Published in final edited form as:

IEEE J Sel Top Signal Process. 2016 October ; 10(7): 1214–1225. doi:10.1109/JSTSP.2016.2601695.

Localizing Sources of Brain Disease Progression with Network Diffusion Model

Chenhui Hu,

Microsoft, Cambridge, MA, 02142 USA

Xue Hua,

M3 Biotechnology, Seattle, WA, 98195 USA

Jun Ying,

Chinese PLA General Hospital (301 Hospital), Haidian, Beijing, 100853 China

Paul M. Thompson,

Neurology & Psychiatry, Imaging Genetics Center, University of Southern California, Los Angeles, CA, 90032 USA

Georges E. Fakhri, and

Center for Advanced Medical Imaging Science, Massachusetts General Hospital, Boston, MA 02114 USA

Quanzheng Li

Center for Advanced Medical Imaging Science, Massachusetts General Hospital, Boston, MA 02114 USA

Abstract

Pinpointing the sources of dementia is crucial to the effective treatment of neurodegenerative diseases. In this paper, we propose a diffusion model with impulsive sources over the brain connectivity network to model the progression of brain atrophy. To reliably estimate the atrophy sources, we impose sparse regularization on the source distribution and solve the inverse problem with an efficient gradient descent method. We localize the possible origins of Alzheimer's disease (AD) based on a large set of repeated magnetic resonance imaging (MRI) scans in Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The distribution of the sources averaged over the sample population is evaluated. We find that the dementia sources have different concentrations in the brain lobes for AD patients and mild cognitive impairment (MCI) subjects, indicating possible switch of the dementia driving mechanism. Moreover, we demonstrate that we can effectively predict changes of brain atrophy patterns with the proposed model. Our work could help understand the dynamics and origin of dementia, as well as monitor the progression of the diseases in an early stage.

Index Terms

Sources of dementia; network diffusion; inverse problem; brain morphology; MRI; Alzheimer's disease

I. Introduction

As a collective term describing various symptoms of cognitive decline that is severe enough to interfere with normal life, *dementia* is expected to affect 75.6 million people worldwide in 2030, compared with the current number 44.4 million. About 50% to 80% of dementia is due to Alzheimer's disease (AD), a progressive neurodegenerative disease with symptoms ranging from memory loss, trouble with language to the loss of bodily functions. So far, the cause and progression of AD are not well understood. Consequently, There is no treatment yet to stop or reverse the progression of the disease.

To reveal the pathology of AD, several hypotheses have been proposed to explain the mechanism of the disease. Among those hypotheses, the most influential ones are *amyloid* and *tau* hypotheses. They postulated that the disease begins in the gray matter with accumulation of misfolded beta-amyloid and/or tau protein and progresses along extant fiber pathways [36], [12]. The progression of dementia results in gross atrophy of the affected brain regions, containing degeneration in the temporal lobe and parietal lobe, and parts of the frontal cortex and cingulate gyrus [46].

In the brain structure and function study, network models have been used to evaluate the variation of brain connectivity due to diseases [33], [42], [5], [14]. Moreover, dynamics of neurological disease progression could be described by diffusion processes on the brain network, inspired by the previous works on information diffusion in social networks [26] and sensor networks [11]. Recently, a network diffusion model [30] and an epidemic spreading model [22] have been put forward to characterize the propagation of brain atrophy and misfolded proteins in dementia. The transmission of disease agents like misfolded beta amyloid and tau protein was modeled as a diffusive mechanism mediated by the *brain connectivity network* [3]. Raj et al. predicted spatially distinct “persistent modes” capturing the patterns of dementia (including AD and frontotemporal dementia in their study) by the network diffusion model in [30]. Moreover, prevalence rates forecasted by the model strongly agreed with the published data. With this model, Raj et al. also predicted longitudinal atrophy patterns based on the current atrophy [31]. It was also demonstrated that network diffusion could accurately model the relationship between structural and functional brain connectivity networks [1], which typically depict the neuron fiber connections and the functional correlation of the neuron units, respectively. In addition, the microscopic substance diffusion and related transport mechanisms in brain tissue justify the macroscopic diffusion over the brain connectivity network [27]. However, it is unclear whether the diffusion model without sources can capture nonlinear trajectories of brain diseases [30], [31].

In this work, we infer sources of dementia progression using a network diffusion model with *sparse impulsive stimulations*. The study of brain disease agents (*i.e.*, prion-like proteins) attributed the cause of dementia to a few seed regions in the brain [10], [23]. Based on this, we propose a new network diffusion model that not only describes the propagative nature of dementia, but also considers the dynamics of the seed effect. Particularly, we adopt a linear diffusive model mediated by the brain connectivity structure as an approximation to the nonlinear dynamics of the brain. Linearization reduces the model complexity and thus

makes the parameter fitting more robust than that in nonlinear models. It also helps gain more insight before we try more complicated models. To account the effect of seeds, we introduce sparse impulsive stimulations arriving at different time points to those seed regions. Given the anatomical structure of a brain network, we are interested in inferring the sources of dementia using longitudinal neuroimaging data of brain atrophy.

In practice, we fit the amplitudes of the input impulses and the diffusion speed to the observed data. By assuming that the stimulations arrive discretely, the fitting problem is transferred to a parametrized sparse linear regression problem. An iterative descent search algorithm is presented with a refinement step added later to improve the estimation of the diffusion speed. Numerical simulations demonstrate that we are able to recover the input impulses at the source nodes reliably. Next, we evaluate the longitudinal MRI data in ADNI using this model. MRI provides a noninvasive way to measure the changes of the brain anatomical structure. The shrinkage/atrophy of certain parts of the brain usually are indicators of dementia, such as AD and frontotemporal dementia. The Jacobian maps of the follow-up scans were generated via measuring the changes with respect to the screening scan. After this, we extract the average atrophy at each brain region from the Jacobian maps and infer the dementia source distributions of both AD patients and Mild Cognitive Impairment (MCI) subjects.

Two primary differences between summarized source distributions of the AD group and MCI group are the lower contrast of the distribution and the denser sources in the temporal lobe for the MCI group. We find that the dementia progression is more evidently driven by a set of leading brain regions in AD than in MCI. Moreover, there is a shift of dominant dementia sources from the temporal lobe and cerebellum for MCI to the central brain regions, frontal lobe and the border between parietal lobe and occipital lobe for AD. This phenomenon is consistent with former findings in [43], [34]. Since MCI subjects have high chances to develop AD, the different patterns of the source distributions between MCI and AD may indicate the evolution of the dementia progression mechanism. Part of the results have been presented in [17]. Since then we have significantly strengthened the theory, the simulation and real experimental results. Specifically, we provide a graph filtering interpretation of our work by defining signals and filtering operator on the brain network with the graph Laplacian matrix of the network. In the simulation study, we have considered noise and obtained Receiver Operating Characteristic (ROC) curves to evaluate the source estimation performance under different noise levels. We also apply the network diffusion model with sources to the prediction of brain atrophy patterns and compare the performance with that of the network diffusion model without sources. In addition, we include a more detailed discussion about the results and the effect of the parameters.

The dynamics of longitudinal brain atrophy or morphology have been studied in literature [43], [34], [8]. Meanwhile, dementia progression has been explored in the field of data mining by various means including association rules, group lasso, high-order feature, and probabilistic event cascades [39], [49], [21], [45]. Our model explicitly solves the inverse problem of inferring the source locations from brain atrophy observations. The results may help better understand the dynamics of the brain and design targeted treatments to dementia.

Moreover, the inference method is also applicable to the diffusion source estimation on other types of networks as we discussed at the end of Section V.

The rest of the paper is organized as follows. In Section II, the network diffusion model, input impulsive stimulation model, and the source inference problem are introduced. Then, in Section III, we present an optimization framework of the sparse source localization. Numerical simulations and the evaluation on ADNI data are performed in Section IV and V, respectively. Finally, we discuss the results in Section VI and sum up in Section VII.

