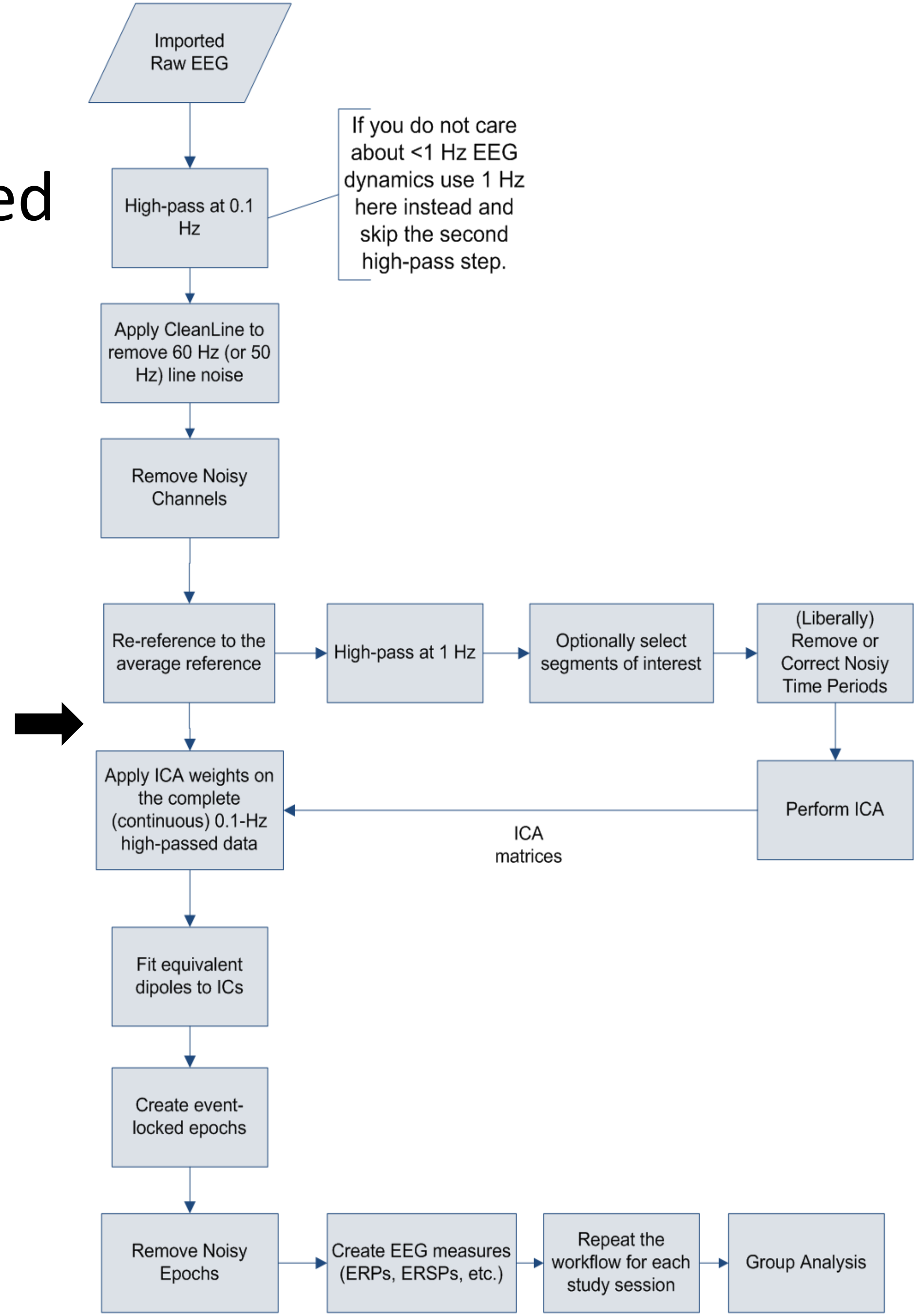
Comparison of Pre-ICA Artifact Correction Methods

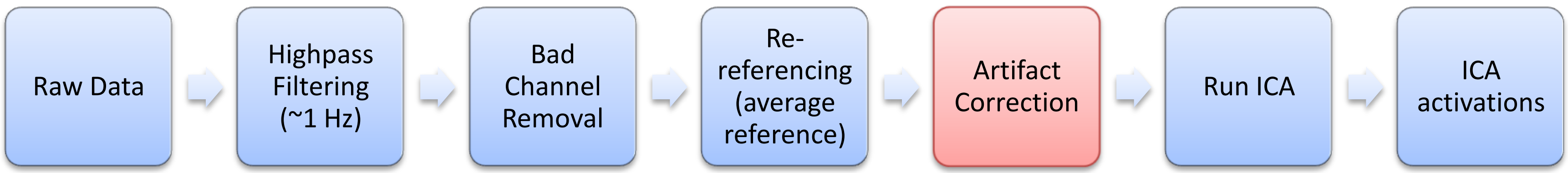
Nima Bigdely-Shamlo, Christian Kothe,
Ken Kreutz-Delgado, Jason Palmer,
Scott Makeig
EEGLAB Workshop 2013

Recommended ICA-based workflow

Clean aggressively before ICA, then apply ICA matrix back to the original data. Perform epoch rejection to deal with artifacts.



ICA Workflow



ICA Artifact Correction:

Removal of or change in data from portions of the EEG recording in order to improve the quality of ICA decomposition.

But, how can we measure “the quality of ICA decomposition”?

1. Number of physiologically plausible (e.g. Dipolar) components.
2. Mutual Information reduction Index: measure of ICA effectiveness on a given data segment.

