Moving away from ICA in multi-echo fMRI denoising

Daniel A Handwerker[1], Javier Gonzalez-Castillo[1], Dylan Nielson[2], Charles Zheng[3], Peter A Bandettini[1,4]

1. Section on Functional Imaging Methods, LBC, 2. Data Science & Sharing Team, 3. Machine Learning Team, 4. Funtional MRI Facility
NIMH, Bethesda, MD

Poster # 2542

INTRODUCTION

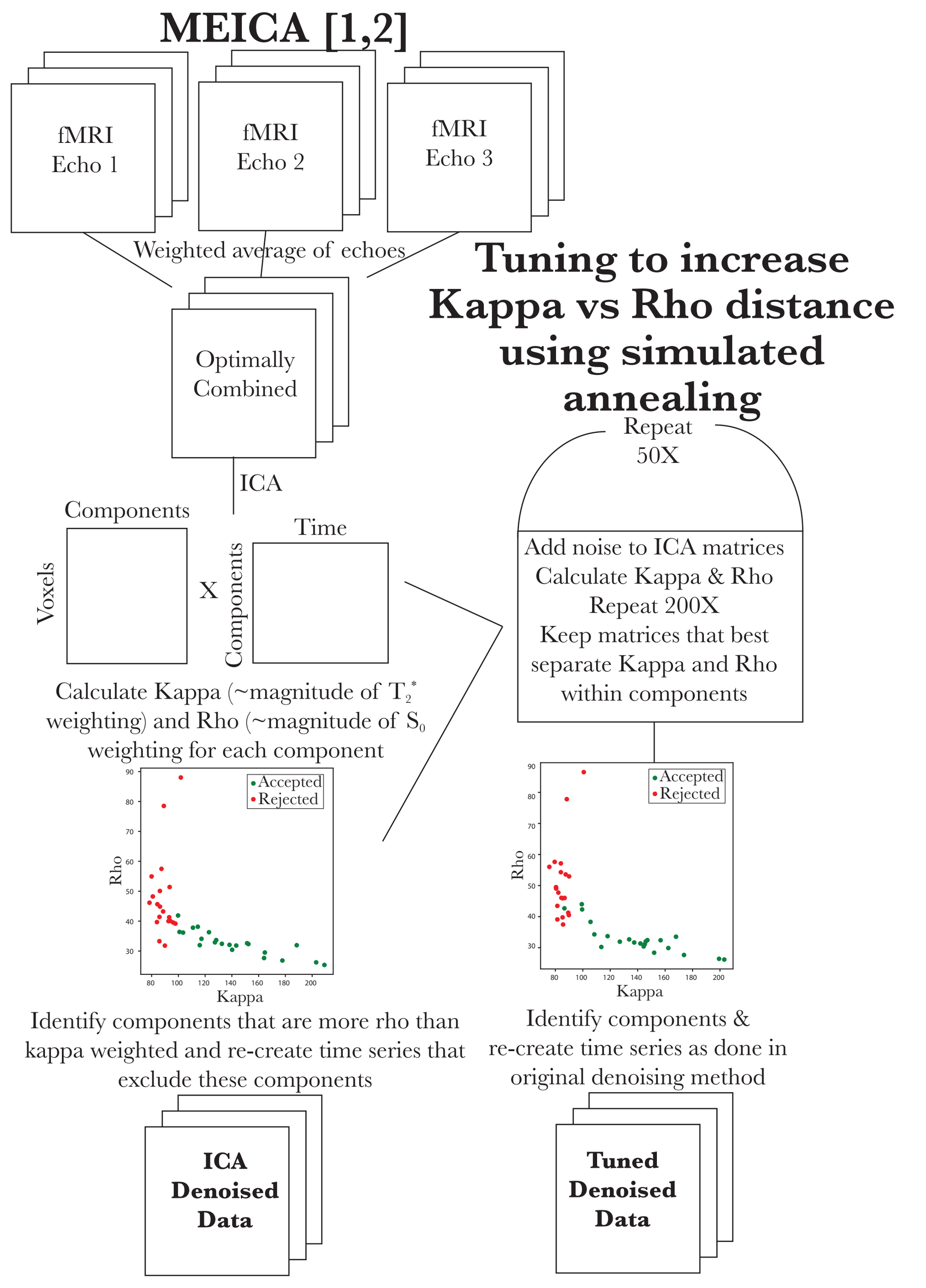
Multi-echo fMRI is used to empirically identify and remove non-T_2^* weighted fluctuations from fMRI data. One common method MEICA[1], uses ICA to break data into components. Components that are classified as insufficiently T_2^* weighted are removed from denoised time series. This method works well, but there is no reason to assume that ICA inherantly parses data into distince T_2^* and non-T_2^* components. Thus every retained and removed component is a combination of potentially neural and non-neural fluctuations.