II. Diffusion Model for Dementia Progression

In this section, we set up the network diffusion model to describe the mechanism of dementia progression and define the inference problem of the dementia sources.

A. Network diffusion model

We model dementia progression as a diffusion process on a hypothesized brain network $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, where $\mathcal{V}$ and $\mathcal{E}$ are the node set and the edge set, correspondingly. Node $v_i \in \mathcal{V}$ represents the i -th brain region (cortical or subcortical gray matter structure) and edge $(i, j) \in \mathcal{E}$ represents a connected region pair (v_i, v_j) by white-matter fiber pathways usually. W is a symmetric weight matrix with W_{ij} quantifying the connection strength between regions v_i and v_j . Connectivity among brain regions are typically selected from parcellation of brain MRI and the weight coefficients are measured by fiber tractography. We constructed the brain connectivity network based on diffusion tensor imaging (DTI) data.

Let us consider the bundle of neuron fibers from an affected region v_i to an unaffected region v_j . The amount of disease agent transmitting from v_i to v_j is proportional to the product of the disease factor concentration x_i and the inter-region connection strength W_{ij} . Adversely, a reverse diffusion from v_j to v_i proportional to $W_{ji}x_j$ exists. Given undirected pathways, the diffusion process could be captured by the first-order differential equation in Eq. (1) of [30]

$$\frac{dx_i}{dt} = -\beta \sum_j W_{ij}(x_i - x_j), \quad (1)$$

where $\beta > 0$ is a constant controlling the speed of diffusion. In [1], a refined version of Eq. (1) is raised by performing the following normalization

$$\mathcal{W} = D^{-1/2} W D^{-1/2}, \quad (2)$$

where D is the diagonal matrix with the i -th diagonal element $D_{ii} = \sum_j W_{ij}$.

Suppose the disease factors at time t in the network are represented by the vector $\mathbf{x}(t) = \{x(v, t), v \in \mathcal{V}\}$ for all the nodes, then as Eq. (2) of [30] the diffusion process can be rewritten into the following “network heat equation”

$$\frac{d\mathbf{x}(t)}{dt} = -\beta \mathcal{L} \mathbf{x}(t), \quad (3)$$

where $\mathcal{L}$ is the (normalized) graph Laplacian matrix defined as $\mathcal{L} \triangleq I - W$. As given by Eq. (4) of [30], the network diffusion Eq. (3) has an explicit solution

$$\mathbf{x}(t) = \exp(-\beta \mathcal{L} t) \mathbf{x}(0), \quad (4)$$

which defines the evolution of the initial configuration $\mathbf{x}(0)$ under subsequent graph diffusion process on $\mathcal{G}$.

B. Graph filtering interpretation

As an efficient way of processing traditional signals, the classical Fourier transform expands a continuous function x of time in terms of the complex exponentials [28], namely

$$\hat{x}(\xi) \stackrel{\text{def}}{=} \langle x, e^{2\pi i \xi t} \rangle = \int_{\mathbb{R}} x(t) e^{-2\pi i \xi t} dt, \quad (5)$$

where $\langle \cdot, \cdot \rangle$ denotes inner product. We adopt $\hat{x}$ to represent the Fourier transform of x . The Fourier basis functions are eigenfunctions of the one-dimensional Laplace operator :

$$-\Delta(e^{2\pi i \xi t}) = -\frac{\partial^2}{\partial t^2} e^{2\pi i \xi t} = (2\pi \xi)^2 e^{2\pi i \xi t}. \quad (6)$$

Analogously, for a signal $\mathbf{x}$ on the vertices of $\mathcal{G}$, we can define its *graph Fourier transform* (GFT) $\hat{\mathbf{x}}$ as the expansion of $\mathbf{x}$ in terms of the eigenvectors of the graph Laplacian [37], [16]:

$$\hat{x}(\lambda_l) \stackrel{\text{def}}{=} \langle \mathbf{x}, \mathbf{f}_l \rangle = \sum_{i=1}^N x(i) f_l^*(i). \quad (7)$$

Accordingly, the *inverse graph Fourier transform* is given by

$$x(i) = \sum_{l=1}^N \hat{x}(\lambda_l) f_l(i). \quad (8)$$

More concisely, the GFT of $\mathbf{x}$ could be written as the projection of $\mathbf{x}$ onto the eigenvectors of the graph Laplacian in a matrix form

$$\hat{\mathbf{x}} = \mathbf{F}^T \mathbf{x}, \quad (9)$$

where elements of $\hat{\mathbf{x}}$ and $\mathbf{x}$ are indexed by graph Laplacian eigenvalues and vertices, respectively. When the context is clear, we could also write $\hat{x}(\lambda_j)$ and $x(i)$ as $\hat{x}_j$ and x_i accordingly. Since $\mathbf{F}\mathbf{F}^T = \mathbf{I}$, the inverse GFT is $\mathbf{x} = \mathbf{F}\hat{\mathbf{x}}$.

The effect of network diffusion can also be interpreted by a filtering process of the initial configuration. According to the eigen decomposition formula $\mathcal{L} = \mathbf{F}\mathbf{\Lambda}\mathbf{F}^T$ where columns of $\mathbf{F}$ are eigenvectors and $\mathbf{\Lambda}$ is a diagonal matrix containing eigenvalues λ_j , we rewrite Eq. (4) into $\hat{\mathbf{x}}(t) = e^{-\beta\mathbf{\Lambda}t}\hat{\mathbf{x}}(0)$ with $\hat{\mathbf{x}}(t) = \mathbf{F}^T\mathbf{x}(t)$, $\hat{\mathbf{x}}(0) = \mathbf{F}^T\mathbf{x}(0)$. The orthogonal projections $\hat{\mathbf{x}}(t)$, $\hat{\mathbf{x}}(0)$ onto the eigenspace of $\mathcal{L}$ are the GFT of $\mathbf{x}(t)$ and $\mathbf{x}(0)$, accordingly. For a given graph $\mathcal{G}$ with the linear diffusion process defining on it, it will transform the original signal to a signal at time step t by multiplying each frequency component of the GFT of the original signal with a different weight $e^{-\beta\lambda_j t}$. Since $\mathcal{L}$ is positive-semidefinite, every eigenvalue λ_j in the diagonal of $\mathbf{\Lambda}$ is nonnegative. When the eigenvalues are increasingly sorted, we obtained the amplitude gain of $\hat{\mathbf{x}}(t)$ relative to $\hat{\mathbf{x}}(0)$ at each frequency (i.e., λ_j), in Fig. 1 for the 100-node random geometric network we generated in Section IV (see Fig. 3). The plotted curves can be understood as *frequency responses* of the network diffusion process after different time length t , if we regard λ_j as frequency of the network. It can be seen that the diffusion process acts like a *low-pass filter* to the initial signal for $\beta t > 0$. For fixed β , as t gets larger, the distortion of $\hat{\mathbf{x}}(0)$ due to the shrinkage of higher frequencies will increase. This means that the recovery of the source information will become harder accordingly.