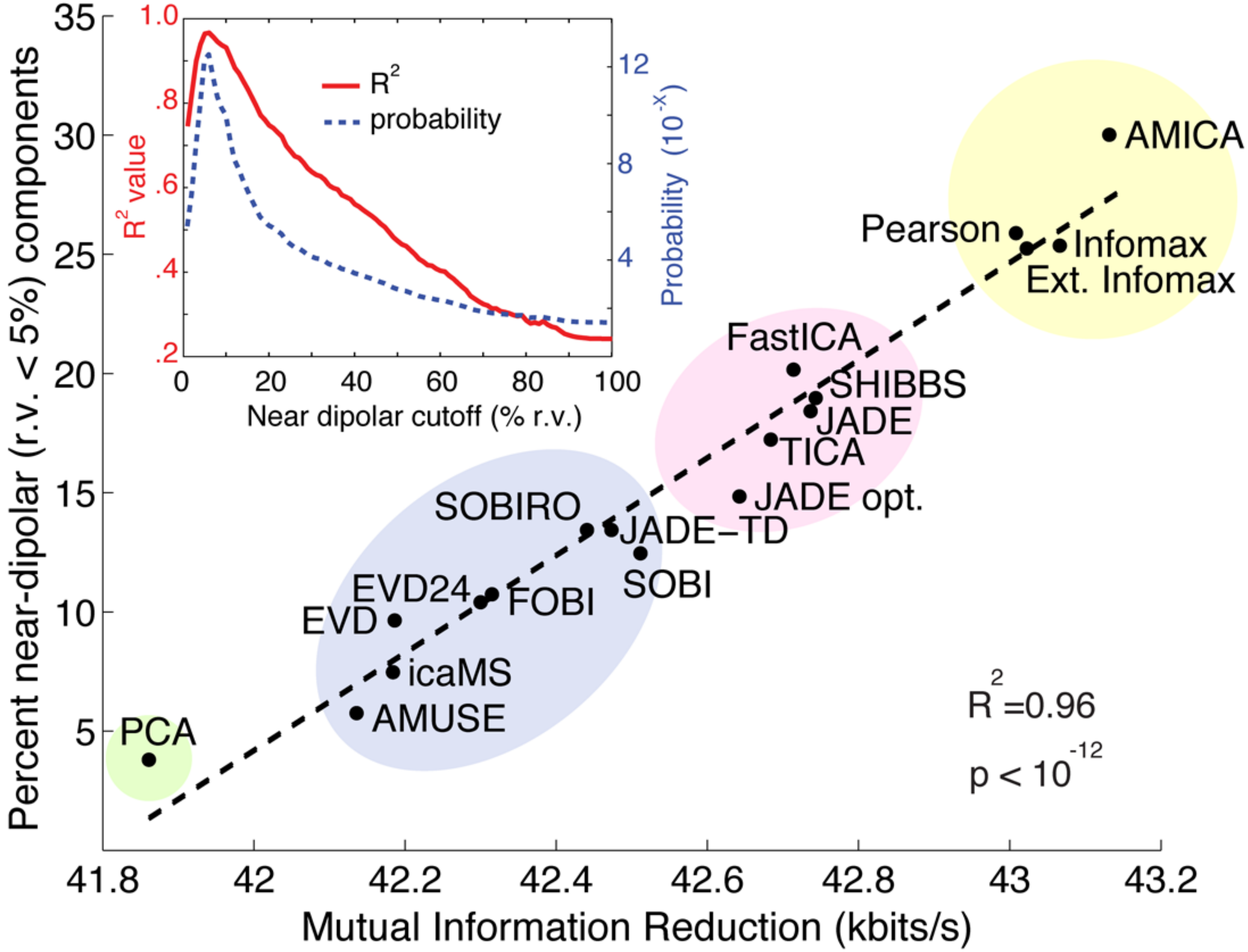
Also remember that often the ICA matrix is used on the whole session and not only the parts that were provided to the ICA algorithm.

ICA Quality Measurement

Mutual Information Reduction index (MIR) tells us how well ICA has done its job for a given portion of the data. Here are the candidates for an ICA quality measure to be used in comparing different pre-processing methods:

1. Number of Dipolar ICs
2. Whole-session MIR
3. Clean-portion MIR
4. Mean Windowed-MIR
5. Median Windowed-MIR

ICA Quality Measurement



From Delorme et al., Independent EEG sources are dipolar PLoS One 2012

ICA Quality Measurement: Number of Dipolar ICs

It is tempting to count the number of 'good looking' or 'dipolar' (low residual variance when fitting by a single dipole or dual symmetric dipoles) and use it as a measure of ICA success. After all we often only use these dipoles in further analysis.

But there are some problem with this approach:

The ICA matrix is often applied to the whole data, but trained only on a portion of the data that assumed to be free of any artifacts. An artifact correction method may remove certain parts of the data that result in a higher number of dipolar ICs but produce low-quality (low MIR) when applied to the whole data.

For example it can reject all eye-blink portions, resulting in more degrees of freedom for brain sources but worst ICA quality on the whole-data

ICA Quality Measurement: Whole-session MIR

Since MIR is generally lower on portions of data not included in the ICA training, this measure is usually higher when no portion of the data is rejected.

So whole-session MIR is **next to useless** as an ICA quality index for comparing different pre-processing methods.

But the method is quite useful for comparing ICA methods, as shown in Delorme, et al. 2012.

