Adaptive Signal Recovery on Graphs via Harmonic Analysis for Experimental Design in Neuroimaging

Won Hwa Kim[†], Seong Jae Hwang[†], Nagesh Adluru[‡],
Sterling C. Johnson[*], and Vikas Singh[§†]

[†]Dept. of Computer Sciences, Univ. of Wisconsin, Madison, WI, U.S.A.
[§]Dept. of Biostatistics & Med. Informatics, Univ. of Wisconsin, Madison, WI, U.S.A.
[*]GRECC, William S. Middleton VA Hospital, Madison, WI, U.S.A.
[‡]Waisman Center, Madison, WI, U.S.A.
http://pages.cs.wisc.edu/∼wonhwa

1 Introduction

In this supplement, we present details of the experimental setup and additional results that were mentioned but not described in the main paper.

2 Additional Experiments with the HCP Study

In this section, we explain more details for our study of the Human Connectome Project (HCP), such as the list of covariates used to generate graphs and provide additional experimental results using various different sampling ratios that were not presented in the main paper. As mentioned in the main text, the goal of this experiment is to utilize economical measurements and a partial observation on expensive measurements to predict the expensive measurements on the full cohort. A full list of 27 non-imaging covariates (i.e., economical measurements) that were used to derive a graph is presented in Table 1. These covariates are known to be closely associated with brain function, and they have a high chance of being correlated with the FA measurements (i.e., expensive measurements). Therefore they are decent ingredients to construct a graph that may resemble underlying structures in the FA data.

Using the graph, we applied our framework to generate sampling probability distribution and predicted the FA values. The prediction involves sampling procedure of measurements for the partial observation of FA measurements. We studied three cases with 20%, 40% and 60% of sampling from the full data and computed the average of estimation across all subjects at each ROI. The result using 40% is presented in the paper, and the remaining 20% and 60% cases are presented here in Fig. 1. As seen in Fig. 1, our estimation (in blue) is much closer to the ground truth (in red) than those from other methods (in green and orange). Notice that our estimation is quite close to the ground truth even in the 20% case, while other methods need larger sample size to reduce the errors.

3 Additional Experiments with the Preclinical AD Study

In this section, we describe more details for our Wisconsin Registry for Alzheimer's Prevention study and provide additional experimental results that were not presented in the paper. Again, the goal here is to make use of economical data (i.e., FA measurements) on the full cohort and partial observations on expensive data (i.e., PiB measurements) from a subset of the 140 participants. Initially, brain networks from 140 subjects were given, where each network was represented as an adjacency matrix of size 162×162 using 162 ROIs. An example of the adjacency matrix is given in Fig. 2, where each element in the adjacency matrix represents the mean of FA values defined over the fibers (consisting a pathway) connecting two different ROIs. There were total of 1704 pathways that were non-zero in the adjacency matrix, and we selected 1% of the total pathways (i.e., 17 pathways) that were considered to be stable with least variations across all subjects to construct a graph.

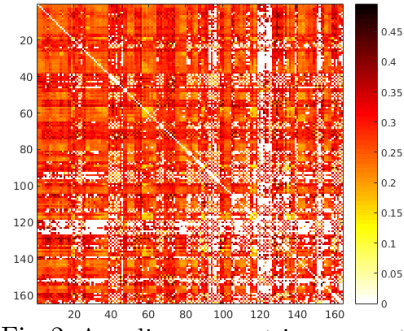

Fig. 2: An adjacency matrix representing brain network. Each element is a mean of FA values defined on fibers that are connecting two different ROIs.

