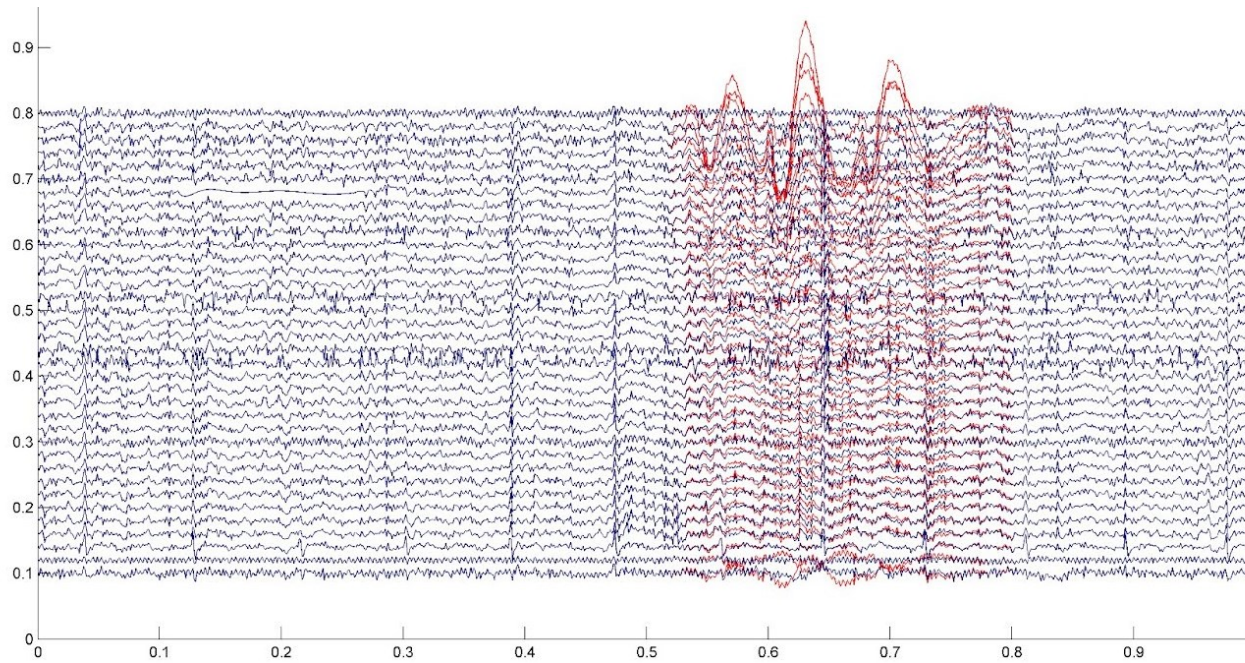


The Artifact Subspace Reconstruction method

Christian A Kothe
SCCN / INC / UCSD
January 2013

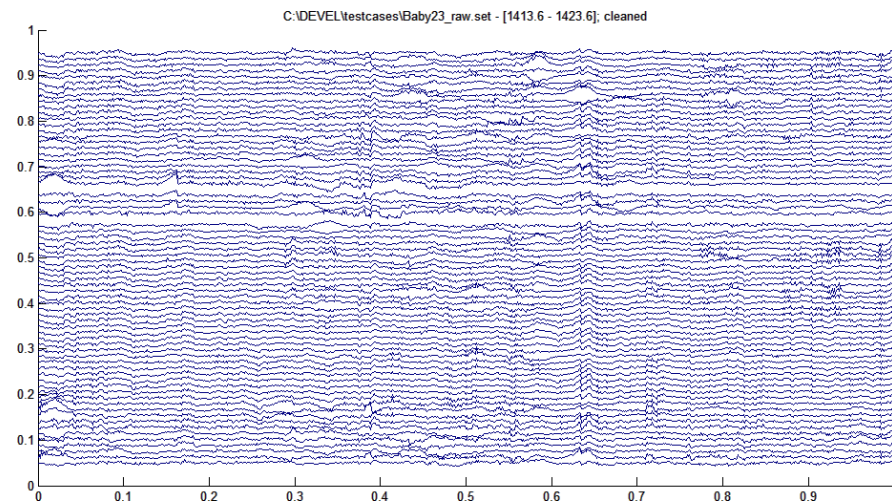
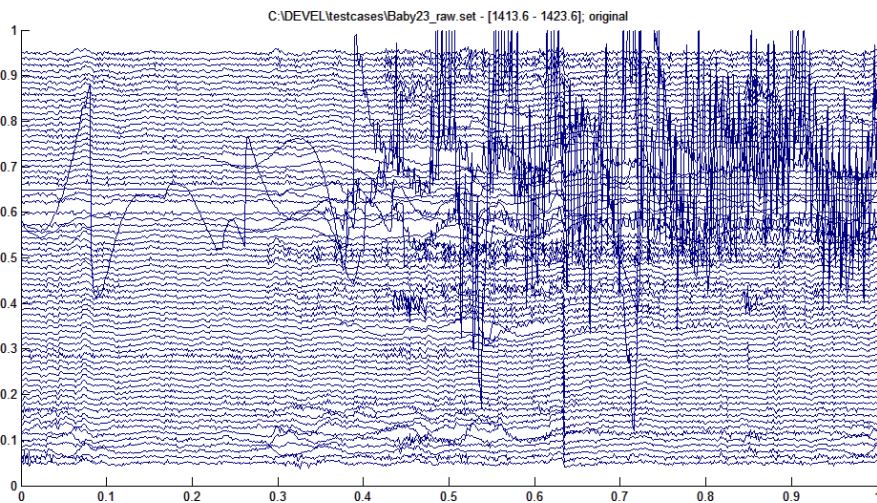
Artifact Subspace Reconstruction

- New algorithm to remove non-stationary high-variance signals from EEG
- Reconstructs the missing data using a spatial mixing matrix (assuming volume conduction)



