

Imaging Genetics Study Based on a Temporal Group Sparse Regression and Additive Model for Biomarker Detection of Alzheimer's Disease

Meiyan Huang^{ID}, Xiumei Chen, Yuwei Yu, Haoran Lai, and Qianjin Feng^{ID}, *Member, IEEE*

Abstract—Imaging genetics is an effective tool used to detect potential biomarkers of Alzheimer's disease (AD) in imaging and genetic data. Most existing imaging genetics methods analyze the association between brain imaging quantitative traits (QTs) and genetic data [e.g., single nucleotide polymorphism (SNP)] by using a linear model, ignoring correlations between a set of QTs and SNP groups, and disregarding the varied associations between longitudinal imaging QTs and SNPs. To solve these problems, we propose a novel temporal group sparsity regression and additive model (T-GSRAM) to identify associations between longitudinal imaging QTs and SNPs for detection of potential AD biomarkers. We first construct a nonparametric regression model to analyze the nonlinear association between QTs and SNPs, which can accurately model the complex influence of SNPs on QTs. We then use longitudinal QTs to identify the trajectory of imaging genetic patterns over time. Moreover, the SNP information of group and individual levels are incorporated into the proposed method to boost the power of biomarker detection. Finally, we propose an efficient algorithm to solve the whole T-GSRAM model. We evaluated our method using simulation data and real data obtained from AD neuroimaging initiative. Experimental results show that our proposed method outperforms several state-of-the-art methods in terms of the receiver operating characteristic curves and area under the curve. Moreover, the detection of AD-related genes and QTs has been confirmed in previous studies, thereby further verifying the effectiveness of our approach and helping understand the genetic basis over time during disease progression.

Index Terms—Imaging genetics, Alzheimer's disease, longitudinal imaging data, nonparametric regression model, magnetic resonance imaging, single nucleotide polymorphism.

Manuscript received October 20, 2020; revised December 13, 2020; accepted February 3, 2021. Date of publication February 8, 2021; date of current version April 30, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 81601562, Grant 81974275, and Grant U1501256, in part by the Science and Technology Planning Project of Guangzhou under Grant 201904010417, in part by the Science and Technology Project of Guangdong Province under Grant 2015B010131011, in part by the ADNI (National Institutes of Health) under Grant U01 AG024904, and in part by the DOD ADNI (Department of Defense) under Grant W81XWH-12-2-0012. (Corresponding authors: Meiyan Huang; Qianjin Feng.)

The authors are with the School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China, and also with the Guangdong Provincial Key Laboratory of Medical Image Processing, Southern Medical University, Guangzhou 510515, China (e-mail: huangmeiyan16@163.com; chenxiumei97@163.com; 2523750358@qq.com; haoranlai@163.com; 1271992826@qq.com).

This article has supplementary downloadable material available at <https://doi.org/10.1109/TMI.2021.3057660>, provided by the authors.

Digital Object Identifier 10.1109/TMI.2021.3057660

I. INTRODUCTION

ALZHEIMER'S disease (AD) is the most common type of dementia that often affects individuals aged more than 65 years [1]. The clinical manifestations of AD are mainly memory impairment, cognitive decline, and behavior disorder [2]. According to a recent study [1], the total number of patients with AD worldwide is over 56 million. In 2050, one person will develop AD in every 33 seconds, and the AD population will increase to over 152 million [3]. Therefore, understanding the cause and mechanism of disease progression of AD is necessary and has gained wide attention in clinical research.

AD risk is affected by genetic variants that an individual carry [4]. Thus, understanding the genetic causes and mechanism of AD is important to achieve early prediction and treatment of this disease [5]. Moreover, neuroimaging data can be used to explore the relationships between genetic variants and brain structures and functions [6], [7]. Research in this emerging field, known as imaging genetics, is aimed at identifying disease-related biomarkers by analyzing the associations between genetic variants, such as single nucleotide polymorphisms (SNPs) and quantitative traits (QTs) of neuroimaging data [8], [9] to reveal the genetic causes of AD [10].

During the last decade, numerous imaging genetics studies have been conducted to identify the associations between QTs and SNPs [5], [9], [11]–[14]. Early imaging genetic studies are based on the mass-univariate linear model, which uses pairwise univariate analysis to identify the relationships between single QT and single SNP. However, the correlation between QTs and SNPs is in fact a complex many to many relationships, i.e., a set of SNPs is probably associated with multiple QTs and a set of QTs is also probably related to multiple SNPs [15]. Therefore, the traditional mass-univariate linear model is insufficient to identify the complex associations between SNPs and QTs. To address this problem, scholars proposed the multivariate sparse regression model [16], [17]. For instance, Wang *et al.* [13] proposed a group sparse multitask regression model to identify associations between SNPs and QTs by using linkage disequilibrium (LD) block and gene grouping to consider the combined effects of SNPs. Zhu *et al.* [18] proposed a structured sparse low rank regression model, which can explicitly consider the correlations within SNPs and QTs simultaneously for imaging genetics study. Sparse canonical correlation analysis (SCCA) is introduced to identify bi-multivariate associations between

multiple SNPs and multiple QTs without pre-selecting candidate biomarkers [14], [19]. Although many efforts have been made by constructing an association model to detect genetic biomarkers, most of these methods are based on a linear model, which may be inappropriate when the associations between QTs and SNPs are nonlinear.

The sparse additive model (SpAM) combined with sparsity constraint is widely used in imaging genetics study, where each additive component is a smooth function of a single SNP [20]. Thus, the nonlinear effect of SNP can be incorporated into the association model to obtain more accurate association analysis between SNP and QT than the traditional linear model [21]. However, most existing SpAM methods only use QT at a single time point. QT changes continuously with time and carries abundant time information when crossing multiple time points. AD is a progressive neurodegenerative disease that develops continuously over time [22]. For instance, a normal control (NC) subject may gradually convert to mild cognitive impairment (MCI) and finally deteriorate to AD, whereas some patients with MCI may remain the same or even revert to NC during this period. The structure and function of the brain change rather than remain the same as the state of the disease [9]. Therefore, longitudinal information should be considered for accurate biomarker detection of AD. To date, few methods have been established for imaging genetics by using longitudinal information in SpAM method. For instance, Marchetti-Bowick *et al.* [8] proposed a time varying group sparse additive model (TV-GroupSpAM) to identify associations between SNP and longitudinal QT. In the TV-GroupSpAM, an l_{21} constrain is used to treat each SNP as a group and each of three SNP genotypes (denoted as AA, Aa, and aa) as a subgroup. Therefore, different and combinational effects of three genotypes in an SNP on QTs at different time points can be detected by using TV-GroupSpAM [8]. However, only a single QT was used in TV-GroupSpAM, i.e., only correlations between hippocampus and multiple SNPs were identified; extending this method to a set of regions of interest (ROIs) in the brain is difficult because it may elide associations between different ROIs and genetic data. Moreover, group information among SNPs is ignored in TV-GroupSpAM, which may weaken the power of biomarker detection [23], [24].

Generally, the current imaging genetics methods face three problems. First, the linear model is used to construct the associations between QTs and SNPs, which is inappropriate when these associations are nonlinear. Second, most of these methods only apply QTs from a single time point (e.g., the baseline visit), thereby missing various associations between QTs and SNPs. Third, these methods consider either each SNP or some SNPs as a group, which may ignore the group or individual information of SNPs. To address these problems, we propose a novel additive model called temporal group sparsity regression and additive model (T-GSRAM), to learn the bi-multivariate associations between longitudinal QTs and SNPs simultaneously. Our contributions are as follows. First, we incorporate brain imaging QTs at different time points to determine various associations between SNPs

and QTs. Although longitudinal QT data are also used in TV-GroupSpAM, only a single QT with different time points is applied in this method, and extending this method to a set of QTs is difficult as mentioned above. Given that different atrophy rates are present across different brain regions in AD, a common progressive pattern is shared among multiple time points instead of all time points in many cases [9], [25]. These regional variations may be caused by the effects of different SNPs at various time points [9]. Therefore, regional variations at different time points and brain regions should be considered in AD imaging genetics studies [9]. To investigate regional variations at different time points, a fuse pairwise group lasso [10] is used to select effective QTs at two adjacent time points. Moreover, the l_1 norm is applied to select effective QTs at a specific time point, and the l_{21} norm is used to select effective QTs at all time points and to detect variations across different brain regions. Therefore, the trajectory of disease progression represented by QTs can be captured by these regularizations.

Second, inspired by TV-GroupSpAM, T-GSRAM uses the l_{21} norm not only to select effective SNPs but also to investigate the different and combinational effects of three genotypes in an SNP on QTs at different time points. However, TV-GroupSpAM only considers single SNP information and ignores group information for SNPs. Moreover, genetic functionalities are usually controlled by multiple SNPs from one gene, and the nonrandom association between alleles at different loci can be described by LD [26]. Therefore, the group l_{21} norm (G_{21} norm) is used to examine the joint effect of SNPs in an inherent LD block [27] or in gene.

Third, the proposed T-GSRAM is a nonparametric regression model where the effects of each SNP on QT are regarded as a smooth function of time. This smooth function may work with a specified parametric form (such as a polynomial of a variable) or nonlinear transformation. Therefore, T-GSRAM is more flexible than a linear method for analyzing the complex associations between SNPs and QTs. To evaluate the performance of T-GSRAM, we perform a set of experiments on simulation data and real data obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) [28]. The real data include longitudinal magnetic resonance imaging (MRI) measurements obtained over a 2-year period and 6452 SNPs from 202 genes. Results of the simulation data experiment show that T-GSRAM can more accurately detect relevant biomarkers compared with four state-of-the-art methods. Meanwhile, results of the real data experiments successfully detect some AD-related risk genes. Therefore, T-GSRAM can be used for the biomarker detection of AD.

The remainder of this paper is organized as follows. Section two describes the proposed T-GSRAM method. Moreover, an efficient algorithm is introduced to solve the T-GSRAM. In section three, the performance of T-GSRAM is evaluated by using simulation studies and a real data analysis. In section four, a comprehensive discussion of the experimental results for the T-GSRAM is presented.