C. Input impulsive stimulations

It is known that dementia is not purely an spontaneous process, implying that there are certain sources driving the disease. We therefore model these sources using a series of impulsive stimulations. Mathematically, the input at all the nodes could be expressed by

$$\mathbf{s}(t) = \sum_{t_{ij} \leq t} c_{ij} \delta_{t, t_{ij}} \mathbf{e}_i, \quad (10)$$

where c_{ij} is the amplitude of the impulse at node i at the j -th time step t_{ij} . Note that $\delta(\cdot)$ is the Kronecker delta function defined as $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$; $\mathbf{e}_i$ is a vector of the standard basis in $\mathbb{R}^{|\mathcal{V}|}$ with the i -th element being nonzero. An example of impulsive stimulations is shown in Fig. 2, where each row corresponds to the sequential impulses imported to a node of a certain network. We restrict the possible arrival time of the impulses on integer units from 0 to 12. By adding the input, we could update the dynamics of the dementia in Eq. (3)

$$\frac{d\mathbf{x}(t)}{dt} = -\beta\mathcal{L}\mathbf{x}(t) + \mathbf{s}(t). \quad (11)$$

The solution to the above equation is

$$\mathbf{x}(t) = e^{-\beta\mathcal{L}t}\mathbf{x}(0) + \sum_{t_{ij} \leq t} c_{ij} e^{-\beta\mathcal{L}(t-t_{ij})} \mathbf{e}_i. \quad (12)$$

D. Source inference problem

The network diffusion model we proposed in Eq. (12) paves the road for inferring properties of the sources that drive the dementia, which still have not been well understood. Solving the inverse problem in a principled way could help us find the causes of the disease and maybe targeted treatments too.

Assuming that the brain network structure is known and we observe the distribution of the disease agent (brain atrophy in experiments) at time 0 and t , we are interested in estimating the set of parameters $\boldsymbol{\theta} = \{\beta, c_{ij}, t_{ij}\}$. This is an inverse problem of determining the diffusion pattern based on the source information. Of course, the problem is not well defined, if we do not bound the number of unknown parameters c_{ij} and t_{ij} . On the other hand, from previous research it is rational to assume that the number of input impulsive stimulations is small [10], [23]. Therefore, we further presume that during time $[0, t]$, there are no more than K impulsive inputs at every node, namely $i = 1, \dots, N$ and $j = 1, \dots, K$ for c_{ij} and t_{ij} with $N = |\mathcal{V}|$. We will introduce how to carry out the inference in the next section.

III. Inferring Sources of Dementia

Next, we present an approach to solve the inverse problem described in Section II-D based on further simplified assumptions.

A. Data fitting framework

In real systems, the dynamics given by Eq. (12) might not be followed exactly due to the unavoidable modeling error. But we could estimate the parameters $\boldsymbol{\theta}$ of the diffusion process by minimizing the difference between the two sides of Eq. (12). It motivates us to fit the observed data $\mathbf{y}(t)$ to the diffusive model by

$$\min_{\boldsymbol{\theta}} \left\| \mathbf{y}(t) - e^{-\beta\mathcal{L}t}\mathbf{y}(0) - \sum_{t_{ij} \leq t} c_{ij} e^{-\beta\mathcal{L}(t-t_{ij})} \mathbf{e}_i \right\|_2^2, \quad (13)$$

where the indicators of the unknown parameters take values $i = 1, \dots, N$ and $j = 1, \dots, K$. The

l_2 vector norm is defined as $\|\mathbf{x}\|_2 = (\sum_{i=1}^N |x_i|^2)^{1/2}$. There are $NK+1$ parameters to be fitted,

but the dimension of the observed data is only N , which may make the optimization result unstable.

B. Sparse source localization

To solve the inverse problem more robustly, we simplified the model of the input disease factor as follows. Instead of assuming that the impulsive stimulations could arrive at any time between $[0, t]$, we restricted the arrive time on integer time steps $1, 2, \dots, K = \lfloor t \rfloor$, meaning that $t_{ik} = k$ for any node i and integer $k \in \{1, \dots, K\}$. This simplification is reasonable, since in practice often we only want to determine the sources up to a certain time resolution, e.g., month, year, and so on. Let $\tilde{\mathbf{y}}_{\beta}(t) = \mathbf{y}(t) - e^{-\beta \mathcal{L}} \mathbf{y}(0)$ and $\mathbf{h}_{ij}(\beta) = e^{-\beta \mathcal{L}(t-t_{ij})} \mathbf{e}_i$, then the minimization in Eq. (13) could be replaced by

$$\min_{\beta, c_{ij}} \left\| \tilde{\mathbf{y}}_{\beta}(t) - \sum_{i=1}^N \sum_{j=1}^K c_{ij} \mathbf{h}_{ij}(\beta) \right\|_2^2. \quad (14)$$

The above procedure can be treated as a linear regression parameterized by β . When $K > 1$, the number of dimensions of the dependent variable $\tilde{\mathbf{y}}_{\beta}(t)$ is smaller than the number of explanatory variables c_{ij} , which leads to many solutions to the regression. To obtain a unique and more meaningful solution, a common way is to enforce the l_1 -sparsity constraint with a sparsity-control parameter α [24]:

$$\min_{\beta, c_{ij}} \left\| \tilde{\mathbf{y}}_{\beta}(t) - H(\beta) \mathbf{c} \right\|_2^2 + \alpha \|\mathbf{c}\|_1, \quad (15)$$

where $H = (\mathbf{h}_{1,1}, \mathbf{h}_{1,2}, \dots, \mathbf{h}_{N,K})$ is a $N \times NK$ matrix and $\mathbf{c} = (c_{1,1}, c_{1,2}, \dots, c_{N,K})^T$ is a column vector storing all the amplitudes of the impulses. If β is known, then the optimization for the sparse linear regression (a.k.a, Lasso [49]) can be effectively solved by a few well-known methods including the coordinate descent learning and the active-set algorithm [35], [41]. However, when β is unknown, the overall estimation task becomes hard. Therefore, we proposed a two-stage algorithm with an initial stage and a refined stage. For the inference problem in Eq. (15), we first adopted the following iterative algorithm:

Algorithm 1

Solving the Parameterized Lasso Problem

-
- 1 Given $0 < \varepsilon_1, \varepsilon_2 < 1$. Initialize any feasible $\beta^{(1)}, \{c_{ij}^{(1)}\}$.
 - 2 For $k = 1, 2, \dots$
 - 2.1 Find a descent direction $\mathbf{d}$ of the Lasso problem in Eq. (15) given $\beta = \beta^{(k)}$ using algorithms in [35].
 - 2.2 Update $\{c_{ij}^{(k+1)}\} = \{c_{ij}^{(k)}\} - \varepsilon_1 \mathbf{d}$.
 - 2.3 Update $\beta^{(k+1)} = \beta^{(k)} - \varepsilon_2 f^T \mathbf{g}$, with

$$f = \tilde{\mathbf{y}}_{\beta^{(k)}}(t) - H(\beta^{(k)}) \mathbf{c}^{(k+1)},$$
-

$$g = e^{-\beta^{(k)} \mathcal{L} t} \mathcal{L} t \mathbf{x}(0) + \sum_{i,j} c_{ij}^{(k+1)} (t - t_{ij}) \mathcal{L} \mathbf{h}_{ij}(\beta^k).$$

Now that we have multiple longitudinal observations, each time when applying the above algorithm, we obtain an estimate of the impulsive inputs after the last observation as well as β . In the end, there will be several estimates of β , namely $\hat{\beta}_1, \dots, \hat{\beta}_K$. Thus, we average them

to get a refined version $\hat{\beta} = \frac{1}{K} \sum_{i=1}^K \hat{\beta}_i$. Next, we treat $\beta = \hat{\beta}$ in Eq. (15) and apply the build-in function LASSO() in Matlab (R2012b) to solve the problem. 5-fold cross-validation is performed for tuning the regularization parameter α at both the initial stage and refined stage. Alternatively, Bayesian information criterion (BIC) or Akaike information criterion (AIC) [15] can also be applied to determine α . These two criteria avoid overfitting by achieving a balance between the model fit and complexity. It is worth mentioning that AIC and leave-one-out cross-validation are asymptotically equivalent [40].

