Medical Image Analysis 70 (2021) 102003

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Identifying associations among genomic, proteomic and imaging biomarkers via adaptive sparse multi-view canonical correlation analysis

Lei Du*, Jin Zhang, Fang Liu, Huiai Wang, Lei Guo, Junwei Han, the Alzheimer's Disease Neuroimaging Initiative[1]

School of Automation, Northwestern Polytechnical University, Xi'an 710072, China

ARTICLE INFO

Article history:
Received 29 September 2020
Revised 10 February 2021
Accepted 15 February 2021
Available online 5 March 2021

Keywords:
Imaging genetics
Multi-omics associations
Sparse canonical correlation analysis
Multi-way bi-multivariate associations

ABSTRACT

To uncover the genetic underpinnings of brain disorders, brain imaging genomics usually jointly analyzes genetic variations and imaging measurements. Meanwhile, other biomarkers such as proteomic expressions can also carry valuable complementary information. Therefore, it is necessary yet challenging to investigate the underlying relationships among genetic variations, proteomic expressions, and neuroimaging measurements, which stands a chance of gaining new insights into the pathogenesis of brain disorders. Given multiple types of biomarkers, using sparse multi-view canonical correlation analysis (SMCCA) and its variants to identify the multi-way associations is straightforward. However, due to the gradient domination issue caused by the naive fusion of multiple SCCA objectives, SMCCA is suboptimal. In this paper, we proposed two adaptive SMCCA (AdaSMCCA) methods, i.e. the robustness-aware AdaSMCCA and the uncertainty-aware AdaSMCCA, to analyze the complicated associations among genetic, proteomic, and neuroimaging biomarkers. We also imposed a data-driven feature grouping penalty to the genetic data with aim to uncover the joint inheritance of neighboring genetic variations. An efficient optimization algorithm, which is guaranteed to converge, was provided. Using two state-of-the-art SMCCA as benchmarks, we evaluated robustness-aware AdaSMCCA and uncertainty-aware AdaSMCCA on both synthetic data and real neuroimaging, proteomics, and genetic data. Both proposed methods obtained higher associations and cleaner canonical weight profiles than comparison methods, indicating their promising capability for association identification and feature selection. In addition, the subsequent analysis showed that the identified biomarkers were related to Alzheimer's disease, demonstrating the power of our methods in identifying multi-way bi-multivariate associations among multiple heterogeneous biomarkers.

1. Introduction

Alzheimer's disease (AD) is a multifactorial neurodegenerative disorder which involves many abnormal alterations happening to the brain. For example, the hippocampus usually exhibits atrophic patterns in AD-affected brain, and simultaneously, the apolipoprotein E (APOE) concentration is also altered in AD patients and thus show relevance to AD pathology (Gupta et al., 2011; Soares et al., 2012). Despite an increasing number of studies during the last decade, the pathological mechanism of AD still remains uncertain (Association, 2019). Therefore, jointly analyzing multiple types of biomarkers, such as magnetic resonance imaging (MRI) derived imaging measurements (Feldman et al., 2020; Fan et al., 2020), blood-based proteomic expression levels and genetic variations, and investigating their associations could deepen our understanding of the pathology of AD. Additionally, a combination of multiple different types of biomarkers as well as their interplays could also increase the reliability and specificity of AD diagnosis, as many biomarkers are not exclusive to AD.

During the last decade, many brain imaging genomic studies arose to investigate the association between two types of biomarkers. A recent systematic review (Shen and Thompson, 2020) showed that most of them were designed to identify the association between the single nucleotide polymorphisms (SNPs) and

* Corresponding author.
E-mail address: dulei@nwpu.edu.cn (L. Du).

[1] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

https://doi.org/10.1016/j.media.2021.102003
1361-8415/© 2021 Elsevier B.V. All rights reserved.

brain imaging quantitative traits (QTs) (Du et al., 2018; 2020a; 2020b; 2020c; Bi et al., 2020a; 2020b). Technically speaking, both the regression methods and sparse canonical correlation analysis (SCCA) methods were widely used. For example, based on regression alone, Wang et al. (2012) proposed the multi-task regression and classification to combine SNPs and imaging QTs to predict the memory deterioration and diagnostic status. Using SCCA alone, Yan et al. (2017) studied the association between proteomic analytes and brain imaging QTs. In addition, the integration of regression and SCCA were also proposed to identify associations among SNPs, imaging QTs and diagnostic outcomes (Zille et al., 2018). To the best of our knowledge, regression methods are typically not designed to directly identify SNP-QT correlations (Wang et al., 2012), and the classical SCCA could only handle two distinct types of biomarkers (Lin et al., 2014; Fang et al., 2016; Yan et al., 2017; Du et al., 2018). Their combination still confronts with the same issue as SCCA. Consequently, it is essential and important to develop novel methods to efficiently and practically identify multi-way associations among more than three different types of biomarkers. By looking into this complex multi-way associations, it would deepen our understanding of the pathological characteristics of AD.

To identify associations among multiple different types of biomarkers, the *results combination* strategy could be an alternative. It first analyzes each kind of biomarkers independently, and then combines the results together to draw a meta conclusion. Obviously, the interplays among different types of biomarkers are overlooked. SMCCA (Witten and Tibshirani, 2009) is another alternative, but directly applying it to identify multi-way associations usually suffers from the gradient domination issue (Kendall et al., 2018) which comes out of the unfair objectives combination (Hu et al., 2017). This is a common problem in imaging genomics since, in general, significantly different correlation levels exhibit among multiple types of biomarkers. For example, the correlation coefficient whose value range is $[-1, 1]$ (or [0,1] in absolute value), obtained by SCCA, between SNPs and structural imaging QTs such as grey matter loss is around [0.2, 0.3] (Du et al., 2021), while that between proteomic markers and structural imaging QTs such as cortical thickness is much higher, with values being around 0.7 for training and 0.38 for testing (Yan et al., 2017). This significant difference incurs gradient domination, and thus leads to the biased optimization. More seriously, as the kinds of biomarkers increase, the gradient domination will get worse. This further deteriorates SMCCA's performance due to its naive fusion strategy. Hu et al. (2017) proposed an adaptive SMCCA, named AdaSMCCA in this paper, which assigns an adaptive weight for each SCCA model. Unfortunately, this method still suffers from the gradient domination. And, since it treats covariance matrices to be identity ones, AdaSMCCA is lacking the theoretical guarantee of consistency and convergence, which might be unreliable (Chen et al., 2013). Therefore, to better identify multi-way bi-multivariate associations, developing more adaptive methods, with solid theoretical properties to handle the gradient domination issue, would be very valuable and meaningful.

In this article, we revisited SMCCA and its limitation in multi-way association identification for imaging genomics. To overcome the gradient domination, we first proposed a robustness-aware AdaSMCCA (rAdaSMCCA) method which adaptively balances between multiple pairwise SCCA models. In addition, to ensure the selection of meaningful biomarkers, we imposed fused pairwise group Lasso (FGL) (Du et al., 2020c) and Lasso to regularize SNPs, and Lasso to both proteomic markers and imaging QTs. We further found that rAdaSMCCA still suffers from the gradient domination issue caused by extreme SCCA model. Therefore, we proposed a novel uncertainty-aware AdaSMCCA (unAdaSMCCA) which resolves the gradient domination issue well with desirable theoretical properties. The contributions of this study were fourfold. First, we proposed two novel AdaSMCCA methods, i.e. rAdaSMCCA and unAdaSMCCA, which could identify multi-way bi-multivariate associations among multiple (≥ 3) types of biomarkers without blindly fusing them. We first introduced rAdaSMCCA since it is an enhancement of AdaSMCCA, and then we introduced unAdaSMCCA which is better than rAdaSMCCA and AdaSMCCA in terms of modeling. Second, both methods overcame the gradient domination issue, and unAdaSMCCA was the best one to overcome this issue. In this study, addressing the gradient domination enabled a better identification of relationships among SNPs, proteomic analytes, and imaging measurements, which could yield interesting findings of AD. Third, the feature grouping penalty for SNPs automatically learnt the grouping structure embedded within neighbouring SNPs. This data-driven regularization could extract SNPs jointly affecting proteomic QTs and imaging QTs. Fourth, to efficiently solve two models, we derived an alternative iteration algorithm with its convergence demonstrated.

In the experiments, we compared rAdaSMCCA and unAdaSMCCA with two state-of-the-art methods, including SMCCA (Witten and Tibshirani, 2009) and adaptive SMCCA (Hu et al., 2017), on four synthetic data sets and one real data set including SNPs, proteomic analyte markers and imaging QTs of 244 subjects from the Alzheimer's disease neuroimaging initiative (ADNI) database. The results on both synthetic and real data sets showed that rAdaSMCCA and unAdaSMCCA identified higher canonical correlation coefficients and better canonical weight patterns indicating enhanced feature selection capability. In particular, unAdaSMCCA performed the best owing to its well-designed loss balancing strategy. In sum, all these results demonstrated that both rAdaSMCCA and unAdaSMCCA held very promising power, with unAdaSMCCA being the best, in identifying multi-way bi-multivariate associations among SNPs, proteomic analytes and imaging QTs. Therefore, our proposed rAdaSMCCA and unAdaSMCCA were promising methods for identifying multi-way associations among multi-omics data in brain imaging genomics.

2. Method

Throughout this article, we denote vectors as lowercase letters, and matrices as uppercase letters. Specifically, $\mathbf{X} = (x_{ij})$ denotes a matrix, and its i-th row and j-th column are separately denoted by $\mathbf{x}^i$ and $\mathbf{x}_j$. The Euclidean norm of $\mathbf{x}$ is denoted as $\|\mathbf{x}\|_2 = \sqrt{\sum x_i^2}$.

2.1. Sparse multi-view canonical correlation analysis (SMCCA)

Given multiple types of data including SNPs, proteomic expression markers and imaging QTs, SMCCA can be applied to find their multi-way associations. Suppose we have n subjects with p SNPs, d proteomic markers and q imaging QTs, and then let $\mathbf{X}_1 \in \mathbb{R}^{n\times p}$ denote the SNP data, $\mathbf{X}_2 \in \mathbb{R}^{n\times d}$ denote the proteomic expression data, and $\mathbf{X}_3 \in \mathbb{R}^{n\times q}$ denote the QT data, SMCCA combines three pairwise SCCA models with respect to these three types of biomarkers, and maximizes this integrated objective as a whole (Witten and Tibshirani, 2009). In particularly, SMCCA is formally defined as

$$\begin{aligned} &\min_{\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_K} \sum_{i<j} -\mathbf{w}_i^\top \mathbf{X}_i^\top \mathbf{X}_j \mathbf{w}_j + \sum_k \Omega(\mathbf{w}_k) \\ &s.t. \quad \|\mathbf{X}_k \mathbf{w}_k\|_2^2 = 1, \ \forall k = 1, \cdots, K. \end{aligned} \tag{1}$$

$\mathbf{w}_k$ is the canonical weight for each kind of biomarkers respectively, and $\Omega(\mathbf{w}_k)$ is the regularization term which introduces sparsity, thereby leading to selection of biomarkers of interest. The tradeoff parameters (explicitly presented later) have been absorbed in $\Omega(\cdot)$ for simplicity.