ICA Quality Measurement: Clean-portion MIR

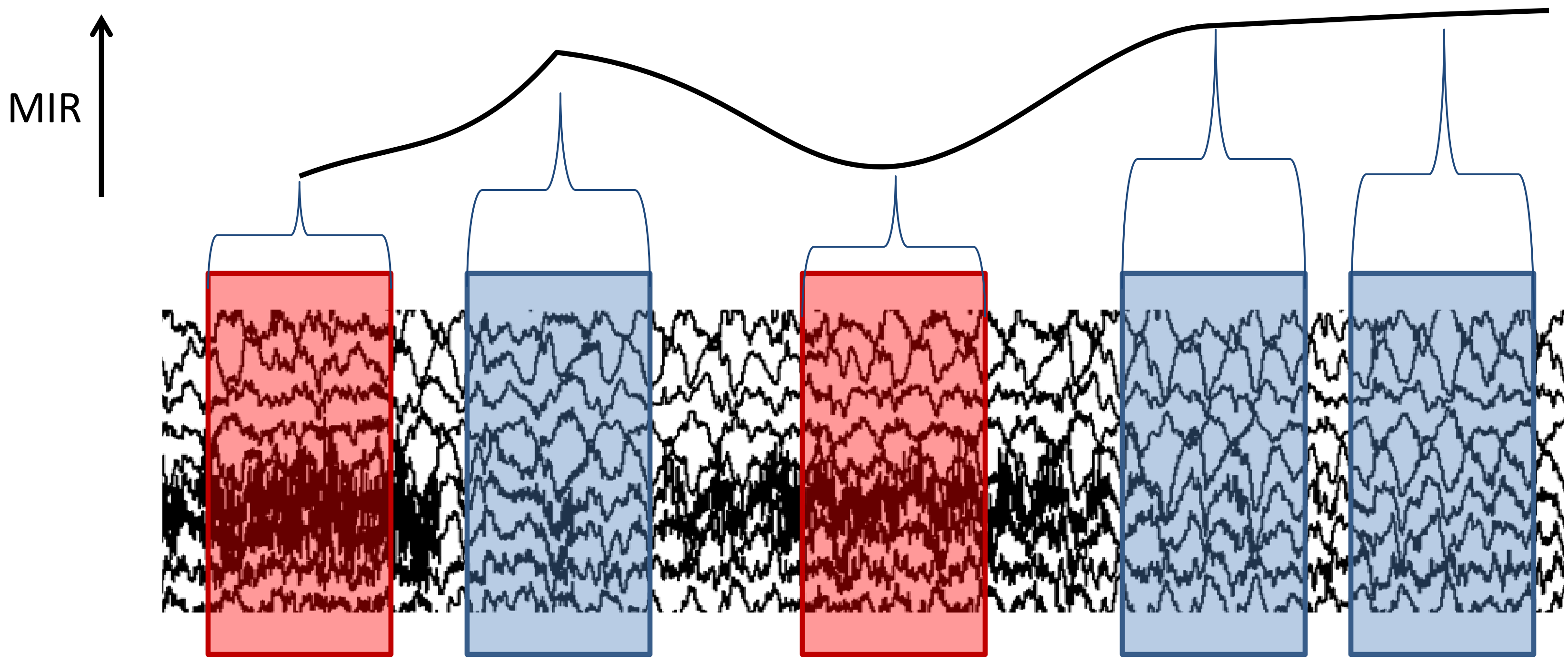
There is no ground truth when it comes to ICA pre-processing. Hand-cleaning is highly variable across individuals and, as we will see, could often do worst than new automatic rejection methods.

Since the portion of the data to be rejected is determined by the pre-processing method, this index could be easily unfair to methods that keep more of the non-stationary (but still valid) portions of the data.

Again, since often the ICA matrix is applied to the whole-data, the same issue raised as in 1.C could arise.

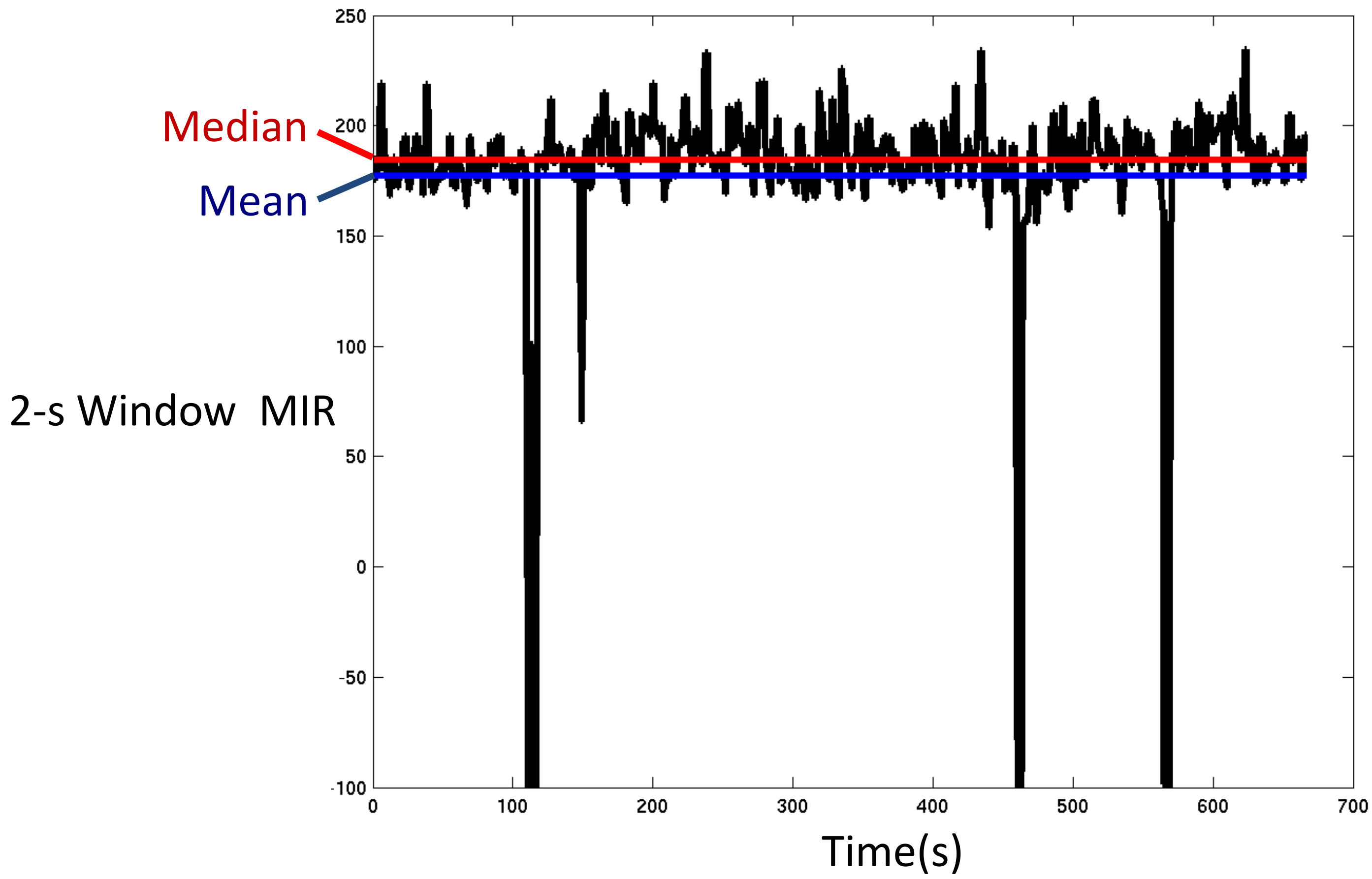
ICA Quality Measurement: Mean Windowed-MIR

MIR may be calculated in windows with lengths of a few seconds. Since most of such short windows may be assumed to be free of artifacts, the average of these values may be less affected by the presence of artifacts than whole-session MIR.