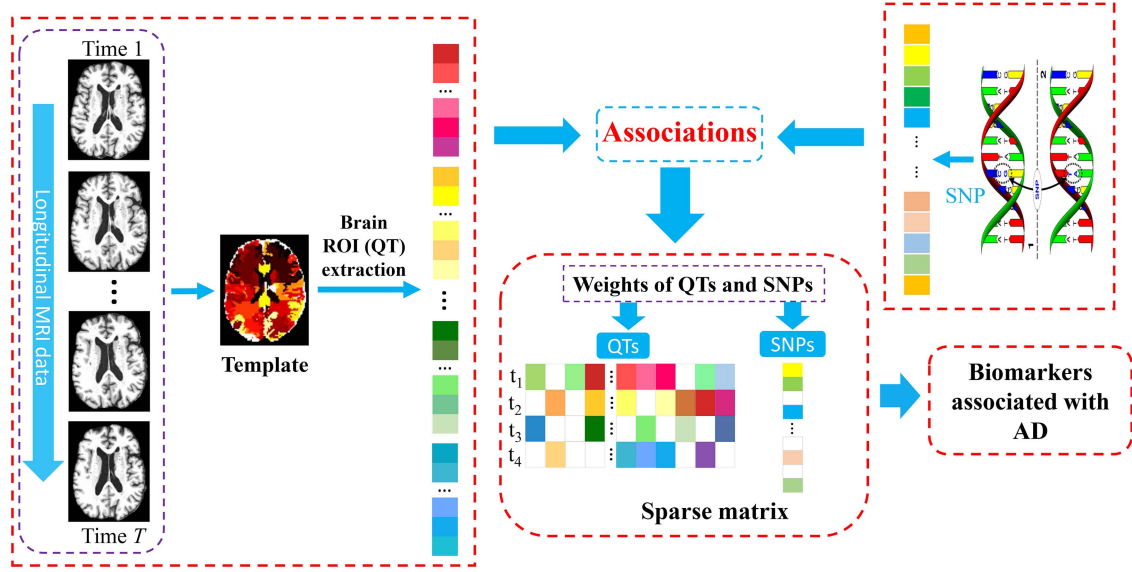


Fig. 1. Flowchart of the proposed T-GSRAM method.

II. METHOD

Fig. 1 presents the T-GSRAM flowchart. This method comprises three main steps, namely, data preprocessing for MRI imaging and SNP data, constructing associations between SNPs and QTs by using T-GSRAM, and detecting AD-associated biomarkers.

In this article, we define a matrix with boldface uppercase letter, a vector with boldface lowercase letter, and a scalar with normal italics. Specifically, let $\mathbf{X} \in \mathbb{R}^{n \times p}$ denotes the SNP data, where n and p are sample number and feature dimension of SNPs, respectively. An SNP takes a value of 0, 1, or 2, which indicate the number of minor alleles at the corresponding chromosome location [20]. Let $\mathbf{Y}_t \in \mathbb{R}^{n \times q}$; $t = 1, \dots, T$ denote QT data at time t , and $\mathbf{Y} \in \mathbb{R}^{n \times q}$ denote QT data at baseline, where q and T are the feature dimension of QT and number of time points, respectively. \mathbf{Y} and \mathbf{Y}_t is normalized. Moreover, we use $\|\mathbf{X}\|_F = \sqrt{\sum_i \|\mathbf{x}_i\|_2^2}$ and $\|\mathbf{X}\|_{2,1} = \sum_i \sqrt{\sum_j x_{ij}^2}$ to denote the Frobenius norm and l_{21} norm, respectively, where x_{ij} is the element of the i -th row and j -th column of matrix \mathbf{X} , \mathbf{x}_i is the i -th row of \mathbf{X} , and \mathbf{x}_j is the j -th column of \mathbf{X} .

A. Sparse Additive Model

The SpAM is a nonparametric setting model, and can be formally defined as

$$\min_f \frac{1}{2} \left\| \mathbf{Y} - \sum_{j=1}^p f_j(\mathbf{x}_j) \right\|_2^2 + \lambda \sum_{i=1}^p |f_j| \quad s.t. \quad \|f_j\|_2 = 1 \quad (1)$$

where $f_j(\mathbf{x}_j)$ is the genetic effect of j -th SNP, and λ is regularization parameter. A sparsity constraint is imposed in SpAM to prevent overfitting because feature dimension is higher than sample size (i.e., $n \ll p$). However, SpAM is a univariate additive model that can only identify the association between

single SNP and QT. To address this problem, Yin *et al.* [29] proposed a group sparse additive model (GroupSpAM), where l_{21} penalty is added into the model to detect joint effects of multiple SNPs. The formula for GroupSpAM is defined as

$$\min_f \frac{1}{2} \left\| \mathbf{Y} - \sum_{k=1}^K \sum_{j=1}^{c_k} f_{kj}(\mathbf{x}_{kj}) \right\|_2^2 + \lambda \sum_{k=1}^K \|f_k\|_2 \quad (2)$$

where K is the group number of SNPs. Although the associations between SNPs and single QT can be identified in GroupSpAM, the structures and functions of the brain changes over time rather than remain stay the same. Thus, developing longitudinal imaging genetics methods is of great importance. Marchetti-Bowick *et al.* [8] proposed a TV-GroupSpAM method, which is a nonparametric regression model with a group LASSO constraint over a given SNP and its corresponding three genotypes. Moreover, longitudinal data were used in TV-GroupSpAM to determine the different effects of three genotypes in an SNPs. The TV-GroupSpAM is denoted as

$$\min_f \frac{1}{2} \sum_{t=1}^T \left\| \mathbf{Y}_t - \sum_{j=1}^p f_j(t, \mathbf{x}_j) \right\|_2^2 + \lambda \|f\|_{21} \quad (3)$$

where $f_j(t, \mathbf{x}_j)$ denotes the genetic effect of SNP \mathbf{x}_j at time t and can be defined as $f_j(t, \mathbf{x}_j) = \sum_{g=0}^2 f_j^g(t) x_j^g$, where $f_j^g(t) \in \{f_j^0, f_j^1, f_j^2\}$ denotes one of the three different SNP effects. x_j is a categorical variable and can be defined as $x_j^g = 1 \Leftrightarrow x_j = g$. Thus, different and combinational effects of three genotypes in an SNP on QTs at different time points can be detected using TV-GroupSpAM [8].

B. T-GSRAM

Although useful longitudinal information associated with an SNP can be detected, regional variations in different brain regions are ignored because only single QT is applied in

TV-GroupSpAM. Moreover, group information among SNPs is neglected in TV-GroupSpAM. The brain structures and functions are affected by SNPs both conjointly at group level and individually. Moreover, different temporal patterns can be exhibited longitudinally in SNPs due to regional variations in cerebral decline of the human brain [30]. Therefore, we propose a novel T-GSRAM method to analyze the complex genetic effects on temporal patterns in imaging genetics

$$\min_{f, \mathbf{V}} \frac{1}{2} \sum_{t=1}^T \left\| \mathbf{Y}_t \mathbf{V}_t - \sum_{k=1}^K \sum_{j=1}^{c_k} \sum_{g=0}^2 f_j^g(t) x_j^g \right\|_2^2 + \Omega(f) + \Omega(\mathbf{V})$$

$$s.t. \quad \|\mathbf{Y}_t \mathbf{V}_t\|_2^2 = 1 \quad \forall t \quad (4)$$

where SNPs are partitioned into K groups, which are genetically linked; c_k is the number of SNPs at k -th group [i.e. $\{f_j\}_{j=1}^{c_k}$]. Moreover, $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_T]$ denotes the weights of QTs.

1) *Sparse Regularization of QT*: Regional variations at different time points and different brain regions caused by disease progression should be considered. Thus, three constraints for QTs are incorporated into the proposed model to detect the trajectory of the disease progression

$$\Omega(\mathbf{V}) = \lambda_{v1} \|\mathbf{V}\|_{FP_{21}} + \lambda_{v3} \sum_{s=1}^q \|\mathbf{v}_s\|_1 + \lambda_{v2} \|\mathbf{V}\|_{21} \quad (5)$$

where λ_{v1} , λ_{v2} and λ_{v3} are the control parameters of regularization terms. First, the fuse pairwise l_{21} norm (FP_{21}) [9, 31] is defined as

$$\|\mathbf{V}\|_{FP_{21}} = \sum_{s=1}^q \sum_{t=1}^{T-1} \sqrt{(v_t^s)^2 + (v_{t+1}^s)^2} \quad (6)$$

This norm first applies the l_2 norm to a single QT at two adjacent time points and then uses the l_1 norm for all QTs across the whole brain. This norm can be applied to investigate changes in QTs at two adjacent time points. Therefore, if a QT remains unchanged at two adjacent time points, then the QT is recognized as irrelevant of AD.

Second, norm $\|\mathbf{v}_s\|_1$ is defined as

$$\|\mathbf{v}^s\|_1 = \sum_{t=1}^T |v^s| \quad (7)$$

The l_1 norm aims to ensure the individual sparsity at a specific time point.

Finally, the l_{21} norm is defined as

$$\|\mathbf{V}\|_{21} = \sum_{s=1}^q \sqrt{\sum_{t=1}^T (v_t^s)^2} \quad (8)$$

This norm first applies the l_2 norm to a single QT at all time points, and then uses the l_1 norm for all QTs across the whole brain. Therefore, imaging QTs that remain stable across all time points would be discarded.

2) *Group and Individual Sparsity for SNP Regularization*: The genetic effects on the brain structure and function are usually carried out conjointly at group level (a gene or a LD block) and individually. Moreover, the different and combinational effects of three genotypes in an SNP should be considered. Therefore, a group sparsity penalty and an individual sparsity penalty on SNPs are introduced into our model, and thus $\Omega(f)$ can be defined as

$$\Omega(f) = \lambda_{f1} \|f\|_{G_{21}} + \lambda_{f2} \|f\|_{l_{21}} \quad (9)$$

where λ_{f1} and λ_{f2} are control parameters of the regularization terms for the group and individual level, respectively. The group sparse norm is defined as

$$\|f\|_{G_{21}} = \sum_{k=1}^K \sqrt{\sum_{j=1}^{c_k} \sum_{g=0}^2 \sum_{t=1}^T (f_j^g)^2} \quad (10)$$

The G_{21} norm defined in (10) enforces group-wise sparsity, which indicates that SNPs at the group level will be selected or unselected simultaneously [13]. However, certain features could be irrelevant within a same group. Thus, we impose an additional individual sparsity to SNPs via l_{21} norm regularization [32], [33]. The l_{21} norm is defined as

$$\|f\|_{l_{21}} = \sum_{j=1}^p \sqrt{\sum_{g=0}^2 \sum_{t=1}^T (f_j^g)^2} \quad (11)$$

The l_{21} norm emphasizes on a single variable sparsity and could help identify a single SNP that plays a role in the longitudinal QTs.

C. Optimization Algorithm

In the section, we describe an optimization algorithm for minimizing (4) to gain optimal f and \mathbf{V} . Directly optimizing (4) is difficult. Therefore, an alternative convex search method is applied to optimize f and \mathbf{V} [34]. First, we initialize f as a random vector and \mathbf{V} as a matrix with all one element. Second, we first fix f to update \mathbf{V} and then fix \mathbf{V} to update f until the algorithm converges.

