

## INTRODUCTION

- Independent Component Analysis (ICA) is used to denoise multi-echo fMRI data<sup>1,2</sup> by removing non-neural signal and can result in better quality data in less acquisition time
- ICA separates data into statistically independent spatial components
- Theoretically, the optimal number of components is the true number of statistically independent sources or dimensions that contribute to total signal
- Over-separation of signal can lead to loss of statistical power from overfitting, while under-separation of signal can make it hard to correctly remove artifacts
- Established dimensionality estimation methods sometimes fail, so we tested whether any commonly used algorithms offered consistent and reliable dimensionality estimates
- Given that the true dimension is difficult to estimate, we then tested a range of numbers of components with respect to a practical definition of optimal: improved contrast-to-noise ratio in task-relevant areas and maintained statistical power

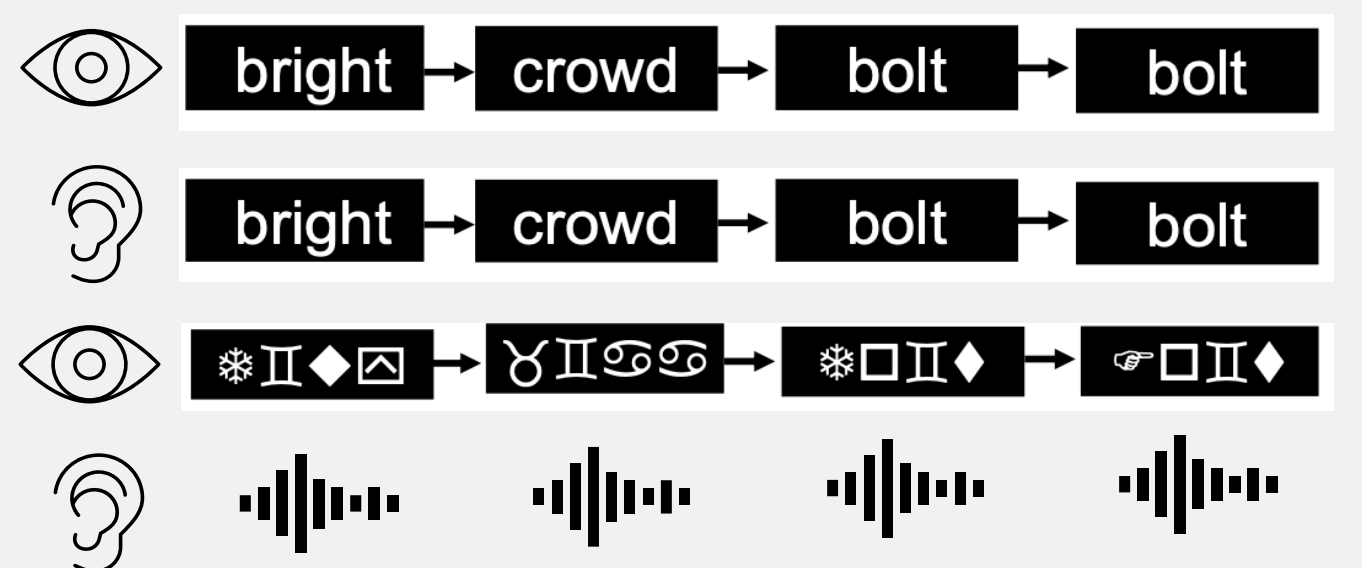
### Central Questions

- Is there a reliable & plausible dimensionality estimation method for fMRI?
- How does dimension affect denoising quality?