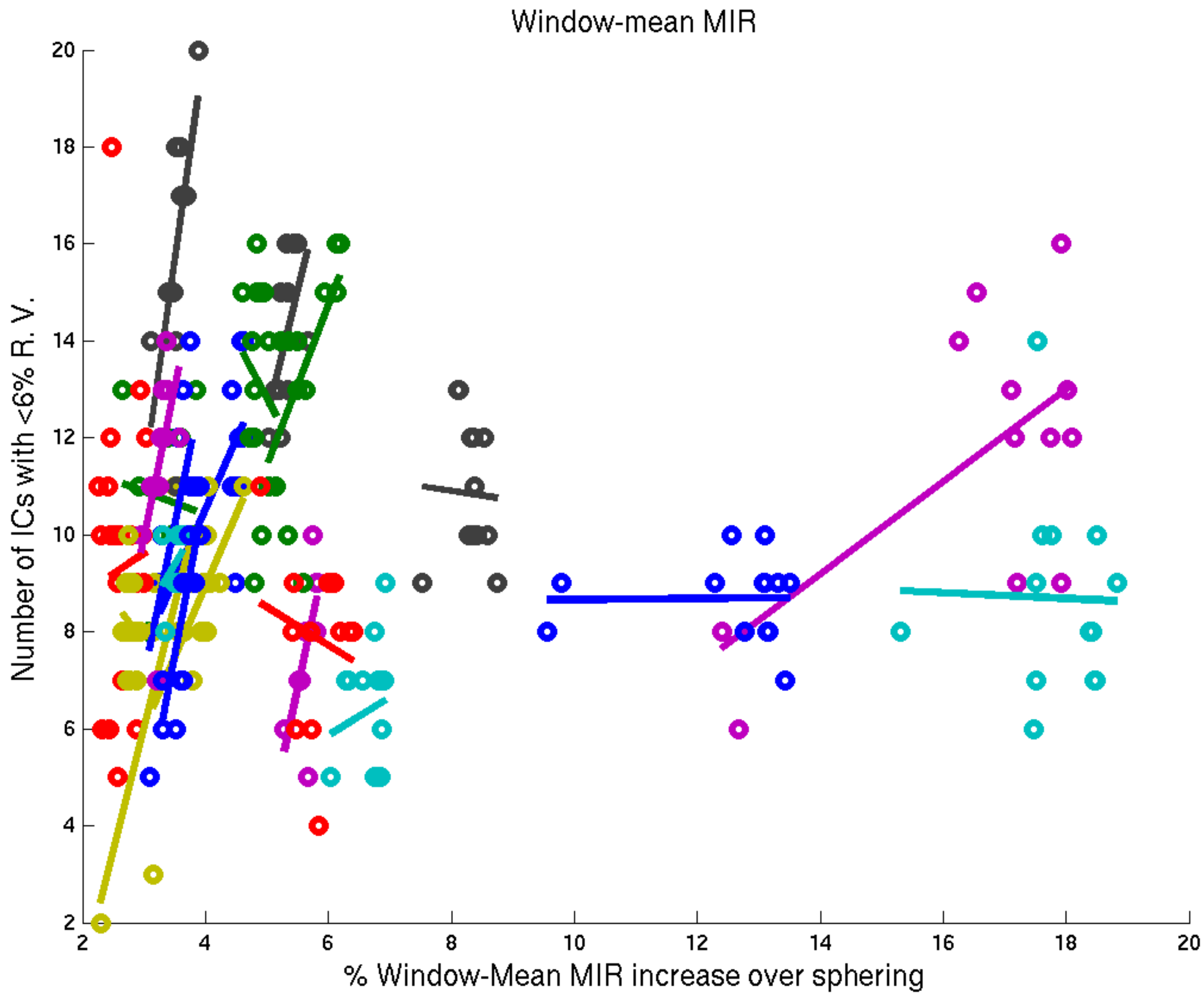


ICA Quality Measurement: Median Windowed-MIR

To obtain a more robust (to artifacts) ICA quality measure, the median of windowed MIR values may be used.

