

Supplementary material: Latent Variable Graphical Model Selection using Harmonic Analysis: Applications to the Human Connectome Project (HCP)

1. Summary

We provide proofs of lemmas and technical details for implementation. Also, additional experimental results are provided which were not included in the main paper due to limited space.

2. Proofs of Lemmas

Lemma 1 *If $\Theta \succ 0$, $\Theta = \Theta^T$ and kernel g satisfies the admissibility condition*

$$\int_0^\infty \frac{g^2(s\sigma)}{s} ds =: C_g < \infty \quad (1)$$

then,

$$\frac{1}{C_g} \int_0^\infty \frac{1}{s} \sum_{\ell=1}^n W_{\Theta,s}(\ell) \psi_{\ell,s}(i,j) ds = \Theta(i,j) \quad (2)$$

Proof: Note that since $\int_0^\infty \frac{g^2(s\sigma)}{s} ds = \int_0^\infty \frac{g^2(x)}{x} dx =: C_g$, the admissibility condition can be written in both ways. Using the definition of $\psi_{\ell,s}$ and $W_{\Theta,s}(\ell)$ for the reconstruction of Θ ,

$$\begin{aligned} & \frac{1}{C_g} \int_0^\infty \frac{1}{s} \sum_{\ell=1}^n W_{\Theta,s}(\ell) \psi_{\ell,s}(i,j) ds \\ &= \frac{1}{C_g} \int_0^\infty \frac{1}{s} \left(\sum_{\ell=1}^n \sigma_\ell g(s\sigma_\ell) g(s\sigma_\ell) V_\ell(i) V_\ell(j) \right) ds \\ &= \frac{1}{C_g} \int_0^\infty \frac{1}{s} \left(\sum_{\ell=1}^n \sigma_\ell g^2(s\sigma_\ell) V_\ell(i) V_\ell(j) \right) ds \\ &= \sum_{\ell=1}^n \left(\frac{1}{C_g} \int_0^\infty \frac{g^2(s\sigma_\ell)}{s} ds \right) \sigma_\ell V_\ell(i) V_\ell(j) \\ &= \sum_{\ell=1}^n \sigma_\ell V_\ell(i) V_\ell(j) = \Theta(i,j) \text{ by the admissibility condition in (1)} \end{aligned}$$

□

Now, we introduce a result for our inverse covariance estimation with a two parameter kernel function $g(s, \sigma)$. Analogously, we can define a basis function $\psi_{\ell,s}$ and wavelet-like coefficients $W_{\Theta,s}$ as (9) and (10) in the main paper, namely,

$$\psi_{\ell,s} = g(s, \sigma_\ell) V_\ell^*(i) V_\ell(j) \text{ and } W_{\Theta,s}(\ell) = \sigma_\ell g(s, \sigma_\ell), \forall \ell \in \{1, \dots, n\}. \quad (3)$$

Then, we can show the admissibility condition for two parameter kernels for the reconstructing the graph structure (a function on edges) represented by the precision matrix as in Lemma 2

Lemma 2 If $\Theta \succ 0$, $\Theta = \Theta^T$ and kernel g satisfies the admissibility condition

$$\int_0^\infty \frac{g^2(s, \sigma)}{s} ds =: C_g < \infty \quad (4)$$

then,

$$\frac{1}{C_g} \int_0^\infty \frac{1}{s} \sum_{\ell=1}^n W_{\Theta, s}(\ell) \psi_{\ell, s}(i, j) ds = \Theta(i, j) \quad (5)$$

This corollary can be proved in the same way as Lemma 1. This result shows that two parameter kernels can be used for the reconstruction of both the graph *structure* (functions on edges) as well as in the classical setting (functions on the nodes) with a well-defined admissibility condition. This gives the ability to define more flexible kernels if needed.

Similar to Lemma 2, we provide the admissibility condition for two parameter kernels for the classical SGWT. We follow the notations of [1].

Lemma 3 If the kernel g satisfies the admissibility condition

$$\int_0^\infty \frac{g^2(s, \sigma)}{s} ds =: C_g < \infty \quad (6)$$

then,

$$\frac{1}{C_g} \int_0^\infty \frac{1}{s} \sum_{n=1}^N W_f(s, n) \psi_{s, n}(m) ds = f^\#(m) \quad (7)$$

where $f^\# = f - \langle \chi_0, f \rangle_{\chi_0}$ if there exists a zero eigenvalue. When all eigenvalues are strictly positive, $f^\# = f$.

