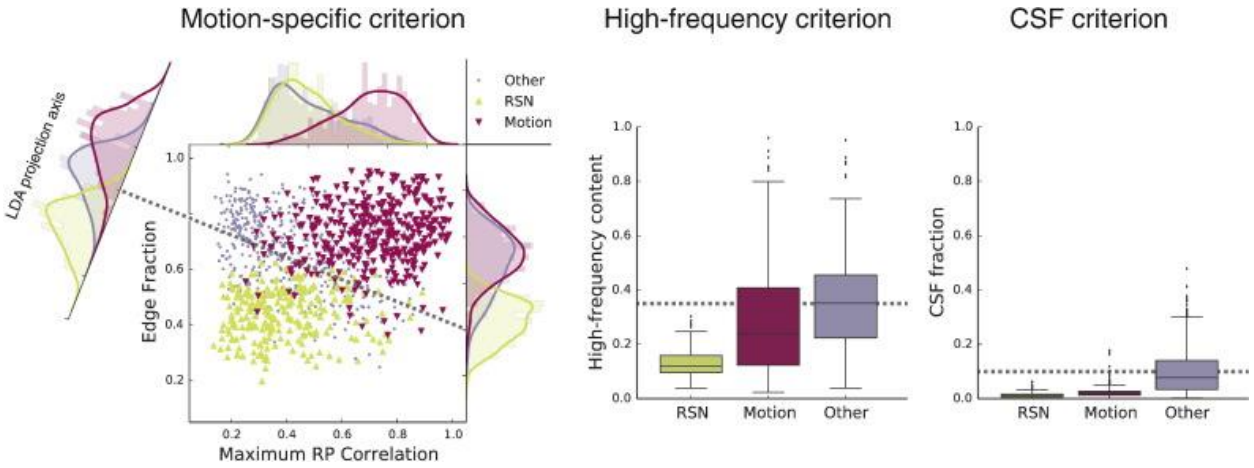
ICA-AROMA: Conclusions and Recommendations

Algorithm Summary (http://doi.org/10.1016/j.neuroimage.2015.02.064***)***

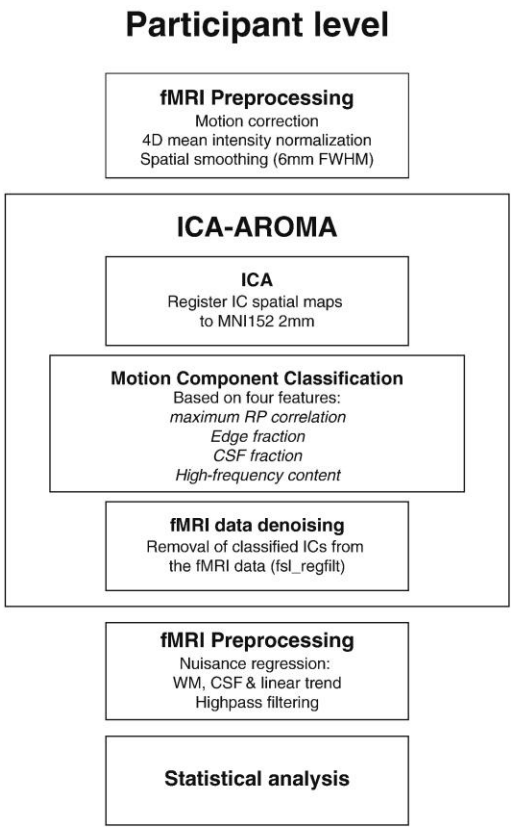
ICA-AROMA is an algorithm, trained on resting state and anatomical data from 30 healthy subjects, which automatically removes independent components that contain motion artifact and then reconstructs the functional image. It has been validated on multiple large resting and task datasets.