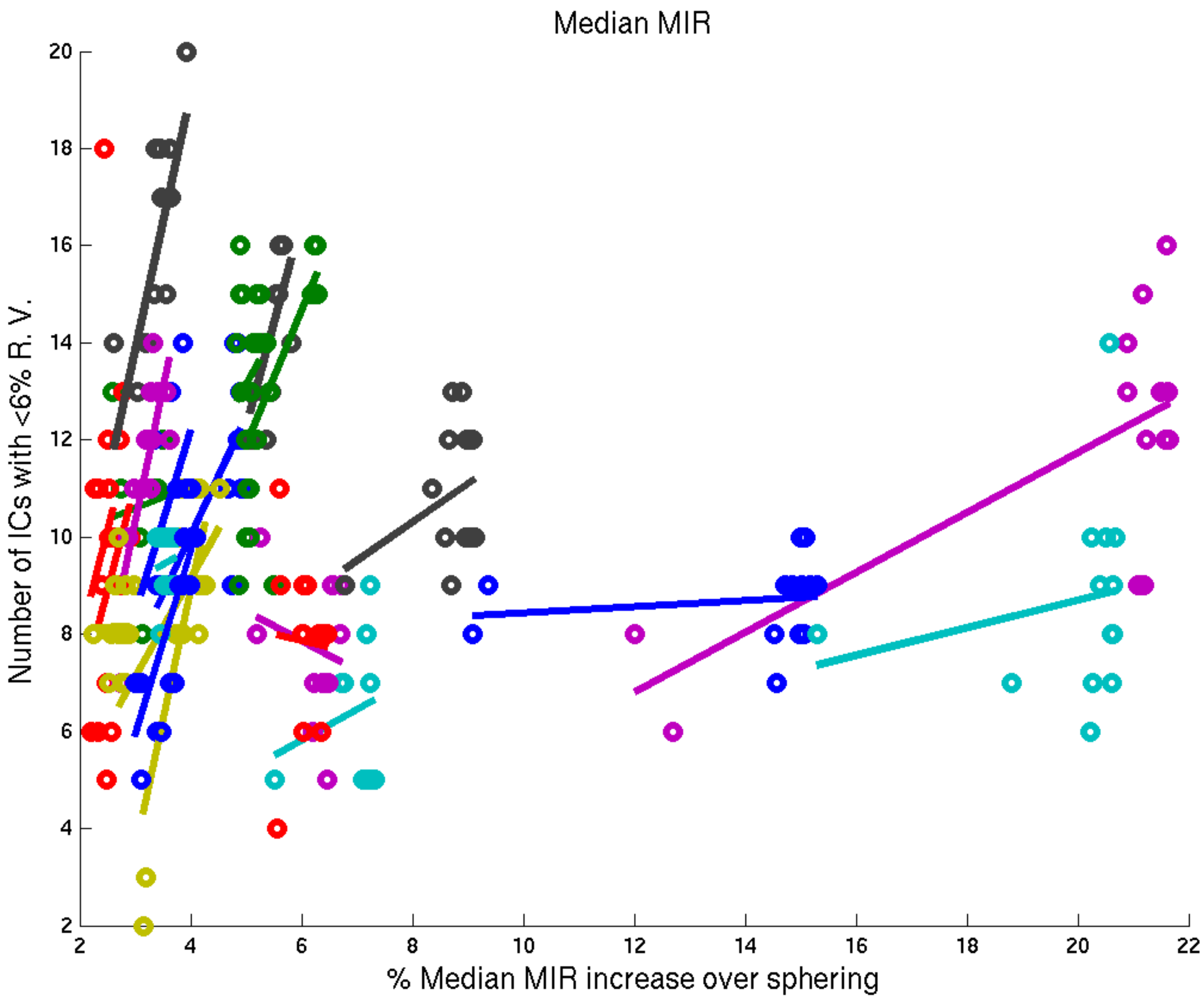


ICA Quality Measurement



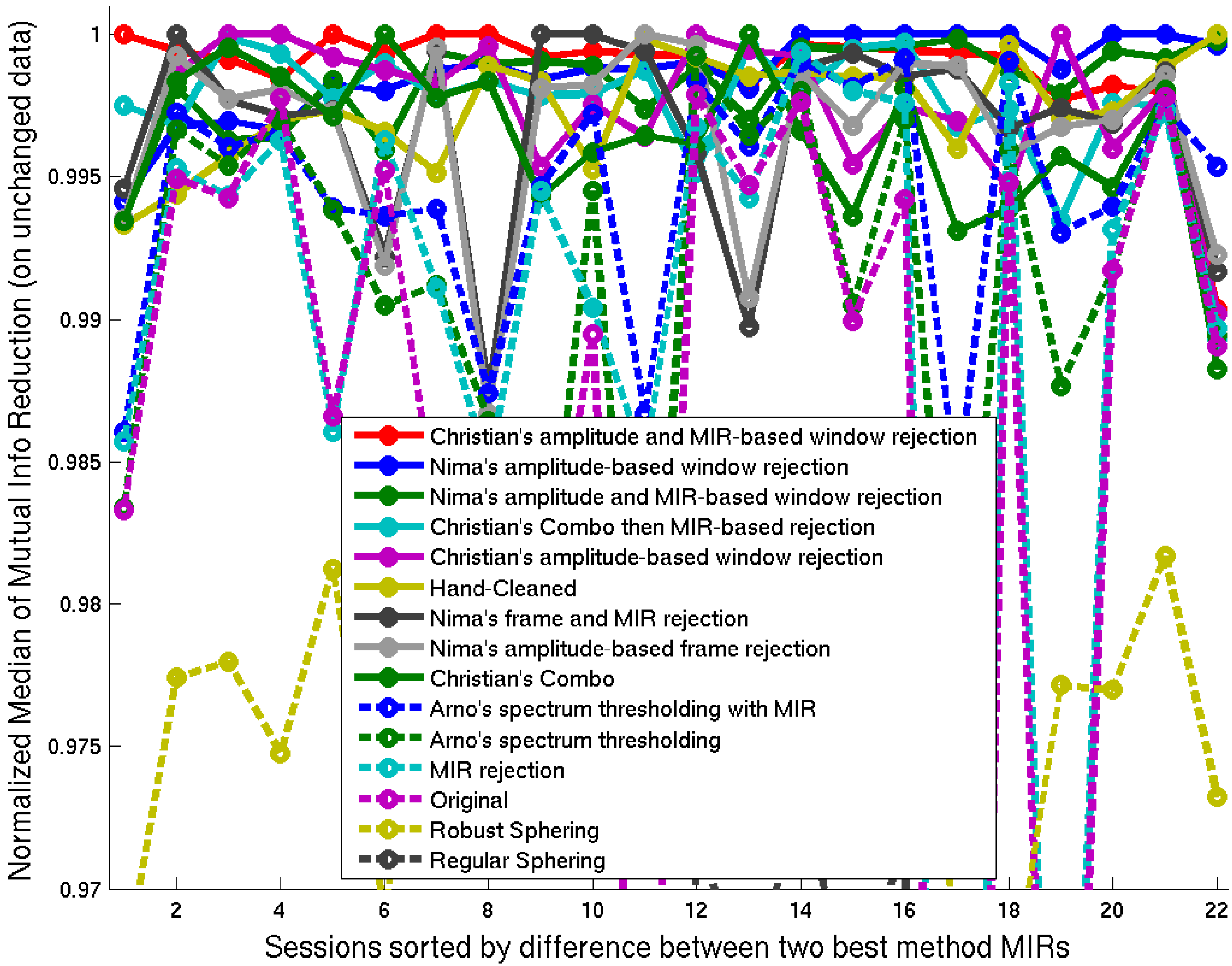
Correlation = 0.388

ICA Quality Measurement

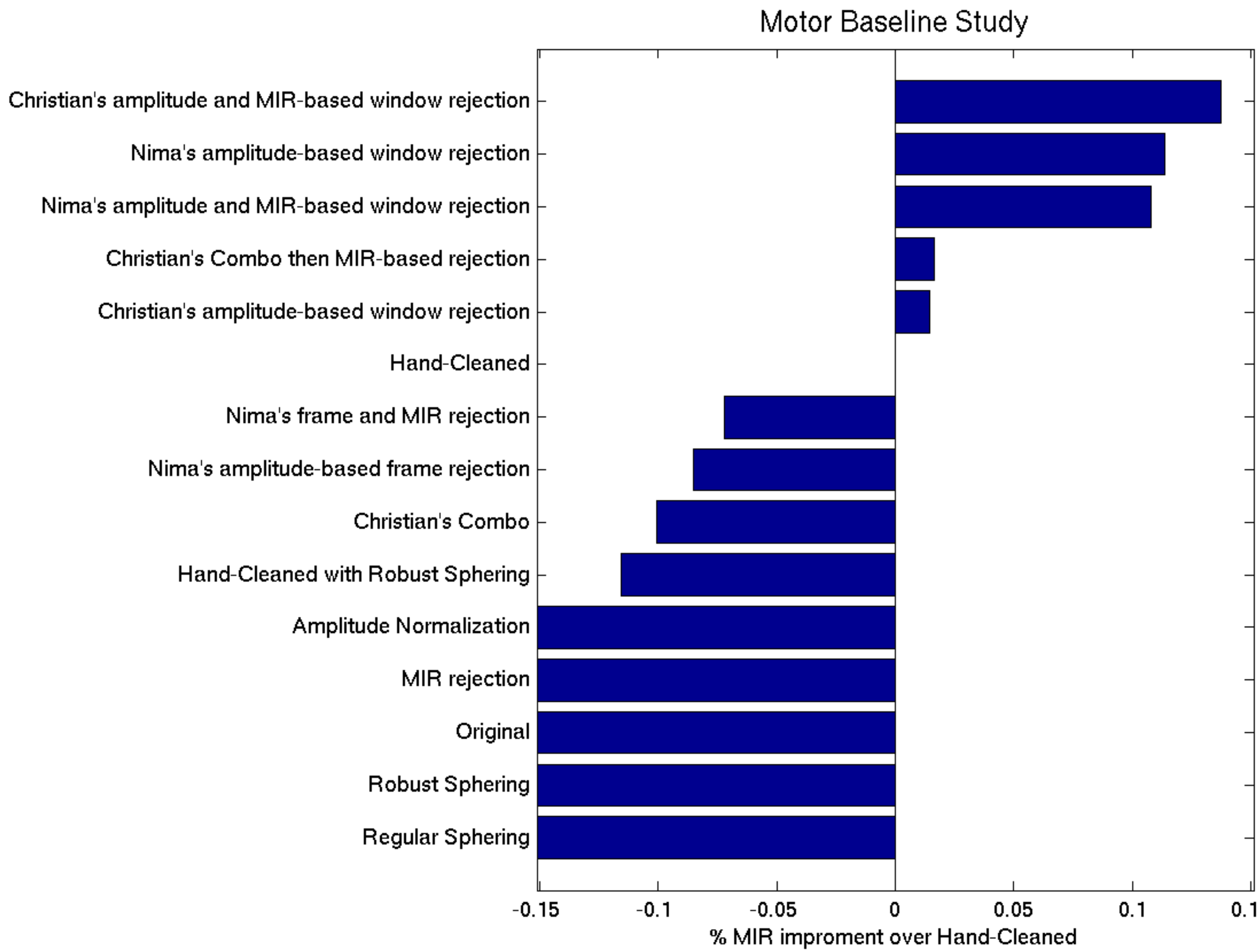


Correlation = **0.431**

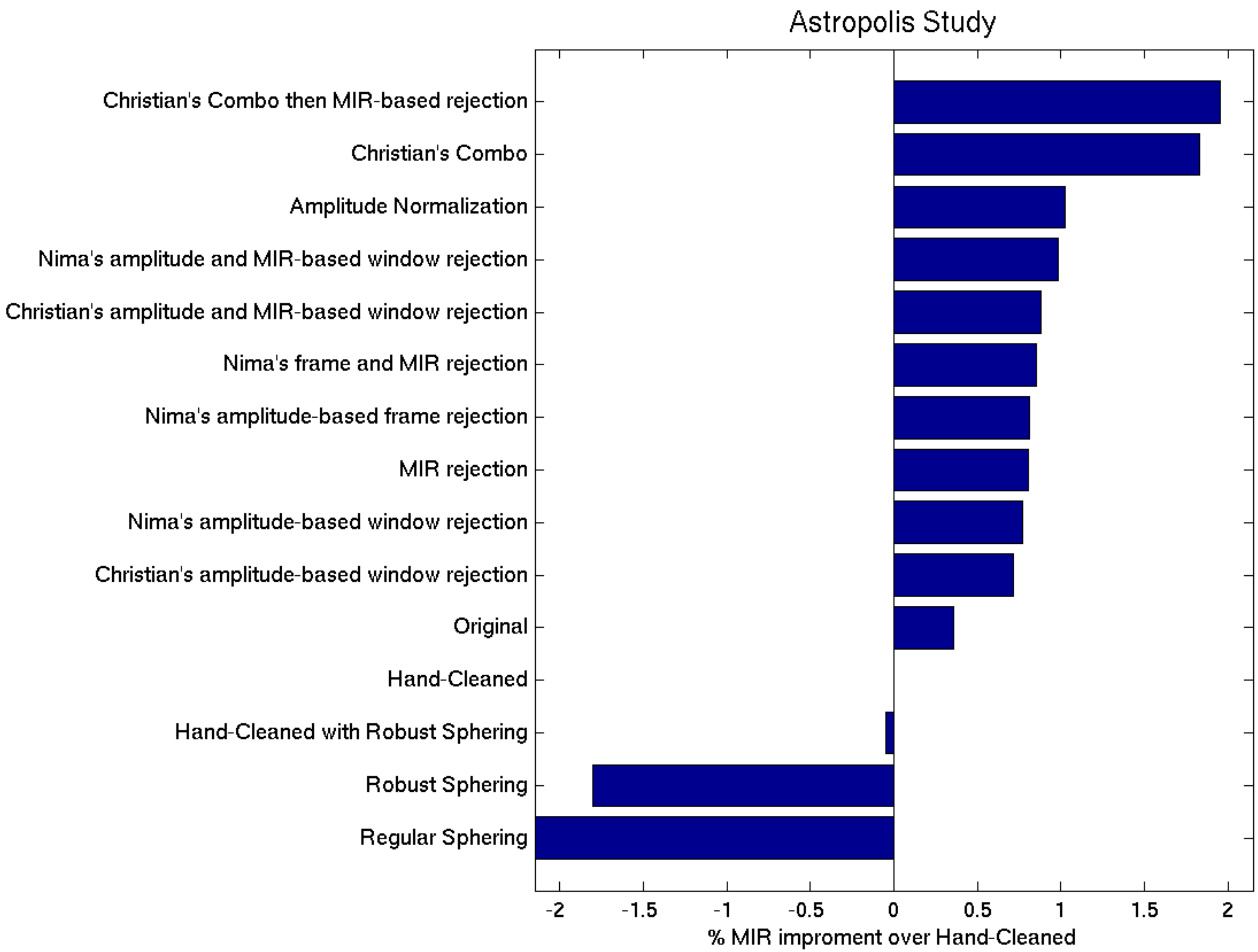