## METHODS

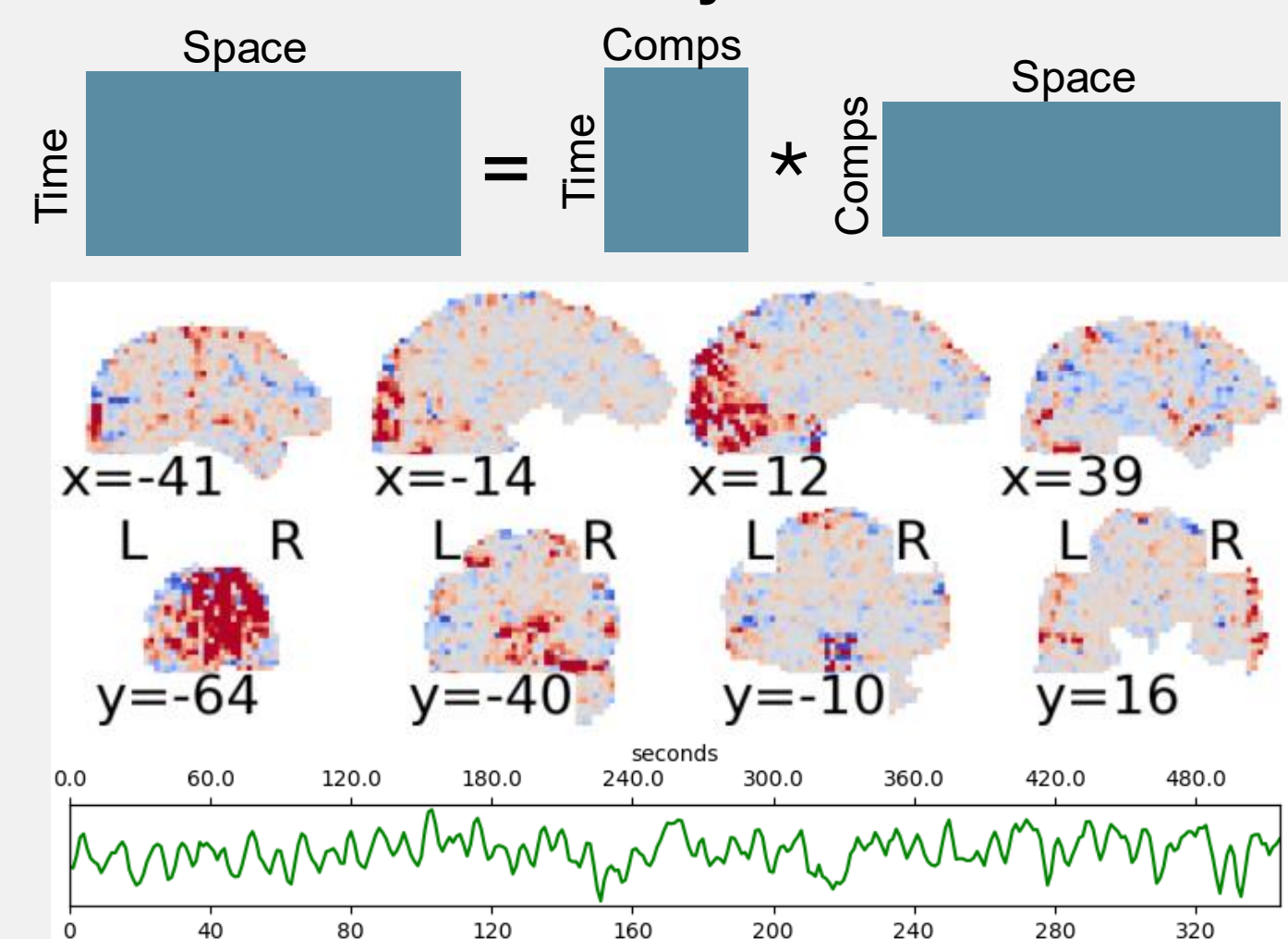
### Data Collection and Task Paradigm

- Collected multi-echo fMRI data (Siemens Prisma 3T; TR=1.5s; 3mm<sup>3</sup> voxels, TEs=13.44, 31.7, & 49.96ms) from 25 subjects
- Subjects completed a visual or audio, word or non-word (VA-WNW) discrimination task<sup>3</sup> (3 8.5 min runs) and a paced breathing task with or without movie viewing (2-3 8 min runs)