ICA-AROMA uses 3 thresholds to label, and subsequently remove, independent components as motion:

1. A 2D combination between Maximum RP correlation (max correlation between component timeseries and 72 realignment parameters) and fraction of activation at edge of brain
2. A 1D threshold for high frequency content
3. A 1D threshold for fraction of activation in CSF



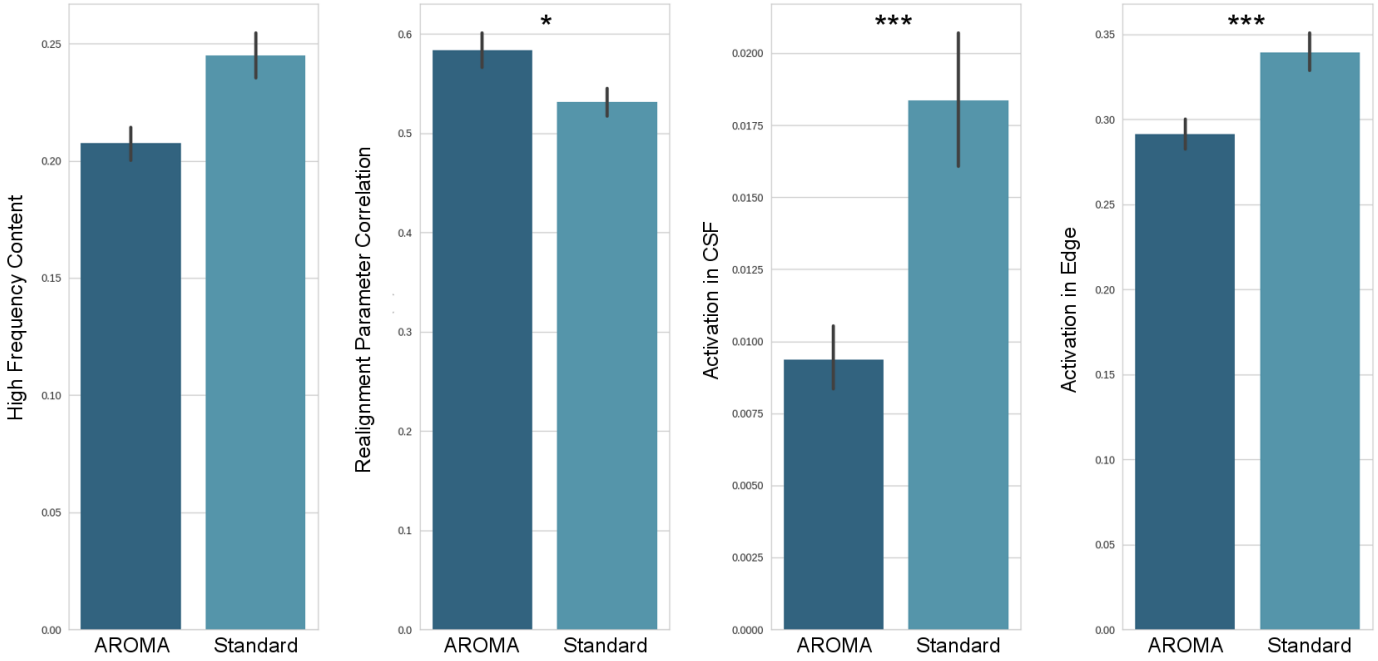
The primary advantage of ICA-AROMA is that it uses a priori knowledge about spatial and temporal artifacts contributing to motion contamination. It is to be used after swar and not in tandem with volume repair:



Efficacy

In the following, we define the “standard” motion correction pipeline as applying volume repair and excluding any subjects with x,y,z,pitch,roll,raw > 6mm or volume repair > 15%

