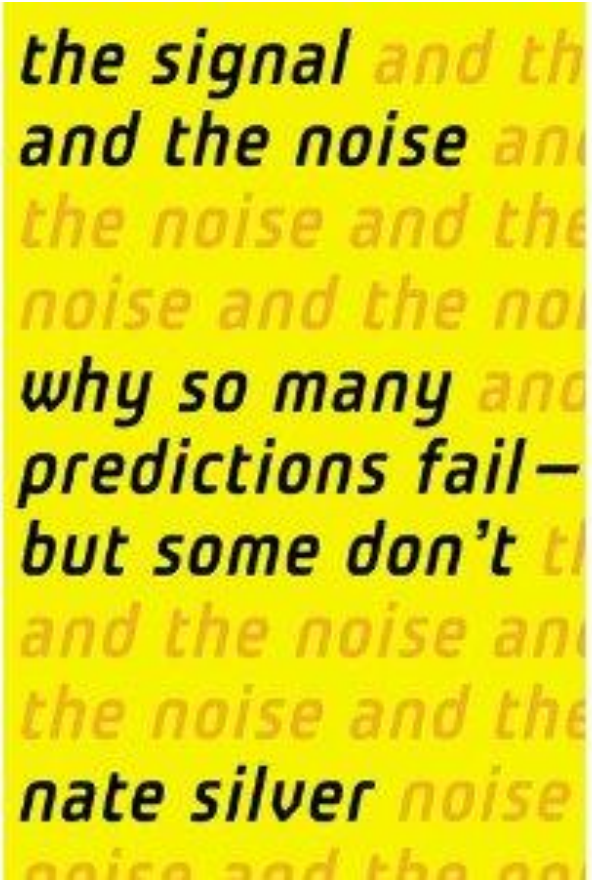
Clean Data

Getting Rid of those Pesky Artifacts in your Data

The Signal and the Noise

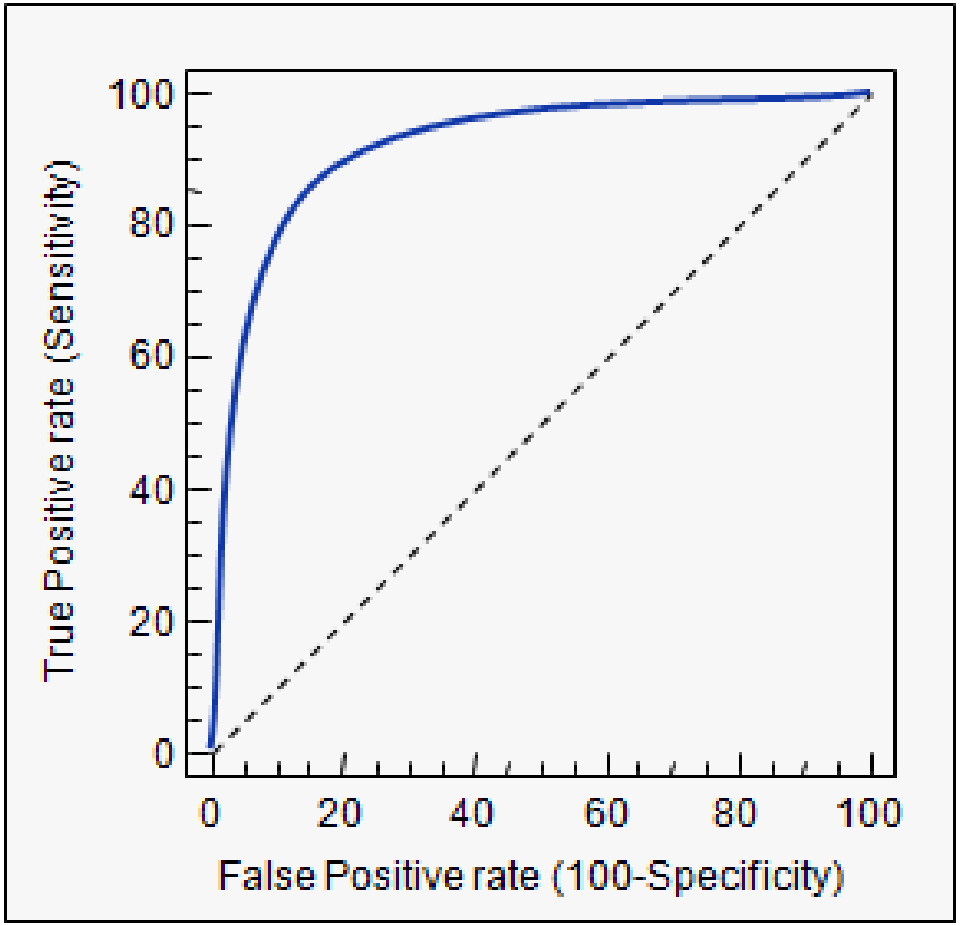
$$\mathrm{SNR} = \frac{P_{\mathrm{signal}}}{P_{\mathrm{noise}}} = \left(\frac{A_{\mathrm{signal}}}{A_{\mathrm{noise}}}\right)^2,$$

$$\mathrm{SNR} = \frac{\mu}{\sigma}$$



Effects of Noise on Decision-Making

ROC Curve



“Noise” vs. “Artifacts”

