

THE FLEXIBLE AUDIO SOURCE SEPARATION TOOLBOX VERSION 2.0

Yann Salaün, Emmanuel Vincent, Nancy Bertin, Nathan Souviraà-Labastie,
Xabier Jaureguiberry, Dung T. Tran, and Frédéric Bimbot



yann.salaun@inria.fr

SCIENTIFIC DESCRIPTION

- Mixture model

$$\mathbf{x}_{fn} = \sum_{j=1}^J \mathbf{y}_{jfn}$$

- Source model

$$\mathbf{y}_{jfn} \sim \mathcal{N}(\mathbf{0}, v_{jfn} \mathbf{R}_{jf})$$

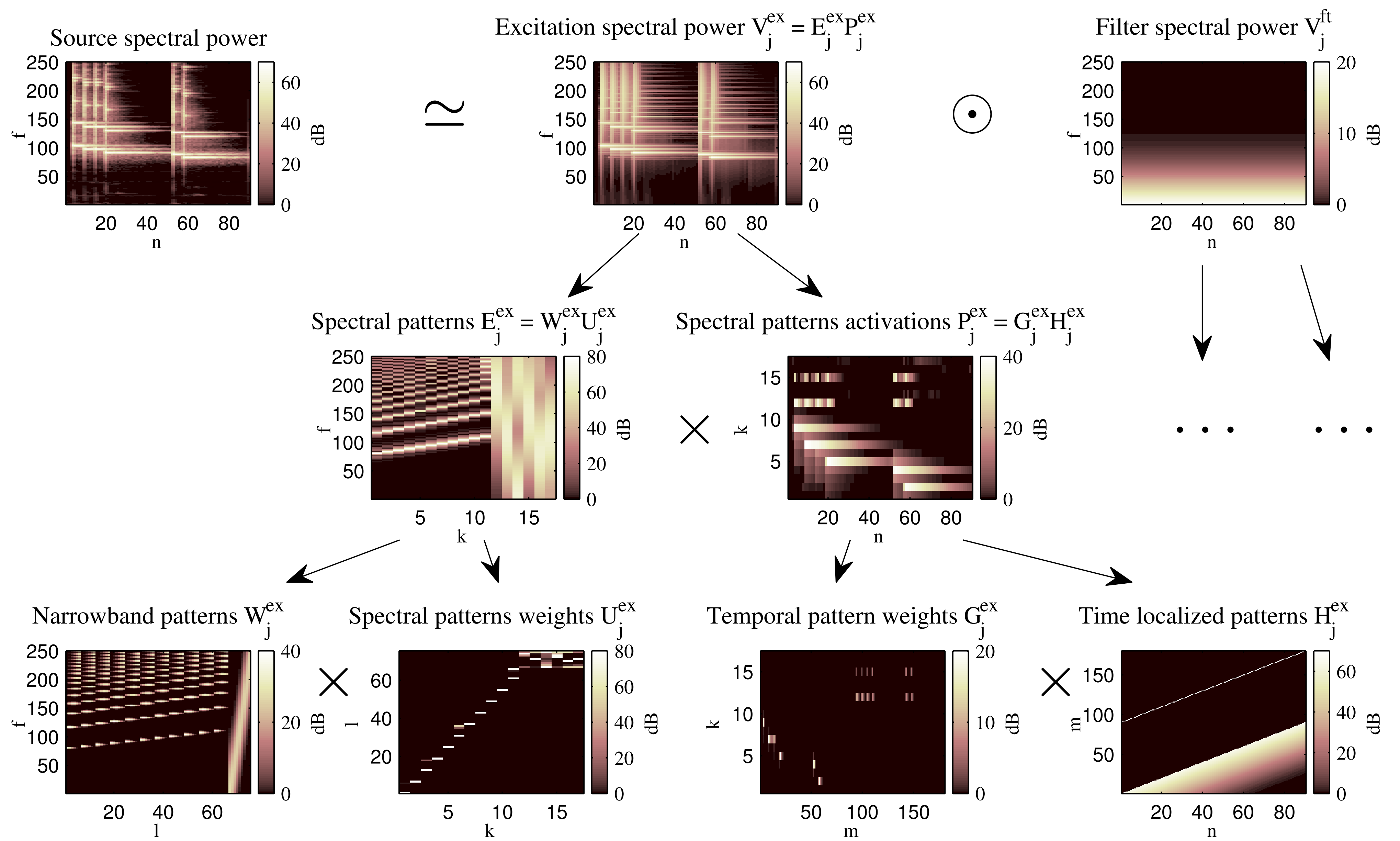
- Spectral parameters

$$\mathbf{V}_j = (\mathbf{W}_j^{\text{ex}} \mathbf{U}_j^{\text{ex}} \mathbf{G}_j^{\text{ex}} \mathbf{H}_j^{\text{ex}}) \odot (\mathbf{W}_j^{\text{ft}} \mathbf{U}_j^{\text{ft}} \mathbf{G}_j^{\text{ft}} \mathbf{H}_j^{\text{ft}})$$