We test whether it is possible to create components with better separation of T_2^* information to improve the fMRI denoising process.

METHODS

Original ICA and Tuned Denoising Algorithmns

MEICA [1,2]



For signal fluctuations in each component:
Kappa is a measure of T_2^* (including BOLD)
Rho is a measure if S_0 (e.g. head motion, signal drift)

A good explanation, with equations, is in [2].

For each voxel in each component, calculate an F statistic across the echo time series for the goodness of fit to an MRI model of T_2^* or S_0. Kappa and Rho are the weighted sum of these voxelwise estimates weighted by the contribution of each voxel to the component

Component Selection: Sort the Kappa & Rho values by magnitude and find an inflection point or elbow, where the slope changes. Remove components with a Kappa lower than the elbow or a Rho higher than the elbow. (The same elbow selection method and no other criteria where used to select components from the ICA and Tuned analyses.

Tuning Cost Function: Calculate the mean Kappa & Rho across components for each run. The κ ρ difference is: κρDiff= sum(κ-ρ*mean(κ)/mean(ρ)) Like z-scoring, this makes sure a relative change in κ or ρ in a component is treated similarly by the cost function

ICAκρDiff = 2-(abs(κρDiff)/max(κρDiff)) for the original ICA κ and ρ values

The cost function is the sum of abs(κρDiff) for each permutation * ICAκρDiff

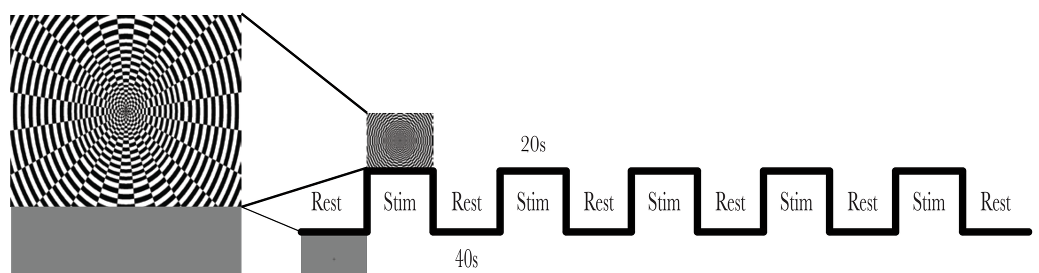
The scaling by ICAκρDiff means that the cost changes more for values with mixed κ and ρ weighting that shift towards being more κ or more ρ.

Data

Two volunteers participated in the same 340s block design task 103 times each over 9 scanning sessions. The massive repetitions make it possible to estimate distributions of effect sizes across runs

Scanning Parameters

GE 3T MR-750 MRI scanner, GE 32 channel head coil.
GRE EPI, TR=2s, TE=15.4, 29.7, & 44.0ms, FA=75°
33 oblique slices, 3.5mm^3 voxels, 0mm gap, 64x64 grid, ASSET=2.
1mm^3 MPRAGE T1 weighted and proton density weighted scans were collected during each session for registration.



5 cycles of a 20s flickering checkerboard followed by 40s of fixation. A letter or number appeared for 0.4s 4 times in each stimulus block. The volunteer would press a button indicating if a letter or number appeared. Same task as: [3]

Preprocessing

Data were processed using AFNI and Python (for the ME-ICA denoising code) in each volunteer's native space. The data were despiked, slice time corrected and motion corrected. The first scan of every session was aligned to the anatomical scan from the same day and then the first day's anatomical scan. Alignment and motion correction parameters were calculated on the middle echo time series and applied to all 3 echoes as a single transform matrix.