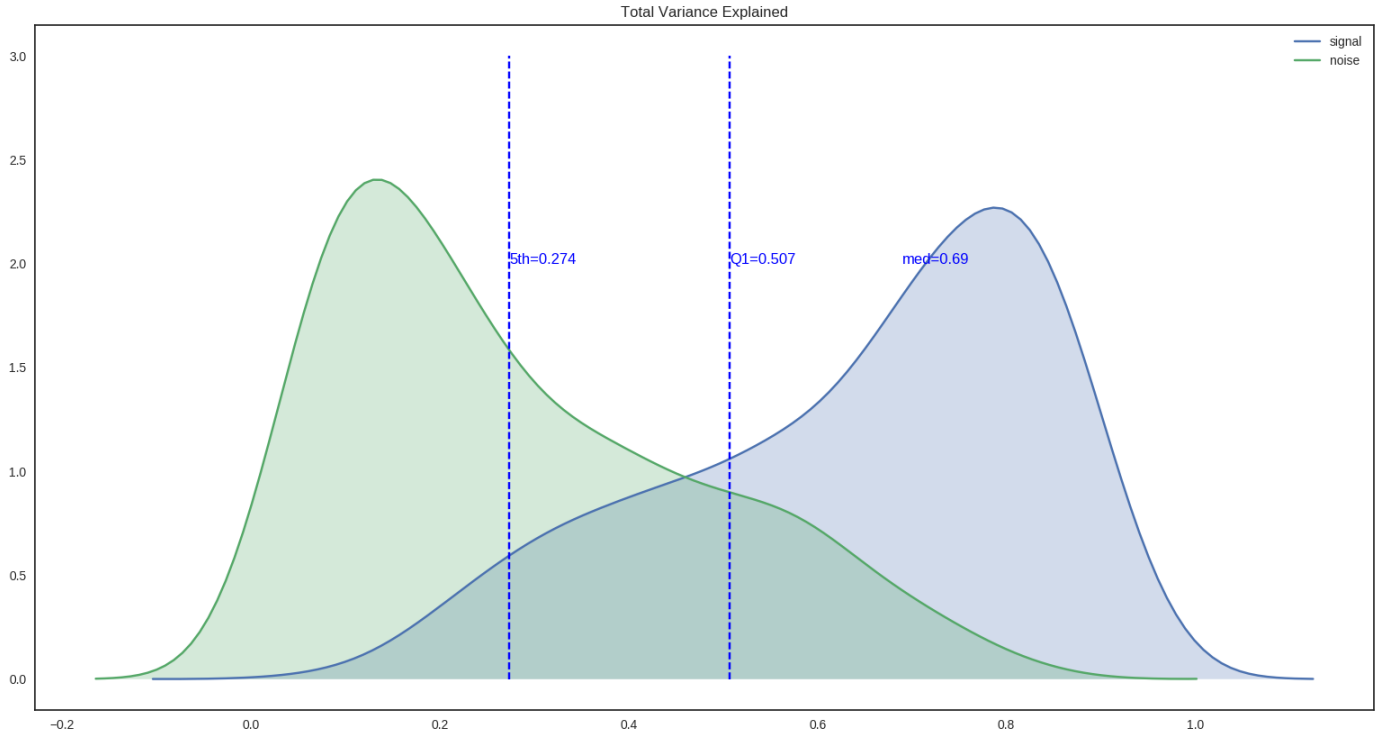
We have shown that group-level results from ICA-AROMA and the standard motion correction pipeline are closely replicable in several datasets. In addition, ICA-AROMA succeeds in removing motion-related artifacts that are present in the standard pipeline. Here, we plot the values for each of the ICA-AROMA exclusion criteria over independent components from 20 subjects (arithmetic addition), where components in the ICA-AROMA case are only those which are classified as non-motion, and those in the Standard case are from an ICA conducted on top of volume repaired data:



We see uniform improvement in each of the criteria not used by the standard pipeline and slightly worse performance for max correlation with realignment parameters. This is expected and, importantly, will become highly relevant for subject exclusion from the ICA-AROMA pipeline.

Thresholding

While the advantages to using ICA-AROMA are clear, there is no typical threshold suggested for excluding a subject. To remedy this, we have attempted to create a threshold for use within SCSNL. “Signal” and “noise” distributions were created over 538 task runs of lab data by summing the % variance explained by independent components labeled as non-motion and motion, respectively. To determine an exclusion threshold, we then take Q1 on the signal distribution and exclude all subjects whose total variance explained by motion components surpasses this value (0.507):