Task design for VA-WNW discrimination task: 4s per trial, 48 trials per run in an event-related design

### ICA and Dimensionality Estimate Methods

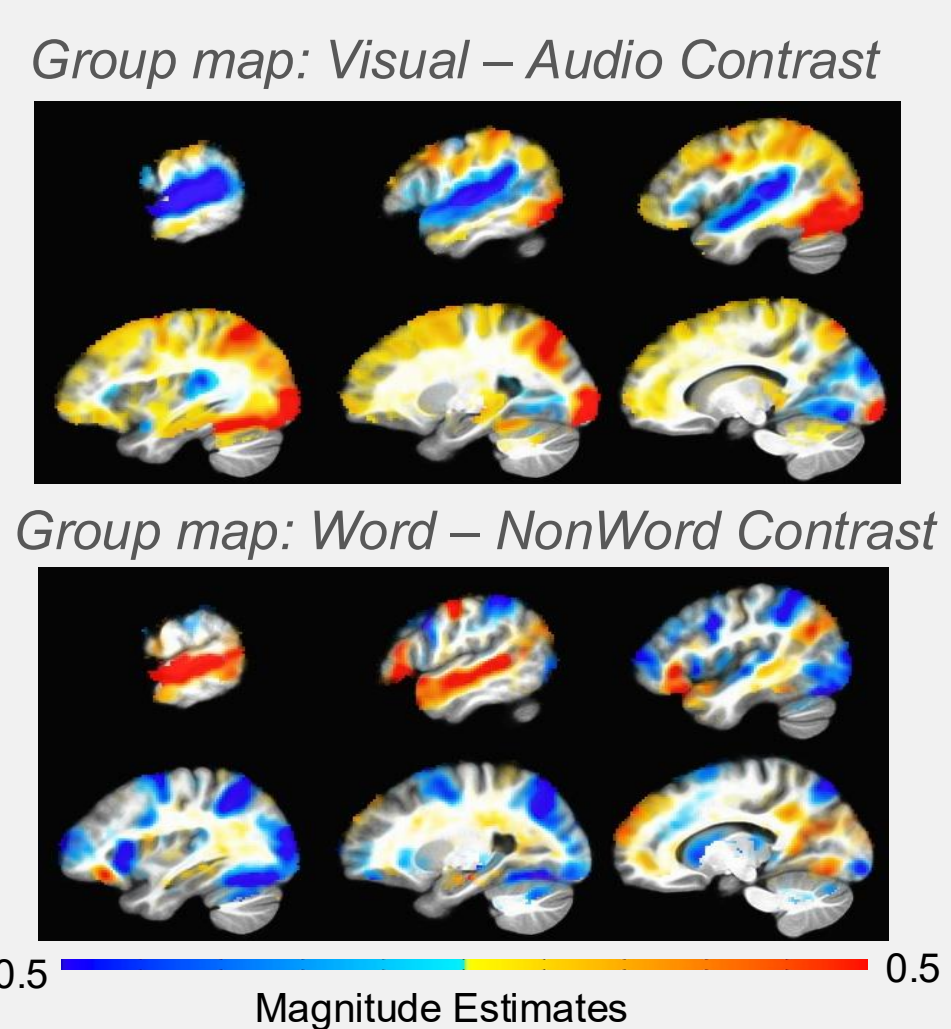


ICA decomposition equation and sample ICA component

- Tested dimensionality estimation methods from MAPCA<sup>1,4</sup> (implemented in tedana) and MELODIC<sup>5</sup> (an FSL fMRI ICA tool) and a modified bootstrapping method (BSA)<sup>6</sup>
- Methods from MELODIC and MAPCA optimize a cost function that relates noise variance to singular value decomposition (SVD) with various number of components ( $n$ )
- BSA compares the stability of SVD components on real and bootstrapped data versus noise
- Compared dimensionality estimates from each method on all runs of data from both tasks