Results for several automated methods



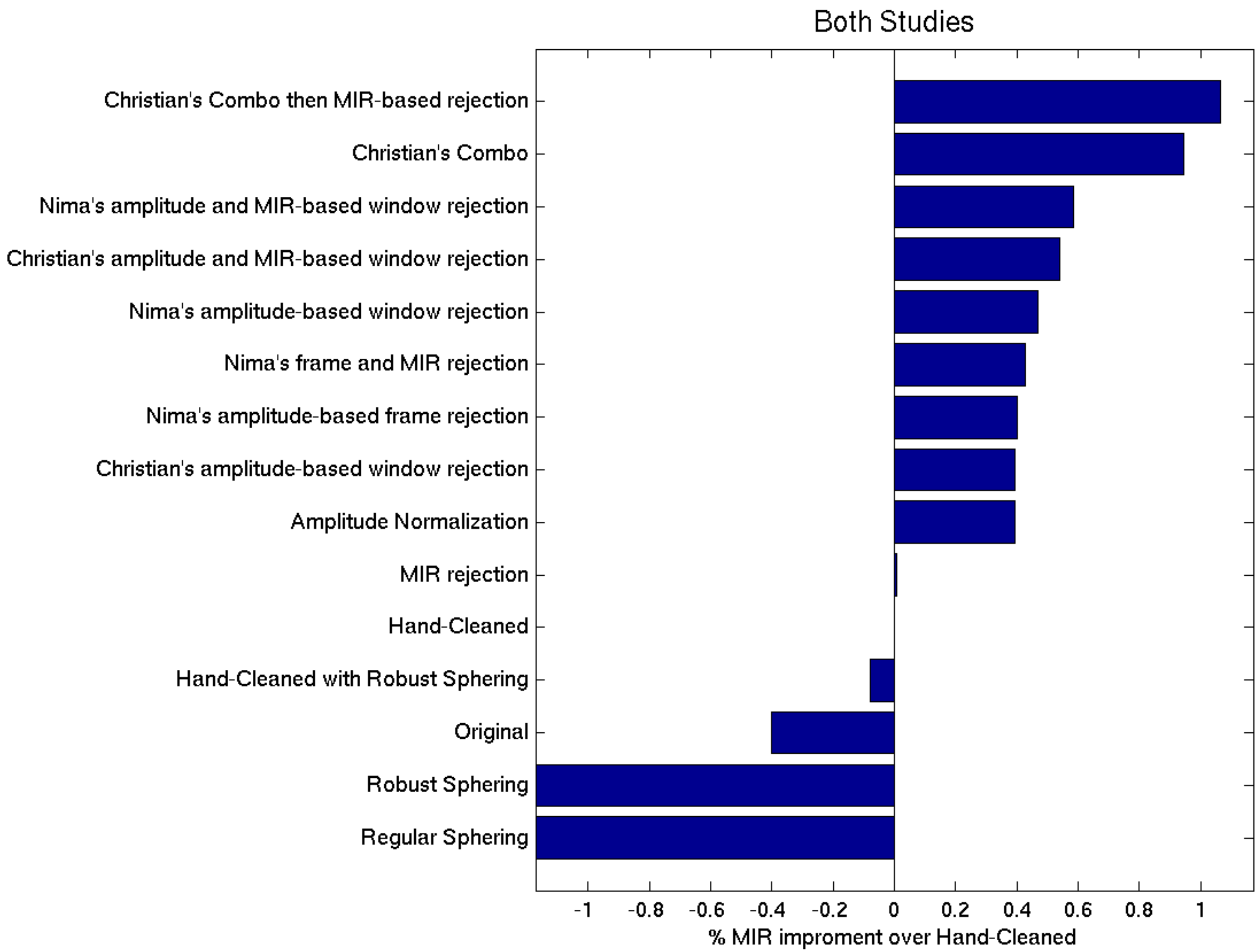
Results for several automated methods



Results for several automated methods



Results for several automated methods



Results for several automated methods

Rank (first on percent session it was the best, then median of MIR) – Method name: (Median of Median Window MIR increase over Hand-Cleaning), percent sessions were it was the best method.

1 - Nima's amplitude-based window rejection: 0.122% MIR, Best in 32%

2 - Christian's amplitude and MIR-based window rejection: 0.102% MIR, Best in 18%