Since $\|\mathbf{X}_k\mathbf{w}_k\|_2^2 = 1$, we easily obtain the following equivalent formula

$$\min_{\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_K} \sum_{i<j} \left\|\mathbf{X}_i\mathbf{w}_i - \mathbf{X}_j\mathbf{w}_j\right\|_2^2 + \sum_k \Omega(\mathbf{w}_k) \tag{2}$$
$$s.t.\quad \|\mathbf{X}_k\mathbf{w}_k\|_2^2 = 1,\ \forall k = 1,\cdots,K.$$

This fusion objective is in least square form which is sensitive to its sub-objective with a very large value, and thus tends to incur biased optimization. As analyzed earlier, pairwise association levels among multiple types of biomarkers are different, and sometimes significantly different. Thus simply fusing multiple SCCA models could be suboptimal. Moreover, in the implementation, SMCCA assumes $\mathbf{X}_k^\top\mathbf{X}_k = \mathbf{I}$ which might further degrade the performance (Chen et al., 2013; Du et al., 2014).

2.2. AdaSMCCA

Hu *et al.* (Hu et al., 2017) proposed the adaptive SMCCA (AdaSMCCA) to alleviate the gradient domination issue. The AdaSMCCA is defined as

$$\min_{\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_K} \sum_{i<j} -\kappa_{ij}\mathbf{w}_i^\top\mathbf{X}_i^\top\mathbf{X}_j\mathbf{w}_j + \sum_k \lambda_k\|\mathbf{w}_k\|_1 \tag{3}$$
$$s.t.\quad \|\mathbf{w}_k\|_2^2 = 1,\ \forall k = 1,\cdots,K,$$

where λ_k is a positive tradeoff parameter which controls the model sparsity. κ_{ij} is an iteratively varying weight rather than a fixed one. Specifically, κ_{ij} is calculated from

$$\frac{\hat{\kappa}_{ij}\hat{\mathbf{w}}_i^\top\mathbf{X}_i^\top\mathbf{X}_j\hat{\mathbf{w}}_j}{corr(\mathbf{X}_i\hat{\mathbf{w}}_i,\mathbf{X}_j\hat{\mathbf{w}}_j)} = \frac{\kappa_{ij}\mathbf{w}_i^\top\mathbf{X}_i^\top\mathbf{X}_j\mathbf{w}_j}{corr(\mathbf{X}_i\mathbf{w}_i,\mathbf{X}_j\mathbf{w}_j)}, \tag{4}$$

with $\hat{\kappa}_{ij}$ denoting the updated κ_{ij} after each iteration.

The key point of AdaSMCCA is that it keeps forcing each SCCA sub-objective to remain its original importance. But this model has two drawbacks. For one thing, it assumes $\mathbf{X}_k^\top\mathbf{X}_k = \mathbf{I}$ which breaks the Pearson correlation coefficient's range ($[-1,1]$, or [0,1] in absolute value), paying the price for introducing unknown risks. According to Chen et al. (2013), this approximation could result in no guarantee of convergence and consistency. For another, this additional weight κ_{ij} is defined to pull its SCCA model back to the value range, and thus it still suffers from the gradient domination issue.

2.3. Robustness-aware AdaSMCCA

To address the gradient domination, we here propose the robustness-aware AdaSMCCA (rAdaSMCCA) which uses the non-squared loss function rather than the squared one,

$$\min_{\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_K} \sum_{i<j} \left\|\mathbf{X}_i\mathbf{w}_i - \mathbf{X}_j\mathbf{w}_j\right\|_2 + \sum_k \Omega(\mathbf{w}_k) \tag{5}$$
$$s.t.\quad \|\mathbf{X}_k\mathbf{w}_k\|_2^2 = 1,\ \forall k = 1,2,\cdots,K.$$

This non-squared loss function is robust to extreme sub-objective which probably dominate the optimization. The penalty $\Omega(\mathbf{w}_k)$ is used to induce the sparsity, which is beneficial to interpretation. Generally, SNPs, both independently and jointly, affect expression levels of proteomic markers and measurements of imaging QTs (Reich et al., 2001). To figure out the isolated influence of a locus, we use the ℓ_1-norm to regularize the weight of SNP data. Meanwhile, to identify the joint impact of multiple loci, we use the fused pairwise group Lasso (FGL) (Du et al., 2020c) to penalize every two neighbouring loci in light of their genomic position. Therefore, $\Omega(\mathbf{w}_1)$ for SNP data takes the following form

$$\Omega(\mathbf{w}_1) = \lambda_1\beta\sum_{i=1}^{p-1}\sqrt{w_{1i}^2 + w_{1(i+1)}^2} + \lambda_1(1-\beta)\|\mathbf{w}_1\|_1, \tag{6}$$

where λ_1 and β are nonnegative tuning parameters. β balances between the effect of joint and individual feature selection, and further λ_1 balances between the whole penalty and the loss function. As a result, this composited penalty encourages a hybrid and useful feature selection.

Besides, neither every proteomic expression marker nor every imaging QT involves in the progression of brain disorders, and thus sparse constraints to select relevant proteomic and imaging markers are necessary too. To accommodate this, we employ ℓ_1-norm to help select both proteomic markers and imaging QTs, i.e. $\Omega(\mathbf{w}_2) = \lambda_2\|\mathbf{w}_2\|_1$ and $\Omega(\mathbf{w}_3) = \lambda_3\|\mathbf{w}_3\|_1$.

Then specifically, to better identify associations among SNPs, proteomic markers and imaging QTs, we propose the rAdaSMCCA as follows

$$\min_{\mathbf{w}_k} \sum_{1\le i<j\le 3} \kappa_{ij}\left\|\mathbf{X}_i\mathbf{w}_i - \mathbf{X}_j\mathbf{w}_j\right\|_2^2 + \lambda_1\beta\|\mathbf{w}_1\|_{FGL}$$
$$+\lambda_1(1-\beta)\|\mathbf{w}_1\|_1 + \lambda_2\|\mathbf{w}_2\|_1 + \lambda_3\|\mathbf{w}_3\|_1$$
$$s.t.\quad \|\mathbf{X}_k\mathbf{w}_k\|_2^2 = 1,\ \forall k = 1,2,3,\quad \text{and}\quad \kappa_{ij} = \frac{1}{\left\|\mathbf{X}_i\mathbf{w}_i - \mathbf{X}_j\mathbf{w}_j\right\|_2}. \tag{7}$$

rAdaSMCCA uses an iteration-changing other than fixed weight to adaptively weigh multiple sub-objectives during the iteration. On this account, rAdaSMCCA has three advantages. First, it assigns adaptive weights to multiple SCCA loss functions, showing higher robustness than SMCCA. In particular, κ_{ij} will be large if $\left\|\mathbf{X}_i\mathbf{w}_i - \mathbf{X}_j\mathbf{w}_j\right\|_2$ is small and vice versa. As expected, this can alleviate the gradient domination. Second, this method is parameter-free, and thus slightly increases the computational burden. Finally, rAdaSMCCA employs FGL and ℓ_1-norm penalties to automatically find out the group structures within SNPs, which is of great meaning owing to the LD in the human genome.

According to robust learning technique (Gao et al., 2015), the non-squared loss function can only weaken the gradient domination issue other than eliminating it. Therefore, rAdaSMCCA is still dominated by the extreme sub-objective. In a word, both rAdaSMCCA and AdaSMCCA might be suboptimal due to the substantial difference among multiple SCCA objectives.

2.4. Uncertainty-aware AdaSMCCA

In this section, we propose a more adaptive method to address the gradient domination issue. According to Eq. (2), in this regression-type model, each SCCA predicts $\mathbf{w}_i$ based on $\mathbf{X}_i$, $\mathbf{X}_j$ and $\mathbf{w}_j$. For ease of presentation, we denote this prediction as $\mathbf{f}_i(\mathbf{X}_i,\mathbf{X}_j,\mathbf{w}_j)$, and $\mathbf{f}_i(\mathbf{w}_j)$ for short. Generally, the output of $\mathbf{f}_i(\mathbf{w}_j)$ follows a Gaussian distribution and we have the probabilistic model

$$p(\mathbf{w}_i|\mathbf{f}_i(\mathbf{w}_j)) = \mathcal{N}(\mathbf{f}_i(\mathbf{w}_j),\sigma_{ij}^2), \tag{8}$$

where σ_{ij}^2 is the variance which measures the noise of the output (Kendall et al., 2018). $\mathbf{f}_i(\mathbf{w}_j)$ and $\mathbf{f}_j(\mathbf{w}_i)$ are symmetric, and thus $\sigma_{ij} = \sigma_{ji}$. In this article, we have three types of biomarkers, and thus three SCCA models corresponding to three probabilistic models,

$$p(\mathbf{w}_1|\mathbf{f}_1(\mathbf{w}_2)) = \mathcal{N}(\mathbf{f}_1(\mathbf{w}_2),\sigma_{12}^2),\ \ p(\mathbf{w}_2|\mathbf{f}_2(\mathbf{w}_3))$$
$$= \mathcal{N}(\mathbf{f}_2(\mathbf{w}_3),\sigma_{23}^2),\ \ p(\mathbf{w}_3|\mathbf{f}_3(\mathbf{w}_1)) = \mathcal{N}(\mathbf{f}_3(\mathbf{w}_1),\sigma_{13}^2). \tag{9}$$

Then we can obtain the SMCCA likelihood by maximizing the product of these probabilities, i.e.

$$\max_{\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3} p(\mathbf{w}_1|\mathbf{f}_1(\mathbf{w}_2))p(\mathbf{w}_2|\mathbf{f}_2(\mathbf{w}_3))p(\mathbf{w}_3|\mathbf{f}_3(\mathbf{w}_1)) \tag{10}$$

Based on the maximum likelihood inference, we take the logarithm of the objective, i.e.

$$\begin{aligned}&\log\left[p(\mathbf{w}_1|\mathbf{f}_1(\mathbf{w}_2))p(\mathbf{w}_2|\mathbf{f}_2(\mathbf{w}_3))p(\mathbf{w}_3|\mathbf{f}_3(\mathbf{w}_1))\right]\\ &\propto -\frac{1}{2\sigma_{12}^2}\|\mathbf{X}_1\mathbf{w}_1-\mathbf{X}_2\mathbf{w}_2\|_2^2-\frac{1}{2\sigma_{23}^2}\|\mathbf{X}_2\mathbf{w}_2-\mathbf{X}_3\mathbf{w}_3\|_2^2\\ &\quad -\frac{1}{2\sigma_{13}^2}\|\mathbf{X}_1\mathbf{w}_1-\mathbf{X}_3\mathbf{w}_3\|_2^2-\log\sigma_{12}-\log\sigma_{23}-\log\sigma_{13}. \end{aligned} \tag{11}$$

Maximizing Eq. (11) is equivalent to minimizing the following objective

$$\begin{aligned}&\frac{1}{2\sigma_{12}^2}\|\mathbf{X}_1\mathbf{w}_1-\mathbf{X}_2\mathbf{w}_2\|_2^2+\frac{1}{2\sigma_{23}^2}\|\mathbf{X}_2\mathbf{w}_2-\mathbf{X}_3\mathbf{w}_3\|_2^2\\ &+\frac{1}{2\sigma_{13}^2}\|\mathbf{X}_1\mathbf{w}_1-\mathbf{X}_3\mathbf{w}_3\|_2^2+\log\sigma_{12}+\log\sigma_{23}+\log\sigma_{13}.\end{aligned}\tag{12}$$

It is interesting that this function takes the estimation variance into account. Additionally, this estimation variance is also regularized to prevent it from increasing too much (Kendall et al., 2018).