Type 2 Error / “Beta” Error

Type 1 Error / “Alpha” Error

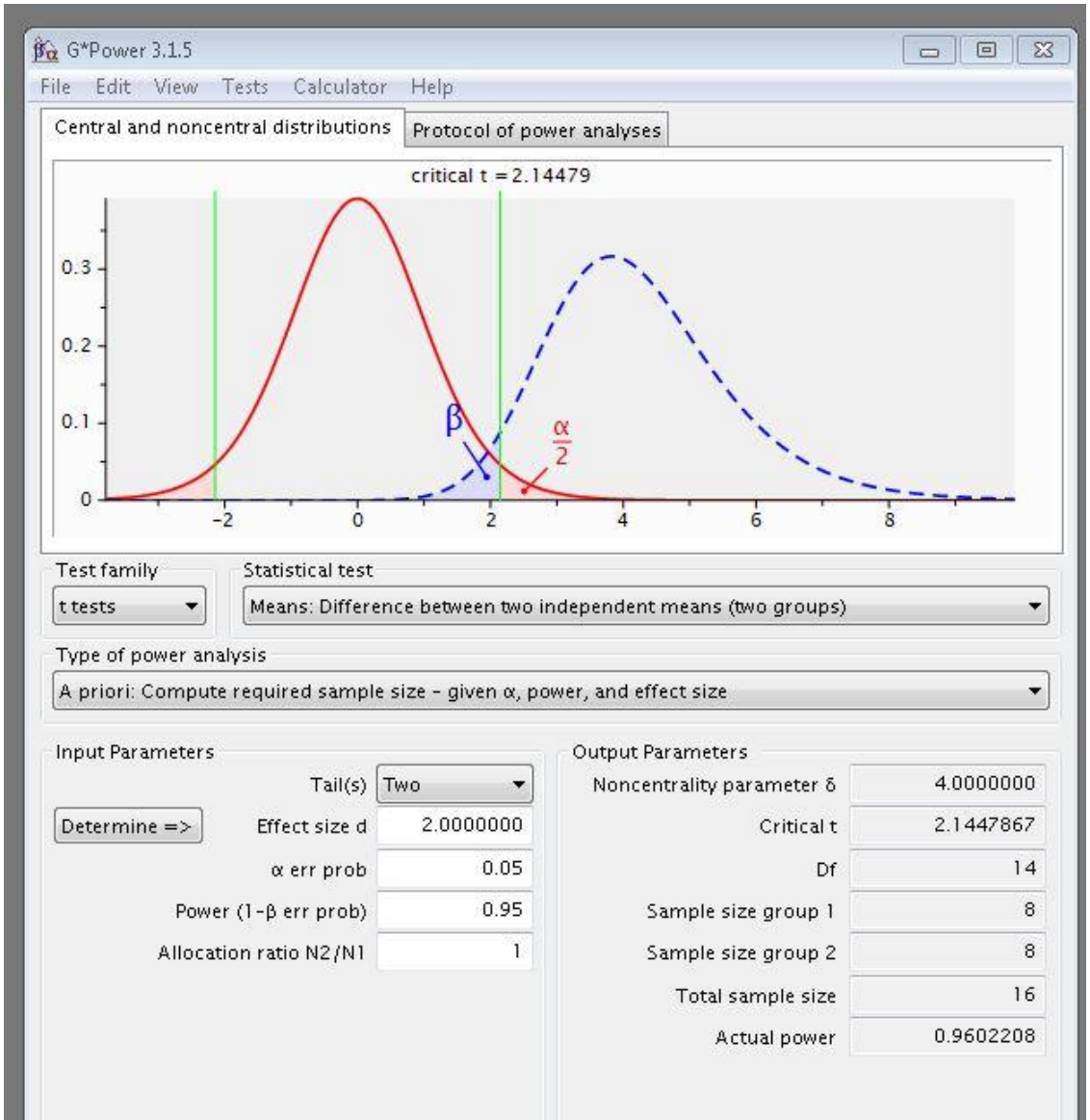
Artifacts can be…

1. Mechanical, Biological, or Psychological, in Origin.

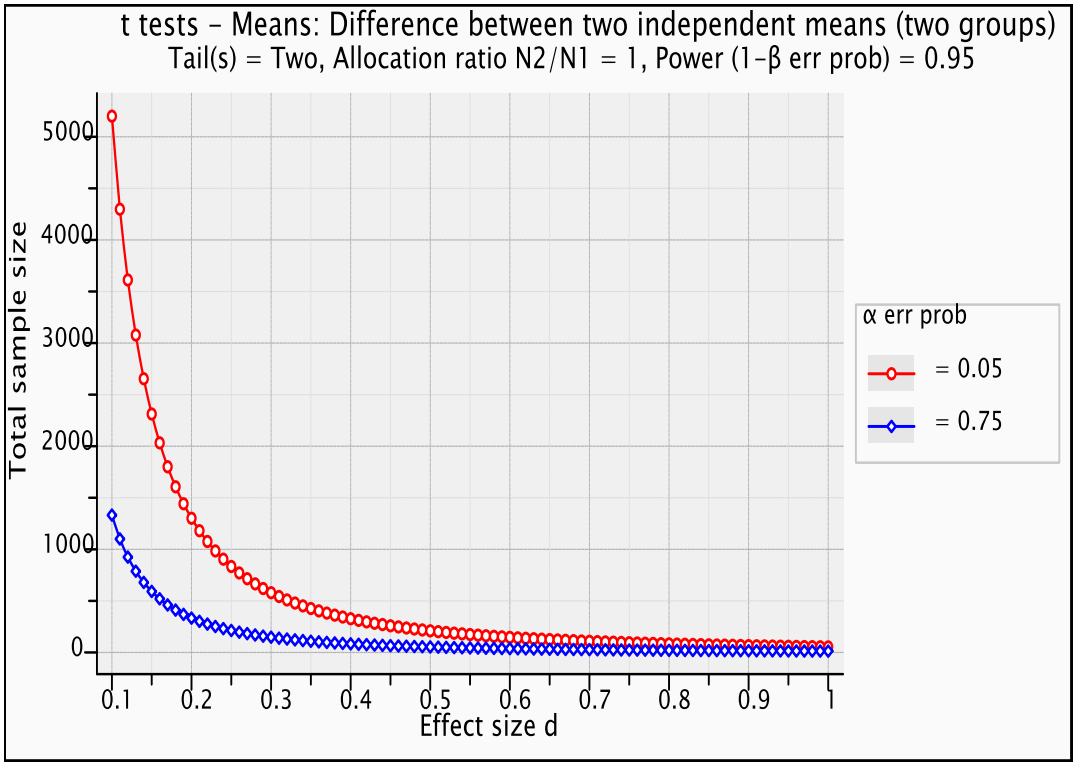
2. Systematic or Random in Time.
   - Can't I just average it out?