C. Atrophy prediction method

With the atrophy source inference algorithm, we can also predict the changes of brain atrophy patterns. Since the distributions of atrophy sources among subjects in a same group might have similarities, potential gain of the prediction power might be obtained by incorporating the information of atrophy sources. Thus, we present the following 10-fold cross-validation approach. First, we randomly divide the subjects in either AD or MCI group into 10 folds as equally as possible. Then, we estimate the atrophy sources for each subject using Algorithm 1 with 9 folds of the data (the rest fold was used for evaluating the prediction) and the brain connectivity network of each group. Note that the source distribution of each subject can be represented by a vector $\mathbf{c} = (c_{1,1}, c_{1,2}, \dots, c_{N,K})^T$, with $N=116, K=24$ indicating the number of ROIs and the number of months between the baseline and the 24th month. The averaged distribution of the atrophy sources was computed by averaging the source distributions of all the training subjects and is denoted by $\mathbf{c}^{\text{train}} = \{\mathbf{c}_{i,j}^{\text{train}}\}$. After that, we predict the atrophy of the subjects in the rest fold according to Eq. (12):

$$\mathbf{x}^{\text{est}}(t) = e^{-\beta \mathcal{L} t} \mathbf{x}(0) + \sum_{t_{ij} \leq t} c_{ij}^{\text{train}} e^{-\beta \mathcal{L} (t - t_{ij})} \mathbf{e}_i, \quad (16)$$

where $\mathbf{x}(0)$ denotes the brain atrophy of the baseline and $\mathbf{x}^{\text{est}}(t)$ is the prediction of the atrophy at time t . By choosing every possible fold as a testing set, we can obtain the prediction of the brain atrophy of each subject at $t = 6, 12, 24$ months. To make the results more stable, we further repeat the above prediction process 100 times by dividing the data into 10 folds randomly each time.

IV. Numerical Simulations

We evaluated the diffusion source inference algorithm by simulations on a random geometric network. To generate the network, we arbitrarily spread 100 nodes in a unit square

area and drew edges between node pairs when their Euclidean distances were less than 0.2 (see Fig. 3). Unit weight was assigned to each edge. Then, we imported 7 impulsive stimulations $c_{10,0} = 1$, $c_{46,3} = 0.6$, $c_{90,7} = 0.6$, $c_{51,9} = 0.2$, $c_{30,10} = 0.6$, $c_{70,18} = 0.3$, $c_{20,20} = 0.7$ with c_{ij} signifying the input at node i at the j -th time step. We mark out the locations of the input sources using blue boxes. Observations with Gaussian noise $y(t) = x(t) + n(t)$ were made at three time steps $t = 6, 12, 24$, where $n(t) \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_{N \times N})$. In Fig. 3, we visualize the observations on the network with heat maps. Warmer colors of the nodes indicate greater measurements and sizes of the nodes are proportional to the absolute values of the measurements. There are several factors which influence the accuracy of sources estimation. Here we mainly discuss the standard deviation of the Gaussian noise σ and the sparsity level regularization parameter α .

We changed the noise level σ from 0.03 to 0.11 with a step size 0.02. The associated estimation performances captured by ROC curves are displayed in Fig. 4. From these ROC curves, we observe that when σ is lower, higher detection rate can be achieved with a same false alarm rate. Thus, higher accuracy of estimation can also be obtained. Here the estimation is correct if the position of impulses are all correctly estimated. Parameter α controls the sparsity of estimation results. When α is bigger, the weight of $|c|_1$ in Eq. (15) is larger, which makes the estimated sources more sparse. On the contrary, there will be more inferred sources if α is smaller. A reasonable value of α could be obtained by cross-validation and grid search. Here we choose $\alpha = 0.1$ for a proper tradeoff between the model fitness and the sparsity of the sources. Furthermore, the amplitudes of source impulses and the locations of sources also affect the accuracy of estimation. Typically, larger impulsive amplitudes make the signal-to-noise ratio higher and thus produce a better inference result. Meanwhile, if the distances between the input sources are larger, it is more easier to localize the sources.

The simulated experiments demonstrate that our method is able to determine the locations of the diffusion sources under moderate assumption of the noise level and source distributions. Next, we will test it on a real dataset in ADNI.

V. Evaluation on ADNI Data

In what follows, we test the diffusion source inference algorithm on the longitudinal brain MRI data collected by ADNI. We present the pipeline of the data processing and then the experiment results.

A. General description

An intermediate state between normal aging and AD is mild cognitive impairment (MCI), a condition with greatly increased risk of developing AD. Almost 10–25% of MCI subjects progress to AD every year [29]. We thus carried out experiments on both AD and MCI data in this section.

High-resolution structural MRI is one of several neuroimaging methods used to track AD. Numerous MRI-derived biomarkers including hippocampal volume, lateral ventricular volumes, gray matter volume or density, *etc.*, have been used to improve the diagnostic

accuracy. Dedicated to promote joint research based on multimodal neuroimaging techniques, the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu) has collected data from 50 sites over the world. ADNI includes three phases: ADNI-1, ADNI-GO, and ADNI-2. In this study, we employed our diffusion source inference method and the tensor-based morphometry (TBM) [19] to analyze the ADNI-1 dataset, containing 3314 MRI scans at screening and follow-up scans.

B. Subject information

The full ADNI-1 dataset contained sequential brain MRI scans from 188 AD patients (age at screening: 75.4 ± 7.5 years, 99 Male (M)/89 Female (F)), 400 individuals with MCI (age: 74.8 ± 7.4 years, 257 M/143 F), and 229 healthy elderly controls (age: 76.0 ± 5.0 years, 119 M/110 F). Subjects were scanned at screening and followed up at 6, 12, 18 (MCI only), 24 months. The complete 2-year visit subset included 98 AD (age: 75.2 ± 7.4 years, 52 M/46 F), 207 MCI (age: 74.9 ± 7.0 years, 139 M/68 F), and 163 healthy subjects (age: 76.0 ± 4.9 years, 83 M/80 F) scanned at screening, 6, 12, 18 (MCI only) and 24 months. In this study, we infer brain atrophy sources using both AD and MCI data.

C. Data acquisition and preprocessing

All subjects were scanned by following a standardized MRI protocol developed for ADNI. High-resolution structural brain MRI scans were attained using 1.5 Tesla MRI scanners. Through a sagittal 3D MP-RAGE scanning protocol, images with 1 mm isotropic voxels were reconstructed. The MP-RAGE was run twice to increase the chance that at least one scan would be available for analysis. To adjust for linear shifts in head position and scale within the same subject, the follow-up scan was linearly registered to its matching screening scan via 9-parameter (9P) registration, using a mutual information (MI) cost function [9]. In addition, to deal with the global discrepancies in brain scale across subjects, the mutually aligned time-series of scans was then linearly registered to the International Consortium for Brain Mapping template (ICBM- 53), applying the same 9P transformation to both mutually aligned scans [18].

As an image analysis tool that measures brain structural differences from the gradients of deformation fields aligning one image to another, TBM was adopted to compute the cumulative brain change in the subjects with respect to the screening scan [32]. By warping the 9P-registered and 'skull-stripped' follow-up scans to match the corresponding screening scan, individual Jacobian maps were produced to estimate 3D patterns of structural brain change over time. Color-coded maps of the Jacobian determinants were created to illustrate regions of ventricular/CSF expansion (*i.e.*, with $\det J(r) > 1$), or brain tissue loss/atrophy (*i.e.*, with $\det J(r) < 1$) over time (see leftmost column of Fig. 5(a)). The above longitudinal maps of tissue change were also spatially normalized across subjects for regional comparisons and group statistical analyses [19].

D. Experiment procedures and results

Next, we constructed the brain connectivity network and extracted the brain atrophy measurements from the Jacobian deformation maps, before inferring the diffusion sources. As illustrated in Fig. 5, the Jacobian maps of the 6, 12, 24 month were registered to the 116

region automated anatomic labeling (AAL) cerebral atlas [44]. Meanwhile, we built the averaged brain connectivity network of each subject group by estimating the connectivity probability between two ROIs with the DTI data in ADNI.