Now plugging sparsity-inducing terms $\Omega(\mathbf{w}_k)$ into this model, we propose the uncertainty-aware AdaSMCCA (unAdaSMCCA) model,

$$\begin{aligned}&\min_{\mathbf{w}_k}\sum_{1\le i<j\le 3}\frac{1}{2\sigma_{ij}^2}\left\|\mathbf{X}_i\mathbf{w}_i-\mathbf{X}_j\mathbf{w}_j\right\|_2^2+\log\sigma_{ij}\\ &+\lambda_1\beta\|\mathbf{w}_1\|_{FGL}+\lambda_1(1-\beta)\|\mathbf{w}_1\|_1+\lambda_2\|\mathbf{w}_2\|_1+\lambda_3\|\mathbf{w}_3\|_1\\ &s.t.\ \ \|\mathbf{X}_k\mathbf{w}_k\|_2^2=1,\ \forall k=1,2,3.\end{aligned}\tag{13}$$

unAdaSMCCA assigns a more adaptive weight for each sub-objective compared with AdaSMCCA and rAdaSMCCA. This fusion strategy has three advantages. First of all, the iteratively changing weight comes from the estimation variance, which is called uncertainty of each SCCA model. This uncertainty can capture the relative confidence among SMCCA's loss functions (Kendall et al., 2018). Second, compared with AdaSMCCA and rAdaSMCCA, unAdaSMCCA addresses the gradient domination surpassingly thanks to its well consideration on the estimation uncertainty. Additionally, unAdaSMCCA also holds the merit of diverse feature selection as rAdaSMCCA.

2.5. The optimization

Both rAdaSMCCA and unAdaSMCCA follow the same modeling paradigm by iteratively reweighing each SCCA loss function. Therefore, we focus on solving unAdaSMCCA, and rAdaSMCCA can be solved in the same way.

The problem Eq. (13) is difficult to solve due to the intertwined multiple canonical weights. Thus we utilize the alternative iteration algorithm. According to Appendix A.2 (Witten et al., 2009), the equality constraint, i.e. $\|\mathbf{X}_k\mathbf{w}_k\|_2^2=1$, can be put aside with focus on the unconstrained problem. Thus we obtain the unconstrained objective with respect to $\mathbf{w}_1$ with those remaining canonical weights and loss weights fixed, i.e.

$$\begin{aligned}&\frac{1}{2\sigma_{12}^2}\|\mathbf{X}_1\mathbf{w}_1-\mathbf{X}_2\mathbf{w}_2\|_2^2+\frac{1}{2\sigma_{13}^2}\|\mathbf{X}_1\mathbf{w}_1-\mathbf{X}_3\mathbf{w}_3\|_2^2\\ &+\lambda_1\beta\|\mathbf{w}_1\|_{FGL}+\lambda_1(1-\beta)\|\mathbf{w}_1\|_1\end{aligned}\tag{14}$$

To minimize this equation, we take the derivative of this objective with respect to $\mathbf{w}_1$, and set it to zero. Then we arrive at

$$\begin{aligned}&\left(\left(\frac{1}{\sigma_{12}^2}+\frac{1}{\sigma_{13}^2}\right)\mathbf{X}_1^\top\mathbf{X}_1+\lambda_1\beta\tilde{\mathbf{D}}_1+\lambda_1(1-\beta)\mathbf{D}_1\right)\mathbf{w}_1\\ &=\frac{1}{\sigma_{12}^2}\mathbf{X}_1^\top\mathbf{X}_2\mathbf{w}_2+\frac{1}{\sigma_{13}^2}\mathbf{X}_1^\top\mathbf{X}_3\mathbf{w}_3.\end{aligned}\tag{15}$$

$\mathbf{D}_1$ is a diagonal matrix with the i-th diagonal entry being $\frac{1}{|w_{1i}|}(i=1,\cdots,p)$. $\tilde{\mathbf{D}}_1$ is also a diagonal matrix, and its i-th diagonal element is $\frac{1}{\sqrt{w_{1(i-1)}^2+w_{1i}^2}}+\frac{1}{\sqrt{w_{1i}^2+w_{1(i+1)}^2}}(i=2,\cdots,p-1)$[2] Guessing initial values for $\tilde{\mathbf{D}}_1$ and $\mathbf{D}_1$, we can proceed on to get the update $\mathbf{w}_1$ as

$$\begin{aligned}\mathbf{w}_1=&\left(\left(\frac{1}{\sigma_{12}^2}+\frac{1}{\sigma_{13}^2}\right)\mathbf{X}_1^\top\mathbf{X}_1+\lambda_1\beta\tilde{\mathbf{D}}_1+\lambda_1(1-\beta)\mathbf{D}_1\right)^{-1}\\ &\left(\frac{1}{\sigma_{12}^2}\mathbf{X}_1^\top\mathbf{X}_2\mathbf{w}_2+\frac{1}{\sigma_{13}^2}\mathbf{X}_1^\top\mathbf{X}_3\mathbf{w}_3\right).\end{aligned}\tag{16}$$

By now, only the equality constraint remains unsolved. Generally, a simple re-scaling step can be applied, i.e.

$$\mathbf{w}_1=\frac{\mathbf{w}_1}{\|\mathbf{X}_1\mathbf{w}_1\|_2}.\tag{17}$$

The parameter σ_{ij} can also be similarly solved in closed-form, i.e.

$$\sigma_{ij}=\left\|\mathbf{X}_i\mathbf{w}_i-\mathbf{X}_j\mathbf{w}_j\right\|_2.\tag{18}$$

Using the same procedure, we can obtain the closed-form solution to both $\mathbf{w}_2$ and $\mathbf{w}_3$

$$\begin{aligned}\mathbf{w}_2=&\left(\left(\frac{1}{\sigma_{12}^2}+\frac{1}{\sigma_{23}^2}\right)\mathbf{X}_2^\top\mathbf{X}_2+\lambda_2\mathbf{D}_2\right)^{-1}\\ &\left(\frac{1}{\sigma_{12}^2}\mathbf{X}_2^\top\mathbf{X}_1\mathbf{w}_1+\frac{1}{\sigma_{23}^2}\mathbf{X}_2^\top\mathbf{X}_3\mathbf{w}_3\right),\\ \mathbf{w}_3=&\left(\left(\frac{1}{\sigma_{13}^2}+\frac{1}{\sigma_{23}^2}\right)\mathbf{X}_3^\top\mathbf{X}_3+\lambda_3\mathbf{D}_3\right)^{-1}\\ &\left(\frac{1}{\sigma_{13}^2}\mathbf{X}_3^\top\mathbf{X}_1\mathbf{w}_1+\frac{1}{\sigma_{23}^2}\mathbf{X}_3^\top\mathbf{X}_2\mathbf{w}_2\right),\end{aligned}\tag{19}$$

where $\mathbf{D}_2$ and $\mathbf{D}_3$ are two diagonal matrices, and their entries take the same form as $\mathbf{D}_1$ for $\mathbf{w}_1$. In particular, the i-th diagonal element of $\mathbf{D}_2$ is $\frac{1}{|w_{2i}|}(i=1,\cdots,d)$, and that of $\mathbf{D}_3$ is $\frac{1}{|w_{3i}|}(i=1,\cdots,q)$. Finally, $\mathbf{w}_2$ and $\mathbf{w}_3$ can be further updated by the following re-scaling steps

$$\mathbf{w}_2=\frac{\mathbf{w}_2}{\|\mathbf{X}_2\mathbf{w}_2\|_2},\quad \mathbf{w}_3=\frac{\mathbf{w}_3}{\|\mathbf{X}_3\mathbf{w}_3\|_2}.\tag{20}$$

Eqs. 16–(20) indicate that we can alternatively obtain $\mathbf{w}_1$, $\mathbf{w}_2$ and $\mathbf{w}_3$, as well as σ_{12}, σ_{13} and σ_{23}. During the iteration, final solutions will be attained once the optimum or predefined stopping conditions are satisfied. In each iteration, those diagonal matrices are easy to calculate, and σ_{12}, σ_{13} and σ_{23} are also easy to obtain. To efficiently solve Eq. (16) and Eq. (19), we solve a system of linear equations instead of calculating the inverse of covariance matrices. Hence our algorithms run fast with desired efficiency. Finally, we present the procedure in Algorithm 1 .

2.6. Convergence analysis

Theorem 1. *The Algorithm 1 monotonously decreases the objective value of Eq. (13) during the iteration.*

Proof 1. We first prove that the objective of Eq. (13) decreases when solving $\mathbf{w}_1$. For ease of presentation, we denote the estimate of parameters at the current iteration t as

[2] The first diagonal element of $\tilde{\mathbf{D}}_1$ is $\frac{1}{\sqrt{w_{11}^2+w_{12}^2}}$, and the p-th diagonal element is $\frac{1}{\sqrt{w_{1(p-1)}^2+w_{1p}^2}}$. Details are in (Du et al., 2020c).

Algorithm 1: Uncertainty-aware AdaSMCCA Algorithm.

```
Result: $\mathbf{w}_1$, $\mathbf{w}_2$ and $\mathbf{w}_3$
Input SNP data $\mathbf{X}_1 \in \mathbb{R}^{n\times p}$, proteomic expression data
$\mathbf{X}_2 \in \mathbb{R}^{n\times d}$, and imaging QT data $\mathbf{X}_3 \in \mathbb{R}^{n\times q}$. The pre-tuned $\lambda_1$,
$\lambda_2$, $\lambda_3$ and $\beta$;
Initialize $\mathbf{w}_1 \in \mathbb{R}^{p\times 1}$, $\mathbf{w}_2 \in \mathbb{R}^{d\times 1}$ and $\mathbf{w}_3 \in \mathbb{R}^{q\times 1}$;
while not converged do
    Calculate the diagonal matrices $\tilde{\mathbf{D}}_1$, $\mathbf{D}_1$, $\mathbf{D}_2$ and $\mathbf{D}_3$;
    Solve $\mathbf{w}_1$ using Eq.~ (16), and scale $\mathbf{w}_1$ so that
    $\|\mathbf{X}_1\mathbf{w}_1\|_2^2 = 1$;
    Solve $\mathbf{w}_2$ and $\mathbf{w}_3$ using Eq.~(19), and scale $\mathbf{w}_2$ and $\mathbf{w}_3$ so
    that $\|\mathbf{X}_2\mathbf{w}_2\|_2^2 = 1$, and $\|\mathbf{X}_3\mathbf{w}_3\|_2^2 = 1$;
    Solve $\sigma_{12}$, $\sigma_{13}$ and $\sigma_{23}$ according to Eq.~(18);
Sorting $\mathbf{w}_1$, $\mathbf{w}_2$ and $\mathbf{w}_3$ in descending order based on the
absolute value respectively.
```

$\left\{\mathbf{w}_1^{(t)}, \mathbf{w}_2^{(t)}, \mathbf{w}_3^{(t)}, \sigma_{12}^{(t)}, \sigma_{13}^{(t)}, \sigma_{23}^{(t)}\right\}$. Besides, we denote the objective of problem (14) as $F(\mathbf{w}_1)$:

$$F(\mathbf{w}_1) \overset{\text{def}}{=} \frac{1}{2\sigma_{12}^2}\|\mathbf{X}_1\mathbf{w}_1 - \mathbf{X}_2\mathbf{w}_2^{(t)}\|_2^2 + \frac{1}{2\sigma_{13}^2}\|\mathbf{X}_1\mathbf{w}_1 - \mathbf{X}_3\mathbf{w}_3^{(t)}\|_2^2 + \lambda_1\beta\|\mathbf{w}_1\|_{FGL} + \lambda_1(1-\beta)\|\mathbf{w}_1\|_1 \tag{21}$$