3. Stationary or Nonstationary in Space.
   - Can't I just regress it out?

A Priori Power Analysis

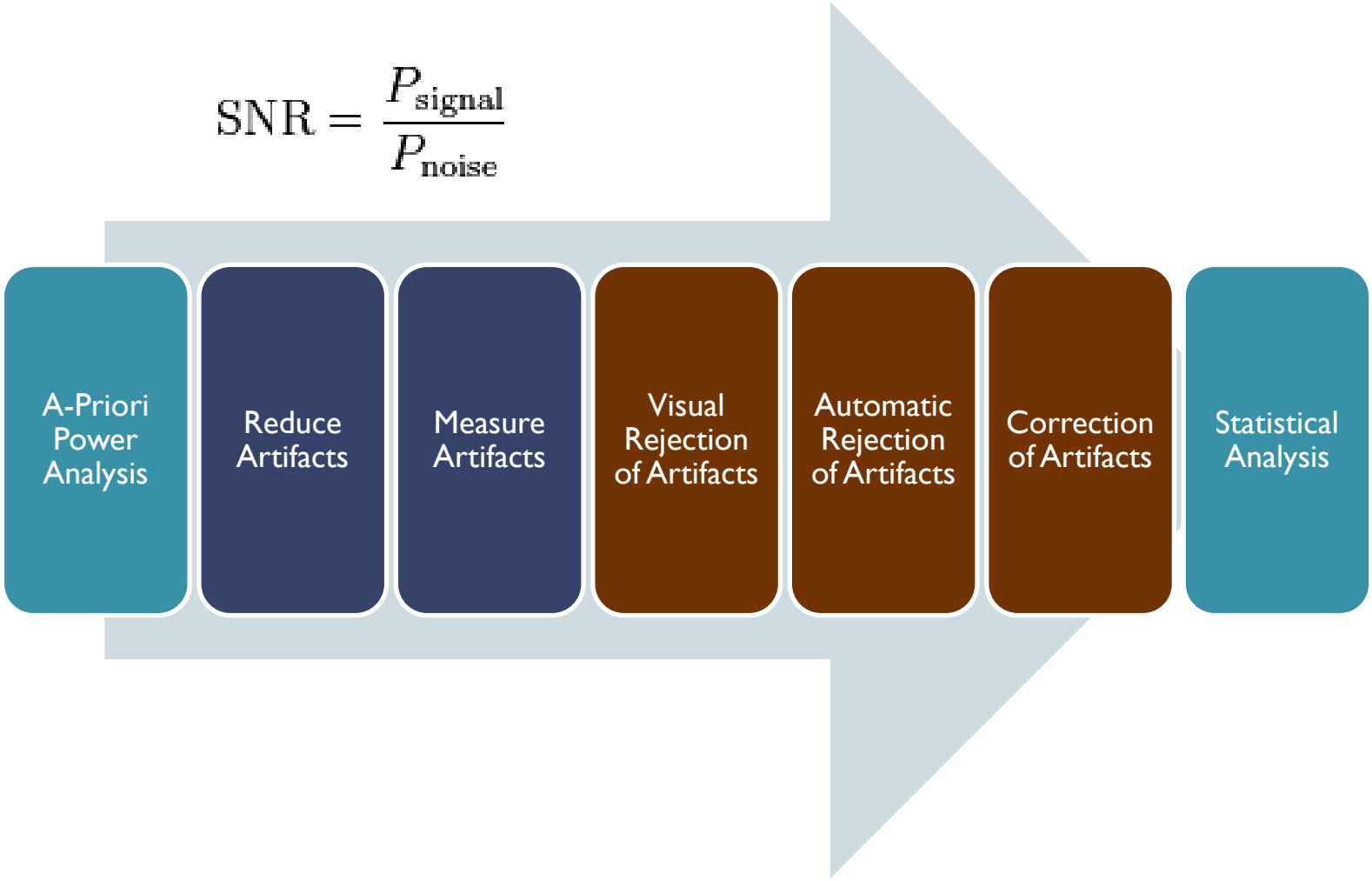


How Powerful is your Data?



The Artifact Action Plan

$$\mathrm{SNR} = \frac{P_{\mathrm{signal}}}{P_{\mathrm{noise}}}$$



1) Collect Powerful Data from the Start

2) Record Your Artifacts Directly

- Makes it easier to interpret your data.

- Examples:
  - Eye Movements: EOG Channels, Eyetracker
  - Muscle Activity: EMG Channels
  - Heart Activity: EKG Channels
  - Head Movements: Head Coils (MEG)

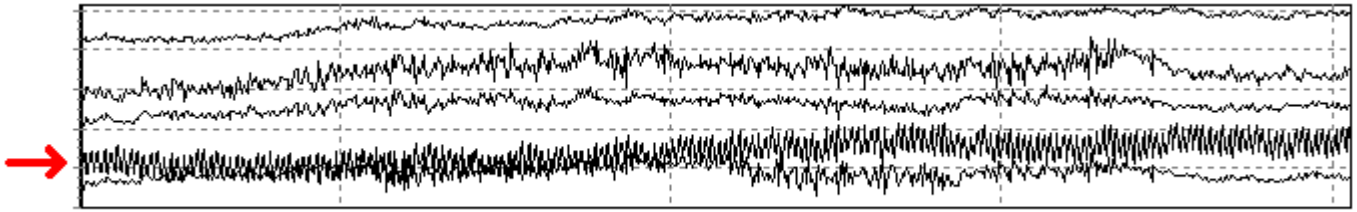
- Don't rely on clear artifact signals in your brain data for source separation!

