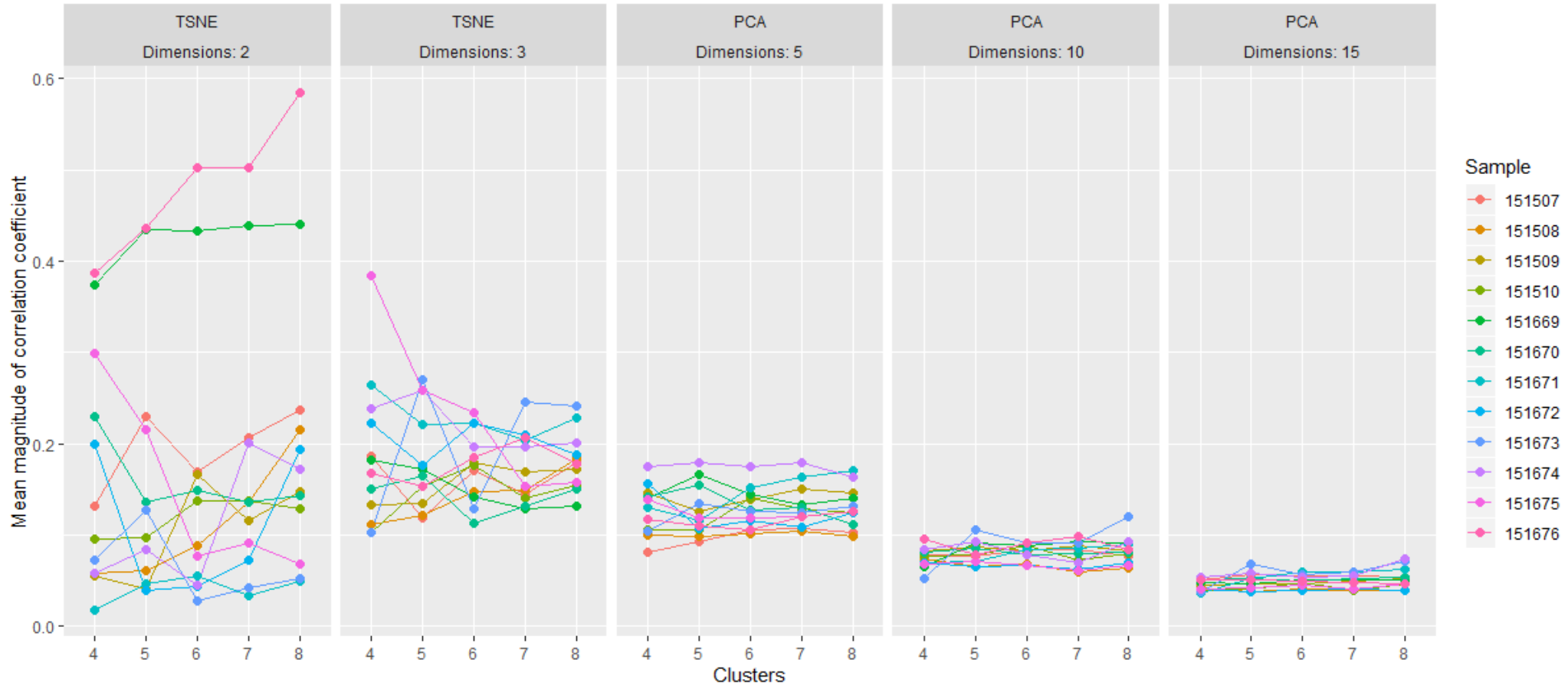
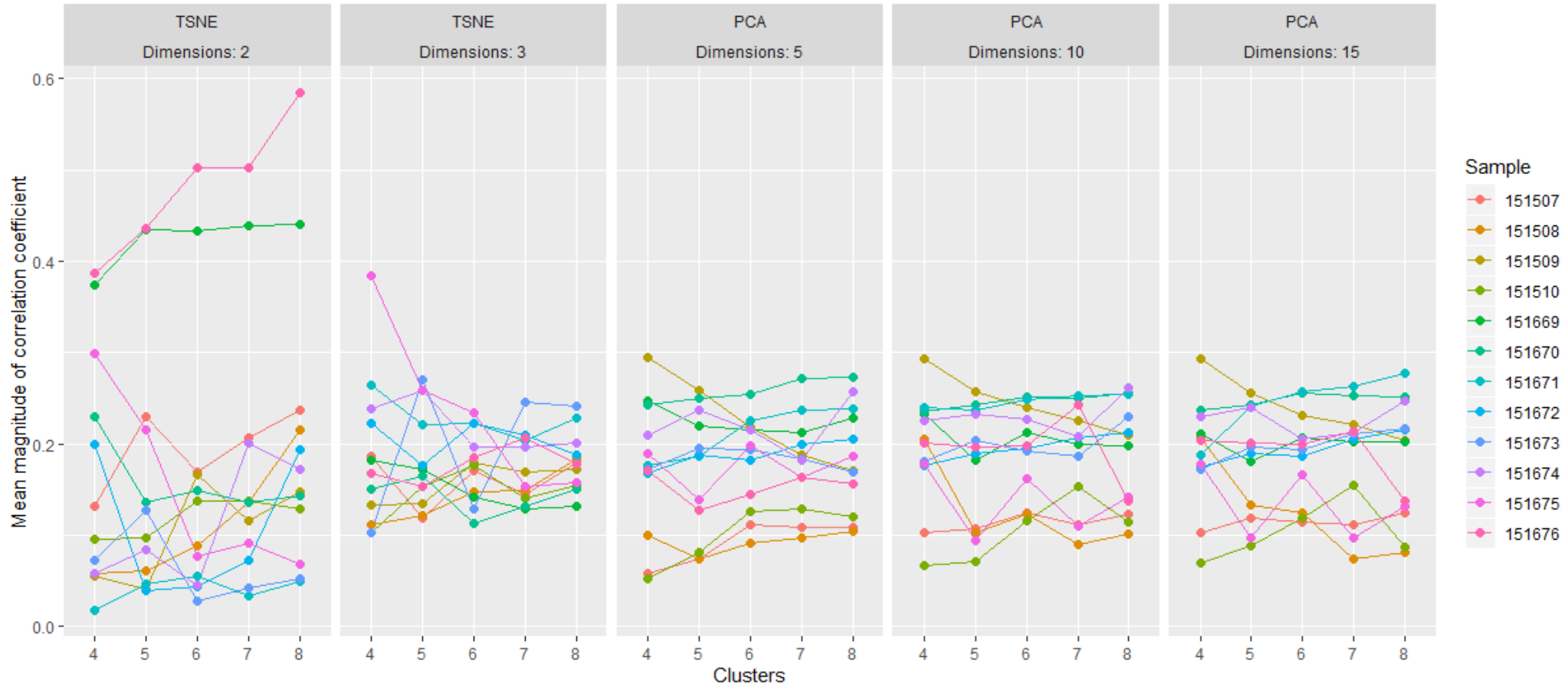
Analysis of Maynard et al., 2020 brain spatial transcriptomic data