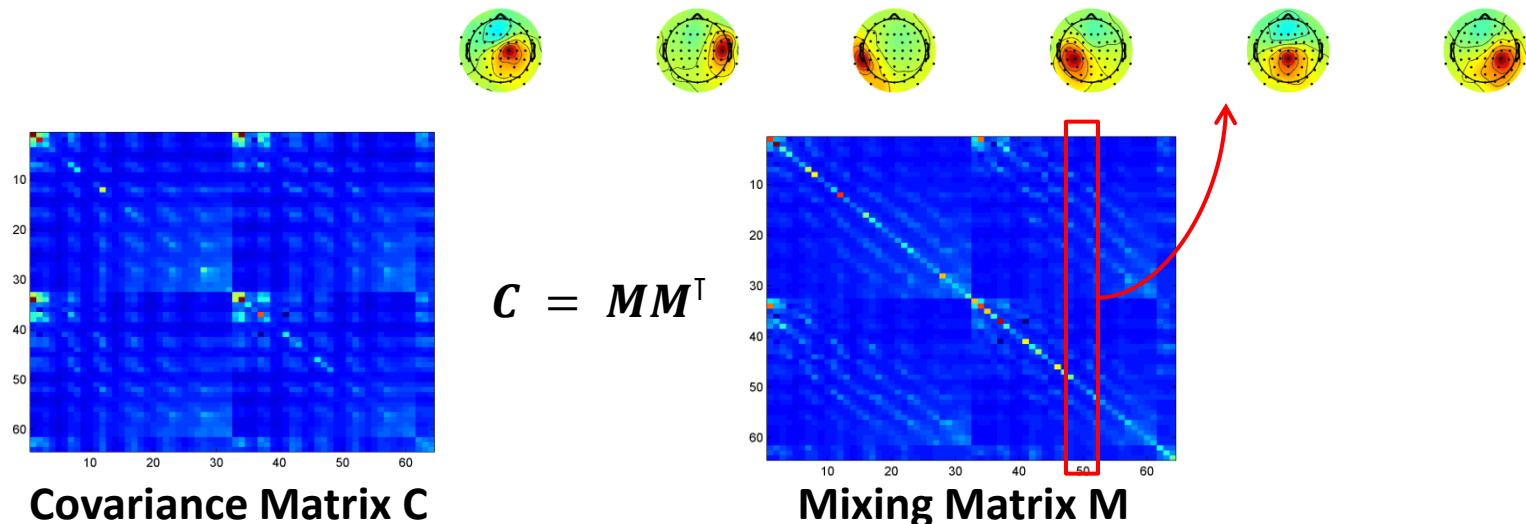
Examples

- **Online:** Recorded by NCTU BRC:
<http://mindoc.com.tw/tw/video.php?act=view&no=9>
- **Offline:** Post-hoc cleaning of various data sets



Basic Algorithm

- Relies on *calibration data* (1 minute clean resting EEG)
- Can alternatively extract clean sections from existing recordings (automatically)
- Derived calibration statistics:



Robust Statistics

- Calibration statistics are estimated in a *robust manner* (to minimize any effect of artifacts)
- Using the Geometric Median $\mathcal{G}$ over windowed (1-second) estimates:

$$\mathcal{G}(X) = \arg \min_y \sum_{i=1}^m \|x_i - y\|_2$$

Geometric Median

$$y_{i+1} = \left(\sum_{j=1}^m \frac{x_j}{\|x_j - y_i\|} \right) / \left(\sum_{j=1}^m \frac{1}{\|x_j - y_i\|} \right)$$

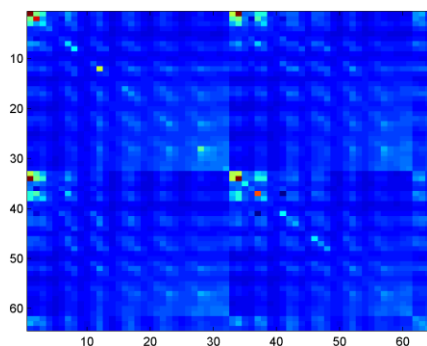
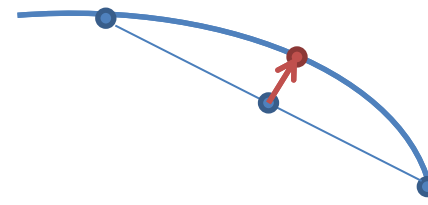
Iterative Formula

(iteratively reweighted least squares)

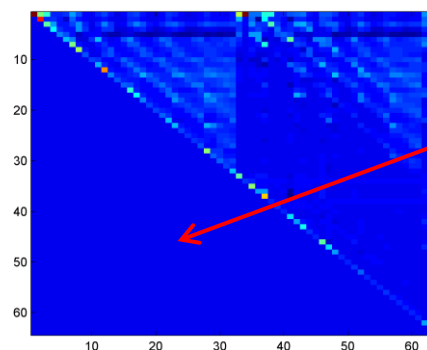
- **Note:** Also a very good drop-in replacement for outlier-sensitive averages in a wide range of statistical procedures (e.g., ICA updates)!

Robust Statistics

- Geometric Median over covariance matrices is not the ideal measure (since covariance matrices lie on a curved manifold)
- Can *re-parameterize* into Cholesky factorizations, take median there, then back-transform:



Covariance Matrix

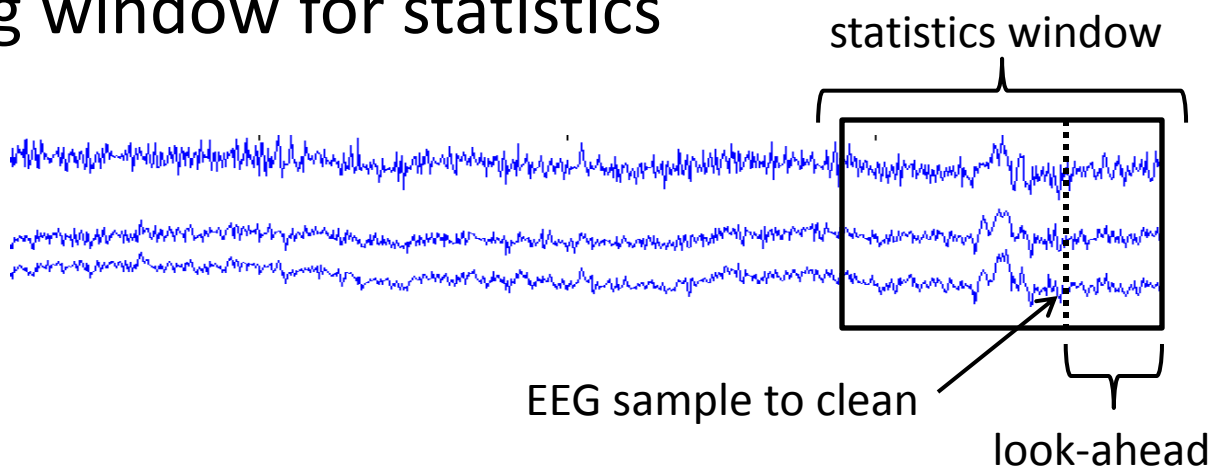


Cholesky Factorization

(median of upper-triangular matrices is still upper-triangular)

