Markov Link Method for combining destructive measurements

Jackson Loper, Trygve Bakken, and Liam Paninski
Columbia University

Destructive measurements

- It is easy to calibrate thermometers
- RNAseq methods? Not so easy

Setup

- ℓ – side information (region of brain, cre line, etc.)
- X – result of experiment under one modality
- Y – result of experiment under another modality

Setup

- ℓ – side information (region of brain, cre line, etc.)
- X – result of experiment under one modality
- Y – result of experiment under another modality

"Calibration"

$$\mathbb{Q}(Y = y | X = x) = ??$$

Setup

- ℓ – side information (region of brain, cre line, etc.)
- X – result of experiment under one modality
- Y – result of experiment under another modality

"Calibration"

$$\mathbb{Q}(Y = y | X = x) = ??$$

Problem: what if can never observe X, Y together?

One solution: Markov Link Method Assumption

If X is *more fine-grained* than ℓ, then we might make approximation

$$\mathbb{P}(X, Y|\ell) = \mathbb{P}(X|\ell)\mathbb{P}(Y|X)$$

One solution: Markov Link Method Assumption

If X is *more fine-grained* than ℓ, then we might make approximation

$$\mathbb{P}(X, Y|\ell) = \mathbb{P}(X|\ell)\mathbb{P}(Y|X)$$

Example:

- ℓ – cre line
- X – fine-grained cell-type (from deeply sequenced scRNA)
- Y – cell-type (from Patch-seq data)

One solution: Markov Link Method Assumption

If X is *more fine-grained* than ℓ, then we might make approximation

$$\mathbb{P}(X, Y|\ell) = \mathbb{P}(X|\ell)\mathbb{P}(Y|X)$$

Observe $\mathbb{P}(X|\ell), \mathbb{P}(Y|\ell)$ and determine that set of calibrations consistent those observables:

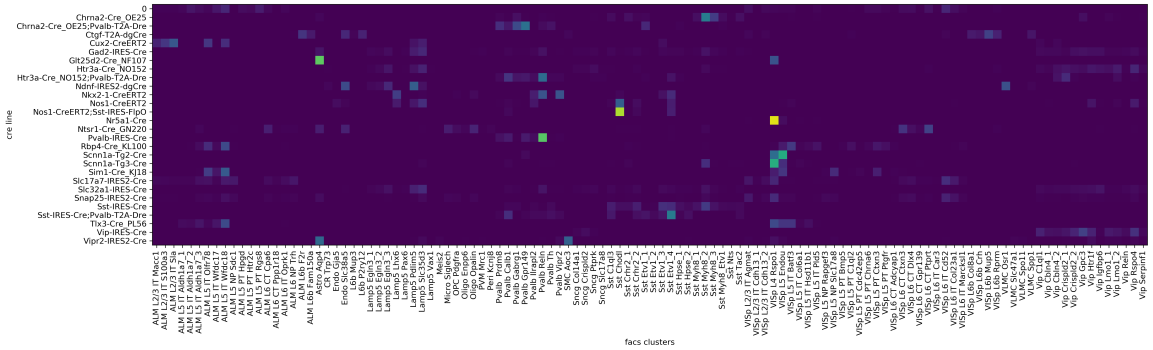
$$\Theta = \left\{ \mathbb{Q} : \ \mathbb{P}(Y|\ell) = \sum_X \mathbb{P}(X|\ell)\mathbb{Q}(Y|X) \right\}$$

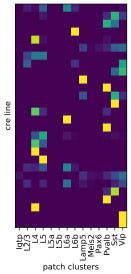
Empirical results

Tasic, Bosiljka, Zizhen Yao, Kimberly A. Smith, Lucas Graybuck, Thuc Nghi Nguyen, Darren Bertagnolli, Jeff Goldy et al. "Shared and distinct transcriptomic cell types across neocortical areas." bioRxiv (2017): 229542.

Empirical results

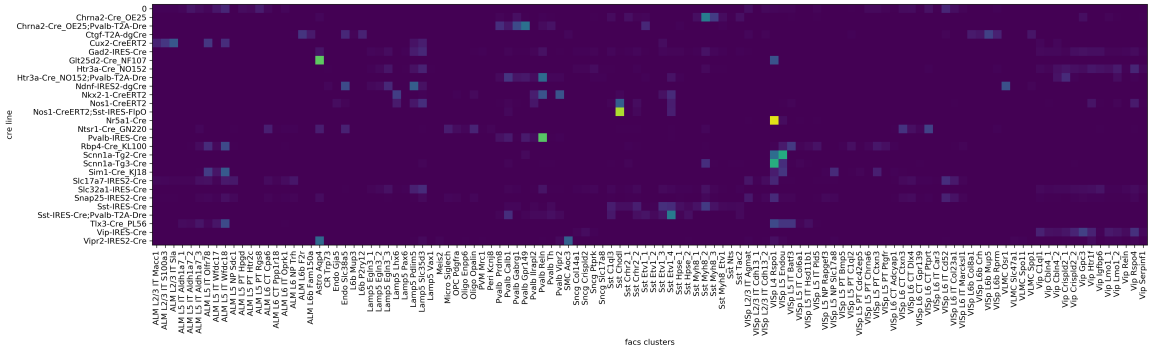
From datasets about two separate experimental methods...