REFERENCES

1. Kundu, P. (2012), 'Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI' NeuroImage
2. Olafsson, V. (2015) 'Enhanced indetification of BOLD-like components with multi-echo simultaneous multi-slice (MESMS) fMRI and multi-echo ICA
Appendix A includes a good description of the underlying MEICA mathematics
3. Gonzalez-Castillo, J. (2012), 'Whole-brain, time-locked activation with simple tasks revealed using massive averaging and model-free analysis' Proceedings of the National Academy of Sciences

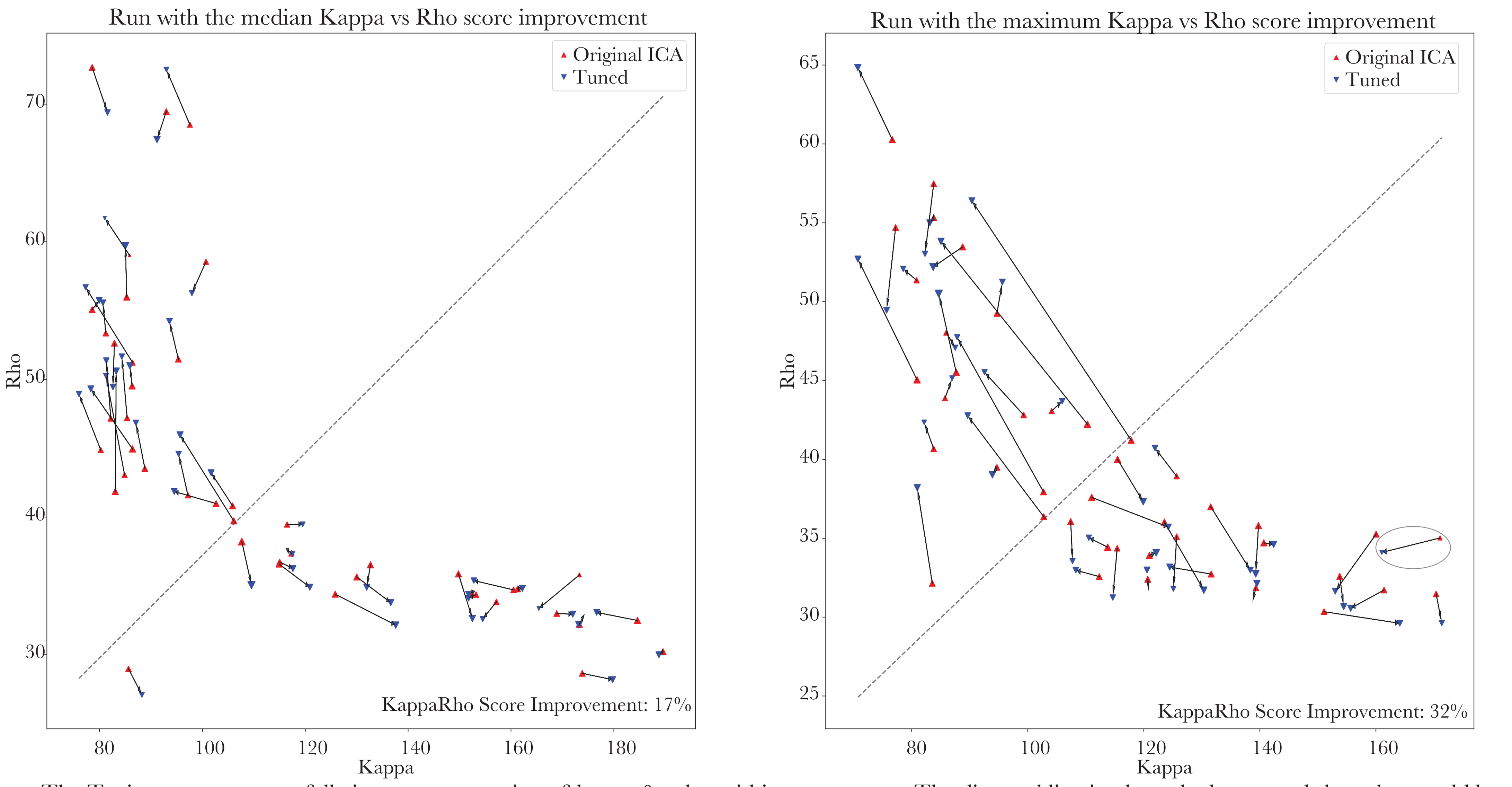
ACKNOWLEDGEMENTS

This work utilized the computational resources of the NIH HPC Biowulf cluster: hpc.nih.gov
Pete Molfese helped with programming quirks
Support: NIMH Intramural Research Program