We reconstructed the whole-brain white matter tracts of 10 AD and 10 MCI subjects using the FMRIB Diffusion Toolbox (FSL, version 5.0; <http://www.fmrib.ox.ac.uk/fsl>). Details of the DTI data processing are described in [4]. Each voxel was selected as a seed region, and its connectivity probability to each of the rest voxels was calculated. Since the probability from voxel i to voxel j and the probability from j to i were highly correlated across the voxels for all subjects, we defined the unidirectional connectivity probability between i and j by averaging these two probabilities. The number of streamlines between every two voxels were estimated by sampling the posterior connectivity probability with Markov chain Monte Carlo (MCMC). Then, we constructed a brain network of the 116 ROIs by computing the number of streamlines between every two ROIs. We visualize the mean brain networks across the subjects of the AD and the MCI group in Fig. 5(b). In addition, to remove suspicious connections, we kept edges with larger connection weights by applying a threshold to each network. We denote the ratio of the edges remained in the thresholded networks as γ , whose effect will be discussed in Section VI-B.

We ran our diffusion source inference algorithm for the AD and MCI groups given the atrophy measurements at time steps $t = 6, 12, 24$ months. For comparison purpose, we set the sparsity regularization factor α in Eq. (15) the same in both cases. From cross-validations, we found that $\alpha \in [10^{-3}, 10^{-5}]$ was favorable for the two groups. In this paper, we presented the results obtained by choosing $\alpha = 10^{-4}$. At $t = 6$ month, we replaced $\mathbf{y}(0) = \mathbf{0}$ and $t = 6$ in Eq. (13). Arrival times and locations of the impulsive stimulations of dementia were estimated via solving this optimization for every individual (the same procedure for AD and MCI groups). The resulted amplitude sequences c_{ij} for $i = 1, \dots, N$ and $j = 1, \dots, K$, were folded up and summed over each ROI to give the cumulative atrophy input $\tilde{c}_i$ at i -th region:

$$\tilde{c}_i = \sum_{j=1}^K c_{ij}. \quad (17)$$

Thus, $\tilde{\mathbf{c}} = (\tilde{c}_1, \dots, \tilde{c}_N)$ describes the dementia source distribution of the subject. In order to obtain a meaningful summary of the atrophy sources across the group, we normalized $\tilde{\mathbf{c}}$ such that the maximum absolute value of its elements was one. The normalized source distributions for the AD and MCI groups were displayed in the first row of Fig. 6. We also inferred the driving sources during the time intervals $t \in [6, 12]$ and $t \in [12, 24]$ by treating $\mathbf{y}(6)$ and $\mathbf{y}(12)$ as initial conditions. These results were in the second and third rows of Fig. 6. From the plots, we observed that the envelopes of the source distributions obtained at different time steps for a certain subject group were close to each other, which might signify that the dementia was due to stimulations at a constant set of regions over the 2 year visit period.

To obtain an atrophy source map for every subject group, we further added up the source distributions $\mathcal{C}$ estimated at the three time steps. These total dementia source distributions were presented on the left panels of Fig. 7. Two major differences between summarized source distributions of the AD and MCI groups were the lower contrast of the distribution and the denser sources in the temporal lobe for MCI. To quantify the contrast of the source distributions, we calculated the ratio between the sum of top 60 leading source regions and that of the rest regions. It turned out that $\rho_{AD} = 8.97$ compared with $\rho_{MCI} = 2.70$, meaning that the dementia progression was more evidently driven by a set of leading brain regions in AD than by those MCI. We also examined the difference of the distributions over brain lobes from the two groups. Using the visualization tool BrainNet Viewer [47] (www.nitrc.org/projects/bnv/), we displayed the center locations of the 60 major dementia sources inside the 3D brain mesh in Fig. 7(b–c) and Fig. 7(e–f).

E. Brain atrophy prediction results

To demonstrate the advantage of the proposed network diffusion model, we applied it to predict patterns of brain atrophy and compared the performance with that of the diffusion model without sources in [30]. Other than our proposed method, we implemented the brain atrophy prediction algorithm in [30], where the authors adopted a diffusion model without sources to predict brain atrophy rate and obtained the absolute atrophy by integration.

In Fig. 8, the atrophy pattern of a representative AD subject at 24 months predicted by the proposed network diffusion model is compared with both the true atrophy and the prediction result of the diffusion model without sources. From the figure, we observe that the estimated distribution of the atrophy is close to the truth and better than the estimation by the diffusion model without sources. The prediction error is probably due to the nonlinear dynamics of the brain atrophy that may not be fully captured by the linear diffusion model. In addition, we plot the estimated mean atrophy of each subject against the truth of the mean atrophy in Fig. 9. From both the results of both the AD and NC group, we find that our model achieves better prediction accuracy compared with the diffusion model without sources. Specifically, for the AD group, the mean square errors (MSEs) of the estimation by the diffusion model with and without sources are 0.9703 and 2.4636, respectively; for the MCI group, the MSEs of the estimation by these two methods are 1.2472 and 2.4330, respectively. This indicates that our method may provide a better support for detecting and monitoring AD in an early stage.

VI. Discussion

A. pathological explanation

In this section, we discuss the dominant regions that drive AD, since they could be good starting points to understand the causes of the disease. In Table I, we summarize part of the significant atrophy sources of the two subject groups (top 11 regions of the AD group and top 12 regions of the MCI group). The above regions demonstrated overlaps with the hubs of the functional brain network revealed in [47], namely the bilateral rolandic operculum, bilateral superior temporal gyrus, right supplementary motor area, right temporal pole, right

supramarginal gyrus, left medial orbital superior frontal gyrus, bilateral insula and bilateral putamen, which are primarily located at the association and subcortical areas.

In addition, there is a shift of dominant dementia sources from the temporal lobe and cerebellum for MCI to the central brain regions, frontal lobe and the border between parietal lobe and occipital lobe for AD. This is in line with previous findings about the dynamics of the brain atrophy in [43], [34]. There may be multiple reasons about the different sources in the MCI group and the AD group. One possible cause is the long-term evolution of AD. Since MCI subjects have high chances to develop AD, the difference of the sources between MCI and AD may indicate the transition of the AD driving mechanism. In [7], Braak et al. demonstrated that different stages of AD corresponded to distinct foci in the brain by studying 2332 subjects from 1- to 100-year-old. Moreover, the diffusion process is spreading on the associated brain network of each subject group. In a long term, the difference between the MCI network and the AD network increases. This also makes the sources of the two groups more distinct. In contrast, the sources of each group over different time intervals are stable as shown in Fig. 6, since we have evaluated the atrophy data in 24 months which is a short period compared to the whole course of AD progression. Notice that the disease stages found by Braak et al. in [7] specify the spread of the disease to new regions. The sources that arise between MCI and AD stages are probably contained in these new regions. Nevertheless, some of the new regions may not be the new sources. It is also possible that some of the new sources are due to the overfitting of the model. Although we have adopted cross-validation to determine the sparsity parameter, the new sources that we have discovered are still dependent on the sparsity parameter in the model. If we increase the sparsity parameter, a few weak sources might disappear. To overcome this issue, we have picked the top 60 atrophy sources by thresholding for each subject group. Another possibility could be that the AD atrophy sources are ubiquitous instead of sparse. In this case, we believe that our model might be helpful to determine the most significant sources of the disease progression.