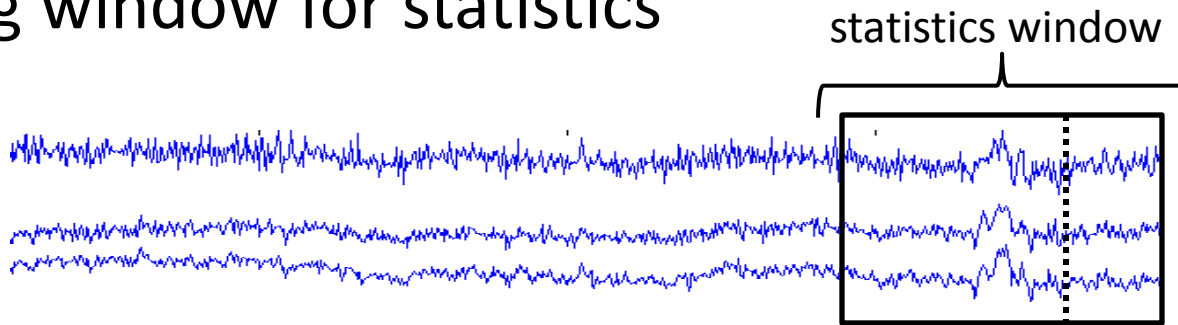
Online Processing

- Done independently sample-by-sample, using a sliding window for statistics

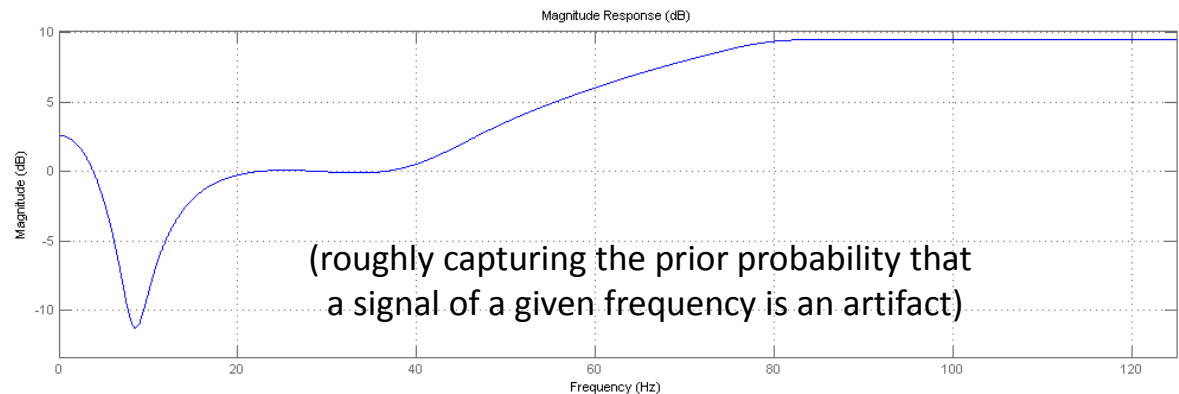


Online Processing

- Done independently sample-by-sample, using a sliding window for statistics

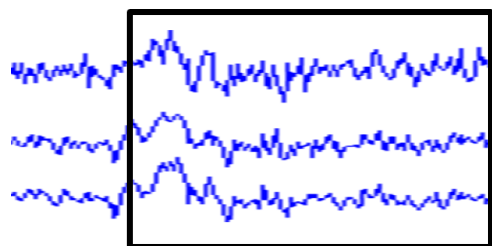


- Statistics are *spectrally weighted* by applying an IIR filter to the data in the statistics window (8'th order Yule-Walker)

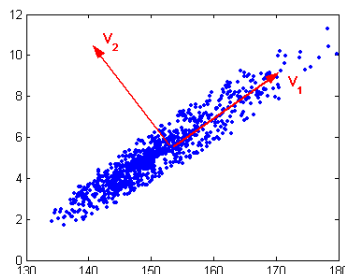


Online Processing

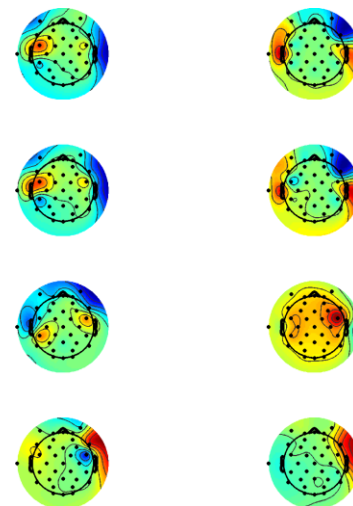
- **Step 2:** Separate *high-amplitude* signal components (potential artifact components) in the statistics window from other components
- Done using Principal Component Analysis (PCA) on a sliding window:



Raw Signal Window W
(spectrally filtered)