Then we define

$$\begin{aligned} G(\mathbf{w}_1) \overset{\text{def}}{=}\ & \frac{1}{2\sigma_{12}^2}\|\mathbf{X}_1\mathbf{w}_1 - \mathbf{X}_2\mathbf{w}_2^{(t)}\|_2^2 + \frac{1}{2\sigma_{13}^2}\|\mathbf{X}_1\mathbf{w}_1 - \mathbf{X}_3\mathbf{w}_3^{(t)}\|_2^2 \\ & + \lambda_1\beta\sum_{i=1}^{p-1}\left(\frac{w_{1i}^2 + w_{1(i+1)}^2}{2\sqrt{{w_{1i}^{(t)}}^2 + {w_{1(i+1)}^{(t)}}^2}} + \frac{\sqrt{{w_{1i}^{(t)}}^2 + {w_{1(i+1)}^{(t)}}^2}}{2}\right) \\ & + \lambda_1(1-\beta)\sum_{i=1}^{p}\left(\frac{w_{1i}^2}{2|w_{1i}^{(t)}|} + \frac{\left|w_{1i}^{(t)}\right|}{2}\right) \\ =\ & \frac{1}{2\sigma_{12}^2}\|\mathbf{X}_1\mathbf{w}_1 - \mathbf{X}_2\mathbf{w}_2^{(t)}\|_2^2 + \frac{1}{2\sigma_{13}^2}\|\mathbf{X}_1\mathbf{w}_1 - \mathbf{X}_3\mathbf{w}_3^{(t)}\|_2^2 \\ & + \lambda_1\beta\left(\frac{1}{2}\mathbf{w}_1^\top\tilde{\mathbf{D}}_1\mathbf{w}_1 + \frac{1}{2}{\mathbf{w}_1^{(t)}}^\top\tilde{\mathbf{D}}_1\mathbf{w}_1^{(t)}\right) \\ & + \lambda_1(1-\beta)\left(\frac{1}{2}\mathbf{w}_1^\top\mathbf{D}_1\mathbf{w}_1 + \frac{1}{2}{\mathbf{w}_1^{(t)}}^\top\mathbf{D}_1\mathbf{w}_1^{(t)}\right), \end{aligned} \tag{22}$$

where $\tilde{\mathbf{D}}_1$ and $\mathbf{D}_1$ are defined in Eq. (15), and the second equality can be easily verified.

It is obvious that $G(\mathbf{w}_1)$ is a convex quadratic function (smooth, differentiable everywhere) that satisfies

$$G\left(\mathbf{w}_1^{(t)}\right) = F\left(\mathbf{w}_1^{(t)}\right), \quad G(\mathbf{w}_1) \geq F(\mathbf{w}_1), \ \forall \mathbf{w}_1 \in \mathbb{R}^p \tag{23}$$

Since the estimate of $\mathbf{w}_1$ at the next iteration $t+1$, expressed in Eq. (16) and denoted as $\mathbf{w}_1^{(t+1)}$, is the (global) minimizer of $G(\mathbf{w}_1)$, we have

$$G\left(\mathbf{w}_1^{(t+1)}\right) \leq G\left(\mathbf{w}_1^{(t)}\right). \tag{24}$$

Putting (23)-(24) together, we have

$$F\left(\mathbf{w}_1^{(t+1)}\right) \leq G\left(\mathbf{w}_1^{(t+1)}\right) \leq G\left(\mathbf{w}_1^{(t)}\right) = F\left(\mathbf{w}_1^{(t)}\right). \tag{25}$$

This proves the convergence, i.e. the objective decreases by fixing $\mathbf{w}_2$ and $\mathbf{w}_3$ to solve $\mathbf{w}_1$. This conclusion remains unchanged after the re-scaling step according to the algorithm.

We can alternatively draw the same conclusion with respect to $\mathbf{w}_2$ and $\mathbf{w}_3$, as well as σ_{12}, σ_{13}, σ_{23} respectively. Denoting our objective as $\mathcal{L}(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3)$ and combining conclusions above together, we have

$$\begin{aligned} & \mathcal{L}(\mathbf{w}_1^{(t+1)}, \mathbf{w}_2^{(t+1)}, \mathbf{w}_3^{(t+1)}) \leq \mathcal{L}(\mathbf{w}_1^{(t+1)}, \mathbf{w}_2^{(t+1)}, \mathbf{w}_3^{(t)}) \\ & \leq \mathcal{L}(\mathbf{w}_1^{(t+1)}, \mathbf{w}_2^{(t)}, \mathbf{w}_3^{(t)}) \leq \mathcal{L}(\mathbf{w}_1^{(t)}, \mathbf{w}_2^{(t)}, \mathbf{w}_3^{(t)}), \end{aligned} \tag{26}$$

which completes the proof.

We further know $\mathcal{L}(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3)$ is lower bounded by zero. Therefore, according to Theorem 1, Algorithm 1 will converge to the optimum.

3. Experiment results

We used the SMCCA (Witten and Tibshirani, 2009) and Adaptive SMCCA (AdaSMCCA) as benchmark methods. SMCCA simply combines multiple SCCA models without consideration on gradient domination. AdaSMCCA combines these SCCA models with each of them associated with an additional weight. By now, AdaSMCCA was the state-of-the-art SMCCA method. Therefore, this comparison study could help show the efficiency and effectiveness of our proposed methods (The Matlab code of our AdaSMCCA methods is publicly available on https://github.com/dulei323/AdaSMCCA). Since the main goal of this study is to identify the multi-way associations among SNPs, proteomic analytes, and imaging measurements, those SCCA methods (Lin et al., 2014; Fang et al., 2016; Du et al., 2016; 2018) that can only identify the pairwise association between two types of data were excluded.

Except SMCCA, there were four parameters for our methods and three ones for AdaSMCCA, which should be fine-tuned before conducting experiments. We employed the nested 5-fold cross-validation method where the inner loop took charge of seeking them from a candidate interval. We used several heuristic rules to reduce the time effort. In particular, β was limited in (0,1), and $\beta < 0.5$ could be used if we preferred individual sparsity, while $\beta > 0.5$ was the choice if group sparsity was preferable. λ_1, λ_2 and λ_3 controlled sparsity levels for SNPs, proteomic biomarkers and imaging QTs. In addition, since these penalty functions were monotone increasing, relative large parameters should be used if the number of biomarkers were large. Based on this, we first searched three λ's from 10^i $(i = -3, -2, -1, 0, 1, 2, 3)$. Once we obtained the winner parameters, we went into the second tuning procedure in a much smaller interval $[0.1, 0.2, \cdots, \cdots, 1]$. All methods ran on the same software platform, and used the same data partition to make the comparison fair.

The stopping condition was set to $\max_{k\in\{1,2,3\}} \max|\mathbf{w}_k^{(t+1)} - \mathbf{w}_k^{(t)}| \leq \epsilon$ with the tolerance error $\epsilon = 10^{-5}$. Experimentally, both rAdaSMCCA and unAdaSMCCA converge within about 20∼30 iterations, and we additionally set the maximum number of iterations to 100 to ensure the efficiency and performance.

3.1. Simulation study

We simulated four data sets with different characteristics following the similar procedure to that used in (Hu et al., 2017). There were three matrices with the same number of subjects n, and different number of features p, d and q respectively. The first three data sets ($n = 200$, $p = 150$, $d = 200$ and $q = 150$) were generated with the same ground truth, where the first two had different levels of gradient domination and the third one did not have. The fourth data set ($n = 500$, $p = 350$, $d = 800$ and $q = 400$) had a distinct number of feature dimensionality. In particular, we first built three sparse canonical weights $\mathbf{w}_1 \in \mathcal{R}^{p\times 1}$, $\mathbf{w}_2 \in \mathcal{R}^{d\times 1}$ and $\mathbf{w}_3 \in \mathcal{R}^{q\times 1}$, as well as a latent variable $\mathbf{z} \in \mathcal{R}^{n\times 1}$. Then $\mathbf{X}_k$ was generated by $(x_{ij})_k \sim \mathcal{N}(z_i w_{kj}, \sigma_k)$ where $k = 1, 2, 3$. Fig. 1 presented the ground truths of four data sets in top row.

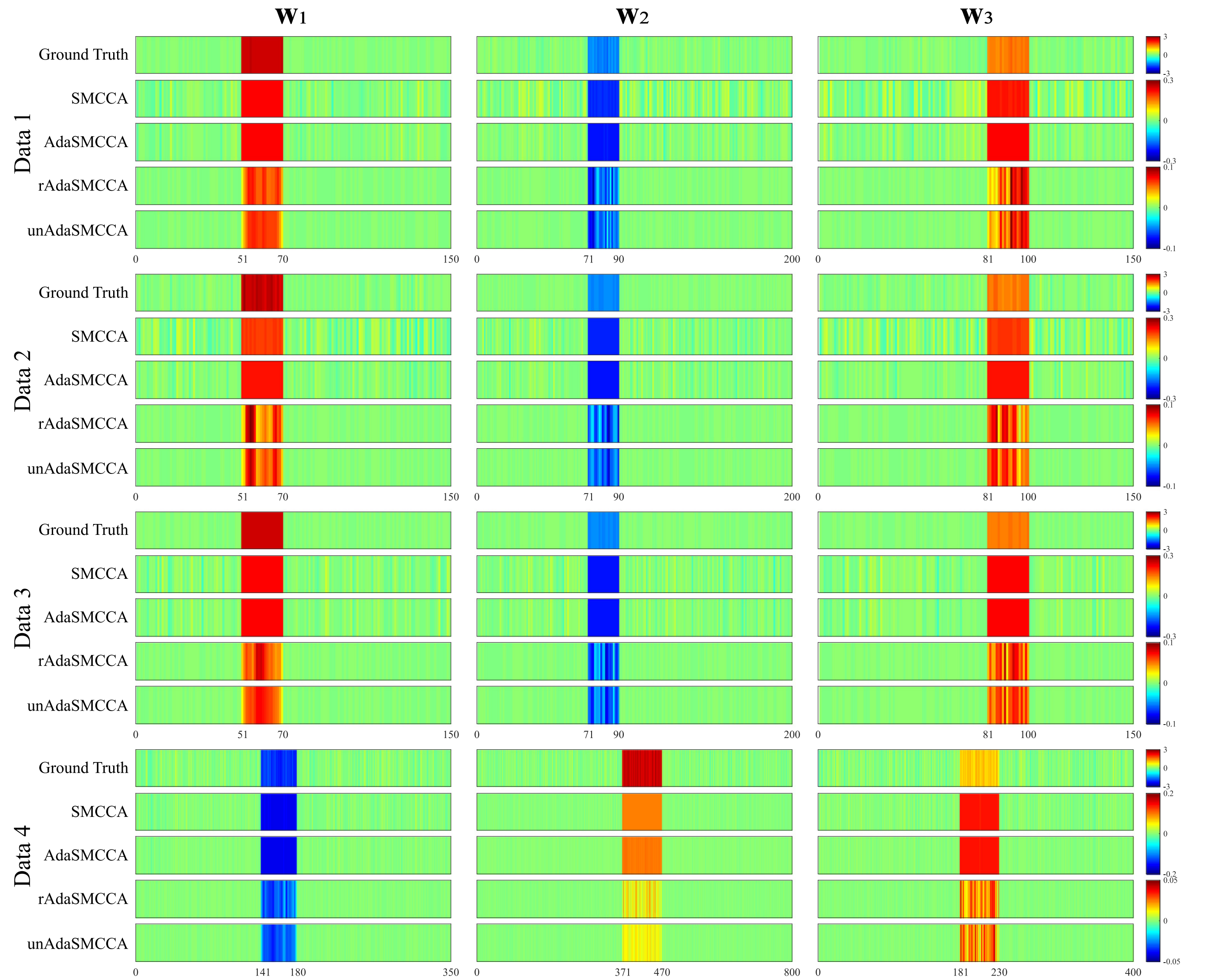

Fig. 1. Canonical weights on synthetic data sets. Results were shown row by row for all data. Within each data set, there were five rows corresponding to ground truth, SMCCA, AdaSMCCA, rAdaSMCCA and unAdaSMCCA respectively. Canonical weights $\mathbf{w}_1$, $\mathbf{w}_2$ and $\mathbf{w}_3$ were shown from left to right for each data set.