fim.nimh.nih.gov
Contact: handwerkerd@nih.gov

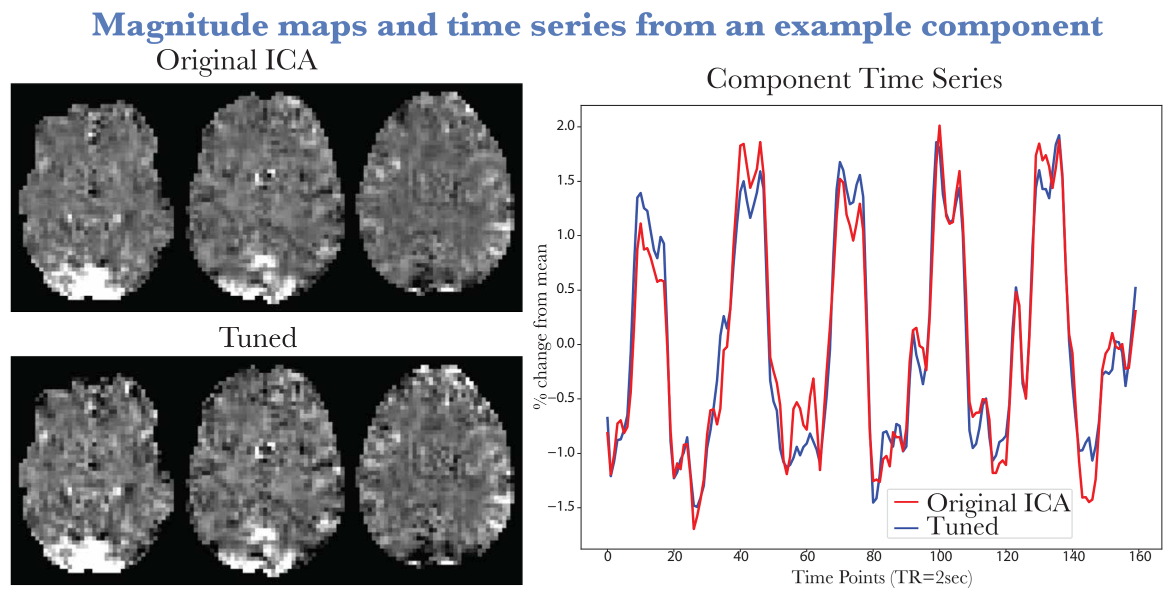
RESULTS

Kappa and Rho Values are more separated across components after Tuning



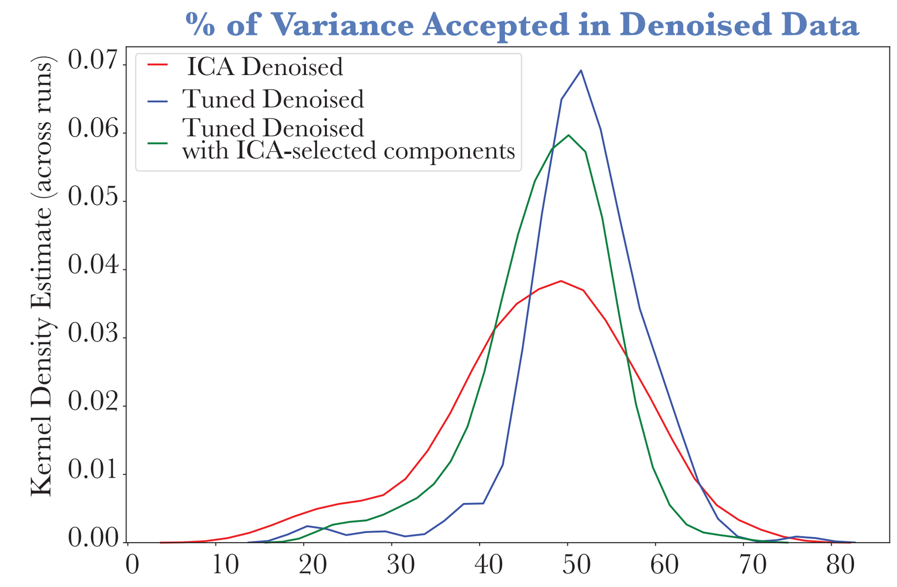
The Tuning process successfully increases separation of kappa & value within components. The diagonal line is where the kappa and rho value would be most mixed within components. After tuning, the components that are closest to this line are farther away. Not every component gets better differentiated

