A. Caprihan, and V.D. Calhoun, "Effect of surrounding vasculature on intravoxel BOLD signal", *Med.Phys.* 37 (4), 1778–1787 (2010).

[48] K.M. Koch, X. Papademetris, D.L. Rothman, and R.A. de Graaf, "Rapid calculations of susceptibility-induced magnetostatic field perturbations for in vivo magnetic resonance", *Physics in Medicine and Biology* 51 (24), 6381–6402 (2006).

[49] D.A. Yablonskiy and E.M. Haacke, "Theory of NMR signal behavior in magnetically inhomogeneous tissues, the static dephasing regime", *Magn. Reson. Med.* 32 (6), 749–63 (1994).

[50] W.M. Spees, D.A. Yablonskiy, M.C. Oswood, and J.J. Ackerman, "Water proton MR properties of human blood at 1.5 Tesla, magnetic susceptibility, T(1), T(2), T*(2), and non-Lorentzian signal behavior", *Magn. Reson. Med.* 45 (4), 533–42 (2001).

[51] A.C. Guyton and J.E. Hall, *Textbook of Medical Physiology,* W.B. Saunders Company, Philadelphia, 1996.

[52] L. Freire, A. Roche, and J.F. Mangin, "What is the best similarity measure for motion correction in fMRI time series?", *IEEE Trans. Med. Imaging* 21 (5), 470–484 (2002).

[53] E.M. Haacke, N.Y. Cheng, M.J. House, Q. Liu, J. Neelavalli, R.J. Ogg, A. Khan, M. Ayaz, W. Kirsch, and A. Obenaus, "Imaging iron stores in the brain using magnetic resonance imaging", *Magn. Reson. Imaging* 23 (1), 1–25 (2005).

[54] S. Khullar, A. Michael, N. Correa, T. Adali, S. Baum, and V.D. Calhoun, "Wavelet-based fMRI analysis, 3-D denoising, signal separate, and validation metrics", *NeuroImage* 54 (4), 2867–2884 (2011).

[55] G. Gilboa, N. Sochen, and Y.Y. Zeevi, "Image enhancement and denoising by complex diffusion processes", *IEEE Trans. Pattern Anal. Mach. Intell* 26 (8), 1020–36 (2004).

[56] T. Adali and H. Li, "A practical formulation for computation of complex gradients and its application to maximum likelihood", *Proc. ICASSP* 1, CD-ROM (2007).

[57] D.B. Rowe, C.P. Meller, and R.G. Hoffmann, "Characterizing phase-only fMRI data with an angular regression model", *J. Neurosci. Methods* 161 (2), 331–41 (2007).

[58] A.J. Bell and T.J. Sejnowski, "An information maximisation approach to blind separation and blind deconvolution", *Neural Computing* 7 (6), 1129–1159 (1995).

[59] P. Comon, "Independent component analysis – a new concept?", *Signal Proc.* 36, 287–314 (1994).

[60] A. Hyvarinen, "One-unit contrast functions for independent component analysis. A statistical analysis", *Proc. NNSP* 1, 388–397 (1997).

[61] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis,* Johns Wiley & Sons, New York, 2001.

[62] Z. Koldovsky, P. Tichavski, and E. Oja, “Efficient variant of algorithm FastICA for independent component analysis attaining the Cramer-Rao lower bound”, *IEEE Trans. Neural Netw*. 17 (5), 1265–77 (2006).

[63] E.G. Learned-Miller, J.W. Fisher III, and T.W. Lee, “ICA using spacings estimates of entropy”, *J. Machine Learning Research* 4, 1271–1295, (2003).

[64] J.F. Cardoso, “Blind signal separation, statistical principles”, *Proc. IEEE* 9 (10), 2009–2025 (1998).

[65] J.A. Palmer, S. Makeig, K. Kreutz-Delgago, and B.D. Rao, “Newton method for the ICA mixture model”, *Proc. IEEE Int. Conf. Acoust. Speech, Signal Processing* 1, CD-ROM (2008).

[66] D.T. Pham and P. Garat, Blind separation of mixture of independent sources through a quasi-maximum likelihood approach, *IEEE Trans. Signal Proc*. 45 (7), 1712–1725 (1997).

[67] T.W. Lee, M. Girolami, and T.J. Sejnowski, “Independent Component Analysis Using an Extended Infomax Algorithm for Mixed Subgaussian and Supergaussian Sources, *Neural Comput.* 11, 417–441 (1999).

[68] D. Erdogmus, K.E. Hild, Y. Rao, and J.C. Principe, “Minimax mutual information approach for independent component analysis”, *Neural Comput*. 16 (1235), 1252 (2004).

[69] A. Hyvarinen, “New approximations of differential entropy for independent component analysis and projection pursuit”, *Advances in Neural Inf. Proc. Sys*. 10, 273–279 (1998).

[70] X. Li and T. Adali, “Complex independent component analysis by entropy bound minimization”, *IEEE Trans. Circuits and Systems* I 57 (7), 1417–1430 (2010).

[71] W. Lu W and J.C. Rajapakse, “Constrained independent component analysis”, in *Adv. Neural Inf. Proc. Sys.* pp. 570–576, MIT Press, Cambridge, 2000.

[72] W. Lu and J.C. Rajapakse, “ICA with reference”, *Proc. Int. Conf. on ICA and BSS* 1, 120–125 (2001).

[73] L. De Lathauwer and B. De Moor, “On the blind separation of non-circular sources”, *Proc. Eur. Signal Process. Conf. (EUSIPCO)* 1, CD-ROM (2002).

[74] J. Eriksson and V. Koivunen, “Complex random vectors and ICA models, Identifiability, uniqueness and separability”, *IEEE Trans. Info. Theory* 52 (3), 1017–1029 (2006).

[75] J.F. Cardoso and A. Souloumiac, “Blind beamforming for non Gaussian signals”, *IEE-Proc.* F 140 (6), 362–370 (1993).

[76] T. Trainini, X.-L. Li, E. Moreau, and T. Adali, “A relative gradient algorithm for joint decompositions of complex matrices”, *Proc. Eur. Signal Process. Conf. (EUSIPCO)* 1, CD-ROM (2010).

[77] J.T.J. Annemueller, T.J. Sejnowski, and S. Makeig, “Complex Independent component analysis of frequency-domain electroencephalographic data”, *Neural Networks* 16, 1311–1323 (2003).

[78] P. Smaragdis, “Blind separation of convolved mixtures in the frequency domain”, *Neurocomputing* 22 (1–3), 21–34 (1998).

[79] E. Bingham and A. Hyvarinen, “A fast fixed-point algorithm for independent component analysis of complex-valued signals”, *Int. J. Neural Syst*. 10 (1), 1–8 (2000).

[80] V.P. Zarzoso and P. Comon, “Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size”, *IEEE Trans. Neural Netw.* 21 (2), 248–61 (2010).

[81] M. Novey and T. Adali, “A complex generalized Gaussian distribution-characterization, generation, and estimation”, *IEEE Trans. on Signal Processing* 58 (3), 1427–1433 (2010).

[82] 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”, *Human Brain Mapping* 6, 160–188 (1998).

[83] A.C. Rencher, *Methods of Multivariate Analysis,* John Wiley & Sons, New York, 1995.

[84] J.A. Mumford and T. Nichols, “Modeling and inference of multisubject fMRI data”, *IEEE Eng Med. Biol. Mag.* 25 (2), 42–51 (2006).

[85] J. Karvanen, J. Eriksson, and V. Koivunen, “Pearson system based method for blind separation”, *Proc. Second Int. Workshop on ICA* 1, CD-ROM (2000).