We showed the canonical correlation coefficients (CCCs) of each method in Table 1, where a higher CCC indicates a better performance. These CCCs were averaged from five folds (repeated 20 runs to enable stability) of four SCCA models and the standard deviations were also contained. For ease of presentation, we denoted the CCC between $\mathbf{X}_1$ and $\mathbf{X}_2$ as CCC_{1-2}, and so on. In the table, unAdaSMCCA and rAdaSMCCA alternatively yielded the highest CCCs, showing their enhanced performance. Interestingly, on Data 1 and Data 2 where the three SCCA sub-objectives were imbalanced, unAdaSMCCA and rAdaSMCCA outperformed both benchmarks by holding higher CCCs. On Data 3 where three SCCA sub-objectives were balanced, all four methods obtained acceptable performance. This demonstrated that our AdaSMCCA methods well addressed the gradient domination issue while both benchmarks cannot. Besides, we exhibited the canonical weight heat maps in Fig. 1. In this figure, we used distinct color bars for four methods since different methods use different scale methods. The color bar can help show the relative importance of features, but will not make the decision for a method. We observed that all four methods successfully identified the true signals. However, if we observed the panel of SMCCA and AdaSMCCA carefully, we can see that there are many irrelevant signals which could mislead the feature selection results. In contrast, both unAdaSMCCA and rAdaSMCCA showed relative cleaner canonical weight profiles than benchmarks, which were in accordance with the ground truth. Combining the results on CCCs and canonical weights together, unAdaSMCCA and rAdaSMCCA performed better than two competitors, especially when the gradient domination issue does exist. This revealed that our method can well address the gradient domination issue. In summary, this simulation study demonstrated the essential and superiority of the adaptive strategy (Kendall et al., 2018), and thus the success of our unAdaSMCCA.

3.2. *Real neuroimaging genetic study*

The genotyping data, quantification of proteomic analytes in plasma and brain imaging data were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The primary goal of the initiative is to test whether serial magnetic resonance imaging (MRI), or other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and

Table 1
The averaged canonical correlation coefficients (CCCs) on synthetic data sets.

	Data 1					
	Training			Testing		
	CCC_{1-2}	CCC_{2-3}	CCC_{1-3}	CCC_{1-2}	CCC_{2-3}	CCC_{1-3}
SMCCA	0.8529 ± 0.1363	0.9924 ± 0.0052	0.8926 ± 0.1061	0.8051 ± 0.1807	0.9884 ± 0.0096	0.8631 ± 0.1309
AdaSMCCA	0.9167 ± 0.1223	0.9958 ± 0.0043	0.9371 ± 0.0964	0.8951 ± 0.1491	0.9934 ± 0.0082	0.9254 ± 0.1075
rAdaSMCCA	0.9999 ± 0.0000	0.9998 ± 0.0000	0.9999 ± 0.0000	0.9998 ± 0.0001	0.9997 ± 0.0001	0.9998 ± 0.0001
unAdaSMCCA	0.9998 ± 0.0000	0.9998 ± 0.0000	0.9999 ± 0.0000	0.9998 ± 0.0001	0.9998 ± 0.0001	0.9998 ± 0.0001
	Data 2					
SMCCA	0.8105 ± 0.1388	0.8057 ± 0.1412	0.9954 ± 0.0025	0.7667 ± 0.1571	0.7594 ± 0.1622	0.9941 ± 0.0036
AdaSMCCA	0.9416 ± 0.0993	0.9375 ± 0.1058	0.9979 ± 0.0021	0.9329 ± 0.1099	0.9273 ± 0.1186	0.9974 ± 0.0027
rAdaSMCCA	0.9999 ± 0.0000	0.9999 ± 0.0000	0.9999 ± 0.0000	0.9999 ± 0.0001	0.9998 ± 0.0001	0.9999 ± 0.0000
unAdaSMCCA	0.9999 ± 0.0000	0.9999 ± 0.0000	0.9999 ± 0.0000	0.9999 ± 0.0001	0.9998 ± 0.0001	0.9999 ± 0.0000
	Data 3					
SMCCA	0.9981 ± 0.0003	0.9982 ± 0.0002	0.9983 ± 0.0002	0.9978 ± 0.0006	0.9980 ± 0.0006	0.9981 ± 0.0005
AdaSMCCA	0.9985 ± 0.0002	0.9985 ± 0.0002	0.9987 ± 0.0002	0.9983 ± 0.0005	0.9984 ± 0.0005	0.9986 ± 0.0004
rAdaSMCCA	0.9998 ± 0.0000	0.9998 ± 0.0000	0.9999 ± 0.0000	0.9998 ± 0.0001	0.9997 ± 0.0001	0.9998 ± 0.0001
unAdaSMCCA	0.9998 ± 0.0000	0.9998 ± 0.0000	0.9999 ± 0.0000	0.9998 ± 0.0000	0.9998 ± 0.0001	0.9998 ± 0.0001
	Data 4					
SMCCA	0.9953 ± 0.0043	0.9953 ± 0.0036	0.9961 ± 0.0039	0.9951 ± 0.0042	0.9950 ± 0.0039	0.9958 ± 0.0044
AdaSMCCA	0.9991 ± 0.0009	0.9986 ± 0.0011	0.9991 ± 0.0005	0.9991 ± 0.0008	0.9986 ± 0.0010	0.9991 ± 0.0005
rAdaSMCCA	0.9999 ± 0.0000	0.9996 ± 0.0000	0.9995 ± 0.0000	0.9998 ± 0.0000	0.9994 ± 0.0001	0.9993 ± 0.0001
unAdaSMCCA	0.9998 ± 0.0000	0.9995 ± 0.0000	0.9994 ± 0.0000	0.9998 ± 0.0000	0.9994 ± 0.0001	0.9992 ± 0.0001

Table 2
Participant characteristics.

	NC	MCI	AD
Num	42	137	65
Gender (M/F, %)	52.38/47.62	69.34/30.66	55.38/44.62
Handedness (R/L, %)	90.48/9.52	92.70/7.30	98.46/1.54
Age (mean±std)	75.40±5.80	74.13±7.22	74.75±7.67
Education (mean±std)	15.88±2.77	16.03±2.98	15.12±3.05

early Alzheimer's disease (AD). For up-to-date information, see www.adni-info.org.

We included 244 non-Hispanic Caucasian subjects with 42 normal controls (NCs), 137 MCIs and 65 ADs in this real study. The detailed participant characteristics were shown in Table 2. The structural magnetic resonance imaging (sMRI) scans were processed with voxel-based morphometry (VBM) in SPM. In general, these scans were aligned to a T1-weighted template image, segmented into gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) maps, normalized to MNI space, and smoothed with an 8 mm^3 FWHM kernel. We subsampled the whole brain and generated 465 GM density measures as imaging QTs.

The blood plasma samples of the same population were measured by Rules Based Medicine, Inc. (RBM) proteomic panel. After quality control (QC), we obtained 146 proteomic markers.

The genotyping data from the ADNI website were genotyped using the Human 610-Quad or OmniExpress Array platforms (Illumina, Inc., San Diego, CA, USA). After standard QC process and imputation by MaCH software, we obtained each subject's SNP data. In this study, we included 827 SNPs around AD risk genes such as *APOE, TOMM40* and *APOC1* (boundary: 170kb) based on ANNOVAR annotation. Given these SNPs, proteomic markers and brain imaging QTs, our aim is to study their multi-way bi-multivariate associations, and to identify biomarkers of relevance which could enable a more targeted and in-depth follow-up analysis.

3.2.1. Multi-way bi-multivariate associations

In Fig. 2, we presented the averaged training and testing CCCs which showed the multi-way bi-multivariate association identification ability. To facilitate the analysis, we denoted association between SNPs and proteomic markers as SNP-Protein, similarly, that between SNPs and imaging QTs as SNP-QT, and that between proteomic markers and imaging QTs as Protein-QT. It is clear that, overall, unAdaSMCCA obtained the highest CCCs. rAdaSMCCA obtained similar CCCs to unAdaSMCCA, and both of them outperformed SMCCA and AdaSMCCA. In addition, for training results, unAdaSMCCA yielded the highest CCCs on SNP-Protein and SNP-QT, and rAdaSMCCA won out on Protein-QT. The similar conclusions could be drawn in testing CCCs as well. More interestingly, in both training and testing results on SNP-QT and Protein-QT, SMCCA showed no obvious difference to those AdaSMCCA meth-

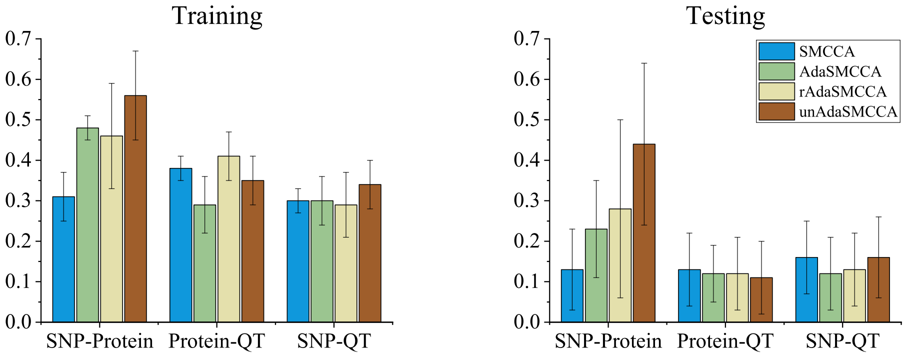

Fig. 2. Comparison of the averaged canonical correlation coefficients (CCCs) across four SCCA models on 5-fold training and testing on ADNI data.