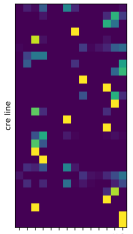




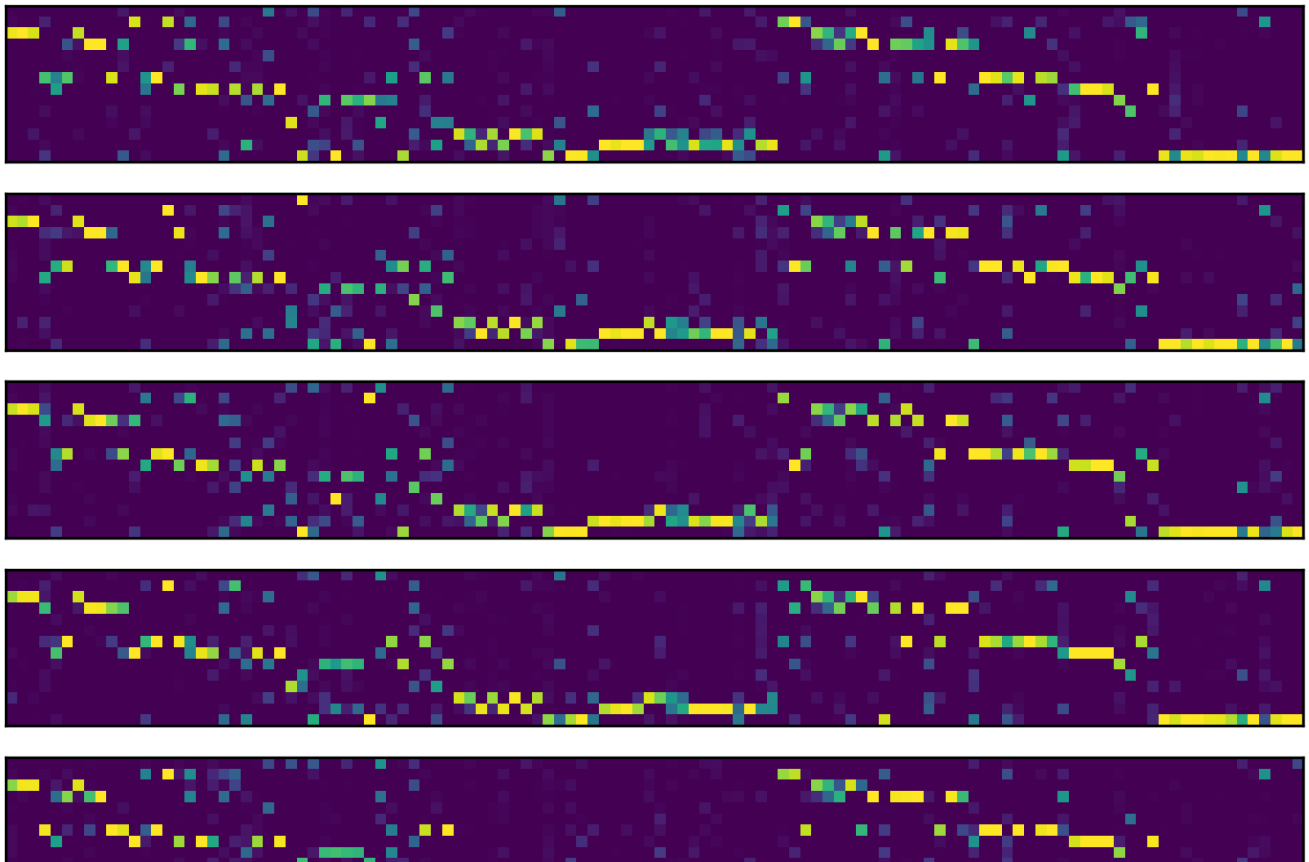
Empirical results

From datasets about two separate experimental methods...





... to a set of possible calibrations between the methods.



Making sense of the set of possible calibrations

- Rotatially Uniform eXtremal distribution
- Uniform distribution
- Diameter estimation
- Center of mass