1) *Update V by Fixing f*: When f is fixed, (4) can be rewritten as

$$\min_v \frac{1}{2} \sum_{t=1}^T \left\| \mathbf{Y}_t \mathbf{V}_t - \sum_{k=1}^K \sum_{j=1}^{c_k} \sum_{g=0}^2 f_j^g(t) x_j^g \right\|_2^2$$

$$+ \lambda_{v1} \|\mathbf{V}\|_{FP_{21}} + \lambda_{v2} \|\mathbf{V}\|_{21} + \lambda_{v3} \sum_{s=1}^q \|\mathbf{v}_s\|_1$$

$$s.t. \quad \|\mathbf{Y}_t \mathbf{V}_t\|_2^2 = 1 \quad (12)$$

where (12) is a convex function. We first take the derivative with respect to \mathbf{V} and then set the derivative to be zero. We can get the solution of \mathbf{V}

$$\mathbf{v}_t = (\lambda_{v1} \mathbf{D} + \lambda_{v2} \tilde{\mathbf{D}} + \lambda_{v3} \bar{\mathbf{D}} + (1 + \gamma_v \mathbf{Y}_t^T \mathbf{Y}_t))^{-1}$$

$$\times \mathbf{Y}_t^T \left(\sum_{k=1}^K \sum_{j=1}^{c_k} \sum_{g=0}^2 f_j^g x_j^g \right) \quad (13)$$

where \mathbf{D} is a diagonal matrix, and $1/\sqrt{(v_{t-1}^s)^2 + (v_t^s)^2} + 1/\sqrt{(v_t^s)^2 + (v_{t+1}^s)^2}$ is the s -th diagonal element in \mathbf{D} . $\tilde{\mathbf{D}}$ and \mathbf{D} are diagonal matrices, where $1/\|\mathbf{v}^s\|_2$ and $1/|v_t^s|$ are the s -th diagonal elements in $\tilde{\mathbf{D}}$ and \mathbf{D} , respectively.

2) *Update f by Fixing \mathbf{V}* : When updating f by fixing \mathbf{V} , we can rewrite (4) as

$$\min_f \frac{1}{2} \sum_{t=1}^T \left\| \mathbf{Y}_t \mathbf{V}_t - \sum_{k=1}^K \sum_{j=1}^{c_k} \sum_{g=0}^2 f_j^g(t) x_j^g \right\|_2^2 + \lambda_{f1} \|f\|_{G_{21}} + \lambda_{f2} \|f\|_{l_{21}} \quad \text{s.t.} \quad \|f\|_2^2 = 1 \quad (14)$$

where (14) is a convex function, and we can use a block coordinate descent algorithm [35] to solve it. (14) can be solved using three steps. First, we take the derivative of (14) with respect to f_k . Second, we derivate the j -th variable of the k -th group. Finally, we set the derivative to be zero and normalize $f_j = f_j / \|f_j\|_2$ in the current iteration to obtain the solution of f . The coordinate-wise solution of f for (14) is defined as

$$f_j^g = (1 + \lambda_{f1} \mathbf{D}_1 + \lambda_{f2} \tilde{\mathbf{D}}_1)^{-1} E(R_j | x_j^g) \quad (15)$$

where $R_j = \mathbf{Y}_t \mathbf{V}_t - \sum_{k \neq j} \sum_{g=0}^2 f_k^g x_k^g$ is the partial residual for f_j . \mathbf{D}_1 is a norm of k -th group and can be represented as $1/\|f_k\|_2$, where $1 \leq k \leq K$. $\tilde{\mathbf{D}}_1$ is also a norm for j -th SNP of k -th group and can be denoted as $1/\|f_j\|_2$. $E(R_j | x_j^g)$ is a conditional expectation operator with respect to R_j . We first assume that $E(\cdot | x_j)$ is a smooth function of x_j and then introduce a smooth matrix S_j^g to replace the conditional expectation operator [8]. S_j^g is defined as

$$S_j^g = \begin{cases} K_\alpha |t^a - t^b| & x_j^a = x_j^b = g \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

where (a, b) is a pair of data points, each corresponding to a particular individual j and time points t , and K_α is a smooth kernel function with bandwidth α . Therefore, (15) can be rewritten as

$$f_j^g = (1 + \lambda_{f1} \frac{1}{\|f_j\|_2} + \lambda_{f2} \frac{1}{\|f_k\|_2})^{-1} S_j^g R_j \quad (17)$$

We use the alternative search algorithm to solve f and \mathbf{V} , respectively, which is described in Algorithm 1. Moreover, the detailed convergence analysis is presented in the Supplementary Materials.

III. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we use simulations and a real example to evaluate the finite-sample performance of T-GSRAM. The code of T-GSRAM is available at the coding sharing site (<https://github.com/Meiyan88/T-GSRAM>).

Algorithm 1 Algorithm for T-AGRAM

Input: SNP data $\mathbf{X} \in \mathbb{R}^{n \times p}$, QT data $\mathbf{Y}_t \in \mathbb{R}^{n \times q}$, time points $t \in \{1, \dots, T\}$. Initialize f_j^g for $j = 1, \dots, p$ and $g \in \{0, 1, 2\}$, and parameters $\lambda_{f1} \lambda_{f2} \lambda_{v1} \lambda_{v2} \lambda_{v3}$.

Output: f_j^g ; $j = 1, \dots, p$, $g = 0, 1, 2$ and $\mathbf{V} \in \mathbb{R}^{q \times T}$.

1. **while** not converge **do**
2. **for** $k = 1, \dots, K$ **do**
3. Compute partial residual for k :
 $\hat{R}_k = \mathbf{YV} - \sum_{\neq k} \sum_j \sum_g f_j^g x_j^g$;
4. Estimate projected residuals: $\mathbf{P}_k^g = S_k^g \hat{R}_k \forall g$
5. Compute group norm: $\hat{w}_k = \sqrt{\sum_j \sum_g \|\mathbf{P}_j^g\|_2^2}$
6. **if** $\hat{w}_k \leq \lambda_{f1} + \lambda_{f2}$ **then** set $f_k^g = 0 \forall g$
7. **else for** $j = 1, \dots, c$
8. Compute partial residual for j : $\hat{R}_{k,j} = \hat{R}_k + f_j$
9. Estimate projected residuals: $\mathbf{P}_j^g = S_j^g \hat{R}_{k,j} \forall g$
10. Compute individual norm: $\hat{w}_j = \sqrt{\sum_g \|\hat{P}_j^g\|_2^2}$
11. **if** $\hat{w}_j \leq \lambda_{f1}$ **then** set $f_j^g = 0 \forall g$.
12. **else** update $f_j^g \forall g$ by iterating until convergence:
 $f_j^g = (1 + \lambda_{f1} \frac{1}{\|f_j\|_2} + \lambda_{f2} \frac{1}{\|f_k\|_2})^{-1} P_j^g$ and scale f_j^g so that $\|f\|_2^2 = 1$
13. **end if**
14. **end for**
15. **end if**
16. **end for**
17. Update \mathbf{D} , $\tilde{\mathbf{D}}$, $\tilde{\mathbf{D}}$
18. Solve \mathbf{v}_t alternatively according to (13) and scale \mathbf{v}_t so that $\|\mathbf{Y}_t \mathbf{V}_t\|_2^2 = 1$
19. **end while**

A. Simulation Studies

1) *Simulation Setting*: In this simulation study, we performed several experiments on simulated data to illustrate the utility of our method. We generated data according to the following procedure. First, we simulated SNP data as follows. We used LD blocks defined by the default method of Haploview and PLINK to form SNP-sets. To calculate LD blocks, we simulated n subjects by randomly combining the haplotypes of HapMap CEU subjects. We used PLINK to determine LD blocks based on these subjects. We randomly selected 20 blocks and combined the haplotypes of HapMap CEU subjects in each block to form genotype variables for these subjects. We randomly selected 10 SNPs in each block, and thus we had $p = 200$ SNPs for each subject. To fully evaluate the influence of the different associations between QTs and SNPs (linear or nonlinear association) on the performance of different methods, we used the following three cases to generate \mathbf{Y} :

Case 1:

$$\mathbf{Y}_t \mathbf{V}_t = \mathbf{XW} + e \quad (18)$$

Case 2:

$$\mathbf{Y}_t \mathbf{V}_t = \sin(\mathbf{X})\mathbf{W} + e \quad (19)$$

Case 3:

$$\mathbf{Y}_t \mathbf{V}_t = \sum_{j=1}^p \sum_{g=0}^2 f_j^g(t) x_j^g + e \quad (20)$$

where $e \sim N(0, \sigma^2)$, $t \sim \text{Unif}(0,1)$, and x_j^g are the simulated SNP data as described above. In Cases 1 and 2, we initially randomly generated sparse matrices $\mathbf{V} \in \mathbb{R}^{q \times t}$ and $\mathbf{W} \in \mathbb{R}^{p \times 1}$ (i.e., the AD-related QTs or SNPs (causal QTs or SNPs) were set to nonzero values, whereas the others were set to zero). Afterward, we simulated \mathbf{Y} by using (18) and (19), respectively. In Case 3, we generated \mathbf{Y} following the additive model, and we initially built a set of realistic imaging data (QTs) by randomly subsampling individuals obtained from the ADNI dataset. Additional information on this ADNI dataset can be found in the ADNI data analysis section. We used MRI scans at the baseline visit (BL), 6 months (M6), 12 months (M12), and 24 months (M24) to construct longitudinal imaging data. For each time point, we used 93 ROIs extracted from MRI scans as QTs. We estimated $f_j^g(t)$ by fitting (20) to real QTs. We randomly generated a sparse matrix $\mathbf{V} \in \mathbb{R}^{q \times t}$ (i.e., the AD-related QTs were set to nonzero values, whereas the others were set to zero) and then chose A SNPs with the highest scores in the estimated $f_j^g(t)$ as the causal SNPs, reserved the corresponding values of the causal SNPs, and set the other values to zeros in $f_j^g(t)$. We eventually calculated \mathbf{Y} by using the above known variables according to the model defined in (20). We built a linear association between \mathbf{X} and \mathbf{Y} in Case 1 and a nonlinear association between \mathbf{X} and \mathbf{Y} in Cases 2 and 3.

The T-GSRAM was evaluated with two nested cross-validation loops (fivefold for each loop). Specifically, for the external fivefold cross-validation, all subject samples were divided into five subsets with the same proportion. For each run, all samples within one subset were successively chosen as the testing set, whereas the remaining samples in the other four subsets were combined and used as the training set for model training. The final results were reported as the mean results from each run. Moreover, parameter tuning was evaluated with the inner fivefold cross-validation on the training set. In particular, the training set can be further split into a training part and a validation part for each run of the external fivefold cross-validation. By varying the values of different parameters, the proposed model was developed using the samples in the training part. The results were obtained during validation. The parameters with the optimized average result during validation were selected. In the simulation studies, five parameters, i.e., λ_{f1} , λ_{f2} , λ_{v1} , λ_{v2} , and λ_{v3} , should be determined, and root mean square errors (RMSE) were used for parameter turning, i.e., the parameters with the lowest mean RMSE would be selected,

$RMSE = \sqrt{\frac{1}{5} \sum_{i=1}^5 \sum_{t=1}^T (\mathbf{Y}_{it} \mathbf{V}_{it} - \sum_{j=1}^p f_{ij})^2}$. Moreover, these parameters were in the range of $\{10^{-5}, 10^{-4}, \dots, 10^4, 10^5\}$. During the experiments, the search algorithm will stop when $\max_j |f_j^{iter+1} - f_j^{iter}| \leq \varepsilon$ and $\max_s |v_s^{iter+1} - v_s^{iter}| \leq \varepsilon$ are satisfied, where ε is the tolerable error and set to 10^{-5} empirically.