We also find several sources in the temporal lobe, which is one of the areas that have the most prominent damage. Both left/right rolandic operculum in the AD group and left/right inferior temporal gyrus in the MCI group are involved with the temporal lobe. Note that part of the rolandic operculum is in the temporal lobe (it lies in between the parietal, frontal, and temporal lobe). However, the hippocampal/entorhinal regions are not found as strong sources. One possible reason is that the subjects we evaluated have already been diagnosed as MCI and AD. During these two late stages, some other regions might play a more significant role in driving the disease. It is possible that we will find hippocampal/entorhinal regions as significant sources if we go far enough to the early stages of AD. Moreover, it has been shown that with the progression of AD the connectivity between these regions and the other brain regions decreases [14], [13]. This makes the hippocampal/ entorhinal regions isolated and reduces the impact of these regions on the rest of the brain. Thus, the importance of hippocampal/entorhinal regions in terms of affecting AD could become less.

In both subject groups, we observe atrophy sources at the cingulate and cingulate gyrus (No. 33–36 in AD group and No. 36 in MCI group). Cingulate areas are abnormal in very early stages of AD. According to Huang et al., those who developed to AD had a significantly

decreased regional cerebral blood flow (rCBF) in the left posterior cingulate cortex [20]. Significantly reduced gray matter density is also discovered in posterior cingulate areas in mild AD patients. This indicates that cingulate disorder is prominent in the early stages of AD. Thus, it is possible that cingulate areas play a leading role in the AD progression process.

Moreover, No. 72–75 in the striatum have been inferred as possible sources of dementia in MCI group. There are evidence showing that striatum suffers more injury during AD. Braak and Braak revealed that AD patients have distinguished amyloid and neurofibrillary changes in the striatum area by sensitive silver methods [6]. Meanwhile, Simic et al. pointed out that AD patients' behavioral and psychological symptoms indicate that brain stem especially locus coeruleus might be the origin of AD [38]. This is supported by some of our discovered regions in the MCI group, namely No. 103–105 and No. 109 in cerebellar hemispheres, which have a close connection with coeruleus.

B. Parameter effect and future work

In Eq. (15), α controls the weight of $|c|_1$ in the minimization. Larger α brings more entries in c to zero, and the estimated sources are more sparse. To better estimate the dementia sources, we have tried different values of α by carrying out grid search in logarithmic scale. We find that when α is too small (less than 10^{-5}), the convergence speed of Algorithm 1 is too slow; while when α is too large (greater than 10^{-1}), the number of the estimated sources reduces to zero. Our results are stable when the regularization parameter $\alpha \in [10^{-5}, 10^{-1}]$. For α in this range, although the sparsity and strength of the inferred sources may vary, the top dementia sources are stable for each subject group.

The parameter γ represents the ratio of the edges retained in the brain networks. We have tested the algorithm by changing it from 0.05 to 1, while other conditions are kept the same. We find that different γ in this range gives very similar source distribution for each subject group. This is probably because the brain networks are dominated by a few strong connections as we can see from Fig. 5(b).

In the experiments, we have used the averaged DTI networks for each subject group. Using individual brain network for each subject may further improve the accuracy of the dementia source estimation and the prediction of the brain atrophy patterns. Nevertheless, we have already achieved satisfactory prediction results in Fig. 9. Thus, the averaged networks serve as good surrogates for the individual networks. In the future, we will test our model based on both the averaged DTI networks obtained from more subjects and individual brain networks.

Our work provides an effective way of estimating the dementia sources and predicting the longitudinal brain changes. The diffusive model and the data fitting framework that we proposed in Eqs. (11) and (13) are applicable to the inference problem of diffusion sources on other network structures as well, for instance, disease spreading networks [2], social networks [48], [25], and sensor networks [11].

VII. Conclusion

We presented a network diffusion model with impulsive sources and an efficient algorithm to solve the inverse problem of tracing back the diffusion sources. Numerical simulations demonstrated that we could estimate the locations of the sources reliably. Possible origins of Alzheimer's disease were found by using this model and longitudinal MRI scans. The average distributions of the dementia sources have different concentrations in the brain lobes for the AD patients and MCI subjects, indicating the evolution of the dementia driving mechanism. We also showed that the proposed method could effectively predict the changes of the brain atrophy patterns. Our method enables the quantitative assessment of the dementia causes, which may help discover targeted treatments of the disease.

Acknowledgments

This work was funded in part by the National Institutes of Health under Grant R01EB013293.

References

1. Abdelnour F, Voss HU, Raj A. Network diffusion accurately models the relationship between structural and functional brain connectivity networks. *Neuroimage*. 2014; 90(15):335–347. [PubMed: 24384152]
2. Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences*. 2009; 106(51):21484–21489.
3. Bassett DS, Bullmore ET. Human brain networks in health and disease. *Current Opinion in Neurology*. 2009; 22(4):340. [PubMed: 19494774]
4. Behrens T, Woolrich M, Jenkinson M, Johansen-Berg H, Nunes R, Clare S, Matthews P, Brady J, Smith S. Characterization and propagation of uncertainty in diffusion-weighted mr imaging. *Magnetic resonance in medicine*. 2003; 50(5):1077–1088. [PubMed: 14587019]
5. Belmonte MK, Allen G, Beckel-Mitchener A, Boulanger LM, Carper RA, Webb SJ. Autism and abnormal development of brain connectivity. *The Journal of Neuroscience*. 2004; 24(42):9228–9231. [PubMed: 15496656]
6. Braak H, Braak E. Alzheimer's disease: striatal amyloid deposits and neurofibrillary changes. *Journal of Neuropathology & Experimental Neurology*. 1990; 49(3):215–224. [PubMed: 1692337]
7. Braak H, Thal DR, Ghebremedhin E, Del Tredici K. Stages of the pathologic process in alzheimer disease: age categories from 1 to 100 years. *Journal of Neuropathology & Experimental Neurology*. 2011; 70(11):960–969. [PubMed: 22002422]
8. Chen R, Resnick SM, Davatzikos C, Herskovits EH. Dynamic bayesian network model for longitudinal brain mophometry. *Neuroimage*. 2012; 59(3):2330–2338. [PubMed: 21963916]
9. Collins D, Neelin P, Peters T, Evans A. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J Comput Assist Tomogr*. 1994; 18:192–205. [PubMed: 8126267]
10. de Calignon A, Polydoro M, Suárez-Calvet M, William C, Adamowicz D, Kopeikina K. Propagation of tau pathology in a model of early Alzheimer's disease. *Neuron*. 2012; 73(4):685–697. [PubMed: 22365544]
11. Dokmani , I., Ranieri, J., Chebira, A., Vetterli, M. Sensor networks for diffusion fields: detection of sources in space and time. *Communication, Control, and Computing (Allerton)*, 2011 49th Annual Allerton Conference on; IEEE; 2011. p. 1552-1558.
12. Goedert M, Clavaguera F, Tolnay M. The propagation of prion-like protein inclusions in neurodegenerative diseases. *Trends in Neurosciences*. 2010; 33(7):317–325. [PubMed: 20493564]