Visual Artifact Rejection

- Advantages:
  - Simple, Powerful
  - Makes other methods more effective

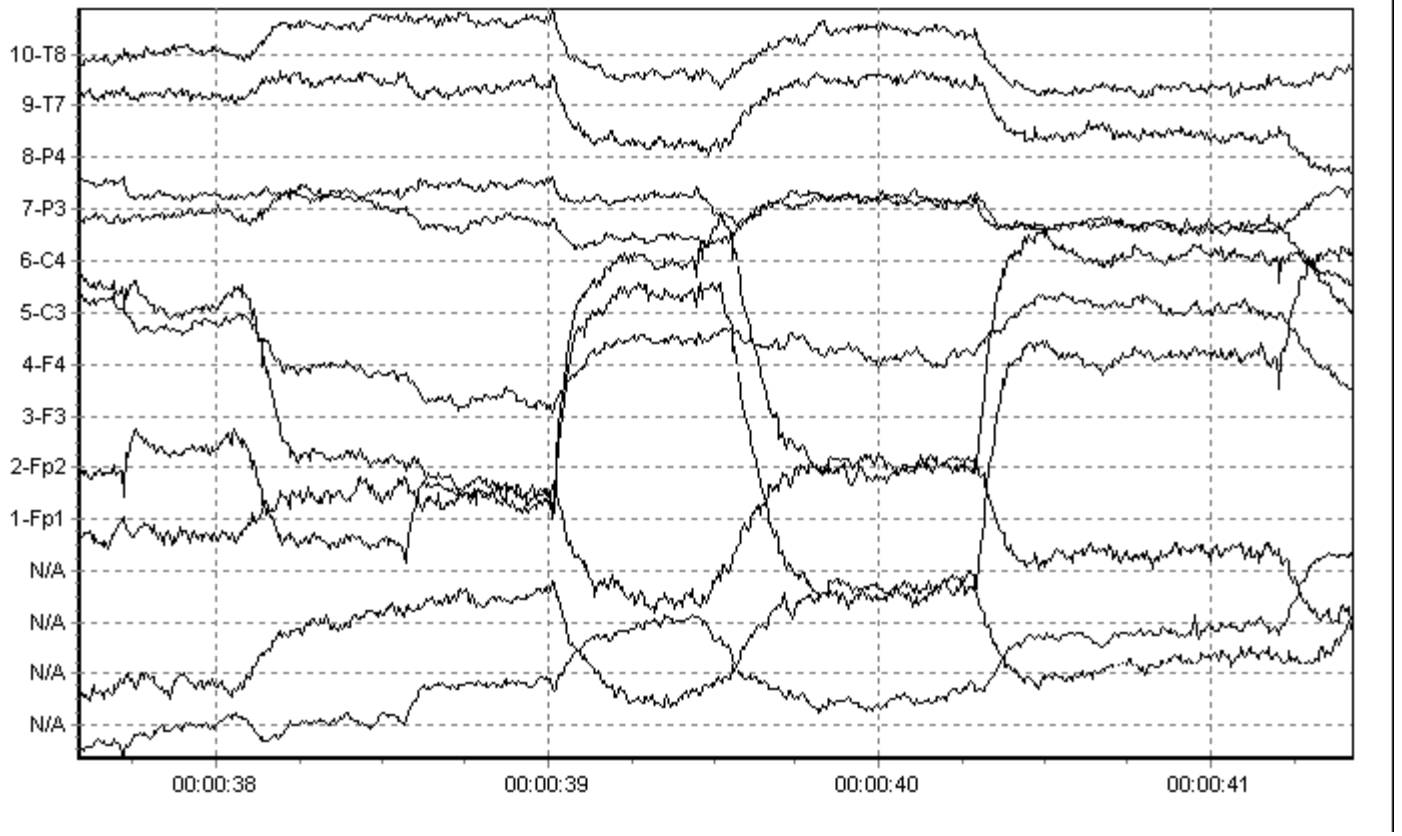
- Disadvantages:
  - Time Consuming
  - Subjective
  - Leaves less data to work with