4/17

TSNE vs. PCA


TSNE
Dimensions: 2
TSNE
Dimensions: 3
PCA
Dimensions: 5
PCA
Dimensions: 10
PCA
Dimensions: 15
Mean magnitude of correlation coefficient
0.0
0.2
0.4
0.6
4
5
6
7
8
Clusters
Sample
151507
151508
151509
151510
151669
151670
151671
151672
151673
151674
151675
151676

TSNE vs. top 3 PCs



Ising Likelihood and pseudolikelihood

- Likelihood:
  - $\mathcal{L}(\gamma; \boldsymbol{z}) = P(\boldsymbol{Z} = \boldsymbol{z}) = \frac{\exp\left(\gamma \sum_{\langle i,j \rangle} z_i z_j\right)}{\sum_{\boldsymbol{z}} \exp\left(\gamma \sum_{\langle i,j \rangle} z_i z_j\right)}$
- Conditional likelihood for spot i:
  - $\mathcal{L}_i(\gamma; \boldsymbol{z}) = P\left(Z_i | \boldsymbol{Z}^{(-i)} = \boldsymbol{z}^{(-i)}\right) = \frac{P(\boldsymbol{Z}=\boldsymbol{z})}{P\left(\boldsymbol{Z}^{(-i)}=\boldsymbol{z}^{(-i)}\right)} = \frac{P(\boldsymbol{Z}=\boldsymbol{z})}{\sum_{z_i} P\left(Z_i=z_i, \boldsymbol{Z}^{(-i)}=\boldsymbol{z}^{(-i)}\right)} =$

$$= \frac{\exp\left(\gamma \sum_{\langle i,j \rangle} z_i z_j\right)}{\sum_{z_i} \exp\left(\gamma \sum_{\langle i,j \rangle} z_i z_j\right)}$$

- Pseudolikelihood:
  - $\mathcal{PL}(\gamma; \boldsymbol{z}) = \sum_{i=1}^{n} \mathcal{L}_i(\gamma; \boldsymbol{z})$

Maximum a posterior probability

- Following the pseudolikelihood framework, we want to find the z_i that will maximize the posterior probability for each spot:
  - $p(z_i|y_i) = \frac{p(y_i|z_i)\pi(z_i)}{p(y_i)} = \frac{p(y_i|z_i)\pi(z_i)}{\sum_{z_i} p(y_i|z_i)\pi(z_i)}$
  - If we fix the domain of z_i, the denominator does not matter
    - So we need the denominator if we want to optimize cluster number, right?

Initialization issues

- Start with weak field (Besag 1986)
  - Start with $\gamma = 0$, gradually increasing to final desired γ in successive iterations
- Estimate γ at each iteration (Besag 1986)
  - Maximize $\prod_i p_i\left(x_i | x^{(-i)}, \gamma\right)$

Next steps

- Speed up code
- Tune γ for deconvolution
- Normality of t-SNE