To fully evaluate the performance of T-GSRAM, we conducted five sets of experiments on simulated data. In the first set of experiments, we aim to assess the influence of different numbers of causal SNPs and different n values on T-GSRAM. We applied this method on data generated from Case 3. We initially fixed T , p , and n at 4, 200, and 100, respectively. We then varied the number of the causal SNPs over $A \in \{10, 20, 30\}$ to assess the effects of different A on the T-GSRAM. Second, we set T , p , and A to 4, 200, and 20, respectively. Then, we varied n from 100, 200, and 300 to 400 to evaluate the influence of different n on the T-GSRAM. In the second set of experiments, we aim to investigate the effectiveness of applying longitudinal QT data on T-GSRAM. We initially fixed T , p , A , and n at 4, 200, 20, and 100, respectively. Afterward, we generated BL and longitudinal data by using the additive model in (20). In the third set of experiments, we aim to assess the effectiveness of using the group information of SNPs in T-GSRAM. We initially fixed T , p , A , and n at 4, 200, 20, and 100 in T-GSRAM, respectively. Afterward, we generated \mathbf{Y} by using the additive model in (20). We then used the following strategy to generate SNP groups. First, the orders of all SNPs were randomly disrupted. Second, based on the disrupted order, the SNPs were grouped, and the group number and SNP number in each group were the same as the original settings mentioned above. After implementing this strategy, the generated SNP groups were inconsistent with those groups having prior knowledge (i.e., different from the original settings). In the fourth set of experiments, we aim to evaluate the effectiveness of using different regularizations in T-GSRAM. We initially fixed T , p , A , and n at 4, 200, 20, and 100 in T-GSRAM, respectively. Afterward, we generated simulation data by using the additive model in (20). We then removed one of the regularizations at each time to investigate the influence of different regularizations on T-GSRAM. In the fifth set of experiments, we initially fixed T , p , A , and n at 4, 200, 20, and 100 in T-GSRAM, respectively, and then compared the performance of this method with that of four state-of-art methods, namely, SpAM [(1)] [20], GroupSpAM [(2)] [29], TV-GroupSpAM [(3)] [8], and Temporal Mutil-task SCCA (T-MTSCCA) [9] on three datasets generated by (18)–(20), respectively. The first three methods have been introduced in the method section. Moreover, T-MTSCCA is bi-multivariate regression model, which imposes two types of regularization for SNP, i.e., one considers the group level structural information within the SNP data, and another one considers the sparsity among the SNP groups. Moreover, longitudinal data were used in T-MTSCCA. We also summarized the objective functions of all the comparison method in Table I. For fair comparison, we used the same procedures in the T-GSRAM to select the optimal parameters of all compared methods listed in Table I. Finally, we used the Receiver Operating Characteristic (ROC) curves and area under the curve (AUC) to evaluate the performance of all methods.

2) Simulation Results: In the first set of experiments, we initially tested different numbers of causal SNPs A from 10, 20, to 30 and varied our sample number n from 100 to 400 to evaluate the finite sample performance of T-GSRAM

TABLE I
OBJECTIVE FUNCTION OF ALL COMPARED METHODS

| Method | Objective functions |
|------------------|--|
| SpAM [20] | $\min_f \frac{1}{2} \left\ \mathbf{Y} - \sum_{j=1}^p f_j(\mathbf{x}_j) \right\ _2^2 + \alpha \sum_{j=1}^p f_j $ |
| GroupSpAM [29] | $\min_f \frac{1}{2} \left\ \mathbf{Y} - \sum_{k=1}^K \sum_{j=1}^p f_j(\mathbf{x}_j) \right\ _2^2 + \beta \sum_{k=1}^K \ f_k\ _2$ |
| TV-GroupSpAM [8] | $\min_f \frac{1}{2} \sum_{t=1}^T \left\ \mathbf{Y}_t - \sum_{j=1}^p f_j(t, \mathbf{x}_j) \right\ _2^2 + \lambda \ f\ _{21}$ |
| T-MTSCCA [9] | $\min_{\mathbf{U}, \mathbf{V}} \sum_{t=1}^T \left\ \mathbf{X}_t \mathbf{U}_t - \mathbf{Y}_t \mathbf{V}_t \right\ _2^2 + \lambda_{u1} \ \mathbf{U}\ _{G_{21}} + \lambda_{u2} \ \mathbf{U}\ _{21} + \lambda_{v3} \sum_{i=1}^p \ \mathbf{u}^i\ _1 + \lambda_{v1} \ \mathbf{V}\ _{FP_{21}} + \lambda_{v2} \ \mathbf{V}\ _{21} + \lambda_{v3} \sum_{j=1}^q \ \mathbf{v}^j\ _1$ |
| T-GSRAM | $\min_{f, \mathbf{V}} \frac{1}{2} \sum_{t=1}^T \left\ \mathbf{Y}_t \mathbf{V}_t - \sum_{k=1}^K \sum_{j=1}^p \sum_{g=0}^2 f_j^g(t) \mathbf{x}_j^g \right\ _2^2 + \lambda_{v1} \ \mathbf{V}\ _{FP_{21}} + \lambda_{v3} \sum_{s=1}^q \ \mathbf{v}_s\ _1 + \lambda_{v2} \ \mathbf{V}\ _{21} + \lambda_{f1} \ \mathbf{f}\ _{G_{21}} + \lambda_{f2} \ \mathbf{f}\ _{21}$ |

in detecting causal SNPs in Case 3. Figs. S1 (a) and (b) show the ROC curves corresponding to different A and n values, respectively. Moreover, the AUC values of different A are listed in Table SI. As predicted, a large A represents major genetic effects and leads to a high probability of detecting causal SNPs (Fig. S1 (a) and Table SI). Moreover, the true positive rate increases with increased sample size, as shown in Fig. S1 (b). As n increases, more potential knowledge of data distribution can be learnt by the model; thus, the detection AUC can be improved using large n .

In the second set of experiments, as shown in Fig. S2, using longitudinal data has a better performance than using BL data. Moreover, the detection AUCs of SNPs and QTs for BL data are 0.72 and 0.41, respectively, whereas those for longitudinal data are 0.81 and 1.0, respectively (the pair t -test p -values are 0.005 and 5.9×10^{-5} for SNP AUC and QT AUC, respectively). These results indicate that the longitudinal data used in the proposed T-GSRAM may provide more information than the BL data to improve biomarker detection performance.

In the third set of experiments, the T-GSRAM with the newly generated SNP groups obtained an AUC of 0.49, which is much lower than that obtained by T-GSRAM with the original SNP group setting (the pair t -test p -value is 0.0001 for SNP AUC), thereby suggesting that having prior knowledge of groups in SNP data is important to the T-GSRAM.

In the fourth set of experiment, Table II shows that the AUC values of SNP detection slightly decrease when one of the regularizations is removed. However, the AUC values of QT detection obviously fluctuate, especially when the FP_{21} and l_1 norms are removed.

In the fifth set of experiments, to evaluate its effectiveness, we compare T-GSRAM with four state-of-art methods in three cases. Fig. 2 and Table III show the ROC curves and mean AUC of these methods, respectively. In detecting causal SNPs, T-GSRAM outperforms the other methods (paired t -test

TABLE II
THE AUC OF REMOVED ONE OF THE REGULARIZATIONS
AT EACH TIME IN T-GSRAM

| Norm of SNPs | | Norm of QTs | | | AUC | |
|--------------|----------|-------------|-------|----------|-----------|-----------|
| G_{21} | l_{21} | FP_{21} | l_1 | l_{21} | SNPs | QTs |
| -- | √ | √ | √ | √ | 0.80±0.03 | 0.79±0.04 |
| √ | -- | √ | √ | √ | 0.79±0.03 | 0.85±0.01 |
| √ | √ | -- | √ | √ | 0.75±0.04 | 0.62±0.03 |
| √ | √ | √ | -- | √ | 0.75±0.05 | 0.75±0.02 |
| √ | √ | √ | √ | -- | 0.74±0.05 | 0.95±0.02 |
| √ | √ | √ | √ | √ | 0.81±0.01 | 1.00±0.00 |

p -value < 0.05), thereby highlighting its flexibility in analyzing the complex associations between SNPs and QTs. In Fig 2, the results of T-MTSCCA for Case 1 (i.e., linear model) are second only to those of T-GSRAM, but its results for Cases 2 and 3 are the worst. Among the four SpAM-based methods (i.e., SpAM, GroupSpAM, TV-GroupSpAM, and T-GSRAM), TV-GroupSpAM and T-GSRAM use longitudinal data and obtain higher AUC scores than the two other methods, which only use imaging data from a single time point. Compared with a linear method (e.g., T-MTSCCA), SpAM-based methods are more flexible in analyzing the nonlinear associations between SNPs and QTs. Therefore, SpAM-based methods outperform T-MTSCCA in Cases 2 and 3. SpAM-based methods also show consistent performance in the three cases, thereby highlighting their stability in dealing with complex associations between SNPs and QTs. SpAM, GroupSpAM, and TV-GroupSpAM use only a single QT in detecting causal QTs. Therefore, we did not detect any causal QT. Compared with T-MTSCCA, T-GSRAM has a higher QT AUC value, thereby highlighting its effectiveness in detecting potential biomarkers. We also calculated the computational time of these methods in Case 3 and present the results in Table SII. Given that only one constraint is used for SNPs in SpAM and TV-GroupSpAM, lower computational cost is required in the two methods compared with that in the other methods. Moreover, QTs and SNPs have more constraints in T-MTSCCA; thus, the highest computational cost is needed in this method. As shown in (2), sparsity constraint between SNP groups rather than within an SNP group is considered in GroupSpAM; smooth matrix on SNP groups is incorporated in this method. Therefore, highest computation cost is observed in GroupSpAM. The computational time of our approach is medial among all methods to balance computational cost and performance.