13. Gómez-Isla T, Price JL, McKeel DW Jr, Morris JC, Growdon JH, Hyman BT. Profound loss of layer ii entorhinal cortex neurons occurs in very mild alzheimer's disease. *The Journal of neuroscience*. 1996; 16(14):4491–4500. [PubMed: 8699259]
14. Greicius MD, Srivastava G, Reiss AL, Menon V. Default-mode network activity distinguishes alzheimer's disease from healthy aging: evidence from functional mri. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101(13):4637–4642. [PubMed: 15070770]
15. Hastie, TJ., Tibshirani, RJ., Friedman, JH. *The elements of statistical learning: data mining, inference, and prediction*. Springer; 2011.
16. Hu C, Cheng L, Sepulcre J, Fakhri GE, Lu YM, Li Q. Matched signal detection on graphs: theory and application to brain network classification. *IPMI*. 2013:1–12.
17. Hu C, Hua X, Thompson PM, Fakhri GE, Li Q. Inferring sources of dementia progression with network diffusion model. *Machine Learning in Medical Imaging*. 2014; 8679:42–49.
18. Hua X, Hibar DP, Ching CR, Boyle CP, Rajagopalan P, Gutman BA. Unbiased tensor-based morphometry: improved robustness and sample size estimates for Alzheimer's disease clinical trials. *Neuroimage*. 2013; 66:648–661. [PubMed: 23153970]
19. Hua X, Lee S, Yanovsky I, Leow A, Chou Y, Ho A. Optimizing power to track brain degeneration in Alzheimer's disease and mild cognitive impairment with tensor-based morphometry: an ADNI study of 515 subjects. *NeuroImage*. 2009; 48:668–681. [PubMed: 19615450]
20. Huang C, Wahlund LO, Svensson L, Winblad B, Julin P. Cingulate cortex hypoperfusion predicts alzheimer's disease in mild cognitive impairment. *BMC neurology*. 2002; 2(1):9. [PubMed: 12227833]
21. Huang J, Alexander DC. Probabilistic event cascades for Alzheimer's disease. *NIPS*. 2012:3104–3112.
22. Iturria-Medina Y, Sotero RC, Toussaint PJ, Evans AC, Initiative ADN, et al. Epidemic spreading model to characterize misfolded proteins propagation in aging and associated neurodegenerative disorders. *PLoS Comput Biol*. 2014; 10(11):e1003956. [PubMed: 25412207]
23. Jucker M, Walker LC. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature*. 2013; 501(7465):45–51. [PubMed: 24005412]
24. Lee H, Lee DS, Kang H, Kim BN, Chung MK. Sparse brain network recovery under compressed sensing. *IEEE Trans on Medical Imaging*. 2011; 30(5):1154–1165. [PubMed: 21478072]
25. Luo W, Tay WP, Leng M. Identifying infection sources and regions in large networks. *IEEE Trans on Signal Processing*. 2013; 61(11):2850–2865.
26. Ma H, Yang H, Lyu MR, King I. Mining social networks using heat diffusion processes for marketing candidates selection. *ACM CIKM*. 2008:233–242.
27. Nicholson C. Diffusion and related transport mechanisms in brain tissue. *Reports on Progress in Physics*. 2001; 64(7):815.
28. Oppenheim, AV., Schafer, RW., Buck, JR., et al. *Discrete-time signal processing*. Vol. 2. Prentice-hall Englewood Cliffs; 1989.
29. Petersen, R. *Mild cognitive impairment: Aging to Alzheimer's disease*. Oxford University Press; New York: 2003.
30. Raj A, Kuceyeski A, Weiner M. A network diffusion model of disease progression in dementia. *Neuron*. 2012; 73(6):1204–1215. [PubMed: 22445347]
31. Raj A, LoCastro E, Kuceyeski A, Tosun D, Relkin N, Weiner M, et al. ADNI (ADNI. Network diffusion model of progression predicts longitudinal patterns of atrophy and metabolism in alzheimers disease. *Cell reports*. 2015; 10(3):359–369.
32. Riddle W, Li R, Fitzpatrick J, DonLevy S, Dawant B, Price R. Characterizing changes in mr images with color-coded Jacobians. *Magn Reson Imaging*. 2004; 22:769–777. [PubMed: 15234445]
33. Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*. 2010; 52(3):1059–1069. [PubMed: 19819337]
34. Sabuncu MR, Desikan RS, Sepulcre J, Yeo BTT, Liu H, Schmansky NJ. The dynamics of cortical and hippocampal atrophy in Alzheimer disease. *Archives of Neurology*. 2011; 68(8):1040–1048. [PubMed: 21825241]

35. Schmidt M, Fung G, Rosales R. Fast optimization methods for l1 regularization: A comparative study and two new approaches. *ECML*. 2007:286–297.
36. Seeley WW, Crawford RK, Zhou J, Miller BL, Greicius MD. Neurodegenerative diseases target large-scale human brain networks. *Neuron*. 2009; 62:42–52. [PubMed: 19376066]
37. Shuman D, Narang SK, Frossard P, Ortega A, Vandergheynst P, et al. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *Signal Processing Magazine, IEEE*. 2013; 30(3):83–98.
38. Simic G, Stanic G, Mladinov M, Jovanov-Milosevic N, Kostovic I, Hof PR. Does alzheimer's disease begin in the brainstem? *Neuropathology and applied neurobiology*. 2009; 35(6):532–554. [PubMed: 19682326]
39. Simon GJ, Li PW, CRJ, Vemuri P. Understanding atrophy trajectories in Alzheimer's disease using association rules on MRI images. *ACM SIGKDD*. 2011:369–376.
40. Stone M. An asymptotic equivalence of choice of model by cross-validation and akaike's criterion. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1977:44–47.
41. Sun J, Bi J, Kranzler HR. Quadratic optimization to identify highly heritable quantitative traits from complex phenotypic features. *ACM KDD*. 2013:811–819.
42. Supekar K, Menon V, Rubin D, Musen M, Greicius MD, et al. Network analysis of intrinsic functional brain connectivity in alzheimers disease. *PLoS Comput Biol*. 2008; 4(6):e1000100. [PubMed: 18584043]
43. Thompson PM, Hayashi KM, Zubicaray GD, Janke AL, Rose SE, Semple J. Dynamics of gray matter loss in Alzheimer's disease. *The Journal of Neuroscience*. 2003; 23(3):994–1005. [PubMed: 12574429]
44. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*. 2002; 15(1):273–289. [PubMed: 11771995]
45. Wang H, Nie F, Huang H, Yan J, Kim S, Risacher SL. High-order multi-task feature learning to identify longitudinal phenotypic markers for Alzheimer's disease progression prediction. *NIPS*. 2012:1286–1294.
46. Wenk GL. Neuropathologic changes in Alzheimer's disease. *J Clin Psychiatry*. 2003; 64(Suppl 9): 7–10.
47. Xia M, Wang J, He Y. BrainNet Viewer: a network visualization tool for human brain connectomics. *PLoS One*. 2013; 8(7):e68910. [PubMed: 23861951]
48. Yang J, Leskovec J. Modeling information diffusion in implicit networks. *IEEE ICDM*. 2010:599–608.
49. Zhou J, Liu J, Narayan VA, Ye J. Modeling disease progression via fused sparse group lasso. *ACM SIGKDD*. Aug.2012 :1095–1103.

Biographies

Chenhui Hu is a Data Scientist at Microsoft in Cambridge, MA. He received his Ph.D. degree from Harvard University in 2016. He received both his B.S. and M.S. degrees in electronic engineering from Shanghai Jiao Tong University in 2007 and 2010 respectively. He was a research assistant in the Gordon Center for Medical Imaging Sciences, Radiology

Department at Massachusetts General Hospital in Boston from 2012 to 2016. He was a recipient of the Outstanding Graduates of Shanghai in 2010 and the 3rd IEEE ComSoc Asia-Pacific Outstanding Paper Award in 2014. His current research interests include machine learning, statistical signal processing, brain network analysis, and mobile wireless communication.

Xue Hua is a Senior Scientist at M3 Biotechnology, designing imaging-based clinical trials to accelerate new Alzheimers drug development. She has over 10 years of research experience in developing brain imaging biomarkers to track Alzheimers disease progression, as part of the Alzheimers Disease Neuroimaging Initiative. Additionally, she has led research projects in neurodegenerative and neurobehavioral disorders such as HIV/AIDS, autism, and schizophrenia. Her research has resulted in over 50 peer-reviewed journal articles with over 3,000 citations. Dr. Hua is an Adjunct Assistant Professor in Neurology at the University of Southern California. Prior to joining M3, she served as a founding faculty member at the Imaging Genetics Center. She earned her Ph.D. in Biomedical Engineering in 2008 from the University of California, Los Angeles.