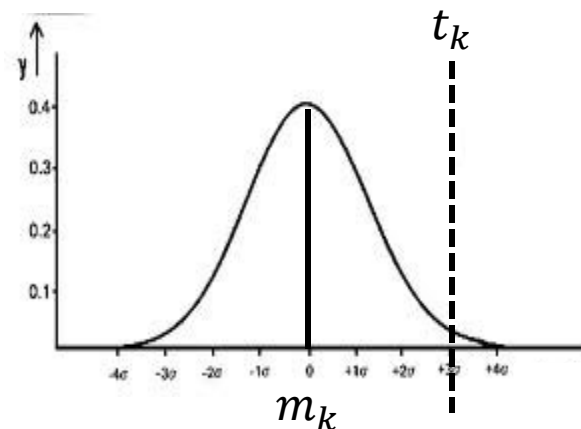
PCA Solution
 $V = \text{eig}(WW^T)$



Random sample of
high-variance components

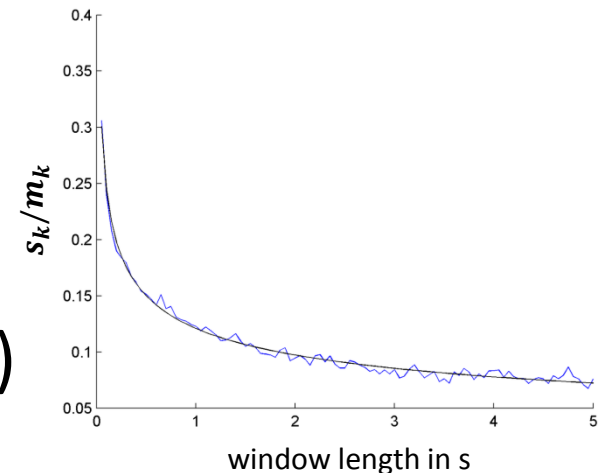
Online Processing

- **Step 3:** Classify each signal component as high variance or nominal variance (= EEG-like)
- Threshold for k 'th component is direction-dependent: depends on mean variance m_k in calibration EEG data along the principal component's direction $\mathbf{v}_k$ ($m_k = \mathbf{v}_k^\top \mathbf{C} \mathbf{v}_k$)
- Actual threshold t_k is at 3 standard deviations s_k above mean m_k



Online Processing

- **Step 3:** Classify each signal component as high variance or nominal variance (= EEG-like)
- Threshold for k 'th component is direction-dependent: depends on mean variance m_k in calibration EEG data along the principal component's direction $\mathbf{v}_k$ ($m_k = \mathbf{v}_k^\top \mathbf{C} \mathbf{v}_k$)
- Actual threshold t_k is at 3 standard deviations s_k above mean m_k (s_k is deduced from m_k using a χ^2 assumption under which these parameters are functionally related)