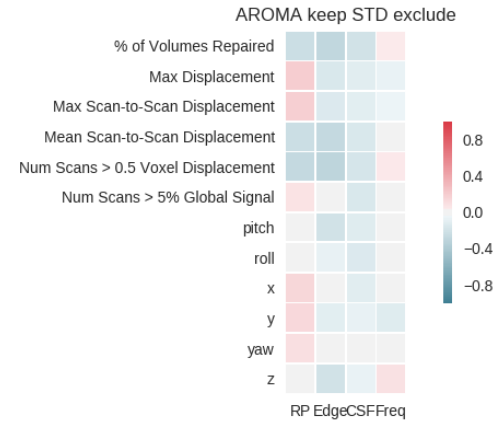


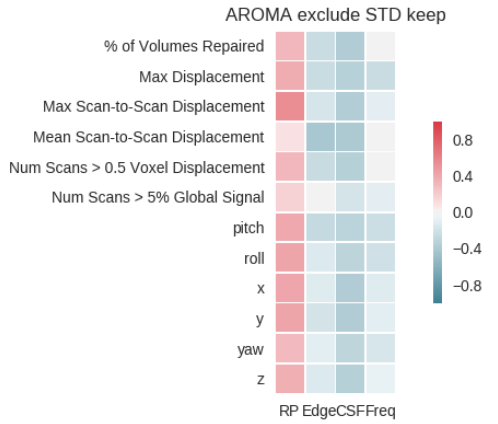
ICA-AROMA excludes 88 runs and the standard pipeline excludes 114 runs. It is highly unlikely that the runs excluded by ICA-AROMA are a subset of those excluded by the standard pipeline, and indeed we see that the two have very little in common:

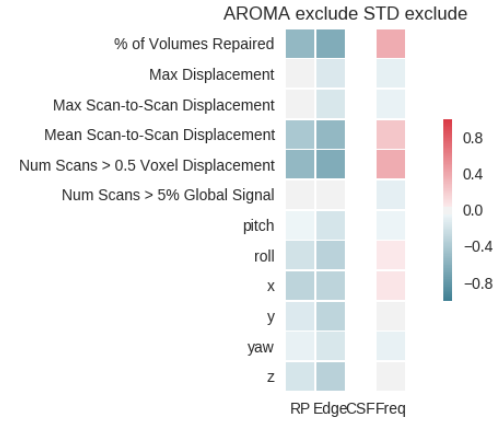
	ICA-AROMA keep	ICA-AROMA exclude
Standard keep	62.5%	13.4%
Standard exclude	21.2%	2.9%

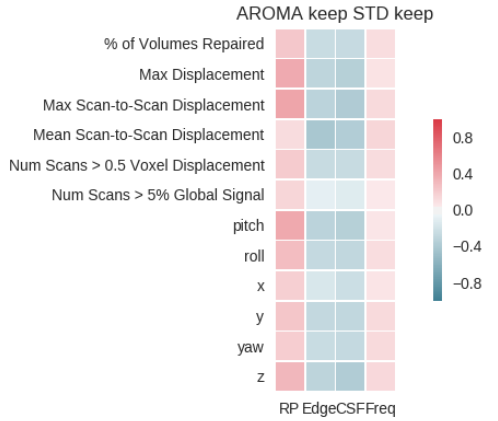
The rest of this analysis will focus on reconciling these differences and making recommendations on the use of ICA-AROMA moving forward.

Looking at correlations between ICA-AROMA criterion and movement summary stats reveals that ICA-AROMA excludes subjects with high correlations for max RP and keeps subjects that show generally moderate correlations across the board:









It is additionally revealing to look at specific subjects in the conditions where disagreement occurs. For the sake of simplicity, we will focus on two subjects, one excluded by ICA-AROMA but not the standard pipeline (12-02-11.1_3T3) and vice-versa (12-05-13.1_3T3).