Bad Channel / Power Line Noise

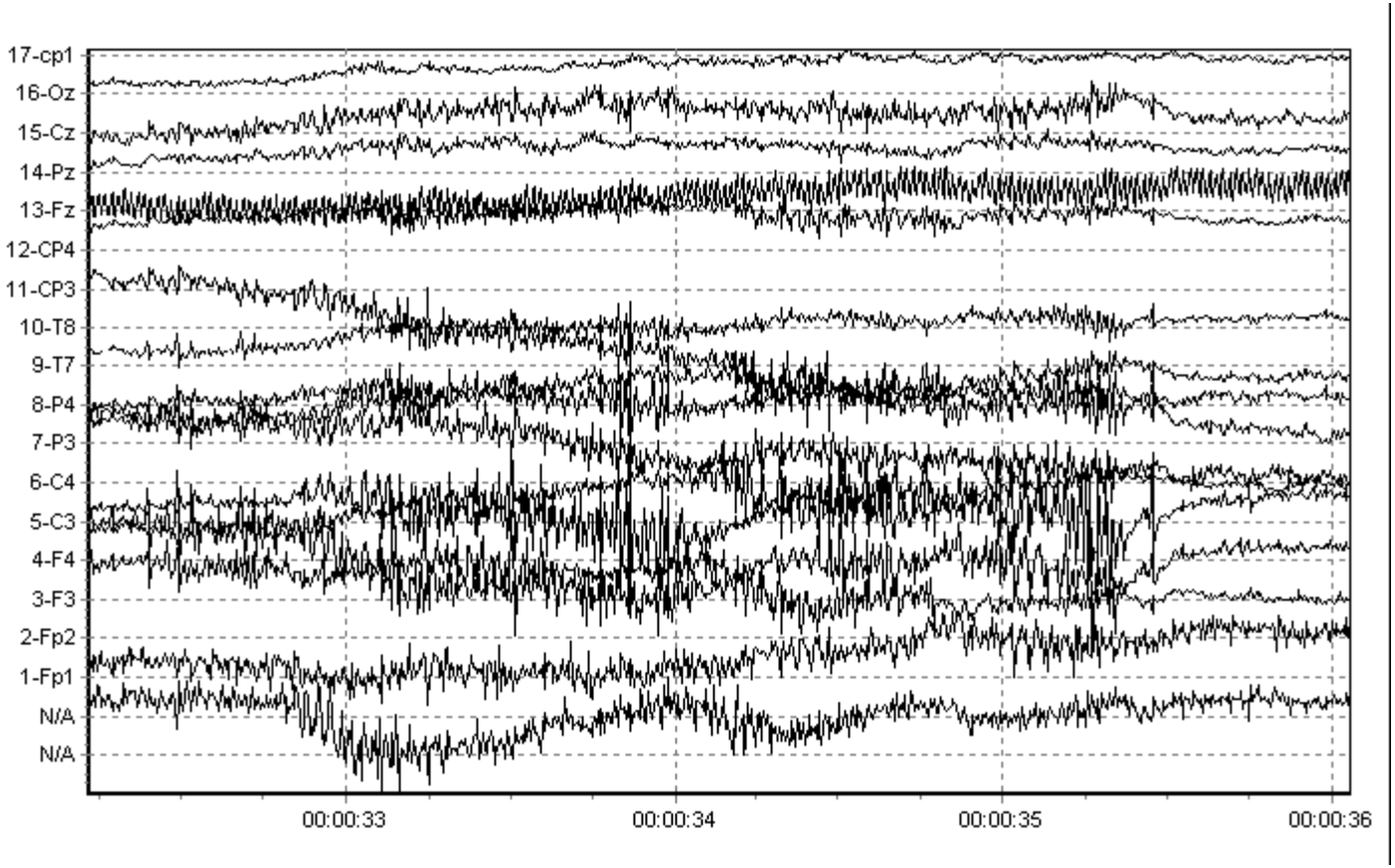


http://www.bci2000.org/

Eye Movements



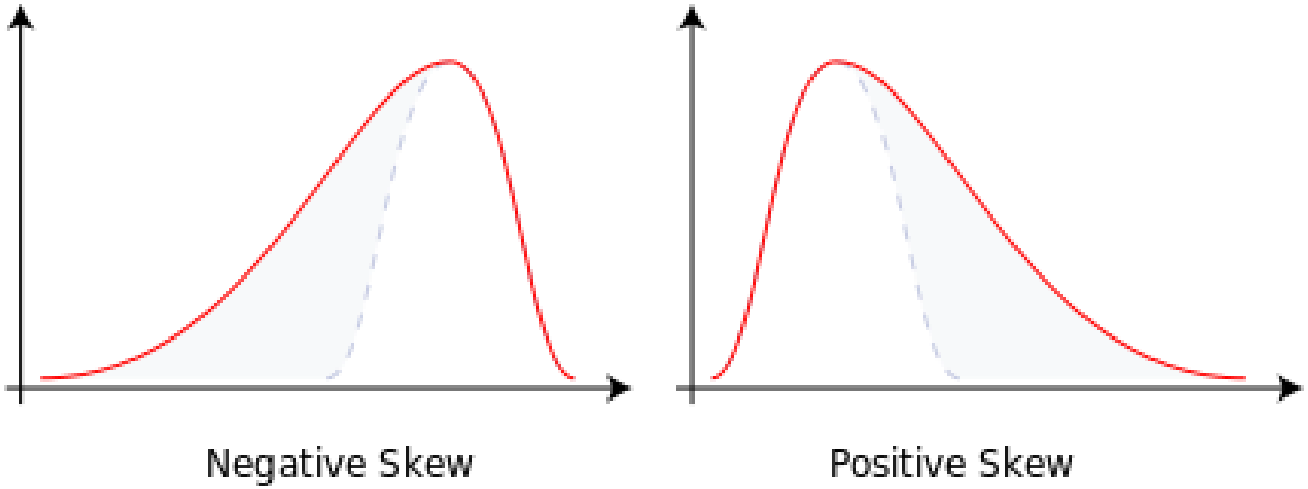
Muscle Activity / Biting



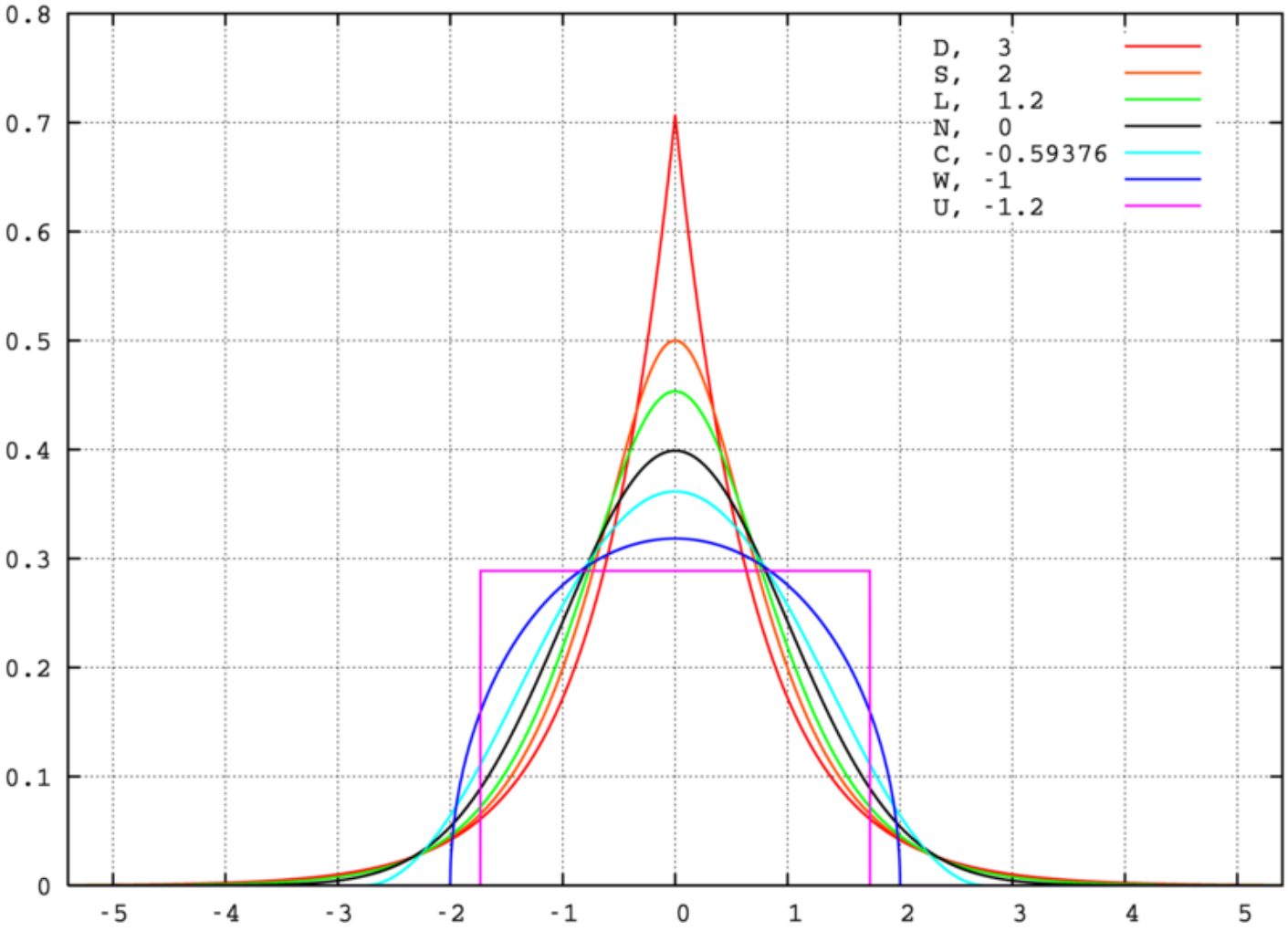
Semi-Automatic Artifact Rejection

- Uses Data Statistics to find artifact-containing time points.
  - The more well-behaved the artifacts, the more powerful this step will be.
- Some Statistics commonly in use:
  - Data Moments:
    - Mean, Variance, Skewness, Kurtosis.
  - Correlation with Artifact Channels
  - Entropy
- Be Conservative:
  - Three small rejections is better than one big one.