### ICA Denoising in Task-Relevant Regions

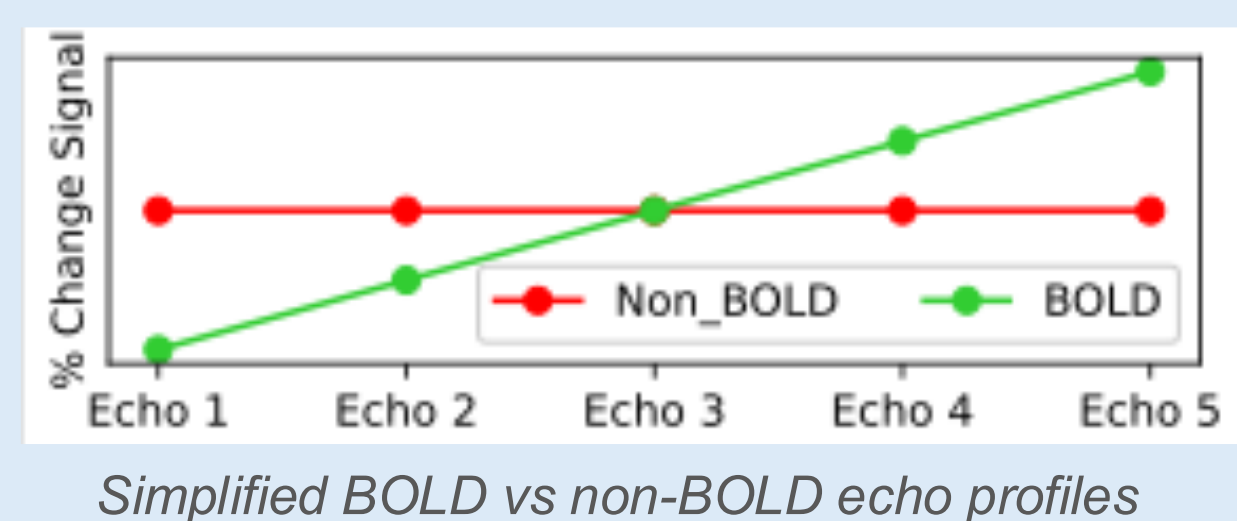
- The VA-WNW task paradigm was fit to the data with no ICA denoising using a GLM
- Significant voxels in predefined anatomical ROIs were used to create subject-specific ROIs
- Data were denoised with tedana using between 20 and 200 components
- For all models, model fit quality was assessed within each ROI using median Fisher Z scores, F statistics, and p-values