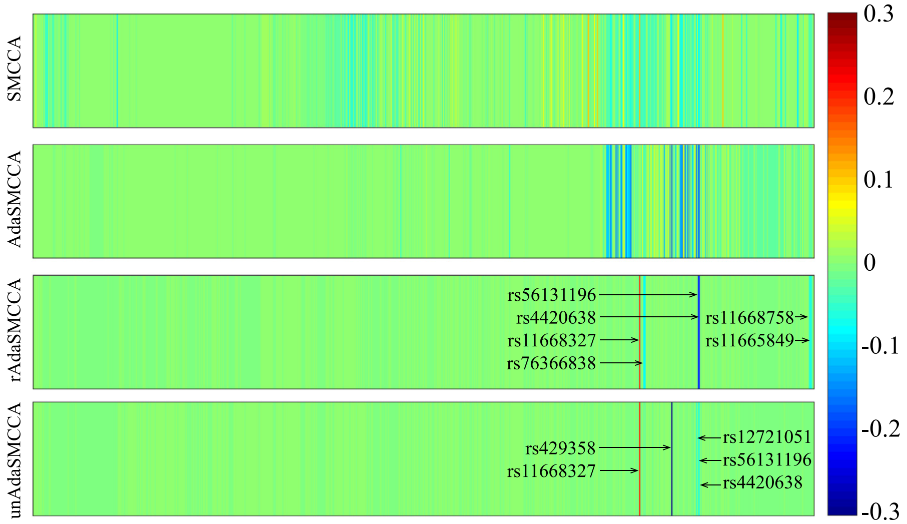

Fig. 3. Canonical weights of SNPs from five-fold cross-validation trials. Each row corresponds to an SCCA method: (1) SMCCA; (2) AdaSMCCA; (3) rAdaSMCCA; and (4) unAdaSMCCA.

Table 3
Top ten SNPs identified by each method according to mean canonical weights. The absolute weight value showed the importance of each SNP.

SMCCA		AdaSMCCA		rAdaSMCCA		unAdaSMCCA	
SNP_ID	Weight	SNP_ID	Weight	SNP_ID	Weight	SNP_ID	Weight
rs11668327	0.1079	rs429358	0.1671	rs56131196	0.2000	rs429358	0.4433
rs73052307	0.0998	rs5117	0.1540	rs4420638	0.2000	rs11668327	0.2000
rs11669338	0.0994	rs12721051	0.1529	rs11668327	0.1883	rs4420638	0.0537
rs11673139	0.0994	rs56131196	0.1529	rs76366838	0.0800	rs56131196	0.0537
rs79429216	0.0981	rs4420638	0.1529	rs11665849	0.0602	rs12721051	0.0537
rs111740474	0.0924	rs483082	0.1527	rs11668758	0.0602	rs76366838	0.0164
rs429358	0.0918	rs438811	0.1527	rs75687619	0.0600	rs114536010	0.0164
rs147901416	0.0879	rs59007384	0.1457	rs114536010	0.0600	rs75687619	0.0164
rs12721051	0.0827	rs283815	0.1453	rs11669609	0.0392	rs188535946	0.0164
rs56131196	0.0827	rs184017	0.1391	rs2142074	0.0392	rs140480140	0.0164

ods. This implies that both sub-objectives of SMCCA, i.e. SNP-QT and Protein-QT, with relative low CCCs dominate the whole objective of SMCCA, which results in a pronounced difference on SNP-Protein. We could also observe that AdaSMCCA alleviated the gradient domination to some extent but not all. In contrast, rAdaSMCCA improved the performance of AdaSMCCA, but it still suffered from gradient domination owing to its halfway strategy of weighing losses. Surprisingly, unAdaSMCCA addressed the gradient domination issue quite well. unAdaSMCCA not only yielded the best or comparable CCCs on SNP-QT and Protein-QT, but also obtained the highest score on the association of SNP-Protein which, without a well-designed reweighting strategy, could probably be missed by the naive SMCCA model. Therefore, by iteratively reweighing each sub-objective, significant improvement could be obtained. In a word, unAdaSMCCA performed the best thanks to its consideration of each sub-objective's uncertainty as well as the elaborated regularization.

3.2.2. *Identification and interpretation of SNPs*

In addition, identifying relevant biomarkers can be another important evaluate criterion, showing a method's potential in feature selection. We showed canonical weights in terms of SNP in Fig. 3 with those top selected SNPs tagged. Both rAdaSMCCA and unAdaSMCCA exhibited a cleaner heatmap patterns than two competitors. To make it clear, we also presented the top ten SNPs identified by each method in Table 3 where SNPs and their estimated weight (absolute) values were contained. It has been known that rs429358 (*APOE*) has a strong influence on the risk of AD. As expected, unAdaSMCCA identified this locus and recognized it as its top identified genotypic marker with a significant higher value. Interestingly, rAdaSMCCA missed rs429358 in its top ten loci. Due to the FGL penalty, rAdaSMCCA identified several groups of loci, and their combined effect could exceed that of rs429358. However, the further investigation should be warranted. AdaSMCCA also identified this locus but its weight value was similar to those remaining loci, which were hard to interpret. In addition, a literature search showed that all selected SNPs, identified by unAdaSMCCA, located in MCI- or AD-related genes such as *APOC1* and *TOMM40* (Gao et al., 2016; Zhou et al., 2019). Six out of ten loci identified by rAdaSMCCA were overlapped with that of unAdaSMCCA, indicating its similar performance to unAdaSMCCA. More interestingly, owing to FGL penalty, both unAdaSMCCA and rAdaSMCCA showed a feature grouping result which was in agreement with linkage disequilibrium (LD). For example, according to canonical weight values, unAdaSMCCA identified two groups, i.e. loci in *APOC1* (rs4420638, rs56131196 and rs12721051) and *TOMM40* (rs76366838, rs114536010 and rs75687619), in the top ten SNPs. rs188535946 was in *APOC1* group but it had not been reported to be an AD risk factor, and thus warranting further investigation. rAdaSMCCA grouped rs4420638 and rs56131196 (*APOC1*) together, and their combined influence could dominate rs429358 in this model, thereby omitting rs429358. However, this combined impact on AD should also be warranted via post hoc analysis. In contrast, though SMCCA and AdaSMCCA also found out interesting SNPs, they identified too many SNPs to be well interpreted. Additionally, they both were lacking feature grouping ability and thus

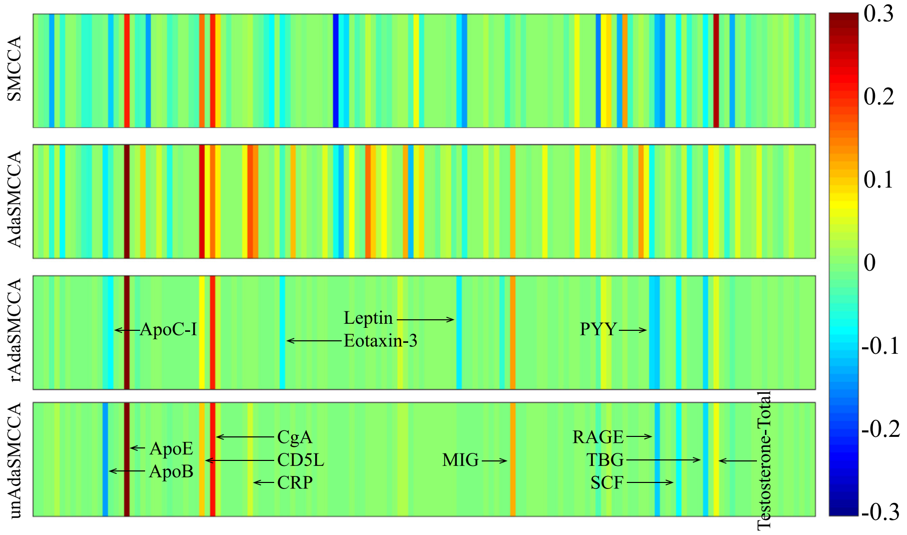

Fig. 4. Canonical weights of proteomic expressions from five-fold cross-validation trials. Each row corresponds to an SCCA method: (1) SMCCA; (2) AdaSMCCA; (3) rAdaSMCCA; and (4) unAdaSMCCA.

Table 4
Top ten proteomic markers identified by each method according to mean canonical weights. The absolute weight value showed the importance of each proteomic marker.

SMCCA		AdaSMCCA		rAdaSMCCA		unAdaSMCCA	
Protein_ID	Weight	Protein_ID	Weight	Protein_ID	Weight	Protein_ID	Weight
Testosterone-Total	0.2699	ApoE	0.4249	ApoE	0.4102	ApoE	0.5980
FSH	0.2236	CD5L	0.2384	CgA	0.2037	CgA	0.2074
CgA	0.2002	CRP	0.1737	MIG	0.1229	ApoB	0.1375
ApoE	0.1975	CgA	0.1516	RAGE	0.1080	MIG	0.1148
CD5L	0.1535	HCC-4	0.1504	PYY	0.0941	RAGE	0.1103
PAI-1	0.1530	Cystatin-C	0.1393	TBG	0.0941	CD5L	0.0970
LH	0.1401	Proinsulin-Intact	0.1307	Leptin	0.0844	TBG	0.0943
BDNF	0.1385	GH	0.1200	SCF	0.0829	SCF	0.0812
RANTES	0.1340	IGM	0.1186	Eotaxin-3	0.0825	Testosterone-Total	0.0609
PLGF	0.1297	IL-13	0.1139	ApoC-I	0.0819	CRP	0.0418

The full name of proteomic analytes is shown in the supplementary.

could not find out structure information embedded among SNPs. In summary, by well considering the impact of gradient domination, unAdaSMCCA held the best capability in identifying meaningful genetic biomarkers, and it, along with rAdaSMCCA, could identify grouping structures within SNPs such as LD.

3.2.3. Identification and interpretation of proteomic markers

Fig. 4 presented the heatmap showing the importance of proteomic markers, and Table 4 presented the top ten identified proteomic biomarkers. Both SMCCA and AdaSMCCA reported too many proteomic signals which were hard to interpret. That is, they could not discriminate relevant proteomic biomarkers from irrelevant ones. In contrast, rAdaSMCCA and unAdaSMCCA yielded sparser canonical weight patterns, showing better feature selection results than benchmarks, thereby providing a more targeted subsequent analysis. In particular, nine out of ten proteomic markers identified by unAdaSMCCA showed high correlation to AD or its prodromal stage. For example, ApoE, CgA (Ciesielski-Treska et al., 1998), ApoB (Wingo et al., 2019), MIG (Soares et al., 2012), RAGE (Deane et al., 2003), CD5L (Hye et al., 2006), SCF (Laske et al., 2011), CRP (Nilsson et al., 2011) and Testosterone-Total (Hall et al., 2015) were all demonstrated to be AD risk proteins. Only TBG has not been reported and thus further investigation should be warranted. These interesting results indicated that unAdaSMCCA could accurately find out AD-related proteomic biomarkers. Besides, rAdaSMCCA had six proteins overlapping with unAdaSMCCA, and three other ones such as PYY, APOC1 were relevant to AD. More importantly, both unAdaSMCCA and rAdaSMCCA decided clear priorities for these proteomic markers, while both benchmarks assigned very similar values for most identified proteins implying a suboptimal protein selection capability.