B. ADNI Analysis

1) Data Processing: In this article, the longitudinal brain imaging data were obtained from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and nonprofit organizations, as a \$60 million, 5 years public private partnership. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the

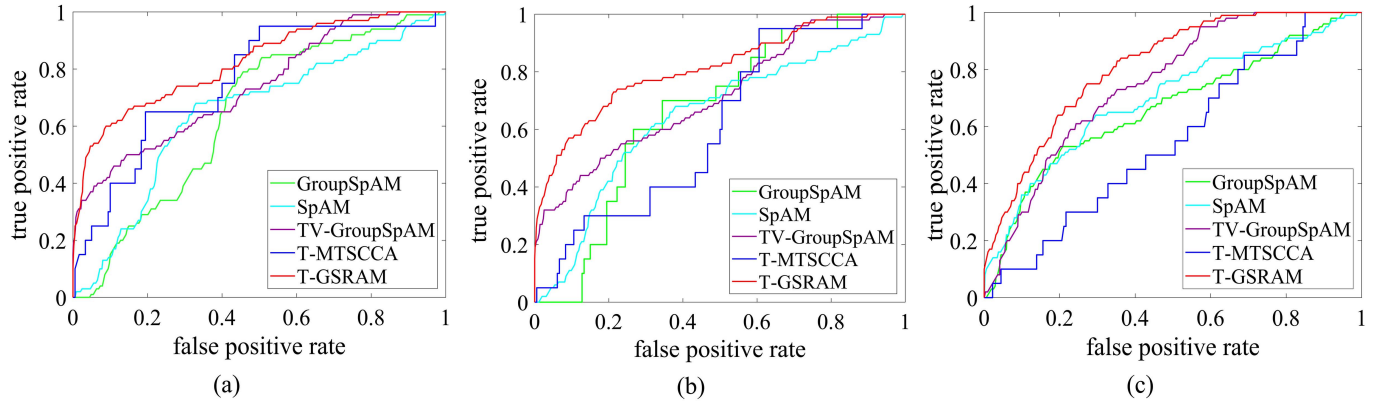


Fig. 2. ROC curves of the performance SNP detection of different methods (a) linear data generated by using (18), (b) nonlinear data generated by using (19), (c) additive model data generated by using (20).

TABLE III

AUC OF DIFFERENT METHODS AND THEIR p -VALUES OF PAIRED t -TESTS COMPARING THE DIFFERENT METHODS WITH T-GSRAM FOR SNPs. “—” IN THE TABLE INDICATES THAT ONLY SINGLE QT IS USED IN THE CORRESPONDING METHOD, NO CAUSAL QT IS DETECTED, AND NO SELF OF PAIRED t -TEST IN T-GSRAM

| AUC | Data 1 | | | Data 2 | | | Data 3 | | |
|--------------|-----------|-----------|-------------|-----------|-----------|-------------|-----------|-----------|----------------------|
| | SNP | QT | p -values | SNP | QT | p -values | SNP | QT | p -values |
| SpAM | 0.65±0.05 | -- | 0.017 | 0.65±0.06 | -- | 0.035 | 0.69±0.08 | -- | 0.032 |
| GroupSpAM | 0.64±0.03 | -- | 0.042 | 0.63±0.06 | -- | 0.024 | 0.66±0.04 | -- | 0.001 |
| TV-GroupSpAM | 0.74±0.03 | -- | 0.011 | 0.74±0.03 | -- | 0.017 | 0.75±0.04 | -- | 0.037 |
| T-MTSCCA | 0.80±0.05 | 0.67±0.02 | 0.010 | 0.61±0.04 | 0.58±0.07 | 0.009 | 0.55±0.07 | 0.40±0.06 | 7.7×10^{-4} |
| T-GSRAM | 0.82±0.13 | 0.96±0.03 | -- | 0.81±0.12 | 0.95±0.05 | -- | 0.81±0.01 | 1.00±0.00 | -- |

progression of MCI and early AD. For up to date information, see www.adni-info.org.

We considered the genotype variables of 818 subjects as acquired by using the Human 610-Quad BeadChip (Illumina, Inc., San Diego, CA, USA) in the ADNI database [36]. To reduce the population stratification effect, we selected 749 Caucasians from these subjects with complete imaging measurements at BL. Our quality control procedures include (i) call rate check per subject and SNP marker, (ii) gender check, (iii) sibling pair identification, (iv) the Hardy–Weinberg equilibrium test, (v) marker removal by the minor allele frequency, and (vi) population stratification. The second line preprocessing steps include removal of SNPs with (i) more than 5% missing values, (ii) minor allele frequencies of below 5%, and (iii) Hardy–Weinberg equilibrium p -values of $<10^{-6}$. The remaining missing genotype variables were imputed as the modal value. After implementing these procedures, only 708 of the 818 subjects remained, which was further reduced to 404 after removing those subjects without MRI images for all four time points.

Following the AlzGene database (www.alzgene.org, as of 4/18/2011), we selected two datasets, with the first dataset including 6283 SNPs extracted from 191 AD candidate genes (boundary: 20 KB) by using the ANNOVAR annotation (<http://www.openbioinformatics.org/annovar/>), and the second dataset including 1144 SNPs belonging to the top 36 AD genes after implementing the standard quality control and imputation procedures. We then combined these two datasets and obtained a total of 6452 SNPs on 202 genes as our SNP data. We used gene information to group our SNPs. The 202 genes and their corresponding number of SNPs are listed in Table SVI in the Supplementary Materials.

T1-weighted MRI images were used in this study. The scanning parameters of 1.5T MRI images can be found in a previous study [37]. A total of 404 subjects, including 140 NC, 181 MCI, and 83 AD subjects, provided by the ADNI 1 dataset were used. For each subject, the MRI scans at the BL, M6, M12, M24 were used in this study when available. The details of imaging data are listed in Table SIV. All MRI data were processed under the following steps to extract ROI-based features: (a) anterior commissure-posterior commissure correction by using MIPAV software (MIPAV in version 10.0, <https://mipav.cit.nih.gov/clickwrap.php>); (b) image intensity inhomogeneity correction by applying N3 algorithm [38]; (c) skull stripping [39] and warping a labeled template to each skull-stripped image for the removal of the cerebellum (aBEAT in version 1.0, <http://www.nitrc.org/projects/abeat/>); (d) tissue segmentation by using FAST method in FSL package to obtain four tissues: gray matter (GM), white matter, ventricle, and cerebrospinal fluid (FAST in FMRIB Software Library version 5.0, <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>); (e) registering all images to the Jacob template [40] by using the 4D-HAMMER method proposed in a previous study [41] (HAMMER in version 1.0, <https://www.nitrc.org/projects/hammer/>); (f) automatically labeling 93 ROIs on the template [42] and projecting ROI labels from the template image to each MRI image; (g) computing the gray matter tissue volume of each ROI in a labeled MRI image. In these processing steps, a feature vector of 93 gray matter tissue volumes was obtained as QT for each time point for subsequent analyses. Moreover, full and short names of 93 ROIs are provided in Table SV in the Supplementary Materials. In order to control for non-genetic effects, we incorporated several static covariates into our model, including intercept, gender, age, and whole brain volume.

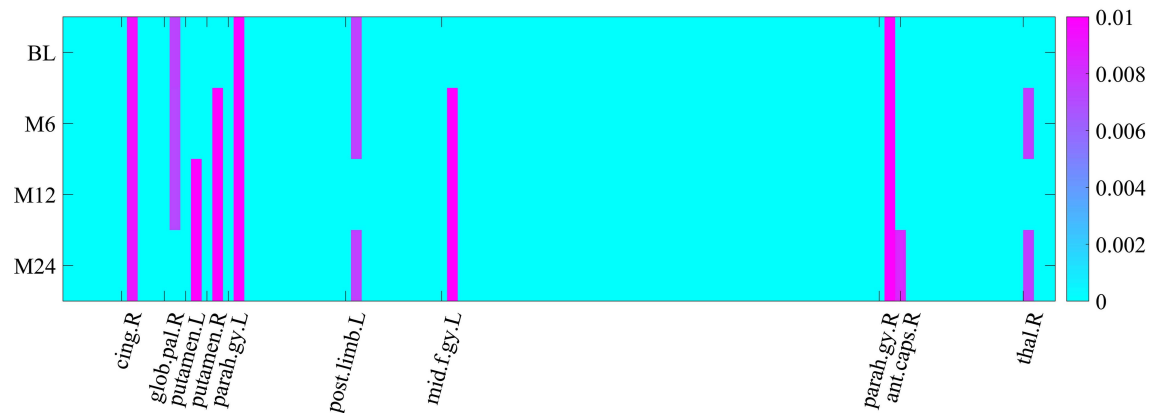


Fig. 3. Heatmaps of weights of top 10 imaging QTs selected by the proposed method. Four rows correspond to four time points of imaging QTs, i.e. BL, M6, M12, and M24.

2) Results on Real Data: In this section, our goal is to examine the genetic effect of each of the 6452 SNPs on the QTs and detect some potential biomarkers associated with AD. We are interested in identifying the brain regions (i.e., ROIs) that are most associated with the top selected SNPs in the T-GSRAM method. We want to know the dynamic patterns of imaging QTs longitudinally to identify the interior behavior of disease progression. In real data, our parameter turning strategy is consistent with that in the simulation data. We averaged each \mathbf{v}_t across fivefold to ensure stable selection. These averaged results at each time point (BL, M6, M12 and M24) of top 10 imaging QTs with the highest absolute weights \mathbf{V} are shown in Fig. 3. We also visualized the top 10 imaging QTs in Fig. 4. As shown in Fig. 3, the weights are sparse at different time points and QTs and can be used to represent the variations in the corresponding imaging QTs during the disease progression trajectory. The degeneration of the putamen leads to a cognitive decline in AD [43]. The remarkable metabolism reduction in the cingulate region at the right hemisphere is associated with memory impairment, which is a feature of early AD [44]. The parahippocampal gyrus is associated with cognitive impairment in AD [45], its volume atrophy is associated with healthy aging and different stages of AD, and its volume sizes significantly differ across groups in the following order: healthy > amnesic MCI > AD [46]. The middle frontal gyrus can affect metabolism and lead to metabolic disorders and cognitive function decline in AD [47]. The high degree of neuroinflammation of the posterior limb of the internal capsule is associated with some cognitive deficits in early AD [48]. The posterior limb of the internal capsule is associated with early AD [48]. The thalamus causes episodic memory loss, which occurs during the early stages of AD [49]. Moreover, the volume of thalamus decreases during the progression of MCI to AD [50]. In sum, T-GSRAM can help uncover those changes that take place during different periods of MCI and AD progression.

We also identified the relevant top SNPs in T-GSRAM study. Table SIII shows the weights of the selected top 20 SNPs. These SNPs are obtained from 10 genes, namely, the APOE gene (e.g., rs429358 and rs405509), CLU gene (e.g., rs2565050, rs2640727, rs4149246, and rs4149247),

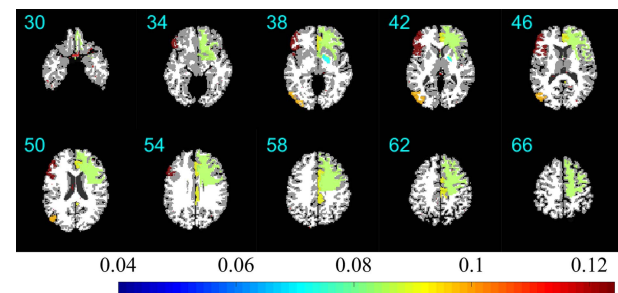


Fig. 4. Top 10 imaging QTs from MRI data selected by T-GSRAM.