Skewness



Kurtosis

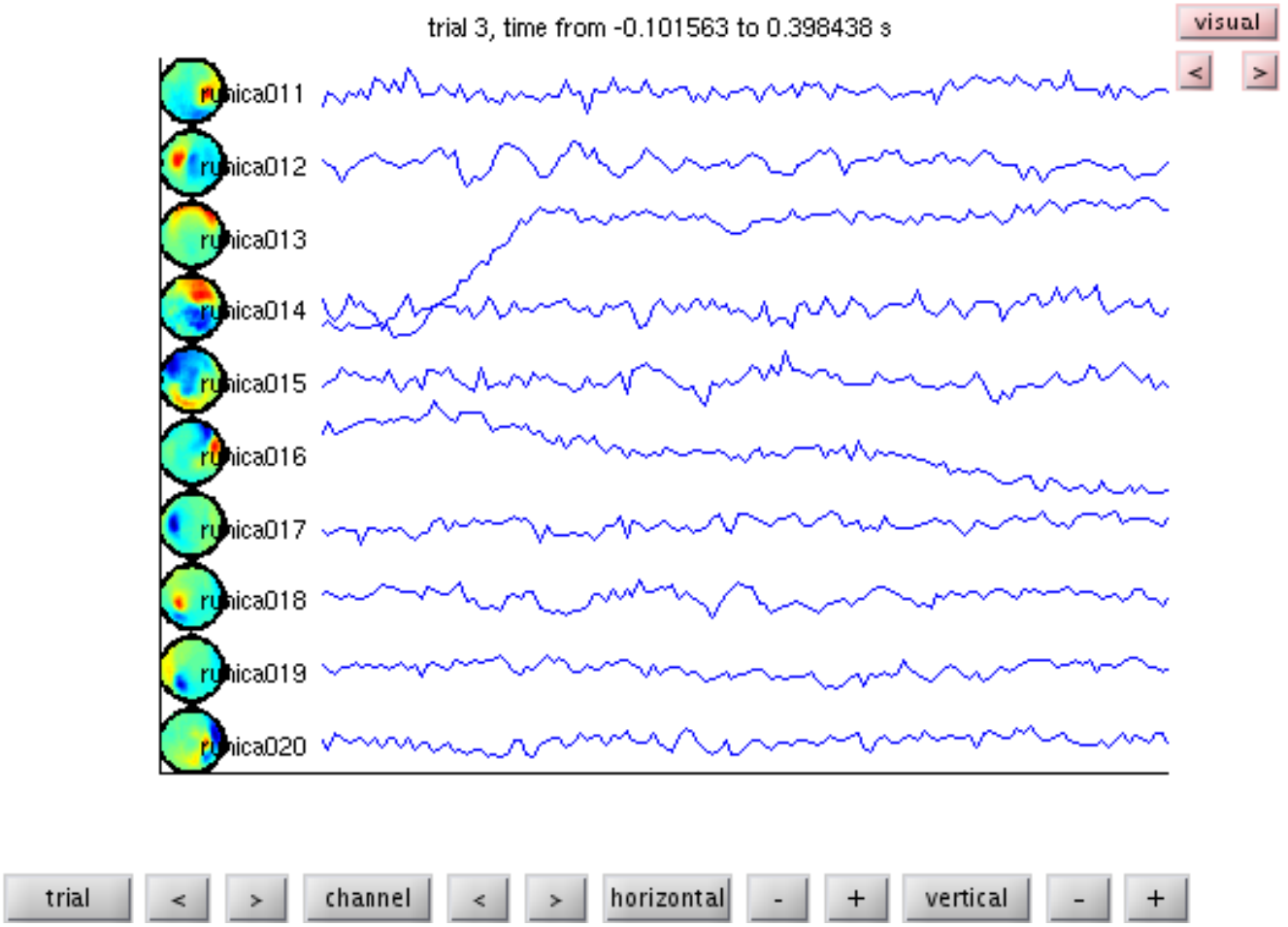


Artifact Correction

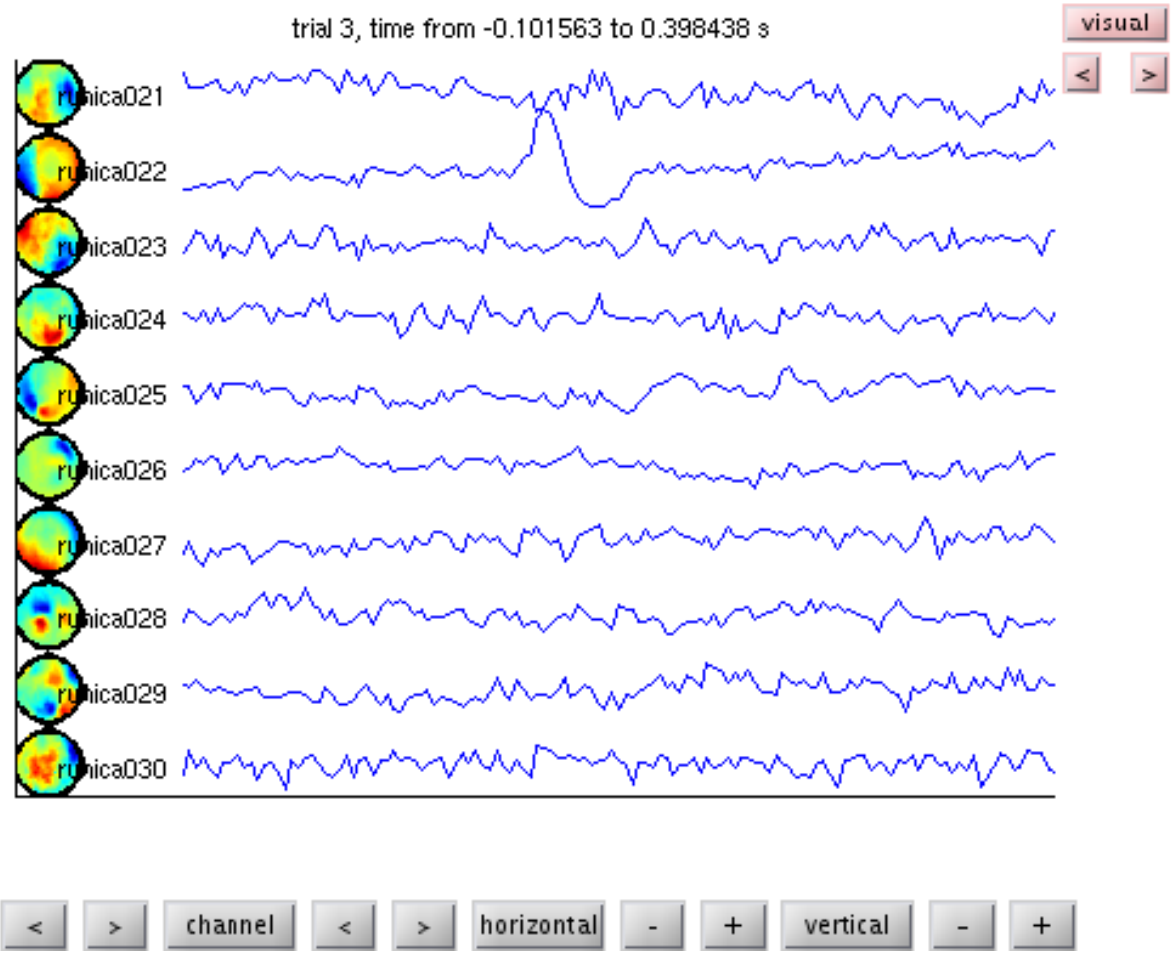
- Linear Regression with Artifact Channels
  - Requires full recording of artifacts
  - Assumes linear relationship between artifact (a.k.a "nuisance") channels and data channels.
  - Assumes a single relationship between artifact and data channels.
  - May helped by adaptive filtering of nuisance channel.

- Data Decomposition
  - Ex) Independent Component Analysis
    - Many ICA algorithms exist: INFOMAX, FASTICA, JADE, etc.
  - Source Separation Requires <u>many</u> data channels
  - Many algorithms are slow, can take a while.