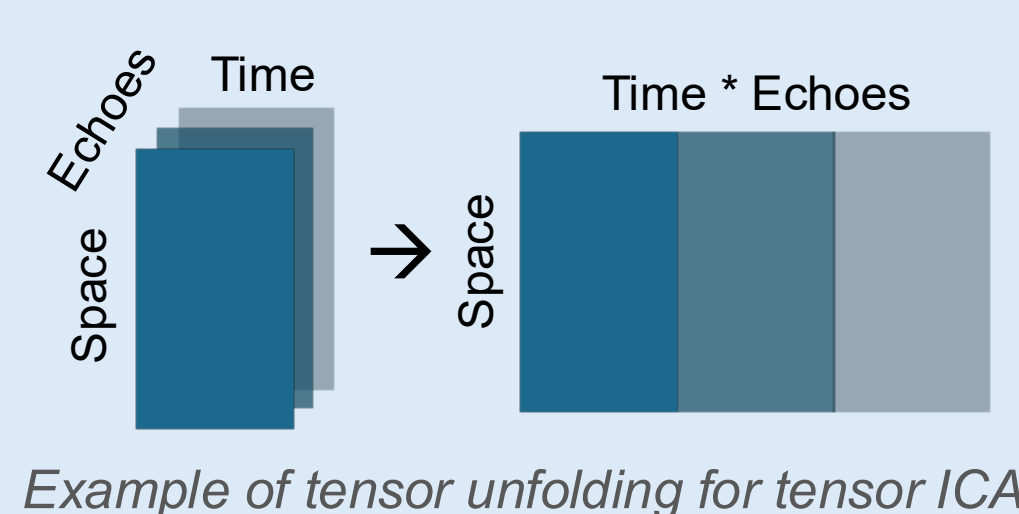
## ONGOING WORK

### Tensor ICA

- Denoising can be improved by better decomposition methods or better inclusion/exclusion criteria
- MRI physics dictates how BOLD and non-BOLD signal should decay across echoes, which can inform inclusion/exclusion criteria
- Tensor based ICA approaches take advantage of all echoes in the component creation process and output components with an additional "echo profile" dimension
- Current fMRI tensor ICA methods<sup>5,6</sup> do not consider physical properties in component creation



Simplified BOLD vs non-BOLD echo profiles



Example of tensor unfolding for tensor ICA

### Echo Constrained Alternating-Projection Tensor ICA Algorithm

- Our work-in-progress Echo Constrained Alternating-Projection Tensor ICA algorithm optimizes for statistical independence (via an ICA-projection step) as well as physics-informed echo-profiles
- Convergence metrics for both characteristics are promising
- Echo slopes are linear but do not have enough variation

Sample component from ECA-P Tensor ICA with 30 Components



Convergence metrics for ECA-P tensor ICA, least squares fit between modeled and real data and negentropy

## RESULTS

### Comparing Methods for Component Number Estimation

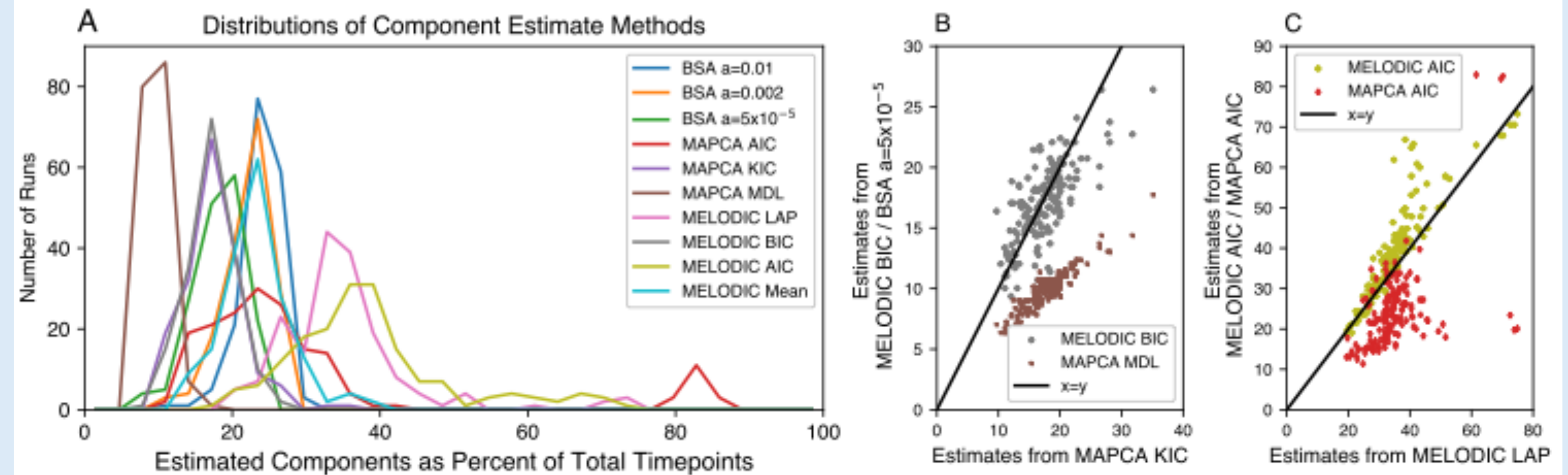


Fig 1. Discrepancies between component number estimates made with various established methods. (a) Histograms show the estimated number of components as a percent of total timepoints. (b,c) Scatterplots show the relationship between percent components estimated by various methods.