Magnitude maps and time series from an example component



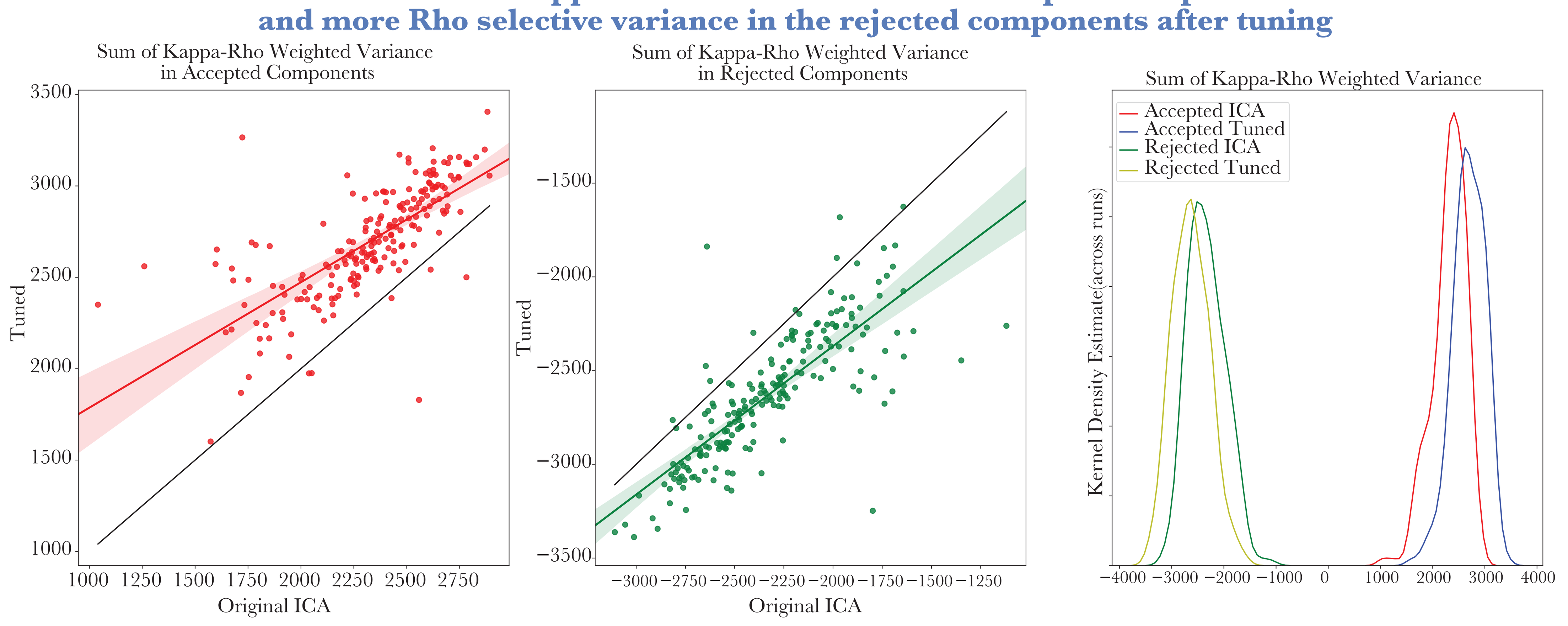
This is an example component from the run with the greatest kappa vs rho score improvement (circled in the above figure). This component is weighted towards the voxels in primary visual cortex and the time series shows the block-design task response. Note that, while the average of all components have better kappa rho differentation, the kappa score decreased with tuning for this component