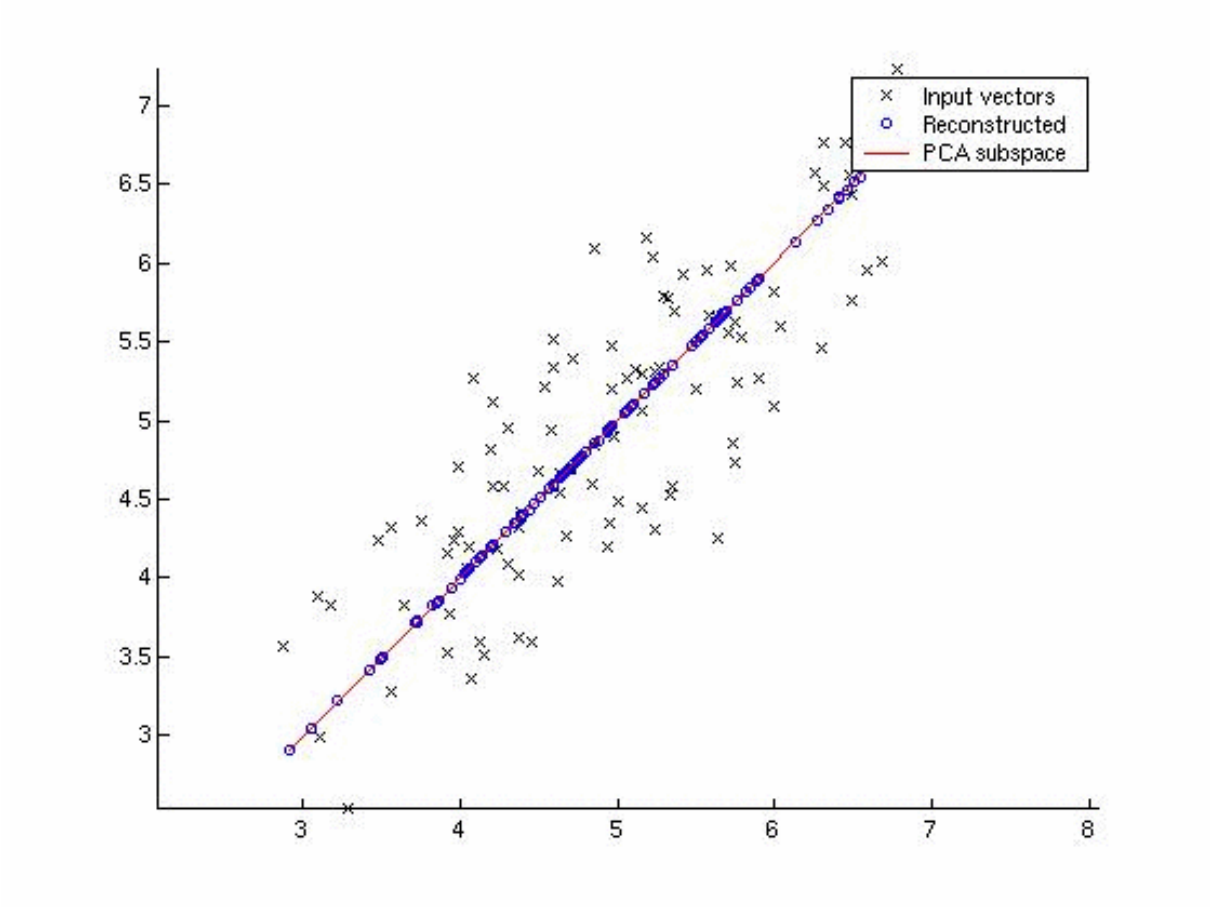
Artifact Components: Vertical Eye Movements



Artifact Components: EKG



Reducing Noise: Principal Component Analysis



Take Home Messages

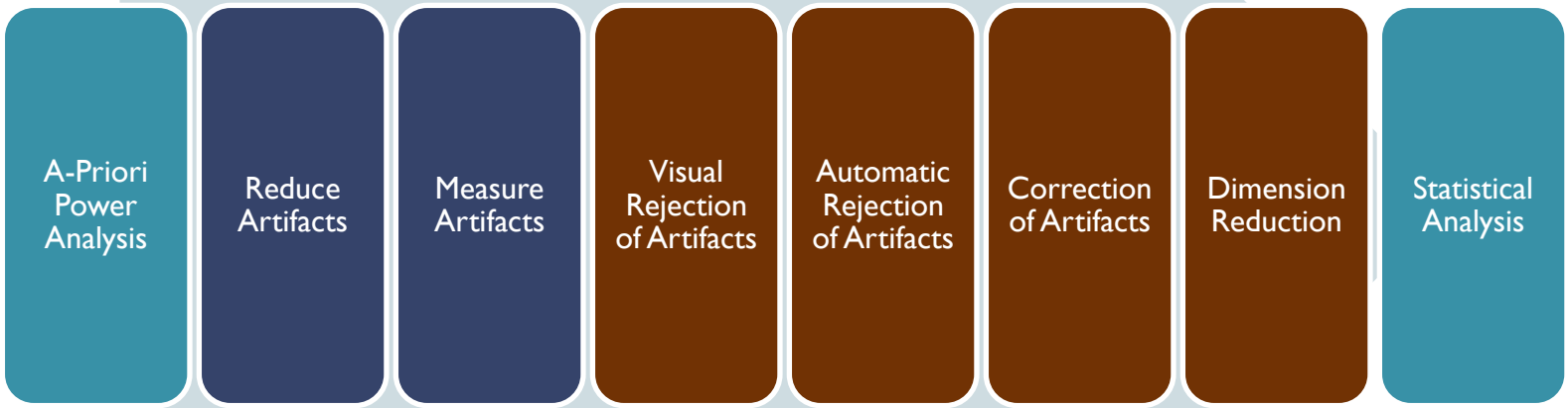
- "Garbage In, Garbage Out".
  - Collect enough high-quality data to detect the signal you are interested in!

- Be Conservative.
  - Be careful of selection bias!
  - Use multiple strategies to remove artifacts.
    - Don't rely too much on a single method.

- Artifact Channel $\neq$ Artifact Type

What is your Artifact Action Plan?

$$\mathrm{SNR} = \frac{P_{\mathrm{signal}}}{P_{\mathrm{noise}}} = \left(\frac{A_{\mathrm{signal}}}{A_{\mathrm{noise}}}\right)^2 = \frac{\mu}{\sigma}$$



Thank You!