Category	**Covariate Name**
Demographics	Age, gender, years of education completed (Edu)
Physical health	Height, weight, BMI
Alertness	Mini mental status exam (MMSE)
Sleep	Pittsburgh sleep questionnaire (PSQI)
Episodic memory	Picture sequence recall (PicSeq)
Cognitive flexibility	Picture matching accuracy and reaction time (CardSort)
Inhibition	Flanking accuracy and reaction time (Flanker)
Fluid intelligence	Correct responses (PMAT_CR),
(Penn progressive matrices)	response time for correct responses (PMAT_RTCR),
	total skipped items (PMAT_SI)
Reading	NIH toolbox reading recognition test (ReadEng)
Vocabulary	NIH toolbox picture vocabulary (PicVocab)
Processing speed	NIH toolbox pattern comparison speed (ProcSpeed)
Spatial orientation	Expected number of correct clicks (VSPLOT_CRTE),
	total off from correct positions (VSPLOT_OFF),
	total count of correct clicks (VSPLOT_TC)
Sustained attention	Short Penn continuous performance test:
	sensitivity (SCPT_SEN), specificity (SCPT_SPEC),
	longest run of non-responses (SCPT_LRNR)
Episodic memory	Penn word memory test: total correct responses
	(IWRD_TOT), reaction time (IWRD_RTC)
Working memory	NIH toolbox sorting working memory (ListSort)

Table 1: Full list of non-imaging covariates used in our analysis spanning a wide range high-level human behavior and highly relevant physiological measurements.

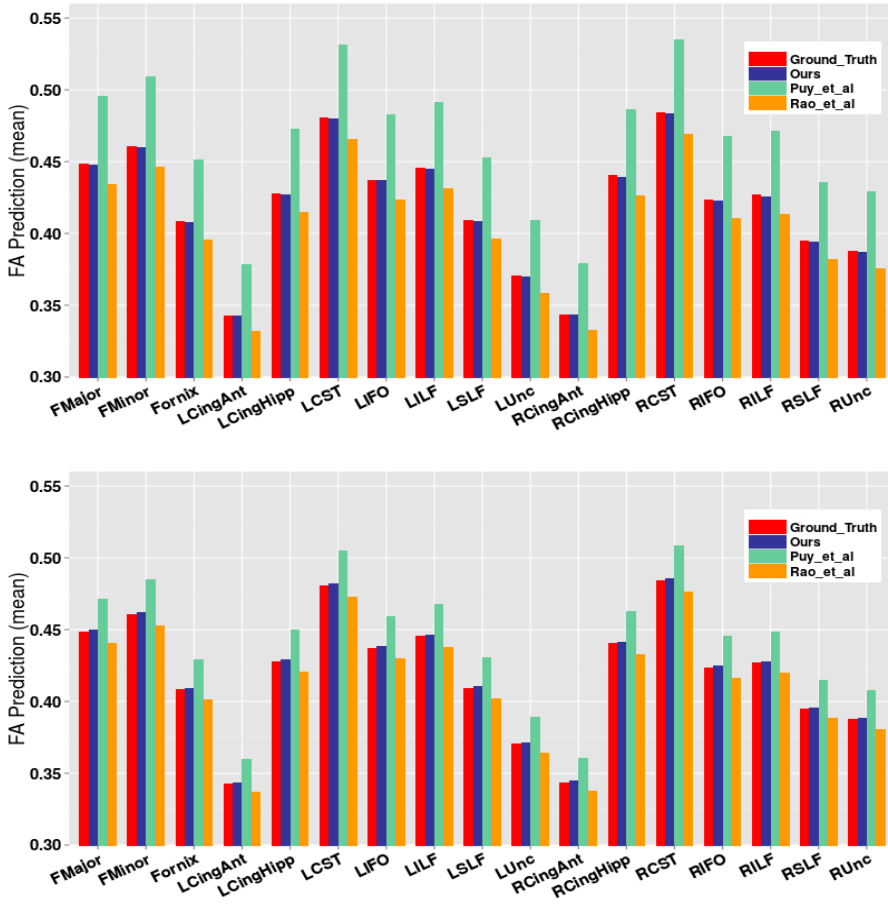

Fig. 1: Estimation of the FA measurements at 17 different ROIs (i.e. Pathways). Top: using 20% of total samples (97 subjects), Bottom: using 60% of total samples (292 subjects). When compared to the ground truth (in red), estimation using our method (in blue) is more accurate than those using Puy et al. (in green) or Rao et al. (in orange) in both cases.