Jun Ying is a Biomedical Engineer of 301 Hospital, China and has been a visiting scholar in Massachusetts General Hospital, Harvard Medical School from 2014 to 2015. He received his Ph.D. degree from Tsinghua University in 2014. In 2006, he joined the biomedical research center of 301 Hospital in Beijing. He is currently a deputy director of the Big Data Center of the hospital leading programs of clinical data analysis and signal processing. His research interests include brain function analysis based on brain network methods and multi-modal medical images as well as deep learning applications in disease prediction.

Paul Thompson is a professor of neurology at the Imaging Genetics Center (IGC) at the University of Southern California. Thompson obtained a bachelor's degree in Greek and Latin languages and mathematics from Oxford University. He also earned a master's degree in mathematics from Oxford and a PhD degree in neuroscience from University of California, Los Angeles. Recent awards include the Fulbright and Hughes Fellowships (1993–1998), SPIE Medical Imaging Best Paper Award (1997), Eiduson Award for Neuroscience Research (1998), and the Turken Award for Alzheimer's Disease research (2003). Thompson specializes in the field of human brain imaging, with research interest in mathematical and computational algorithm development for human brain mapping, and has contributed to more than 1300 publications. He currently leads the Enhancing Neuro Imaging Genetics Through Meta-Analysis (ENIGMA) project, a global data collection and sharing effort designed to understand how brain structure changes during the trajectory of brain atrophy, mental illness and Alzheimer's and the underlying genetic landscape.

Georges El Fakhri is a Professor of Radiology at Harvard Medical School and the founding Director of the Gordon Center for Medical Imaging (GCMi). He also directs the Massachusetts General Hospital (MGH) Gordon PET Core, and is Co-Director of the Division of Nuclear Medicine and Molecular Imaging. Dr. El Fakhri is an internationally recognized expert in quantitative SPECT, PET-CT, and PET-MR. He has pioneered novel approaches to compensate for many of the physical factors affecting quantitative SPECT, PET/CT and PET/MR and objectively assessing the achieved improvement in image quality, specifically in PET oncologic, neurologic and cardiac imaging as well as in the development of novel approaches to quantitative cardiac and brain modeling. Recent work includes the development of novel approaches to mapping mitochondrial complex I and membrane potential non invasively in the heart as well as developing synergistic approaches in PET/MR for motion compensation, anatomical priors, PET/fMRI and PET/MRS. Dr El Fakhri has been recognized with many awards including the 2004 Dana Foundation Award, the 2005 Mark Tetalman Award, the Howard Hughes Medical Institute Training Innovation

Award, a 2014 Distinguished Investigator Award and was elevated to IEEE Fellow for contributions to biological imaging.

Quanzheng Li is Assistant Professor of Radiology at Massachusetts General Hospital, Harvard Medical School. He received his M.S. degree from Tsinghua University in 2000, and his Ph.D degree in Electrical Engineering from the University of Southern California (USC) in 2005. He did his post-doctoral training at USC from 2006 to 2007, and was a Research Assistant Professor from 2008 to 2010. In 2011, he joined the Gordon Center for Medical Imaging Sciences, Radiology Department at Massachusetts General Hospital in Boston where he is currently the assistant Director of Gordon Center and the director of image reconstruction and analysis program in Gordon Center. Dr. Li is the recipient of 2015 IEEE Nuclear and Plasma Sciences Society (NPSS) early achievement award. He is an associate editor of IEEE Transaction on Image Processing and editorial board member of Theronostics. His research interests include image reconstruction methods in PET, SPECT, CT and MRI as well as applications combining image formation, analysis and statistical inference together to optimize task-based performance.

Fig. 1.

A frequency responsive view of the network diffusion process. The x-axis indicates the indices of increasingly sorted eigenvalues λ_i of $\mathcal{L}$; the y-axis represents the gain of the observed signal relative to the original signal at each eigenvalue (or frequency of the network) after diffusion.

Fig. 2.

Schematic representation of the impulsive sequences imported to the nodes of a network as indicated by $i = 1, 2, 3$. The horizontal and vertical axes correspond to the arrival time (t_{ij}) and the amplitude (c_{ij}) of the impulsive disease stimulation.

Fig. 3.

Observations of a diffusion process on a 100-node random geometric network at time steps $t = 0, 6, 12, 24$. Edges which were built by connecting nearby nodes have unit weights. The blue boxes mark out the input sources between every two consecutive time intervals with the numbers aside being the node indices. Note that the sizes and colors of the nodes both indicate the amplitudes of the noisy measurements.

Fig. 4.

ROC curves of inferencing results under different noise levels. Noise standard deviations varied from 0.03 to 0.11 with a step size 0.02. We recovered locations of the input sources between $t = 0$ and $t = 24$. False Positive (*a.k.a.*, $1 - \text{specificity}$) means the percentage of nodes that were not sources but estimated to be sources; True Positive (*a.k.a.*, *sensitivity*) means the ratio of nodes that were sources and indeed estimated as sources.

Fig. 5.

(a) Schematic of the postprocessing procedures. The leftmost are the color-code Jacobian maps of the observations at the 6th, 12th, 24th months (warmer colors indicate ventricle/CSF expansion and cooler colors denote tissue loss [18]). The maps are registered to the AAL116 template to extract the average brain deformation within each ROI, giving the plots in the 3rd column; meanwhile, the brain connectivity network is estimated from DTI data. Afterwards, we remove the positive values in the average deformation resulting the 4th column. (b) Weight matrices of average DTI brain networks for the AD and MCI group, respectively. The cross between every row and column corresponds to a certain edge. The color of each edge represents the logarithm of the streamlines between every two ROIs.

Fig. 6.

Average source distributions recovered at three time steps ($t = 6, 12, 24$ months) for the AD group and the MCI group. Horizontal axes of the subplots denote the ROI index of the 116-region AAL atlas (see http://neuro.imm.dtu.dk/wiki/Automated_Anatomical_Labeling for the names and descriptions of the ROIs); vertical axes of the charts signify the normalized cumulative atrophy imported to a certain brain region after the last observation averaged over population. The color of each bar represents the brain network module that the region belongs to, as illustrated by the color of the ball in Fig. 7(b–c) and (e–f) as well.

Fig. 7.

Estimated total dementia source distributions averaged over the AD group (1st row) and the MCI group (2nd row). Dotted lines in (a) and (d) delineate the thresholds above which there are 60 brain regions for both groups. (b–c) and (e–f) show the sagittal and axial views of these ROIs with significant atrophy, where the sizes of the color-balls are proportional to the atrophy levels and the colors denote different brain modules.

Fig. 8.

The prediction results of the brain atrophy at 24 months of a representative subject (AD subject, No. 5). The atrophy of the ROIs has been projected to the cortex. The color shows the absolute value of the atrophy.

Fig. 9.

The atrophy prediction performance using diffusion model with and without sources. Each point represents a subject at a certain time step.

TABLE I

Part of the significant atrophy sources of the AD group and MCI group, respectively.

Atrophy sources of the AD group		Atrophy sources of the MCI group	
17: Left rolandic operculum	18: Right rolandic operculum	6: Right superior orbitofrontal gyrus	36: Right posterior cingulate gyrus
33: Left middle cingulate	34: Right middle cingulate	73: Left putamen	74: Right putamen
35: Left posterior cingulate gyrus	36: Right posterior cingulate gyrus	75: Left globus pallidus	76: Left globus pallidus
46: Right cuneus	51: Left middle occipital gyrus	89: Left inferior temporal gyrus	90: Right inferior temporal gyrus
64: Right supramarginal gyrus	65: Left angular gyrus	103: Left lobule VIII of cerebellum	104: Right lobule VIII of cerebellum
66: Right angular gyrus		105: Left lobule IX of cerebellum	109: Lobule I, II of vermis