CD33 gene (e.g., rs8111536 and rs10424255), CR1 gene (e.g., rs2294938, rs6684558, rs1120599, rs12138784, rs10746420, and rs11811428), ABCA7 gene (e.g., rs10416031), MS4A6A gene (e.g., rs4939364), CD2AP gene (e.g., rs17217010), BCAM gene (e.g., rs10402038), CELF2 (e.g., rs1324317), and PICALM gene (e.g., rs7931249). These selected SNPs show an increased risk of AD or MCI progression in previous studies. For instance, APOE is identified in many AD-related studies as one of the top 10 associated genes, indicating a strong association between APOE and QTs in AD. Previous evidence shows that APOE is highly correlated with the atrophy patterns of the parahippocampal gyrus and the medial temporal regions in AD [6], [51], [52]. CLU gene expression shows a white matter integrity that may increase the risk of AD [53]. CD33 inhibits the microglial uptake of amyloid beta, thereby greatly accelerating cognitive decline and leading to early onset dementia [54]. CR1 is associated with rapid cognitive decline in AD and an aging population [55]. Cognitive and memory decline, which is strongly associated with AD, is affected by methylation at several CpG sites in the ABCA7 gene [56]. Soluble trigger receptors expressed in myeloid cells 2 in cerebrospinal fluid will affect the age of AD onset and can be influenced by the NS4A6A gene [57]. CD2AP is reported to be associated with late onset AD [58]. PICALM may play an important role in AD pathology by participating in altering synaptic vesicle cycling or APP endocytosis [59]. In sum, the SNP selection results indicate that T-GSRAM can identify effective SNPs.

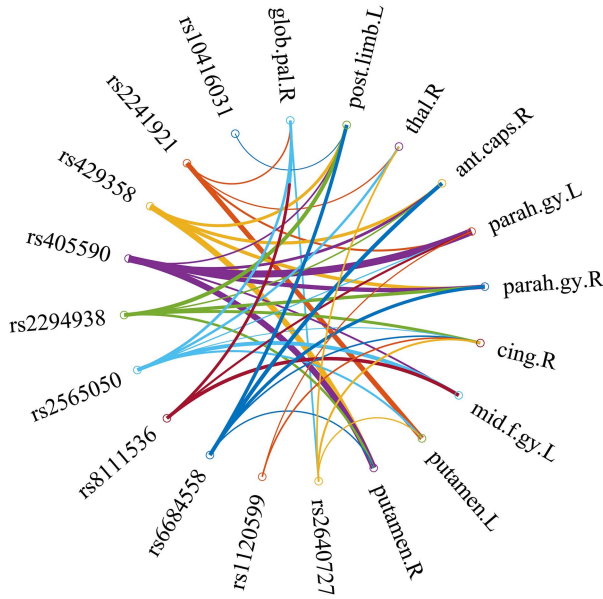


Fig. 5. Associations between top 10 SNPs and top 10 QTs, where nodes denote SNPs or QTs and edges denote association strengths between SNPs and QTs.

TABLE IV

RMSE OF DIFFERENT METHODS ON REAL DATA. “–” IN THE TABLE INDICATES THAT A NON- p VALUE IS OBTAINED

| Method | SpAM | Group SpAM | TV-Group SpAM | T-MTSCCA | T-GSRAM |
|------------|-----------|------------|---------------|-----------|-----------|
| RMSE | 1.33±0.18 | 1.33±0.21 | 1.14±0.01 | 0.65±0.04 | 0.15±0.01 |
| p -value | 0.0292 | 0.0029 | 0.0026 | 0.0387 | -- |

To illustrate the complex associations between SNPs and QTs, we first averaged the weights of SNPs and QTs. We then selected the top 10 SNPs and QTs with the largest averaged weights. Finally, we multiplied the weights of the top 10 SNPs by the weights of the top 10 QTs to obtain the associations between the top 10 SNPs and QTs, and the results are shown in Fig. 5. Both rs429358 and rs405590 (APOE) have strong associations with putamen and parahippocampal gyrus. Hashimoto *et al.* [60] found a significant interaction between the ApoE e4 genotype and progressive morphological changes (a more rapid reduction of volume is observed in ApoE e4 carriers than in noncarriers) in the parahippocampal gyrus. The remaining SNPs and QTs have been verified to be associated with AD [44], [52], [61]–[63], and their associations are worthy for further investigation and verification.

To evaluate its effectiveness, we compared the performance of T-GSRAM on real data with that of four other state-of-art methods. Given that the biomarkers associated with AD in QTs and SNPs for real data have no ground truth, we used RMSE in the performance evaluation. The results are shown in Table IV. From this table, T-GSRAM obtains the lowest RMSE, while the RMSE obtained by T-MTSCCA is lower than that obtained by the other SpAM-based methods (except for T-GSRAM).

IV. DISCUSSION

In this article, we propose a novel T-GSRAM method to analyze the associations between SNPs and longitudinal QTs to uncover the genetic basis of the brain structure, function,

and disorder associated with AD. In the proposed T-GSRAM method, regression analysis was performed to select relevant QTs and SNPs in AD. Different from traditional regression methods that ignored useful information embedded in the longitudinal imaging data across multiple time points, the T-GSRAM incorporated longitudinal imaging data and applied three constraints on the longitudinal imaging data for better discovering potential information in the data. First, the intermediate temporally stable pattern of one QT as well as the sparsity among all QTs are considered by applying a FP_{21} norm on longitudinal QTs. Second, an l_1 norm is applied to select effective QTs at a specific time point. Third, an l_{21} norm is used to select effective QTs at all time points to detect regional variations in different brain regions. For the SNP data, a G_{21} norm is applied to discovery group and individual information in SNP groups. Moreover, an l_{21} norm is used in our model to jointly select SNPs relevant to important imaging QTs. Moreover, we assume that the effects of each SNP on QT are regarded as a smooth function of time, and SNP effects vary smoothly over time. Therefore, compared with a linear method, the proposed T-GSRAM method is more flexible for analysis of the complex associations between SNPs and QTs.

A. Effect of Parameters

To evaluate the effects of five parameters (i.e., λ_{f1} , λ_{f2} , λ_{v1} , λ_{v2} , and λ_{v3}) on the performance of T-GSRAM, we initially changed their values within the range of $\{10^{-5}, 10^{-4}, \dots, 10^4, 10^5\}$, and then turned the optimal parameters through grid search. However, we only varied one parameter and fixed the other parameters at their optimal values to display the mean AUC of SNP detection achieved by T-GSRAM in the three data-generated cases. The results are shown in Fig. S3. The AUC slightly fluctuates within a small range along with an increasing λ_{f1} , λ_{f2} , λ_{v1} , λ_{v2} , and λ_{v3} , thereby suggesting that the AUC values obtained by T-GSRAM are generally stable with respect to these five parameters. Following the same parameter turning strategy as that used in T-GSRAM, the AUC values of SNP detection obtained by all the other methods with varying parameters can be obtained and shown in Figs. S4–S7. Fig. S4 shows that the AUC values obtained by T-MTSCCA greatly fluctuate within a small range along with an increasing λ_{u1} , λ_{u2} , and λ_{u3} . Compared with those of T-MTSCCA, the AUC values obtained by SpAM-based methods with different parameters are more stable (Figs. S3–S7), thereby indicating that these methods may be less influenced by different parameters.

B. Effect of Using Longitudinal QT Data

The ROC curves of T-GSRAM obtained by using longitudinal and BL data are shown in Fig. S2. Using longitudinal data shows a better performance than using BL data, thereby highlighting the importance of using longitudinal QT data for T-GSRAM. Table II shows that when longitudinal QT-related regularizations are removed, the performance of T-GSRAM remarkably decreases. For example, the FP_{21} norm is used to select effective QTs at two adjacent time points to investigate regional variations at different time points, whereas the l_1 norm is applied to select effective QTs at a specific time point. These results indicate that the variations across brain regions and the

disease progression trajectory in specific brain regions across multiple time points are important and should be considered when analyzing the associations between longitudinal QT and SNP data.

C. Effect of Group Information of SNP

The SNP information at the group and individual levels is considered in T-GSRAM. The group level information is based on genes or LD blocks (constrained by G_{21}), whereas the individual level information is used to consider the different and combinational effects of three genotypes in an SNP (constrained by l_{21}). When one of these two regularizations are removed, the AUC values of SNP detection slightly decrease, whereas the AUC values of QT detection remarkably decrease (Table II). These results highlight the importance of investigating the joint effects of SNPs in a group and considering the different and combined effects in a single SNP to accurately model the complex influence of SNPs on QTs. Moreover, when the group information is inconsistent with prior knowledge (i.e., different from the original settings), the AUC value of SNP detection is only 0.49, which suggests that having prior knowledge of groups in SNP data is important to T-GSRAM.

D. Comparison With Previous Studies

We compared T-GSRAM with four state-of-the-art methods on three sets of simulated data and a real ADNI 1 dataset. For both causal QT and SNP detection, the detection AUC value of T-GSRAM is higher than that of the other methods in three simulated datasets (linear or nonlinear model), thereby indicating that T-GSRAM is more flexible than the other methods in analyzing the complex associations between SNPs and QTs. Fig. 2 and Table III show that T-MTSCCA has a better performance in the linear model than in the nonlinear model (the AUC values are 0.8, 0.61, and 0.55 in cases 1, 2, and 3, respectively). Moreover, the AUC value of T-MTSCCA is higher than that of SpAM-based methods (except for T-GSRAM) in the linear model (Case 1). However, SpAM-based methods obtain a higher AUC than the linear model (e.g., T-MTSCCA) by using the smooth function in Cases 2 and 3 (as shown in Figs. S1 (b) and (c)). Therefore, T-MTSCCA is more suitable in the linear case than in the nonlinear case, whereas the proposed T-GSRAM is more flexible than T-MTSCCA and can be used in both linear and nonlinear cases. As shown in Fig. 2 and Table III, T-GSRAM and TV-GroupSpAM outperform both SpAM and GroupSpAM, thereby highlighting the effectiveness of using longitudinal imaging data. TV-GroupSpAM uses only a single QT and ignores the group information of SNPs. By contrast, T-GSRAM incorporates the group information of SNPs and the different effects for three genotypes of single SNPs. Moreover, T-GSRAM considers multiple QTs and QT variations over different time points. Fig. 2 and Table III show that T-GSRAM outperforms TV-GroupSpAM, thereby indicating that the variable information among QTs and the group information of SNPs are useful in biomarker detection. From Table SII, highest computation cost is observed in GroupSpAM. In addition to the T-MTSCCA method, smooth matrix is introduced in the solution of the SpAM-based method, SpAM and GroupSpAM

TABLE V
RMSE OF DIFFERENT METHODS ON REAL DATA PROCESSED BY THE FREESURFER PIPELINE. “—” IN THE TABLE INDICATES THAT A NON- p VALUE IS OBTAINED

| Method | SpAM | Group SpAM | TV-Group SpAM | T-MTSCCA | T-GSRAM |
|------------|-----------|------------|---------------|-----------|-----------|
| RMSE | 1.31±0.05 | 1.67±0.15 | 1.12±0.07 | 0.56±0.10 | 0.43±0.31 |
| p -value | 0.0021 | 0.0013 | 0.0054 | 0.0041 | -- |

introduce smooth matrix for SNPs, whereas TV-GroupSpAM and our method introduce smooth matrix for time. Compared with the smooth matrix for SNPs, the smooth matrix for time is smaller and the amount of calculation is smaller. Moreover, the smooth matrix used in GroupSpAM is for a group of SNPs, which requires more calculation.