Using the constructed graph, we applied our framework to predict the PiB values of the full cohort based on partial observations. For the amount of partial observation, three different cases were studied with 20%, 40% and 60% of sampling from the full data. The average of estimation across all subjects at each ROI was computed to evaluate the performance of our method and others. The result using 40% is presented in the main paper, therefore we present the remaining 20% and 60% cases here in Fig. 3. As seen in Fig. 3, our estimation (in blue) and the ground truth (in red) are similar while other methods (in green and orange) either overestimate or underestimate the PiB measurements with larger errors. Notice that our estimation is quite close to the ground truth even in the 20% case, while other methods need larger sample size to reduce the er-

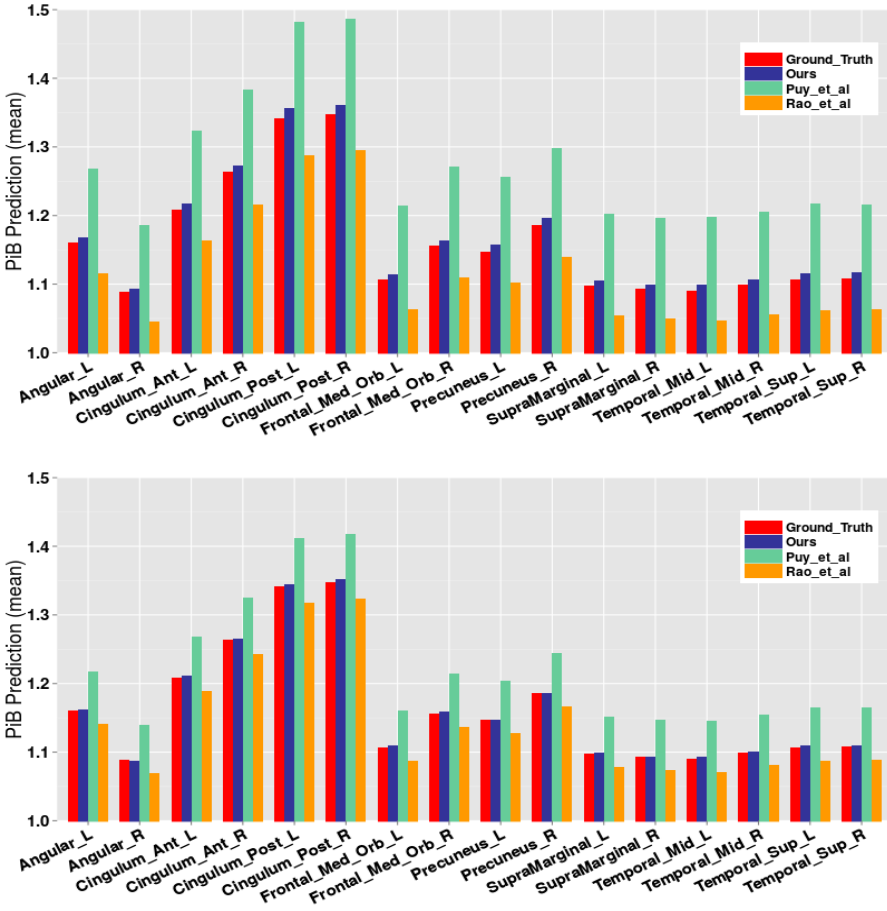

Fig. 3: Estimation of the PiB measurements at 16 different ROIs. Top: using 20% of total samples (28 subjects), Bottom: using 60% of total samples (84 subjects). When compared to the ground truth (in red), estimation using our method (in blue) is more accurate than those using Puy et al. (in green) or Rao et al. (in orange) in both cases.

rors. The cost of acquiring PiB DVR measurements can cost up to 10 times of that for requiring FA measurements, but we are able to successfully estimate the PiB DVR measurements of the entire participants using only 20% of the PiB DVR data and the cheap FA values.