3.2.4. Identification and interpretation of imaging QTs

Finally, we investigated the identification of imaging QTs and showed canonical weights in Fig. 5. Both benchmarks again obtained non-sparse canonical weights while rAdaSMCCA and unAdaSMCCA yielded cleaner patterns. This indicated that our methods hold stronger feature identification ability. Table 5 contained the top ten imaging QTs. The prominent imaging QTs of unAdaSMCCA were hippocampus and parahippocampus, suggesting that structural changes to these areas were an indicator of AD. This is in line with previous studies that AD patients suffer from severe atrophy in hippocampus. unAdaSMCCA also captured signals from frontal, cingulum, temporal and cerebelum, demonstrating its power in detecting AD relevant imaging QTs, given widely reported correlations to AD pathology of these areas. rAdaSMCCA reported similar results with different priorities for imaging QTs. It additionally found out an AD-related deep area, i.e. thalamus (De Jong et al., 2008), showing its enhanced information mining ability. The top ten imaging QTs of SMCCA and AdaSMCCA were also of interest. However, they estimated such many QTs of very similar weight values that we could not confidently pick out relevant ones. This is unwelcome in practice since a clinician needs to figure out relevant imaging QTs in light of personal experi-

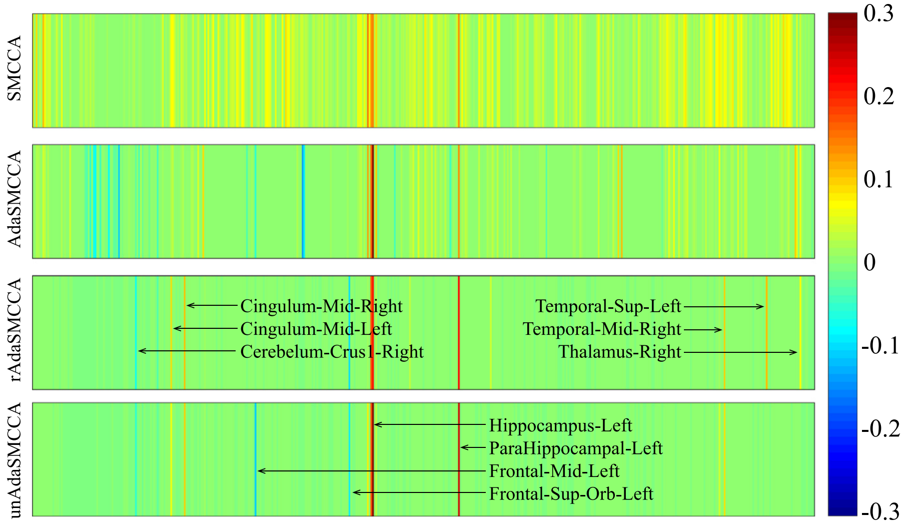

Fig. 5. Canonical weights of imaging QTs from five-fold cross-validation trials. Each row corresponds to an SCCA method: (1) SMCCA; (2) AdaSMCCA; (3) rAdaSMCCA; and (4) unAdaSMCCA.

Table 5
Top ten imaging QTs identified by each method according to mean canonical weights. The absolute weight value showed the importance of each QT.

SMCCA		AdaSMCCA		rAdaSMCCA		unAdaSMCCA	
QT_ID	Weight	QT_ID	Weight	QT_ID	Weight	QT_ID	Weight
HIP.L	0.1528	HIP.L	0.3107	PHG.L	0.2171	HIP.L	0.3535
HIP.L	0.1439	HIP.L	0.1599	HIP.L	0.2084	PHG.L	0.2526
PHG.L	0.1356	PHG.L	0.1376	HIP.L	0.1712	HIP.L	0.1693
HIP.L	0.1317	SFGdor.L	0.1324	STG.L	0.1086	MFG.L	0.1110
ANG.R	0.1079	HIP.L	0.1221	DCG.R	0.1063	DCG.R	0.1014
HIP.R	0.0884	PUT.R	0.1103	MTG.R	0.0959	MTG.R	0.0884
STG.R	0.0866	Cbe9.L	0.1074	ORBsup.L	0.0940	ORBsup.L	0.0852
ORBmid.L	0.0859	THA.L	0.0998	DCG.L	0.0878	HIP.L	0.0755
ANG.L	0.0855	HIP.R	0.0966	THA.R	0.0706	DCG.L	0.0745
MOG.L	0.0843	IFGoperc.R	0.0929	CbeCru1.R	0.0659	CbeCru1.R	0.0529

The full name of imaging QTs is shown in the supplementary.

ence. To sum up, combining results on SNPs, proteomic makers and imaging QTs together, our two AdaSMCCA methods outperformed both state-of-the-art SMCCA methods, revealing their great potential in multi-way association identification and feature selection for multi-omics data fusion.

3.3. Refined analysis

By now we have independently shown the relevance of identified SNPs, proteomic markers and imaging QTs. To explain the relationships among these three types of markers, we conducted refined analysis in this subsection. For ease of presentation, we only showed the results of unAdaSMCCA, and those of other methods can be analyzed similarly.

We first presented the pairwise correlation of SNPs and proteomic markers, as their associations irrespective of diagnosis could be a useful screening tool for AD diagnosis and intervention (Soares et al., 2012). Fig. 6 showed the heatmap of pairwise correlations of the top ten selected SNPs and proteomic markers, where blocks labeled with ”×” indicated that this SNP-protein pair reached the significance level ($p < 0.01$). When looking vertically, we can clearly observe that the monokine induced by gamma interferon (MIG) level was significantly correlated with all ten SNPs. Besides, apolipoprotein E (ApoE) also showed a significant correlation with most (nine out of ten) SNPs. When looking horizontally, rs429358 was the most noticeable locus, implying that its high correlation with plasma concentrations of ApoB, ApoE, CD5L, CgA, CRP and MIG. Another interesting thing is that rs56131196, rs4420638 and rs188535946 exhibited the same correlation pattern with proteomic markers, indicating a grouping effect of them. Recall the results in subsection 3.2.2, unAdaSMCCA grouped these

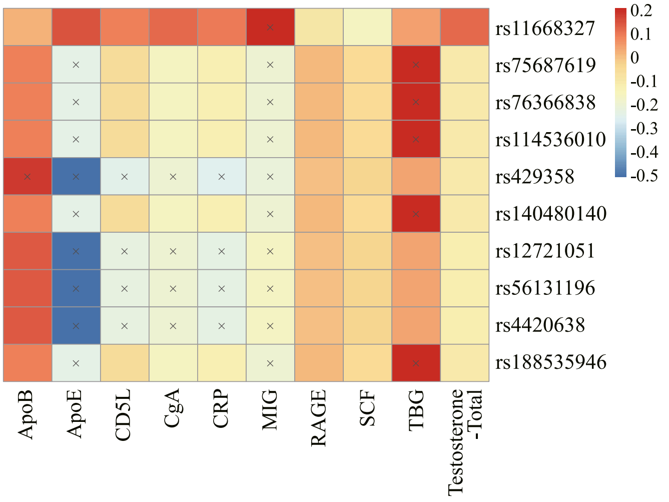

Fig. 6. Pairwise correlation between the identified SNPs and plasma proteomic analytes, and ”×” indicated that this SNP-protein pair reached the significance level ($p < 0.01$).

three loci together by assigning them the same weight values. The same correlation patterns demonstrated the excellent performance of unAdaSMCCA. These correlations irrespective of diagnosis showed that ApoE and MIG, coupled with the locus rs429358 and loci (rs56131196, rs4420638 and rs188535946) of *APOC1*, could be potential analytes for screening AD (Soares et al., 2012).

Further, we intend to uncover the impact of SNP-protein pair on brain imaging QTs. The two-way analysis of variance (ANOVA) was applied to investigate the effects of genotype, protein expression levels and their SNP-by-protein interaction on imaging QTs with age, gender, years of education and handedness being included as covariates. We looked into two SNP-protein pairs, i.e. (rs429358, ApoE) and (rs11668327, MIG), since (rs429358, ApoE) held the highest negative correlation and (rs11668327, MIG) held the highest positive correlation. In addition, since the left hippocampus was the most relevant imaging QT identified by unAdaSMCCA, we took it as the response in the ANOVA analysis. The two-way ANOVA results showed that only the main effect of rs429358 genotype ($p = 9.03^{-6}$) was significant on the left hippocampus, while the main effect of ApoE concentrations (0.35) and SNP by protein interaction (0.65) were insignificant. This is very interesting and meaningful since it indicates that although the *ApoE* genotype was significantly correlated to AD-altered imaging QTs, the ApoE concentrations were not. We also analyzed the (rs11668327, MIG)'s effect on the left hippocampus. The ANOVA results showed that the main effects of both rs11668327 genotype ($p = 2.20 \times 10^{-4}$) and MIG concentrations ($p = 0.03$) reached the significant level, but their SNP by protein interaction (0.55) was insignificant.

In summary, these statistical analysis results above confirmed the value of the multi-way association among SNPs, proteomic and neuroimaging biomarkers. In the situation where one type of biomarkers malfunction, the other types of biomarkers can be added as a supplement. This indicates that using multi-omics biomarkers could deepen our understanding of the pathogenesis of AD.

3.4. Discussion

We proposed two adaptive SMCCA methods to analyze the complicated associations among genetic, proteomics, and neuroimaging measurements. All the above results demonstrated that both rAdaSMCCA and unAdaSMCCA performed better than existing methods in identifying this complex multi-way association. The theory of rAdaSMCCA is simple since it naively uses the reciprocal of the square root of the sub-objective as an additional weight for each sub-objective. This technique has been widely used in machine learning to remove the impact of outliers and, in this paper, this can help remove the influence of outliers in terms of sub-objective. unAdaSMCCA is more complicated but its performance is the best. In practice, we suggest employing rAdaSMCCA when one faces a moderate gradient domination issue or has no idea of the gradient domination issue. In addition, if the gradient domination is severe, we suggest using unAdaSMCCA attributing to its well-designed counter-gradient-domination mechanism.

4. Conclusion

Alzheimer's disease is a multifactorial neurodegenerative disorder which could incur many abnormal alterations to the brain. Brain imaging genomics jointly analyzes genetic variations, imaging QTs and other biomarkers such as proteomic expressions. Multiple heterogeneous markers carry valuable complementary information and fusing them might yield interesting findings. However, directly fusing multiple SCCA models might be suboptimal due to undesired gradient domination. We proposed two AdaSMCCA methods, i.e. the robustness-aware AdaSMCCA and uncertainty-aware AdaSMCCA which could well address gradient domination. We also armed our methods with an automatic feature grouping penalty. An efficient algorithm is derived to solve both novel models and its convergence to a local optimum is provided.

We used both synthetic data and real data in our experiments. The conventional SMCCA (Witten and Tibshirani, 2009) and state-of-the-art one (AdaSMCCA (Hu et al., 2017)) were used as benchmarks. unAdaSMCCA not only obtained the highest CCCs, but also identified cleaner and meaningful genetic variations, proteomic markers and neuroimaging QTs. rAdaSMCCA also performed better than AdaSMCCA. In a word, all adaptive methods performed better than conventional SMCCA, demonstrating that it is the right direction to design intelligently fusing methods in multi-omics studies. Though rAdaSMCCA and unAdaSMCCA were proposed for imaging genetics, they can also be applied to other real applications such as analyzing the relationship among multimodal brain imaging data. In addition, the statistical analysis demonstrated the identification capability of the proposed methods. As a solid fusion strategy, it is interesting to apply our AdaSMCCA to genome wide association study, or to include more than three types of biomarkers.

Declaration of Competing Interest

None.

CRediT authorship contribution statement

Lei Du: Conceptualization, Methodology, Writing - original draft. **Jin Zhang:** Software, Writing - review & editing. **Fang Liu:** Software, Visualization, Investigation. **Huiai Wang:** Validation, Writing - review & editing. **Lei Guo:** Conceptualization. **Junwei Han:** Writing - review & editing.

Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. HoffmannLa Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

This work was supported by the National Natural Science Foundation of China [61973255, 61936007, 61602384]; Natural Science Basic Research Program of Shaanxi [2020JM-142]; the China Postdoctoral Science Foundation [2020T130537] at Northwestern Polytechnical University. This work was also supported by the Shanghai Municipal Science and Technology Major Project [2018SHZDZX01] at LCNBI and ZJLab.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.media.2021.102003

References

Association, A., 2019. 2019 Alzheimer's disease facts and figures. Alzheimer's & Dementia 15 (3), 321–387.