% of Variance Accepted in Denoised Data



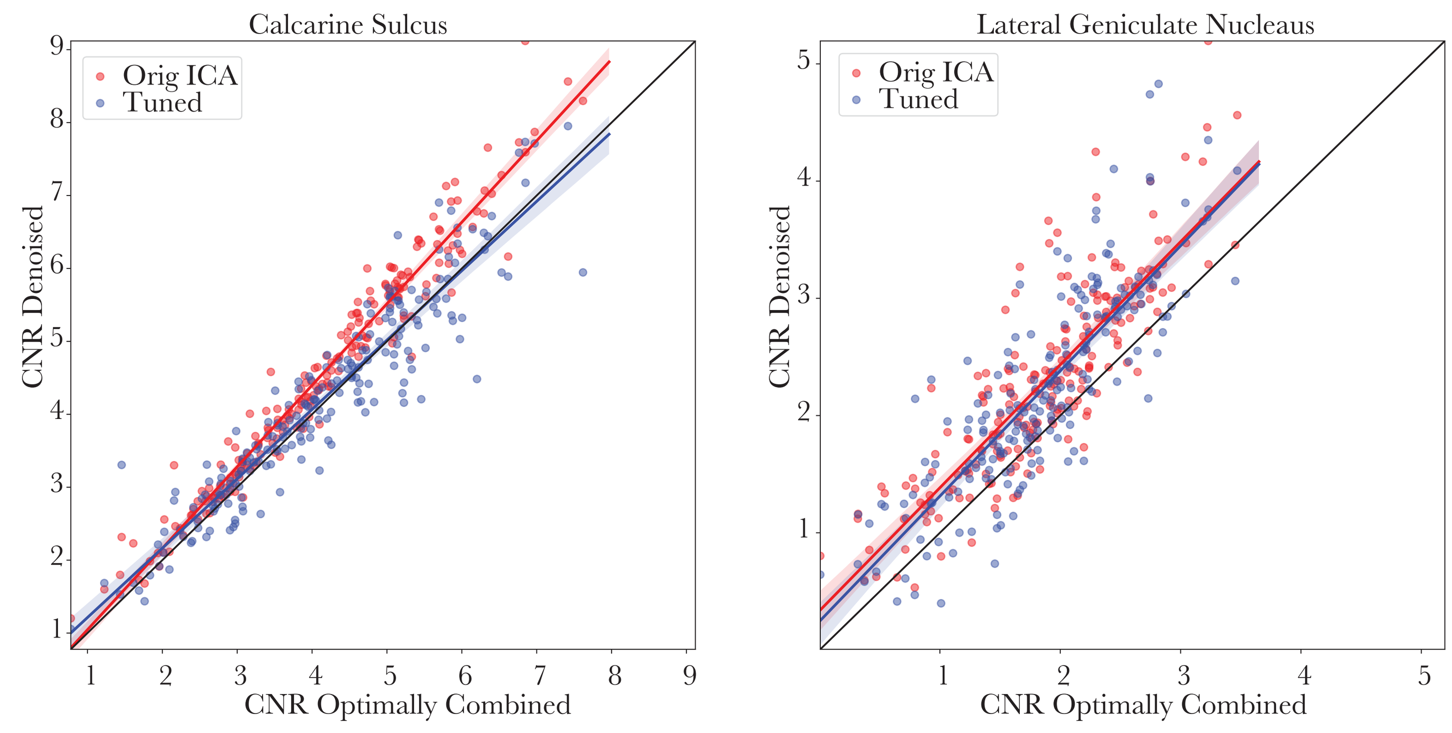
The tuning process results in runs where more of the total variance remains in the denoised data. While this is not inherantly good or bad, it's noteworthy that the distribution is also narrower after tuning. The narrowing remains even when selecting the same components as used in the original ICA. Tuning may increase the consistency of the result of the denoising process.

There is more Kappa selective variance in the accepted components and more Rho selective variance in the rejected components after tuning