Online Processing

- **Step 4:** Reconstruct content of high-variance subspace from content of nominal-variance subspace (i.e., estimate missing data)
- **Basic idea:** EEG is highly correlated; can estimate content of one channel based on its neighbors
- The same works not just for a channel but also for a linear combination of channels (e.g., sum or difference of 2), i.e., can estimate content of artifact subspace from non-artifact subspace

Geometric Approach

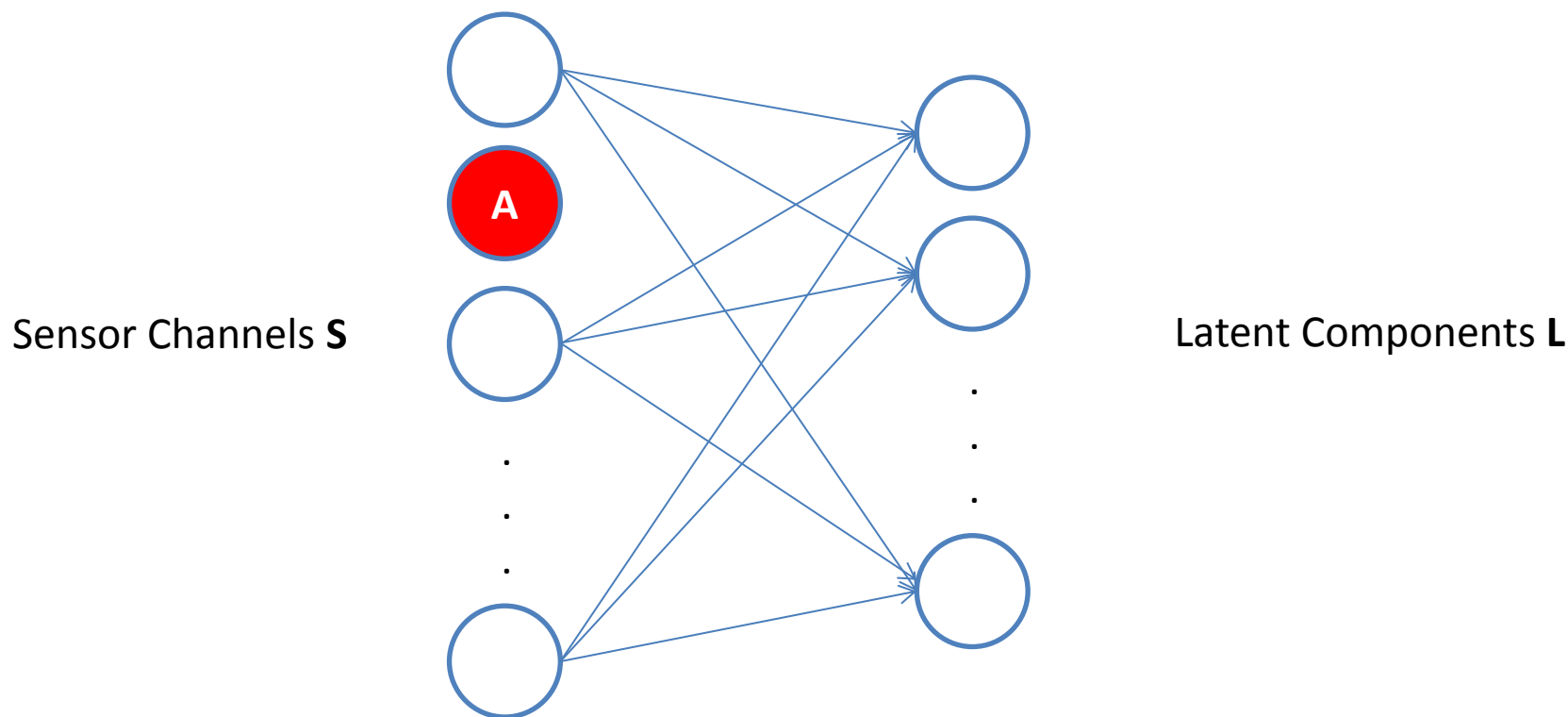
- Mixing matrix $\mathbf{M}$ represents linear mapping from orthog. latent components $\mathbf{L}$ onto sensors $\mathbf{S}$: $\mathbf{S} = \mathbf{M}\mathbf{L}$
- Component activation can be estimated as $\mathbf{L} = \mathbf{M}^{-1}\mathbf{S}$ (but fails to keep artifacts out of $\mathbf{L}$)
- To estimate a clean $\mathbf{L}$, the inverse of a truncated mixing matrix (artifact *channels* zeroed out) can be used: $\mathbf{M}_{trunc} = \mathbf{M} \circ \mathbf{T}$,
$$\mathbf{L} = \mathbf{M}_{trunc}^+ \mathbf{S}$$