Proof: By the definition of $\psi_{s, n}$ and $W_f(s, n)$ with graph Fourier bases, the left hand side of (7) is given as

$$\begin{aligned} & \frac{1}{C_g} \int_0^\infty \frac{1}{s} \sum_{n=1}^N \left(\sum_{l=1}^N g(s, \sigma_l) \chi_l(n) \hat{f}(l) \sum_{l'=1}^N g(s, \sigma_{l'}) \chi_{l'}^*(n) \chi_{l'}(m) \right) ds \\ &= \frac{1}{C_g} \int_0^\infty \frac{1}{s} \left(\sum_{l, l'} g^2(s, \sigma_l) \hat{f}(l) \chi_{l'}(m) \sum_{n=1}^N \chi_{l'}^*(n) \chi_l(n) \right) ds \\ &\stackrel{(a)}{=} \frac{1}{C_g} \int_0^\infty \frac{1}{s} \left(\sum_{l=1}^N g^2(s, \sigma_l) \hat{f}(l) \chi_l(m) \right) ds \\ &= \frac{1}{C_g} \sum_{l=1}^N \left(\int_0^\infty \frac{g^2(s, \sigma_l)}{s} ds \right) \hat{f}(l) \chi_l(m) \\ &= \sum_{l=1}^N \hat{f}(l) \chi_l(m) \text{ by (6)} \\ &= f^\#, \end{aligned}$$

Equality (a) holds since $\sum_n \chi_{l'}^*(n) \chi_l(n) = \delta_{l, l'}$ (Kronecker's delta) with orthonormal basis χ_l . □

3. Choice of the Kernel Function

For the choice of kernel function $g()$, we used the popular Gaussian function $\exp(-\frac{1}{2}sx)$. This kernel function models diffusion or a random walk process, and is used to define diffusion type of wavelets. The kernel function itself may not satisfy the admissibility condition, which is not an issue because we work with a single scale. However, to work with all scales concurrently, we will be limited to only that class of kernels which directly satisfies the admissibility condition.

$$K(s, \sigma) = \sigma g^2(s\sigma) = \sigma e^{-s\sigma} \quad (8)$$

is able to perfectly reconstruct the original sample precision matrix Θ . Intuitively, in our optimization problem, we should be able to reconstruct the exact Θ in the non-regularized setting. This is because of σ in front of $g(\cdot)$ in (8). That is, our estimation $\tilde{\Theta}$

$$\tilde{\Theta} = \sum_l \sigma_l g^2(s\sigma_l) V_l V_l' \quad (9)$$

becomes P when $s = 0$ as

$$\tilde{\Theta} = \sum_l \sigma_l e^{-s\sigma_l} V_l V_l' \quad (10)$$

$$= \sum_l \sigma_l V_l V_l' = \Theta. \quad (11)$$

4. Details of Our Experiments

4.1. Experimental Setup and Additional Analysis of Synthetic Pathways

To provide more details for the experiments with synthetic brain pathways, we first demonstrate how the data were generated as in Fig. 1. First, arbitrary relations between the 50 connectivity pathways were randomly generated assuming and their relation with the 10 covariates were defined as shown in Fig. 1 a). Then, 10 latent variables were added with random associations with the observed variables (i.e., pathways and covariates) as in Fig. 1 b), which is a precision matrix with all variables and demonstrates the effects from the latent variables to our observation. Finally, a covariance matrix was derived from the precision matrix as in Fig. 1 c). The covariance matrix was used to generate 2000 multi-variate Gaussian samples. For our algorithm, we chose the kernel function $g(\cdot)$ as in section 3, and $\gamma = 2$ for the sparsity parameter which yields a good

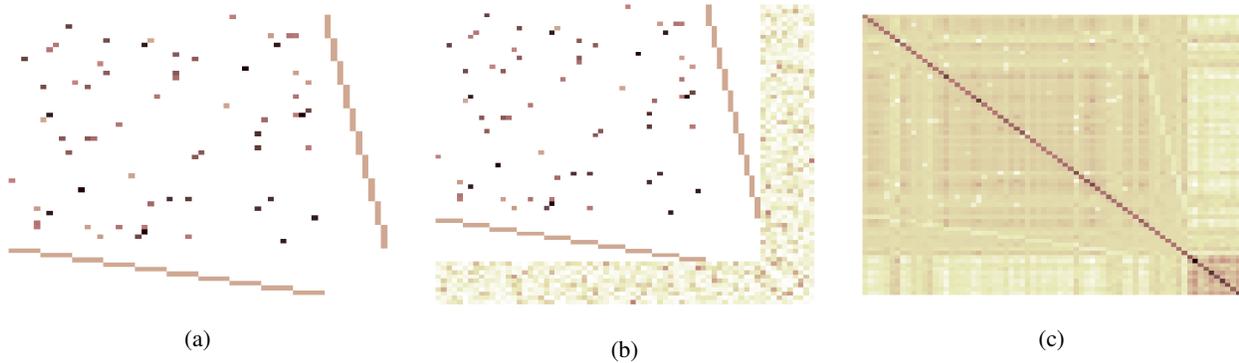


Figure 1: a) Precision with 50 observed variables only (Ground truth), b) Precision matrix with both 50 observed and 10 latent variables, c) Covariance matrix derived from b).

result as shown in the main paper.

In the following, we demonstrate additional analysis of precision matrix estimation. In Fig. 2, results from experiments with different number of latent variables (i.e., 0, 5, 10, 20) are displayed. The top row in Fig. 2 shows the inverse covariance matrix and the bottom row shows our result. At a glance, we can easily see that the increase in the number of latent variables makes the inverse covariance matrix denser, on the other hand, our estimation results yield correct sparse graphical models despite the increase. In Fig. 3, we show the precision matrix estimation in different scales with 5 latent variables. Fig. 3 a), b) and c) are the estimation results with $s = 0, 0.5, 0.7$, and d) shows a result with the optimal scale $s = 0.2089$ obtained from our optimization scheme. The estimation result significantly varies depending on the scale parameter with dense non-zero elements, but our optimization is able to find the optimal scale that provides an estimation that is sparse and similar to the inverse covariance matrix.

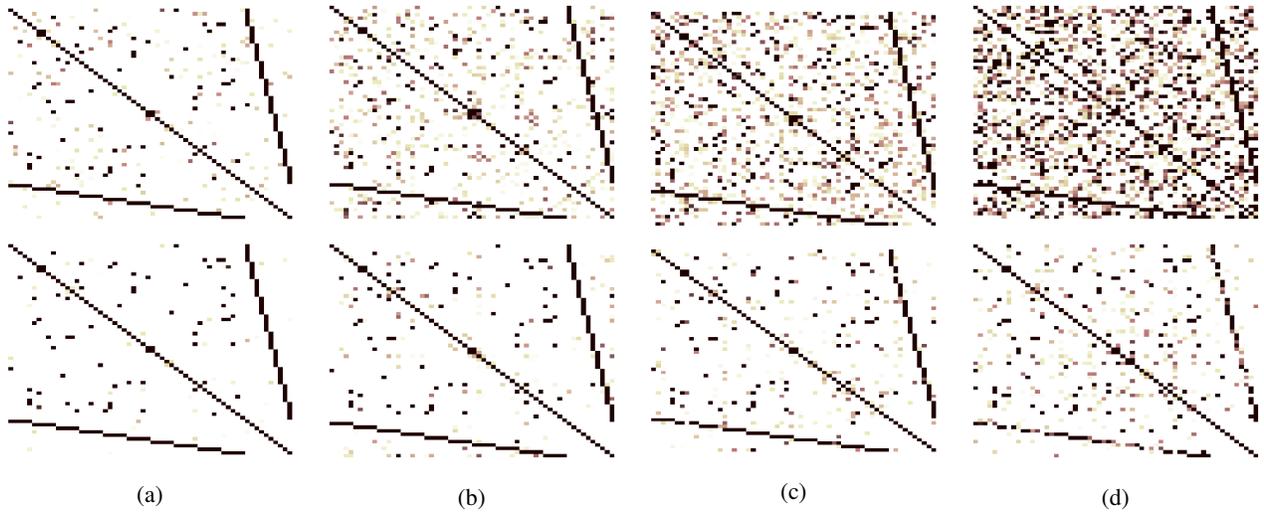


Figure 2: Precision matrix estimation with different numbers of latent variables. Top row shows the inverse covariance matrix, and the bottom row shows our estimation result. a) no latent variables, b) 5 latent variables, c) 10 latent variables, d) 20 latent variables. We can easily see that the inverse covariance matrix becomes denser as the number of latent variables increases, while our method yields good estimation of the sparse graphical model.

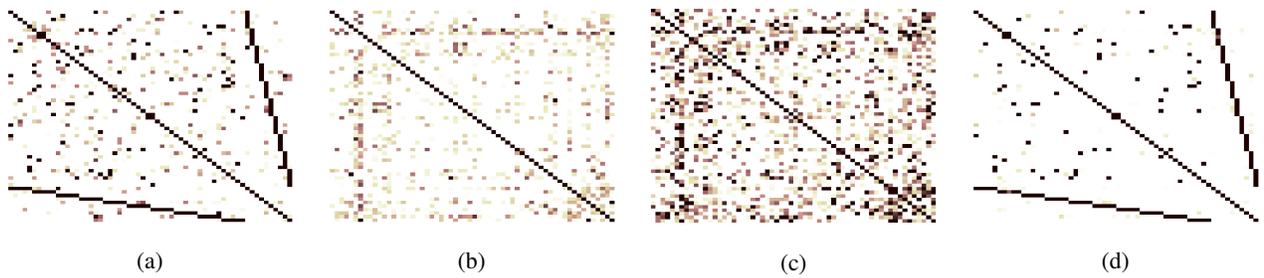


Figure 3: Precision matrix with 5 latent variables in different scales. a) estimation with $s = 0$, b) estimation with $s = 0.5$, c) estimation with $s = 0.7$, d) estimation with optimal scale $s = 0.2089$. Our optimization scheme find the optimal scale that gives a sparse graphical model.

4.2. Details of non-Imaging Covariates

In this section, we provide a full list of non-imaging covariates that we used in our HCP brain connectivity pathways analysis. The list consists of various measurements from physical, mental, and cognitive status of participants, and are commonly used for analysis of structural and functional behavior of brains.

References

- [1] D. K. Hammond, P. Vandergheynst, and R. Gribonval. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2):129–150, 2011. [2](#)

Category	Covariate name
Demographics	Age, gender, years of education completed (Edu)
Physical health	Height, weight
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 in Penn progressive matrices (PMA_CR)
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 (VSPLIT_CRTE), total off positions (VSPLIT_OFF)
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.