For the real ADNI dataset, T-GSRAM obtains the lowest RMSE value, followed by T-MTSCCA. For these two methods, three constraints are included in the longitudinal QT data (i.e., FP_{21} , l_1 , and l_{21}) to capture the trajectory of disease progression represented by QTs, which may partially account for their low RMSE values. In contrast to T-MTSCCA, T-GSRAM uses individual level information to take into consideration the different and combinational effects of three genotypes in an SNP, which may be conducive to further reducing the RMSE value on the ADNI data in this study.

Publicly available ROI data processed by the FreeSurfer pipeline, which can be downloaded from the ADNI website (<http://adni.loni.usc.edu/updated-cross-sectional-freesurfer-5-1-data-available-8/>), were used to investigate the influence of different data preprocessing methods on T-GSRAM. Among these data, we identified 761 subjects with MRI images at BL. After implementing the quality control procedures for SNP, 665 of these 761 subjects remained, which was further reduced to 437 after removing those subjects without MRI images for all four time points. Given that the biomarkers associated with AD in QTs and SNPs for real data have no ground truth, we used RMSE to evaluate the performance of different methods on publicly available data. Table V shows that the RMSE value of T-GSRAM is lower than that obtained by the other methods. Similar results are obtained from the ROI data processed by our pipeline (Table IV). The top 10 SNPs and QTs selected by using the FreeSurfer processed data are listed in Table SVII. The top 10 SNPs were obtained from six genes, namely, APOE, PICLAM, CR1, MS4A6A, CLU, and CD33, which were the same genes selected by using our processed data (Table SIII). Among the top 10 QTs, the hippocampus, cerebral cortex, lateral ventricle, insula, supramarginal, and superior frontal are closely related to AD [45], [64]–[68]. These results highlight the effectiveness of T-GSRAM on data subjected to different processing procedures.

E. Biomarker Detection

For the real imaging genetic data, the famous APOE (e.g., rs429358, rs405509) is detected using the T-GSRAM. This result may be attributed to the group information of SNPs considered in the proposed method. Thus, the adjacent SNPs, which are associated with AD, can be detected by

the T-GSRAM. Moreover, our method detects brain regions related to AD, such as hippocampal and fornix. Fig. 3 shows that the sparsity of the QTs at each time point and the sparsity of a QT at all time points. As shown in Fig. 3, the weights of the parahippocampal gyrus at all time points are consistently high probably due to the usage of the l_{21} norm in selecting effective QTs at all time points. Moreover, the other selected associated SNPs and ROIs are consistent with the findings in the literature, thereby confirming the effectiveness of our proposed method. Therefore, our proposed method can identify potential AD-related biomarkers for early prediction and diagnosis of AD and possibly can be extended to other neuropsychological diseases.

F. Limitations and Future Work

Several issues should be addressed in our future research. First, we only used small samples. The small sample size may lead to the overfitting problem for various penalized regression methods. Therefore, more samples should be used in our experiments in future. Second, we ignore the functional connectivity information among brain regions. However, the brain network connections of different individuals reflect the comprehensive characteristics of different brain systems. Therefore, we will consider the association between brain regions in future studies. Third, our method only analyzed one modal imaging information. However, various characteristics of image phenotypes can be obtained from different neuroimaging modalities (e.g., functional MRI, structure MRI, and diffusion tensor imaging). In the future, we will apply multi-modal data combined with our research method to obtain better results. Moreover, non-genetic effects, including intercept, gender, age, and whole brain volume, are considered in this study. However, other effects, such as multiple sites and scanner vendor, are ignored in this study and should be incorporated into methods proposed in future work to construct a highly accurate association model.

We only used a specific atlas, which may limit the improvements in biomarker detection performance because QTs are extracted based on this atlas. However, we used another publicly available dataset to assess the performance of T-GSRAM. The lowest RMSE value is obtained by T-GSRAM, thereby highlighting its flexibility in this dataset. Future studies may consider using some other widely used atlases for brain regions extraction [69], such as the automatic anatomical labeling atlas [42] and Craddock's spatially constrained spectral clustering atlas [70], to further investigate the influence of different atlases on T-GSRAM.

In this study, some AD-related QT and SNP biomarkers are detected by T-GSRAM. For example, the volume atrophy of parahippocampal gyrus has a high risk of progressing from MCI to AD [50], and may be used as a potential biomarker for AD prediction. Moreover, to evaluate the effectiveness of the detected biomarkers in AD prediction, we can use these biomarkers as features to train a classifier for classifying MCI subjects into MCI conversion (MCI subjects who will convert to AD in 36 months) or MCI non-conversion (MCI subjects who will not convert to AD in 36 months).

REFERENCES

- [1] 2020 Alzheimer's Disease Facts and Figures, Alzheimer's Association, Chicago, IL, USA, Mar. 2020.
- [2] H. Chui, "Alzheimer disease," *Alzheimer Disease Associated Disorders*, vol. 10, no. 1, p. 53, 1996.
- [3] A. S. Disease. (2019). *Alzheimer's Disease International: The World Alzheimer Report 2019: Attitudes to Dementia*. [Online]. Available: <https://www.alz.co.uk/research/WorldAlzheimerReport2019.pdf>
- [4] J. C. Lambert *et al.*, "Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease," *Nature Genet.*, vol. 45, no. 12, pp. 1452–1458, Dec. 2013.
- [5] M. Huang, Y. Yu, W. Yang, and Q. Feng, "Incorporating spatial-anatomical similarity into the VGWAS framework for AD biomarker detection," *Bioinformatics*, vol. 35, no. 24, pp. 5271–5280, Dec. 2019.
- [6] X. Wang *et al.*, "Longitudinal genotype-phenotype association study via temporal structure auto-learning predictive model," *Res Comput Mol Biol*, vol. 10229, pp. 287–302, May 2017.
- [7] A. R. Hariri, E. M. Drabant, and D. R. Weinberger, "Imaging genetics: Perspectives from studies of genetically driven variation in serotonin function and corticolimbic affective processing," *Biol Psychiatry*, vol. 59, no. 10, pp. 888–897, May 15 2006.
- [8] M. Marchetti-Bowick, J. Yin, J. A. Howrylak, and E. P. Xing, "A time-varying group sparse additive model for genome-wide association studies of dynamic complex traits," *Bioinformatics*, vol. 32, no. 19, pp. 2903–2910, Oct. 2016.
- [9] L. Du *et al.*, "Identifying progressive imaging genetic patterns via multi-task sparse canonical correlation analysis: A longitudinal study of the ADNI cohort," *Bioinformatics*, vol. 35, no. 14, pp. i474–i483, Jul. 2019.
- [10] A. J. Saykin *et al.*, "Genetic studies of quantitative MCI and AD phenotypes in ADNI: Progress, opportunities, and plans," *Alzheimer's Dementia*, vol. 11, no. 7, pp. 792–814, Jul. 2015.
- [11] M. Huang, C. Deng, Y. Yu, T. Lian, W. Yang, and Q. Feng, "Spatial correlations exploitation based on nonlocal voxel-wise GWAS for biomarker detection of AD," *NeuroImage, Clin.*, vol. 21, 2019, Art. no. 101642.
- [12] M. Huang *et al.*, "FVGWAS: Fast voxelwise genome wide association analysis of large-scale imaging genetic data," *NeuroImage*, vol. 118, pp. 613–627, Sep. 2015.
- [13] H. Wang *et al.*, "Identifying quantitative trait loci via group-sparse multitask regression and feature selection: An imaging genetics study of the ADNI cohort," *Bioinformatics*, vol. 28, no. 2, pp. 229–237, Jan. 15 2012.
- [14] T. Zhou, K.-H. Thung, M. Liu, and D. Shen, "Brain-wide genome-wide association study for Alzheimer's disease via joint projection learning and sparse regression model," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 1, pp. 165–175, Jan. 2019.
- [15] O. Kohannim *et al.*, "Predicting temporal lobe volume on MRI from genotypes using L^1 - L^2 regularized regression," in *Proc. 9th IEEE Int. Symp. Biomed. Imag. (ISBI)*, May 2012, pp. 1160–1163.
- [16] M. Silver, G. Montana, and A. Disease Neuroimaging In, "Fast identification of biological pathways associated with a quantitative trait using group lasso with overlaps," *Stat. Appl. Genet. Mol. Biol.*, vol. 11, no. 1, pp. 1–43, Jan. 2012.
- [17] E. C. Chi, G. I. Allen, H. Zhou, O. Kohannim, K. Lange, and P. M. Thompson, "Imaging genetics via sparse canonical correlation analysis," in *Proc. IEEE Int. Symp. Biomed. Imag.*, Dec. 2013, pp. 740–743.
- [18] X. Zhu, H. I. Suk, H. Huang, and D. Shen, "Structured sparse low-rank regression model for brain-wide and genome-wide associations," *Med. Image Comput. Comput. Assist. Interv.*, vol. 9900, pp. 344–352, Oct. 2016.
- [19] D. M. Witten and R. J. Tibshirani, "Extensions of sparse canonical correlation analysis with applications to genomic data," *Stat. Appl. Genet. Mol. Biol.*, vol. 8, no. 1, pp. 1–27, Jan. 2009.
- [20] P. Ravikumar, J. Lafferty, H. Liu, and L. Wasserman, "Sparse additive models," *J. Roy. Stat. Soc., B Stat. Methodol.*, vol. 71, no. 5, pp. 1009–1030, 2009.
- [21] X. Q. Wang *et al.*, "Quantitative trait loci identification for brain endophenotypes via new additive model with random networks," *Bioinformatics*, vol. 34, no. 17, pp. 866–874, Sep 1 2018.
- [22] B. Cheng, M. Liu, D. Zhang, and D. Shen, "Robust multi-label transfer feature learning for early diagnosis of Alzheimer's disease," *Brain Imag. Behav.*, vol. 13, no. 1, pp. 138–153, Feb. 2019.
- [23] D. H. Ballard, J. Cho, and H. Zhao, "Comparisons of multi-marker association methods to detect association between a candidate region and disease," *Genet Epidemiol.*, vol. 34, no. 3, pp. 201–212, Apr. 2010.