3 - Christian's amplitude-based window rejection: 0.075% MIR, Best in 18%

4 - Nima's frame and MIR rejection: 0.011% MIR, Best in 14%

5 - Christian's Combo: -0.071% MIR, Best in 9%

6 - **Hand-Cleaned: 0.000% MIR, Best in 5%, performed by a veteran ICA user.**

7 - Nima's amplitude-based frame rejection: -0.008% MIR, Best in 5%

8 - Nima's amplitude and MIR-based window rejection: 0.064% MIR, Best in 0%

9 - Christian's Combo then MIR-based rejection: 0.059% MIR, Best in 0%

10 - Arno's spectrum thresholding with MIR: -0.194% MIR, Best in 0%

11 - Arno's spectrum thresholding: -0.382% MIR, Best in 0%

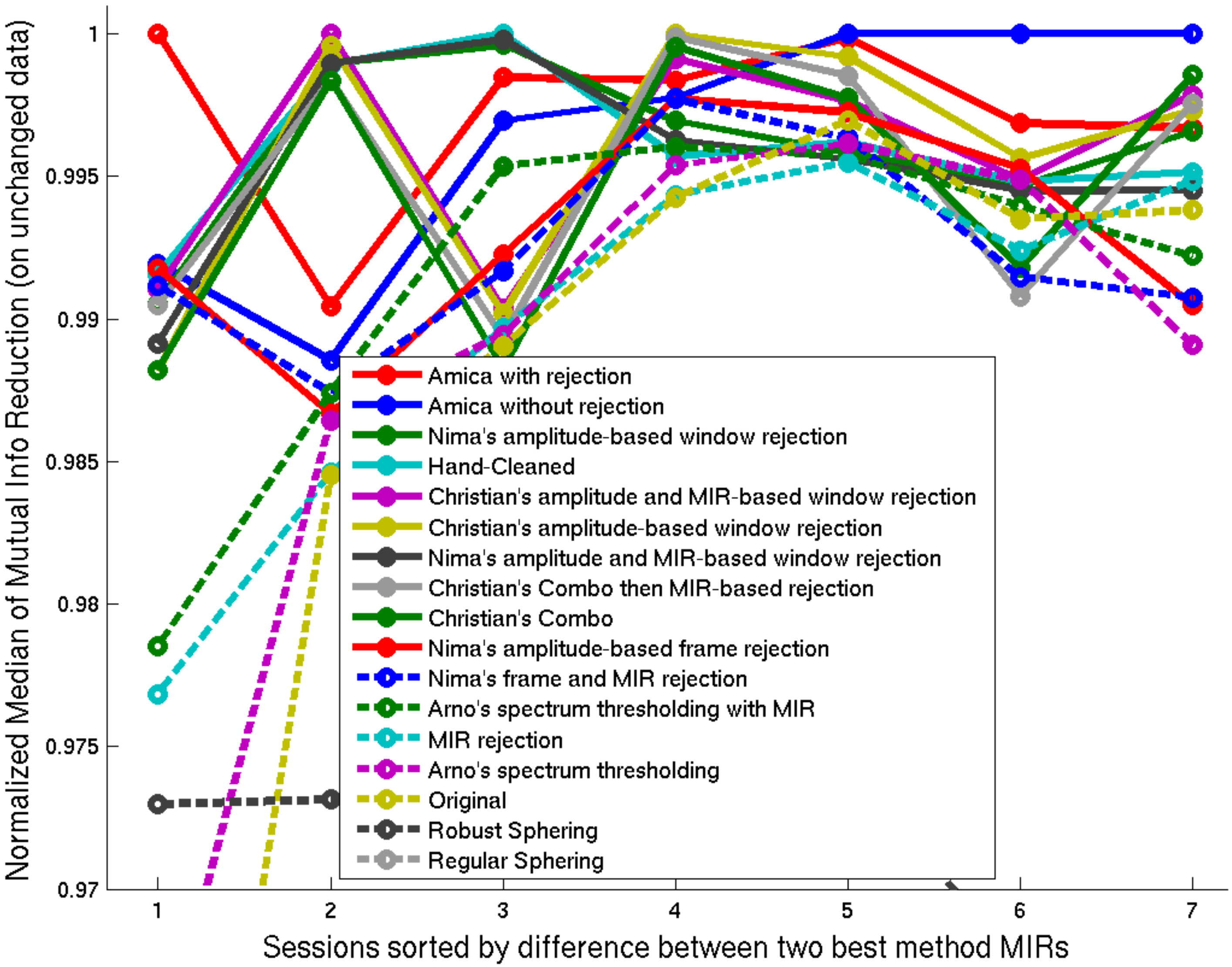
12 - MIR rejection: -0.397% MIR, Best in 0%

13 - Original: -0.572% MIR, Best in 0%

14 - Robust Sphering: -1.932% MIR, Best in 0%

15 - Regular Sphering: -4.117% MIR, Best in 0%

Results with 7 AMICA sessions



Results with 7 AMICA sessions

1 - Amica without its rejection: 0.205% MIR, Best in 43%

2 - Amica with its rejection: 0.205% MIR, Best in 14%

3 - Christian's amplitude and MIR-based window rejection: 0.109% MIR, Best in 14%

4 - Christian's amplitude-based window rejection: 0.081% MIR, Best in 14%

5 - Hand-Cleaned: 0.000% MIR, Best in 14%

6 - Nima's amplitude-based frame rejection: 0.020% MIR, Best in 0%

7 - Nima's amplitude-based window rejection: -0.015% MIR, Best in 0%

8 - Nima's amplitude and MIR-based window rejection: -0.032% MIR, Best in 0%

9 - Christian's Combo then MIR-based rejection: -0.053% MIR, Best in 0%

10 - Christian's Combo: -0.054% MIR, Best in 0%

11 - Original: -0.146% MIR, Best in 0%

12 - MIR rejection: -0.243% MIR, Best in 0%

13 - Arno's spectrum thresholding with MIR: -0.296% MIR, Best in 0%

14 - Nima's frame and MIR rejection: -0.338% MIR, Best in 0%

15 - Arno's spectrum thresholding: -0.611% MIR, Best in 0%

16 - Robust Sphering: -2.579% MIR, Best in 0%

17 - Regular Sphering: -4.714% MIR, Best in 0%

Conclusions

- In most (~95% of) cases, an automated method exists that does a better job at pre-ICA cleaning than manual cleaning.
- There are 2-3 good automated methods for this job.
- AMICA does a better job of dealing with artifacts (in 43% of cases no rejection was the best option).

Pre-ICA Cleaning (PIC) Index

We can normalize the median windowed MIR by the MIR value of median based sphering. This tell us how well ICA has done its job, compared to a robust sphering-only signal separation.

$$PIC = 100 \times \left(\frac{MIR(\text{whole data, ICA matrix from cleaning method})}{MIR(\text{whole data, robust sphering matrix})} - 1\right)$$

PIC from different cleaning methods can then be compared to find the best one.
For a good ICA decomposing PIC has to be greater than 0.

Rule of thumb: PIC over 1.5% often indicates a good ICA.

You can download PICI function long with several automatic pre-ica cleaning functions from:

http://sccn.ucsd.edu/wiki/PICI