### Comparing Current Dimensionality Estimate Algorithms

- No method offers reliable and plausible dimensionality estimates across all data
- For each method, estimates for runs within the same dataset fell within a wide range (1a)
- All methods gave some implausibly low (less than 15% of total timepoints) or implausibly high (more than 60% of total timepoints) estimates
- For a single run of data, estimates from different methods varied (1b,c)

### Visual - Audio in Regions of Interest

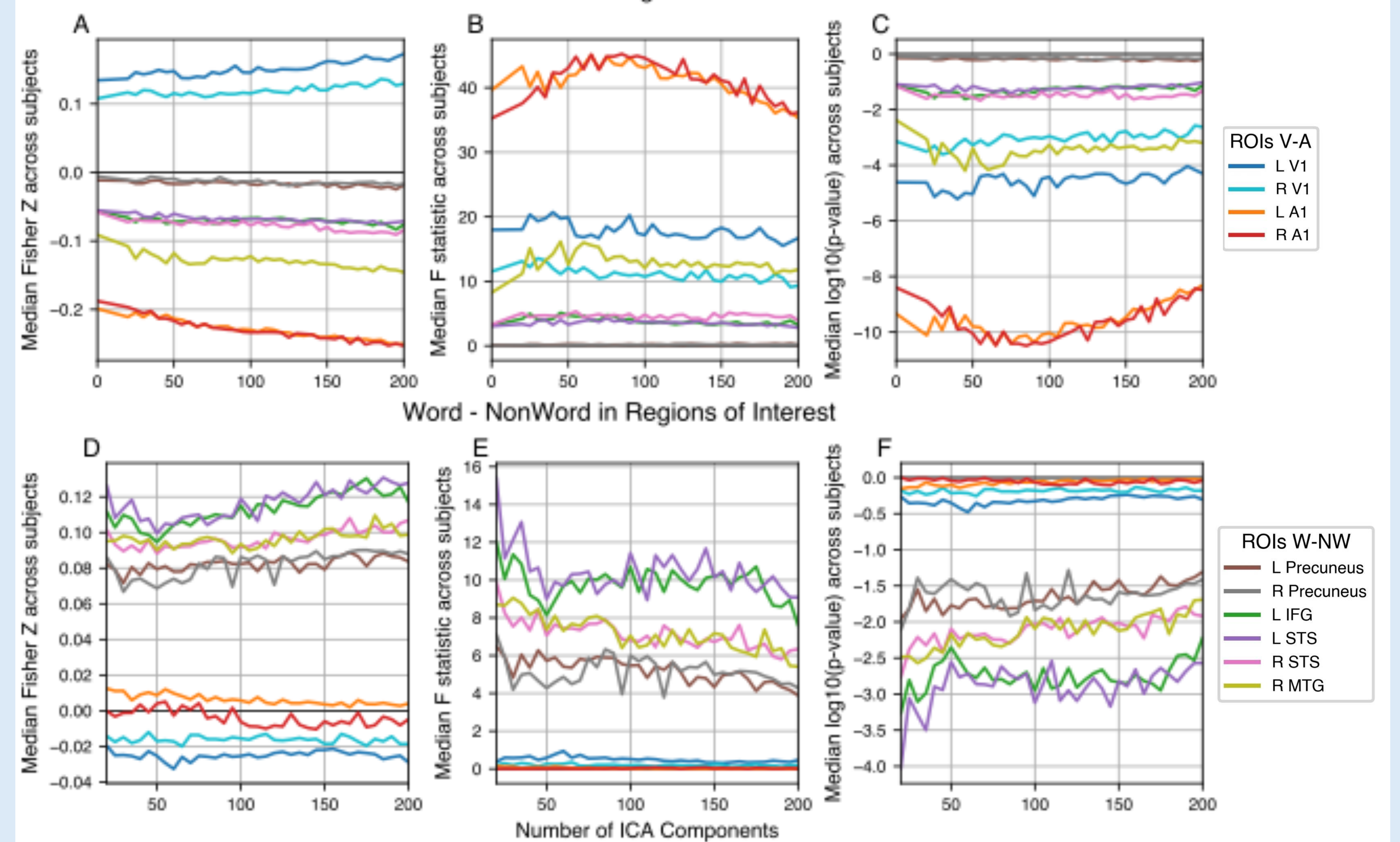


Fig 2. Statistics from tedana denoising run on data from the VA-WNW discrimination task with 20-200 components. Statistical measures of (a-c) visual - audio contrast across ROIs (effects expected in V1 and A1) and (d-f) word - nonword contrast across ROIs (effects expected in STS, IFG, MTG, Precuneus). (a,d) Median Fisher Z values indicate a positive association between  $n$  and CNR, (b,e) median F statistics and (c,f) median log(p-values) show little effect on statistical power.

### Measuring the Effect of Component Number on Denoising Quality in ROIs

- Fisher Z scores, a measure of contrast-to-noise ratio (CNR), showed a positive association between number of components and CNR in most task-related ROIs (2a,d)
- F statistics and log(p-values) had local maxima in A1 in Visual - Audio contrast, but all values were well beyond the threshold for statistical significance (2b,c)
- F statistics and log(p-values) were fairly stable as number of components increased, showing little effect on statistical power despite loss of degrees of freedom from additional components (2b,c,e,f)
- Across subjects and regions, the statistical measures show that choices of  $n$  between 20% and 50% of total timepoints (60-150 components) have little effect on data quality

## CONCLUSIONS

- Dimensionality estimates vary widely within and across methods
- As dimension increases, CNR increases in some ROIs and statistical power is relatively stable in most ROIs, despite a decrease in degrees of freedom
- Tensor ICA may provide a novel approach to component separation and inclusion/exclusion

### Key Takeaways

- Current methods make inconsistent dimensionality estimates
- ICA denoising quality is relatively stable for a wide range of dimensions

## REFERENCES

- DuPre E, Salo T, Ahmed Z, et al. TE-dependent analysis of multi-echo fMRI with tedana. *J Open Source Softw.* 2021;6(66):3669. doi: 10.21105/joss.03669
- Kundu P, et al. "Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI" *NeuroImage.* 2012;60(3):1759-70. doi: 10.1016/j.neuroimage.2011.12.028
- Malins JG, Gumkowski N, Buis B, et al. Dough, Tough, Cough, Rough: A "fast" fMRI Localizer of Component Processes in Reading. *Neuropsychologia.* 2016;91:394-406. doi: 10.1016/j.neuropsychologia.2016.08.027
- Li YO, Adali T, Calhoun VD. Estimating the number of independent components for functional magnetic resonance imaging data. *Hum Brain Mapp.* 2007;28(11):1251-1266. doi: 10.1002/hbm.20359
- Beckmann CF, Smith SM. Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging. *IEEE Trans Med Imaging.* 2004;23(2):137-152. doi: 10.1109/TMI.2003.822821
- Majeed W, Avison MJ. Robust Data Driven Model Order Estimation for Independent Component Analysis of fMRI Data with Low Contrast to Noise. *PLoS One.* 2014;9(4):e94943. doi: 10.1371/journal.pone.009494

## ACKNOWLEDGEMENTS

Research supported by NIH Grants: ZIAMH002783 & ZICMH002968  
 This work utilized the resources of the NIH HPC Biowulf cluster