- [24] M. Silver *et al.*, "Pathways-driven sparse regression identifies pathways and genes associated with high-density lipoprotein cholesterol in two asian cohorts," *PLoS Genet.*, vol. 9, no. 11, Nov. 2013, Art. no. e1003939.
- [25] L. Harper *et al.*, "Patterns of atrophy in pathologically confirmed dementias: A voxelwise analysis," *J Neurol. Neurosurg. Psychiatry*, vol. 88, no. 11, pp. 908–916, Nov. 2017.
- [26] J. C. Barrett, B. Fry, J. Maller, and M. J. Daly, "Haploview: Analysis and visualization of LD and haplotype maps," *Bioinformatics*, vol. 21, no. 2, pp. 263–265, Jan. 2005.
- [27] J. K. Pritchard and M. Przeworski, "Linkage disequilibrium in humans: Models and data," *Amer. J. Human Genet.*, vol. 69, no. 1, pp. 1–14, Jul. 2001.
- [28] S. G. Mueller *et al.*, "The Alzheimer's disease neuroimaging initiative," *Neuroimag. Clin.*, vol. 15, no. 4, pp. 869–877, Nov. 2005.
- [29] J. Yin, X. Chen, and E. P. Xing, "Group sparse additive models," in *Proc. Int. Conf. Mach. Learn.*, 2012, pp. 871–878.
- [30] N. C. Fox and J. M. Schott, "Imaging cerebral atrophy: Normal ageing to Alzheimer's disease," *Lancet*, vol. 363, no. 9406, pp. 392–394, Jan. 2004.
- [31] L. Du *et al.*, "Identifying associations between brain imaging phenotypes and genetic factors via a novel structured SCCA approach," *Inf. Process. Med. Imag.*, vol. 10265, pp. 543–555, Jun. 2017.
- [32] S. Lee, J. Zhu, and E. P. Xing, "Adaptive multi-task lasso: With application to eQTL detection," in *Proc. Adv. Neural Inf. Process. Syst.*, Dec. 2010, pp. 1306–1314.
- [33] K. Puniyani, S. Kim, and E. P. Xing, "Multi-population GWA mapping via multi-task regularized regression," *Bioinformatics*, vol. 26, no. 12, pp. i208–i216, Jun. 2010.
- [34] J. Gorski, F. Pfeuffer, and K. Klamroth, "Biconvex sets and optimization with biconvex functions: A survey and extensions," *Math. Methods Oper. Res.*, vol. 66, no. 3, pp. 373–407, Nov. 2007.
- [35] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *J. Optim. Theory Appl.*, vol. 109, no. 3, pp. 475–494, Jun. 2001.
- [36] A. J. Saykin *et al.*, "Alzheimer's Disease neuroimaging initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans," *Alzheimers Dement*, vol. 6, no. 3, pp. 265–273, May 2010.
- [37] C. R. Jack *et al.*, "The Alzheimer's Disease neuroimaging initiative (ADNI): MRI methods," *J. Magn. Reson. Imag.*, vol. 27, no. 4, pp. 685–691, Apr. 2008.
- [38] J. G. Sled, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Trans. Med. Imag.*, vol. 17, no. 1, pp. 87–97, Feb. 1998.
- [39] Y. Wang *et al.*, "Knowledge-guided robust MRI brain extraction for diverse large-scale neuroimaging studies on humans and non-human primates," *PLoS ONE*, vol. 9, no. 1, Jan. 2014, Art. no. e77810.
- [40] N. J. Kabani, D. J. MacDonald, C. J. Holmes, and A. C. Evans, "3D anatomical atlas of the human brain," *NeuroImage*, vol. 7, no. 4, p. S717, May 1998.
- [41] D. Shen and C. Davatzikos, "Measuring temporal morphological changes robustly in brain MR images via 4-dimensional template warping," *NeuroImage*, vol. 21, no. 4, pp. 1508–1517, Apr. 2004.
- [42] N. Tzourio-Mazoyer *et al.*, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *NeuroImage*, vol. 15, no. 1, pp. 273–289, Jan. 2002.
- [43] L. W. de Jong *et al.*, "Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: An MRI study," *Brain*, vol. 131, no. 12, pp. 3277–3285, Dec. 2008.
- [44] S. Minoshima, B. Giordani, S. Berent, K. A. Frey, N. L. Foster, and D. E. Kuhl, "Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease," *Ann. Neurol.*, vol. 42, no. 1, pp. 85–94, Jul. 1997.
- [45] M. J. De Leon *et al.*, "Frequency of hippocampal formation atrophy in normal aging and Alzheimer's disease," *Neurobiol. Aging*, vol. 18, no. 1, pp. 1–11, Jan. 1997.
- [46] C. Echavarri *et al.*, "Atrophy in the parahippocampal gyrus as an early biomarker of Alzheimer's disease," *Brain Struct. Function*, vol. 215, nos. 3–4, pp. 265–271, Jan. 2011.
- [47] S. G. Snowden *et al.*, "Association between fatty acid metabolism in the brain and alzheimer disease neuropathology and cognitive performance: A nontargeted metabolomic study," *PLOS Med.*, vol. 14, no. 3, Mar. 2017, Art. no. e1002266.
- [48] I. Suridjan *et al.*, "In-vivo imaging of grey and white matter neuroinflammation in Alzheimer's disease: A positron emission tomography study with a novel radioligand, [18 F]-FEPPA," *Mol. Psychiatry*, vol. 20, no. 12, pp. 1579–1587, 2015.
- [49] J. P. Aggleton, A. Pralus, A. J. Nelson, and M. Hornberger, "Thalamic pathology and memory loss in early Alzheimer's disease: Moving the focus from the medial temporal lobe to Papez circuit," *Brain*, vol. 139, no. Pt 7, pp. 1877–1890, Jul. 2016.
- [50] H. A. Yi *et al.*, "Relation between subcortical grey matter atrophy and conversion from mild cognitive impairment to Alzheimer's disease," *J. Neurol. Neurosurg. Psychiatry*, vol. 87, no. 4, pp. 425–432, Apr. 2016.
- [51] J. B. Pereira *et al.*, "Influence of age, disease onset and ApoE4 on visual medial temporal lobe atrophy cut-offs," *J. Internal Med.*, vol. 275, no. 3, pp. 317–330, Mar. 2014.
- [52] S. Risacher, "The role of apolipoprotein e (APOE) genotype in early mild cognitive impairment (E-MCI)," *Frontiers Aging Neurosci.*, vol. 5, p. 11, 2013.
- [53] M. N. Braskie *et al.*, "Common Alzheimer's disease risk variant within the CLU gene affects white matter microstructure in young adults," *J. Neurosci.*, vol. 31, no. 18, pp. 6764–6770, May 4 2011.
- [54] A. Griciuc *et al.*, "Alzheimer's disease risk gene CD33 inhibits microglial uptake of amyloid beta," *Neuron*, vol. 78, no. 4, pp. 631–643, May 22 2013.
- [55] L.-N. Hazrati, C. Van Cauwenberghe, P. L. Brooks, N. Brouwers, M. Ghani, C. Sato, M. Cruts, K. Sleegers, P. S. George-Hyslop, C. Van Broeckhoven, and E. Rogaeve, "Genetic association of CR1 with Alzheimer's disease: A tentative disease mechanism," *Neurobiol. Aging*, vol. 33, pp. 2949.e5–2949.e12, Dec. 2012.
- [56] A. De Roeck, C. Van Broeckhoven, and K. Sleegers, "The role of ABCA7 in Alzheimer's disease: Evidence from genomics, transcriptomics and methylomics," *Acta Neuropathologica*, vol. 138, no. 2, pp. 201–220, Aug. 2019.
- [57] C. Cruchaga, B. Benitez, C. M. Karch, and L. Piccio, "O3-13-06: The ms4a gene cluster is a key modulator of soluble trem2 and alzheimer disease risk," *Alzheimer's Dementia*, vol. 15, p. P922, Jul. 2019.
- [58] S. L. Rosenthal and M. I. Kamboh, "Late-onset Alzheimer's disease genes and the potentially implicated pathways," *Current Genetic Med. Rep.*, vol. 2, no. 2, pp. 85–101, Jun. 2014.
- [59] J.-T. Yu *et al.*, "Genetic association of PICALM polymorphisms with Alzheimer's Disease in han Chinese," *J. Neurol. Sci.*, vol. 300, nos. 1–2, pp. 78–80, Jan. 2011.
- [60] R. Hashimoto *et al.*, "Effect of the brain-derived neurotrophic factor and the apolipoprotein e polymorphisms on disease progression in preclinical Alzheimer's disease," *Genes, Brain Behav.*, vol. 8, no. 1, pp. 43–52, Feb. 2009.
- [61] X. Li, H. Wang, Y. Yu, Y. Tian, S. Zhou, and L. Xu, "Emotional memory related grey matter volume changes of Alzheimer's disease," *Chin. J. Med. Imag. Technol.*, vol. 32, pp. 1015–1019, Jul. 2016.
- [62] C. T. Watson *et al.*, "Genome-wide DNA methylation profiling in the superior temporal gyrus reveals epigenetic signatures associated with Alzheimer's disease," *Genome Med.*, vol. 8, no. 1, p. 5, Jan. 2016.
- [63] R. C. Cumming and D. Schubert, "Amyloid-beta induces disulfide bonding and aggregation of GAPDH in Alzheimer's disease," *FASEB J.*, vol. 19, no. 14, pp. 2060–2062, Dec. 2005.
- [64] R. N. Kalaria and S. I. Harik, "Reduced glucose transporter at the blood-brain barrier and in cerebral cortex in alzheimer disease," *J. Neurochem.*, vol. 53, no. 4, pp. 1083–1088, Oct. 1989.
- [65] T. Ertekin *et al.*, "Total intracranial and lateral ventricle volumes measurement in Alzheimer's disease: A methodological study," *J. Clin. Neurosci.*, vol. 34, pp. 133–139, Dec. 2016.
- [66] A. L. Foundas, C. M. Leonard, S. M. Mahoney, O. F. Agee, and K. M. Heilman, "Atrophy of the hippocampus, parietal cortex, and insula in Alzheimer's disease: A volumetric magnetic resonance imaging study," *Neuropsychiatry Neuropsychol. Behav. Neurol.*, vol. 10, no. 2, pp. 81–89, Apr. 1997.
- [67] M. D. Ikonomovic, E. E. Abrahamson, B. A. Isanski, J. Wu, E. J. Mufson, and S. T. J. A. O. N. DeKosky, "Superior frontal cortex cholinergic axon density in mild cognitive impairment and early Alzheimer disease," *Arch. Neurol.*, vol. 64, no. 9, pp. 1312–1317, 2007.
- [68] J. R. Gatchel *et al.*, "Regional 18F-fluorodeoxyglucose hypometabolism is associated with higher apathy scores over time in early alzheimer disease," *Amer. J. Geriatric Psychiatry*, vol. 25, no. 7, pp. 683–693, Jul. 2017.
- [69] B. Lei *et al.*, "Adaptive sparse learning using multi-template for neurodegenerative disease diagnosis," *Med. Image Anal.*, vol. 61, Apr. 2020, Art. no. 101632.
- [70] R. C. Craddock, G. A. James, P. E. Holtzheimer, X. P. Hu, and H. S. Mayberg, "A whole brain fMRI atlas generated via spatially constrained spectral clustering," *Hum. Brain Mapping*, vol. 33, no. 8, pp. 1914–1928, Aug. 2012.