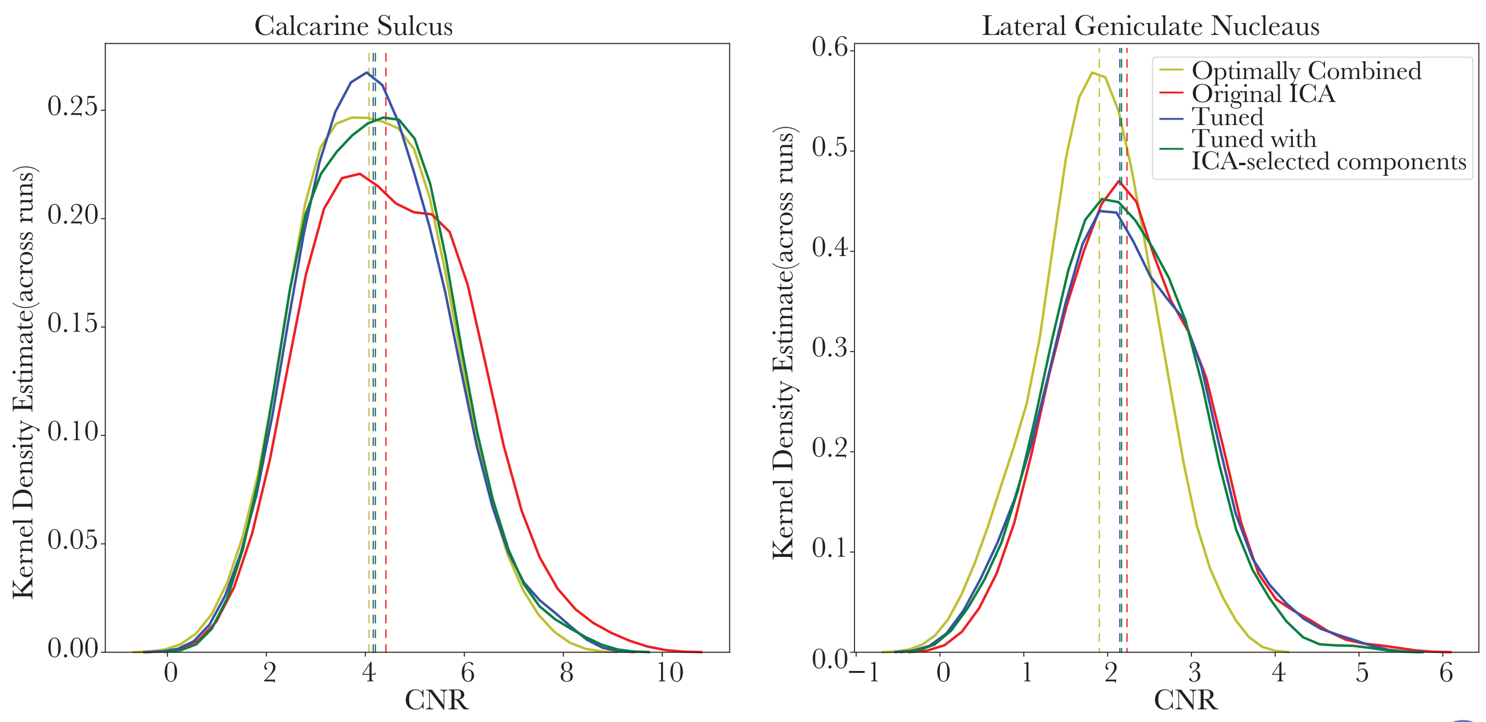


These values are the κρDiff for each component, multipled by the variance for the component and summed across all accepted or rejected components. The more positive a value, the more kappa-weighted variance there is in the components and negative values are more rho weighted. Each dot is a run. The black line has a slope of 1 (original=tuned). After tuning, nearly every run has more kappa-weighted variance in the accepted components and more rho-weighted variance in the rejected components.

Contrast-to-Noise Ratio Does Not Improve with Component Tuning



Scatter plots show the CNR for the optimally combined data vs the denoise time series for two regions-of-interest, the Calcarine sulcus and the LGN. Each dot is a run and the black line marks no change in CNR. For both ROIs, the CNR increases from optimally combined to ICA-based denoising. For the Calcarine Sulcus, which usualy has a large CNR for this task, the tuned denoised time series showa little improvement over Optimally Combined in CNR. For the LGN, the tuned time series is similar, but not better than the original ICA-desnoised time series.



These show the distributions of the CNR values across runs for both ROIs. The dashed vertical lines are the median CNR for each processing method. The CNR distribution for the original ICA denoised time series are shifted larger than the optimally combined data, but the tuned denoised time series have slightly worse or equal CNR to the ICA time series. Using the tuned components, but rejecting only the components selected by the original ICA doesn't shift these distributions.

DISCUSSION

We demonstrate that ICA is not the optimal way to separate data by T_2^* weighting.

The presented tuning process resulted in accepted components whose total variance was more kappa-weighted and rejected components whose total variance was more rho-weighted

This opens up the possibility to more selectively remove non-BOLD-weighted noise from multi-echo fMRI data

The Contrast-to-Noise Ratio in two task-specific ROIs improved with ICA Denoising, but the added tuning did not improve CNR

OPEN QUESTIONS

Is there a different cost function for tuning that will improve CNR by better balancing the kappa rho difference improvements across all components?

Are kappa & rho nonideal proxies for T_2^* and S_0 weighting?

Does CNR improve with tuning in other brain areas with lower initial CNR?

Can other methods, like IVA, more efficiently estimate kappa vs rho differentiated components?