(Note: here we frame it in terms of channels, moving to components later)

Geometric Approach

- Given a clean estimate of $\mathbf{L}$, can back-project onto channels again using the full $\mathbf{M}$:

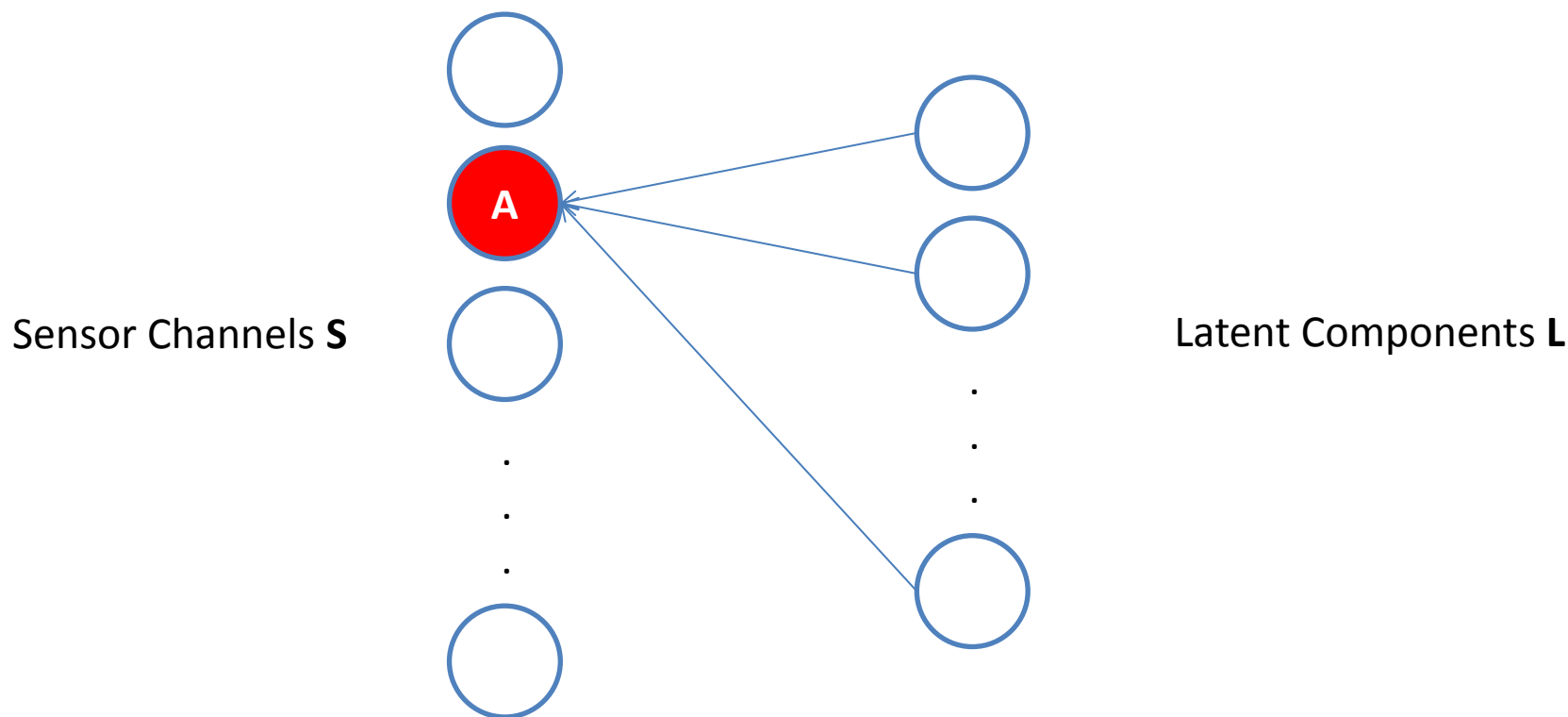
$$\mathbf{S}_{clean} = \mathbf{M}(\mathbf{M} \circ \mathbf{T})^+ \mathbf{S}$$



Geometric Approach

- Given a clean estimate of $\mathbf{L}$, can back-project onto channels again using the full $\mathbf{M}$:

$$\mathbf{S}_{clean} = \mathbf{M}(\mathbf{M} \circ \mathbf{T})^+ \mathbf{S}$$



Geometric Approach

- Given a clean estimate of $\mathbf{L}$, can back-project onto channels again using the full $\mathbf{M}$:

$$\mathbf{S}_{clean} = \mathbf{M}(\mathbf{M} \circ \mathbf{T})^+ \mathbf{S}$$

- Doing the same in artifact/non-artifact principal component space $\mathbf{V}$ requires a rotation into $\mathbf{V}$, followed by back-rotation:

$$\mathbf{S}_{clean} = \mathbf{V}\mathbf{M}_V(\mathbf{M}_V \circ \mathbf{T})^+ \mathbf{V}^\top \mathbf{S}$$

- ... using a rotated mixing matrix $\mathbf{M}_V = \mathbf{V}^\top \mathbf{M}$
- All steps can be baked into a re-projection matrix $\mathbf{R}$ such that $\mathbf{S}_{clean} = \mathbf{R}\mathbf{S}$:

$$\mathbf{R} = \mathbf{M}(\mathbf{V}^\top \mathbf{M} \circ \mathbf{T})^+ \mathbf{V}^\top$$

Speed-up

- Calculate reprojection matrix **R** for every n 'th sample ($n=32$)
- Then interpolate **R** for intermediate samples (using a raised-cosine window)
- Runs in real time on up to 256 channels

Clusters of Rejected Components

- Most rejections resembles the projection of a small number of equivalent dipoles:

