1 4-component graph

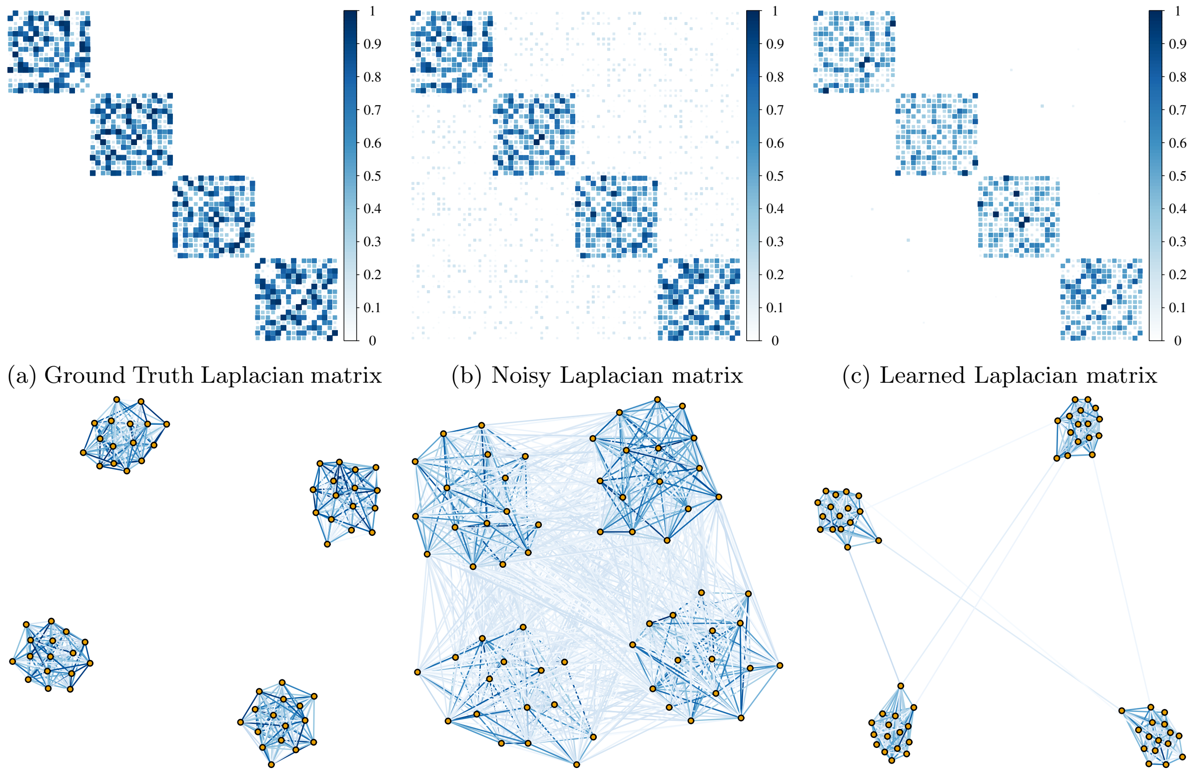
An experiment with synthetic data generated by a 4-component graph was conducted. We consider the following model:

$$\mathbf{L}_{\mathsf{noisy}} = \kappa \mathbf{L}_{\mathsf{true}} + (1-\kappa)\mathbf{L}_{\mathsf{ER}}, \tag{1}$$

where $\mathbf{L}_{\mathsf{true}}$ represents the Laplacian matrix of a K-component graph (for this example, $K=4$) denoted as $\mathcal{G}_K^{(p_1,p_2)}$, in which p_1 and p_2 represent the probabilities of node connections across components and within components, respectively; $\mathbf{L}_{\mathsf{ER}}$ represents the Laplacian of an Erdos-Renyi graph $\mathcal{G}_{\mathsf{ER}}^{(p)}$, in which p is the probability of a node connecting to any other node; and $\kappa \in (0,1)$ controls how much noise is added into $\mathbf{L}_{\mathsf{true}}$ by the Erdos-Renyi model. Additionally, the weighted edges of both $\mathcal{G}_K^{(p_1,p_2)}$ and $\mathcal{G}_{\mathsf{ER}}^{(p)}$ were drawn from $\mathsf{Uniform}(0,1)$. Finally, we set $p_1 = 0$, $p_2 = 1$, $p = 0.35$, and $\kappa = 0.8$.

Then, data were sampled in the form of $\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, \mathbf{L}_{\mathsf{noisy}}^{\dagger})$, where $\mathbf{A}^{\dagger}$ denotes the generalized inverse of the matrix $\mathbf{A}$. The total number of nodes N and the number of drawn samples T were set to $N = 64$ (16 nodes per component) and $T/N = 30$.

Figure 1 illustrates the ground truth model, its noisy version, and the model learned by our spectral topology algorithm with $\beta = 10$. We compute the performance of the learning process by means of the relative error (RE) and the F-score (FS). For this example, our algorithm achives $(\mathsf{RE}, \mathsf{FS}) = (0.131, 0.993)$ which means almost perfect clustering accuracy even in a noisy model that heavily suppress the ground truth weights when κ is much larger than zero.



(a) Ground Truth Laplacian matrix

(b) Noisy Laplacian matrix

(c) Learned Laplacian matrix

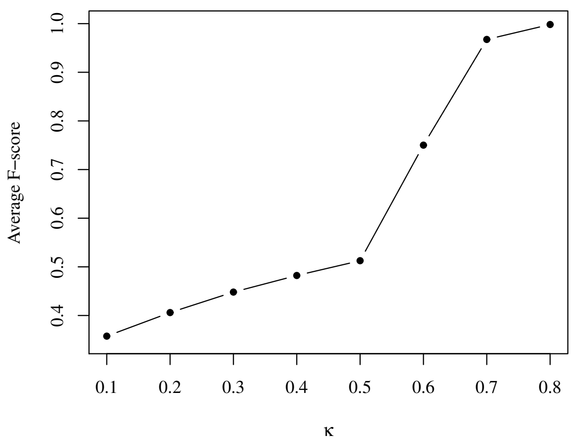
(d) Ground Truth graph

(e) Noisy graph

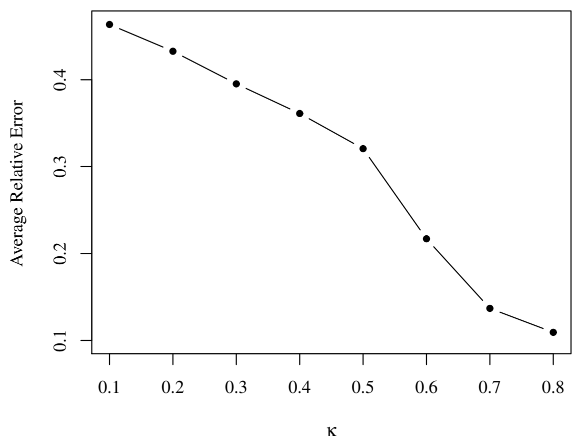
(f) Learned graph

Figure 1: An example of estimating a 4-component graph. (a) the ground truth graph Laplacian matrix ($\mathbf{L}_{\mathsf{true}}$), (b) $\mathbf{L}_{\mathsf{true}}$ after being corrupted by noise, (c) the learned graph Laplacian with a performance of $(\mathsf{RE}, \mathsf{FS}) = (0.131, 0.993)$. The panels (d), (e), and (f) correspond to the graphs represented by the Laplacian matrices in (a), (b), and (c), respectively.

Figure 2 shows the performance of our algorithm, using the same settings as the experiment above, but for different noise regimes.



(a) Average F-score vs noise factor



(b) Relative error vs noise factor

Figure 2: Average performance results, as a function of the noise factor, for learning Laplacian matrix of a modular graph embedded in noise.