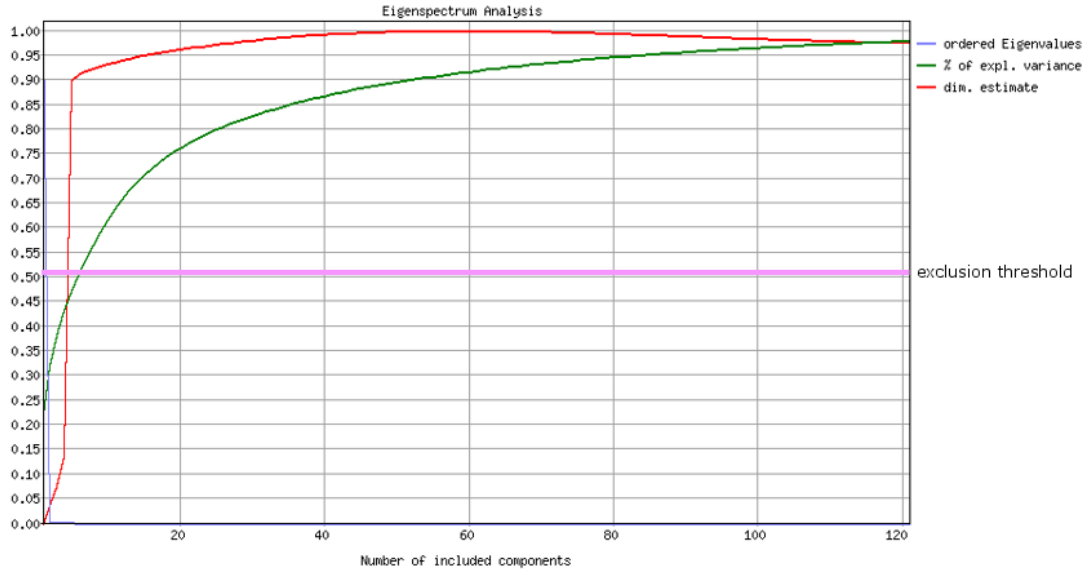
ID	Task	x	y	z	roll	pitch	yaw	% vol repair
12-02-11.1_3T3	comparison_number	0.28151	0.30368	0.9742	0.24567	0.55221	0.16574	0.5555
12-05-13.1_3T3	comparison_number_redo	7.20083	9.57536	18.667	7.55997	26.9800	11.9380	33.333

12-02-11.1_3T3 has fantastic movement stats, but was excluded by ICA-AROMA largely because the top three independent components were removed as noise:

IC	Motion/Noise	Max RP corr	Edge fraction	High freq cont	CSF fraction	Tot % Var Exp
1	True	0.86	0.61	0.40	0.00	0.22
2	True	0.68	0.23	0.40	0.01	0.31
3	True	0.72	0.52	0.55	0.03	0.38

These components were clearly removed due to high max RP correlation, which was likely found in one of the 72 parameters we do not typically consider, edge fraction, and high frequency content which are all artifacts missed by the

standard exclusion pipeline. Importantly, and obviously, this demonstrates that the top few components will dominate our threshold calculation:



12-05-13.1_3T3, on the other hand, has abhorrent movement stats and would easily be excluded using the standard criterion, but because none of the top components were classified as noise, we keep the subject using ICA-AROMA:

IC	Motion/Noise	Max RP corr	Edge fraction	High freq cont	CSF fraction	% Var Exp
1	False	0.79	0.3	0.07	0.01	0.14
2	False	0.79	0.25	0.21	0.01	0.25
3	False	0.55	0.37	0.12	0.02	0.34

Interestingly, we see very high Max RP correlation in the top two components, but because the edge fraction is within a normal range and these measures are combined to create a 2d threshold in ICA-AROMA's classification algorithm, the components are not classified as noise. **Looking at this final case reveals an important shortcoming of ICA-AROMA: it can easily miss issues with high motion stats due to the combined edge fraction/Max RP threshold**.

Conclusions

ICA-AROMA is highly effective at removing motion artifacts related to the edge fraction, high frequency content, and CSF fraction, but cannot be relied upon to exclude subjects with outrageous realignment parameters. **Subjects with glaring motion issues, i.e. that are well beyond the standard displacement threshold criterion, should first be excluded prior to running ICA-AROMA.** Only after removing these subjects can we be sure that ICA-AROMA will effectively remove problematic subjects. **This dual approach means that ICA-AROMA will likely result in more stringent or similar exclusion, and thus more or equal data loss, due to motion**. Overall, this means higher quality data and findings, but if the goal is to try and "save" data, it appears unlikely that ICA-AROMA will allow for this.