Bi, X.-a., Hu, X., Wu, H., Wang, Y., 2020. Multimodal data analysis of alzheimer's disease based on clustering evolutionary random forest. IEEE J Biomed Health Inform 24 (10), 2973–2983.

Bi, X.-a., Liu, Y., Xie, Y., Hu, X., Jiang, Q., 2020. Morbigenous brain region and gene detection with a genetically evolved random neural network cluster approach in late mild cognitive impairment. Bioinformatics 36 (8), 2561–2568.

Chen, M., Gao, C., Ren, Z., Zhou, H.H., 2013. Sparse cca via precision adjusted iterative thresholding. arXiv preprint arXiv:1311.6186.

Ciesielski-Treska, J., Ulrich, G., Taupenot, L., Chasserot-Golaz, S., Corti, A., Aunis, D., Bader, M.-F., 1998. Chromogranin a induces a neurotoxic phenotype in brain microglial cells. J. Biol. Chem. 273 (23), 14339–14346.

De Jong, L.W., Der Hiele, K.V., Veer, I.M., Houwing, J.J., Westendorp, R.G.J., Bollen, E.L.E.M., De Bruin, P.W., Middelkoop, H.A.M., Van Buchem, M.A., Der Grond, J.V., 2008. Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: an MRI study. Brain 131 (12), 3277–3285.

Deane, R., S, D.Y., Submamaryan, R.K., Larue, B., Jovanovic, S., Hogg, E., Welch, D., Manness, L., Lin, C., Yu, J., et al., 2003. Rage mediates amyloid-beta peptide transport across the blood-brain barrier and accumulation in brain.. Nat. Med. 9 (7), 907–913.

Du, L., Huang, H., Yan, J., Kim, S., Risacher, S.L., et al., 2016. Structured sparse canonical correlation analysis for brain imaging genetics: an improved graphnet method. Bioinformatics 32 (10), 1544–1551.

Du, L., Liu, F., Liu, K., Yao, X., Risacher, S.L., Han, J., Guo, L., Saykin, A.J., Shen, L., 2020. Identifying diagnosis-specific genotype-phenotype associations via joint multi--task sparse canonical correlation analysis and classification. Bioinformatics 36 (S1), i371–i379.

Du, L., Liu, F., Liu, K., Yao, X., Risacher, S.L., Han, J., Saykin, A.J., Shen, L., 2020. Associating multi-modal brain imaging phenotypes and genetic risk factors via a dirty multi-task learning method. IEEE Trans Med Imaging 39 (11), 3416–3428. doi:10.1109/TMI.2020.2995510.

Du, L., Liu, K., Yao, X., Risacher, S., Han, J., Saykin, A., Guo, L., Shen, L., 2021. Multi–task sparse canonical correlation analysis with application to multi-modal brain imaging genetics. IEEE/ACM Trans. Comput. Biol. Bioinf. 18 (1), 227–239.

Du, L., Liu, K., Yao, X., Risacher, S.L., Han, J., Saykin, A.J., Guo, L., Shen, L., 2020. Detecting genetic associations with brain imaging phenotypes in Alzheimer's disease via a novel structured SCCA approach. Med Image Anal 61, 101656.

Du, L., Liu, K., Zhang, T., Yao, X., Yan, J., Risacher, S.L., Han, J., Guo, L., Saykin, A.J., Shen, L., 2018. A novel SCCA approach via truncated ℓ_1-norm and truncated group lasso for brain imaging genetics. Bioinformatics 34 (2), 278–285.

Du, L., Yan, J., Kim, S., Risacher, S.L., et al., 2014. A novel structure-aware sparse learning algorithm for brain imaging genetics. In: International Conference on Medical Image Computing and Computer Assisted Intervention, pp. 329–336.

Fan, C., Cheng, Y., Gou, H., Liu, C., Deng, S., Liu, C., Chen, X., Bu, J., Zhang, X., 2020. Neuroimaging and intervening in memory reconsolidation of human drug addiction. Science China Information Sciences 63 (7), 1–11.

Fang, J., Lin, D., Schulz, S.C., Xu, Z., Calhoun, V.D., Wang, Y., 2016. Joint sparse canonical correlation analysis for detecting differential imaging genetics modules.. Bioinformatics 32 (22), 3480–3488.

Feldman, D.E., McPherson, K.L., Biesecker, C.L., Wiers, C.E., Manza, P., Volkow, N.D., Wang, G.-J., 2020. Neuroimaging of inflammation in alcohol use disorder: a review. Science China Information Sciences 63 (7), 1–19.

Gao, H., Nie, F., Cai, W., Huang, H., 2015. Robust capped norm nonnegative matrix factorization. In: the 24th ACM International on Conference on Information and Knowledge Management, p. 871C880.

Gao, L., Cui, Z., Shen, L., Ji, H.-F., 2016. Shared genetic etiology between type 2 diabetes and Alzheimer's disease identified by bioinformatics analysis. J. Alzheimers Dis. 50 (1), 13–17.

Gupta, V., Laws, S.M., Villemagne, V.L., Ames, D., Bush, A.I., Ellis, K.A., Lui, J.K., Masters, C., Rowe, C.C., Szoeke, C., et al., 2011. Plasma apolipoprotein e and Alzheimer disease risk: the aibl study of aging. Neurology 76 (12), 1091–1098.

Hall, J., Wiechmann, A., Cunningham, R.L., Johnson, L.A., Edwards, M., Barber, R., Singh, M., Winter, S., Obryant, S.E., 2015. Total testosterone and neuropsychiatric symptoms in elderly men with Alzheimer's disease. Alzheimer's Research & Therapy 7 (1), 24.

Hu, W., Lin, D., Cao, S., Liu, J., Chen, J., Calhoun, V.D., Wang, Y.-P., 2017. Adaptive sparse multiple canonical correlation analysis with application to imaging (epi) genomics study of schizophrenia. IEEE Trans. Biomed. Eng. 65 (2), 390–399.

Hye, A., Lynham, S., Thambisetty, M., Causevic, M., Campbell, J., Byers, H.L., Hooper, C., Rijsdijk, F., Tabrizi, S.J., Banner, S., et al., 2006. Proteome-based plasma biomarkers for Alzheimer's disease. Brain 129 (11), 3042–3050.

Kendall, A., Gal, Y., Cipolla, R., 2018. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 7482–7491.

Laske, C., Sopova, K., Hoffmann, N., Stransky, E., Hagen, K., Fallgatter, A.J., Stellos, K., Leyhe, T., 2011. Stem cell factor plasma levels are decreased in Alzheimer's disease patients with fast cognitive decline after one-year follow-up period: the pythia-study. J. Alzheimers Dis. 26 (1), 39–45.

Lin, D., Calhoun, V.D., Wang, Y., 2014. Correspondence between fmri and snp data by group sparse canonical correlation analysis. Med Image Anal 18 (6), 891–902.

Nilsson, K., Gustafson, L., Hultberg, B., 2011. C-Reactive protein level is decreased in patients with Alzheimer's disease and related to cognitive function and survival time.. Clin. Biochem. 44 (14), 1205–1208.

Reich, D.E., Cargill, M., Bolk, S., Ireland, J., Sabeti, P.C., Richter, D.J., Lavery, T., Kouyoumjian, R., Farhadian, S.F., Ward, R., et al., 2001. Linkage disequilibrium in the human genome. Nature 411 (6834), 199.

Shen, L., Thompson, P.M., 2020. Brain imaging genomics: integrated analysis and machine learning. Proc. IEEE 108 (1), 125–162.

Soares, H., Potter, W.Z., Pickering, E.H., Kuhn, M., Immermann, F.W., Shera, D., Ferm, M., Dean, R.A., Simon, A.J., Swenson, F., et al., 2012. Plasma biomarkers associated with the apolipoprotein e genotype and Alzheimer disease. JAMA Neurol 69 (10), 1310–1317.

Wang, H., Nie, F., Huang, H., Risacher, S.L., Saykin, A.J., Shen, L., Initiative, A.D.N., 2012. Identifying disease sensitive and quantitative trait-relevant biomarkers from multidimensional heterogeneous imaging genetics data via sparse multimodal multitask learning. Bioinformatics 28 (12), i127–i136.

Wingo, T.S., Cutler, D.J., Wingo, A.P., Le, N.-A., Rabinovici, G.D., Miller, B.L., Lah, J.J., Levey, A.I., 2019. Association of early-onset alzheimer disease with elevated low-density lipoprotein cholesterol levels and rare genetic coding variants of APOB. JAMA Neurol 76 (7), 809–817.

Witten, D.M., Tibshirani, R., Hastie, T., 2009. A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. Biostatistics 10 (3), 515–534.

Witten, D.M., Tibshirani, R.J., 2009. Extensions of sparse canonical correlation analysis with applications to genomic data. Stat Appl Genet Mol Biol 8 (1), 1–27.

Yan, J., Risacher, S.L., Nho, K., Saykin, A.J., Shen, L., 2017. Identification of discriminative imaging proteomics associations in Alzheimer's disease via a novel sparse correlation model. In: Pacific Symposium on Biocomputing 2017. World Scientific, pp. 94–104.

Zhou, X., Chen, Y., Mok, K.Y., Kwok, T.C., Mok, V.C., Guo, Q., Ip, F.C., Chen, Y., Mullapudi, N., Giusti-Rodríguez, P., et al., 2019. Non-coding variability at the apoe locus contributes to the Alzheimer's risk. Nat Commun 10 (1), 1–16.

Zille, P., Calhoun, V.D., Wang, Y., 2018. Enforcing co-expression within a brain-imaging genomics regression framework. IEEE Trans Med Imaging 37 (12), 2561–2571.

Update

Medical Image Analysis

Volume 72, Issue , August 2021, Page

DOI: https://doi.org/10.1016/j.media.2021.102081

Medical Image Analysis 72 (2021) 102081

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Erratum

Corrigendum to Identifying associations among genomic, proteomic and imaging biomarkers via adaptive sparse multi-view canonical correlation analysis [Medical Image Analysis 70 (2021) 1–12/102003]

Lei Du*, Jin Zhang, Fang Liu, Huiai Wang, Lei Guo, Junwei Han

School of Automation, Northwestern Polytechnical University

Dear Sir or Madam,

Sorry to bother you.

I am writing to ask for a Corrigendum regarding our recently accepted paper (Reference No. MEDIMA 102003. Title: Identifying Associations among Genomic, Proteomic and Imaging Biomarkers via Adaptive Sparse Multi-view Canonical Correlation Analysis). I find that we have missed one funding acknowledgments (National key R&D Program of China [2017YFB1002201]) in the ACKNOWLEDGEMENTS section, which is very important to me. So, may I ask that could you please do me a favour to add this funding information? For your convenience, I have copied the full contents listed below, and if possible, you could just replace the second paragraph of the ACKNOWLEDGEMENTS section by this new content below. The text with the yellow background is the newly added text.

The authors acknowledge that This work was supported in part by the National Key R&D Program of China [2017YFB1002201]; National Natural Science Foundation of China [61973255, 61936007, 61602384]; Natural Science Basic Research Program of Shaanxi [2020JM-142]; the China Postdoctoral Science Foundation [2020T130537] at Northwestern Polytechnical University.

The authors would like to apologise for any inconvenience caused.

DOI of original article: 10.1016/j.media.2021.102003

* Corresponding author.

E-mail address: dulei@nwpu.edu.cn (L. Du).

https://doi.org/10.1016/j.media.2021.102081