- Spatial parameters

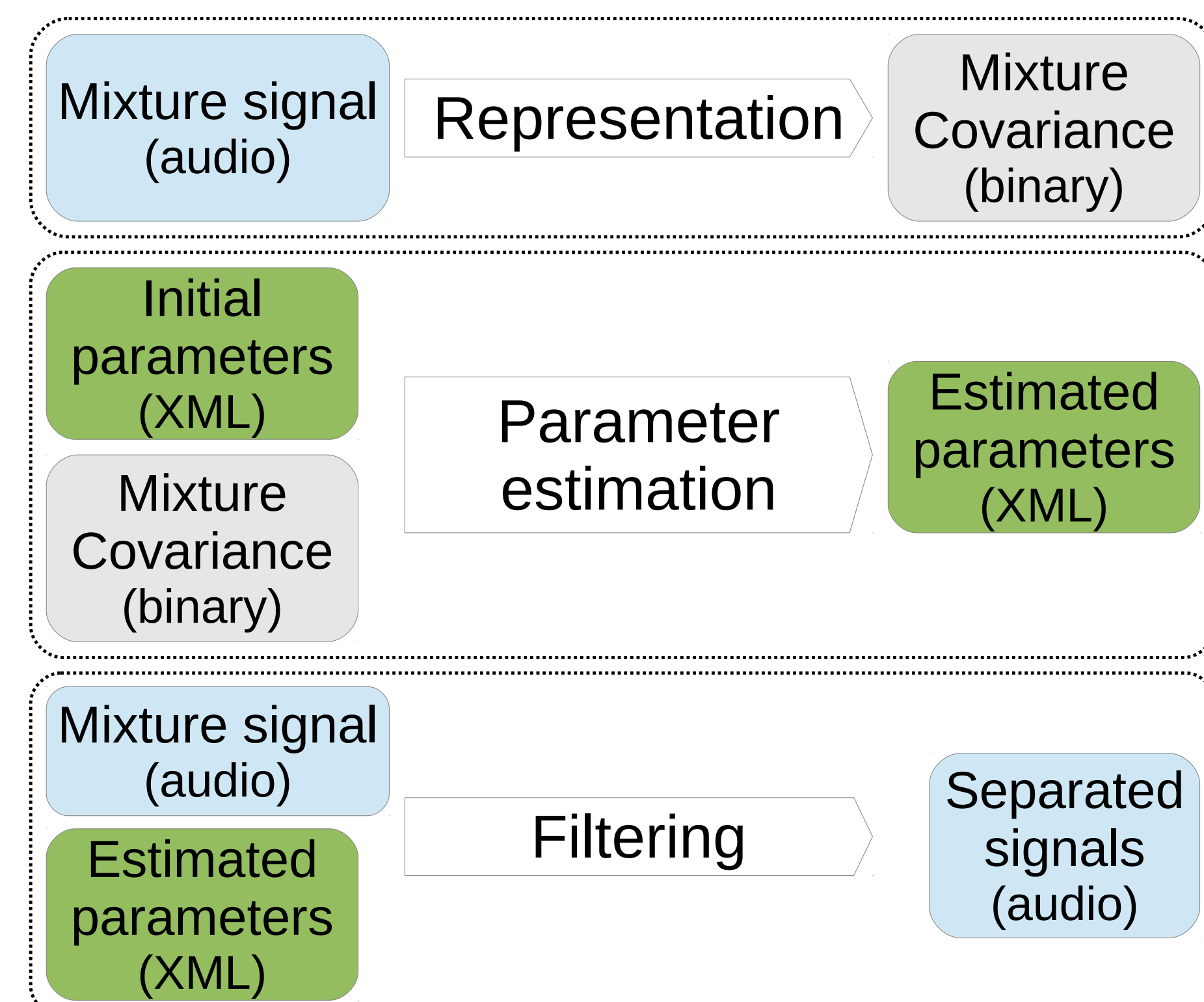
$$\mathbf{R}_{jf} = \mathbf{A}_{jf} \mathbf{A}_{jf}^H$$

- Parameters estimated from the empirical mixture covariance matrix $\hat{\mathbf{R}}_{\mathbf{x},fn}$ in the maximum likelihood (ML) sense using an expectation-maximization (EM) algorithm
- \mathbf{y}_{jfn} subsequently obtained by multichannel Wiener filtering



IMPLEMENTATION

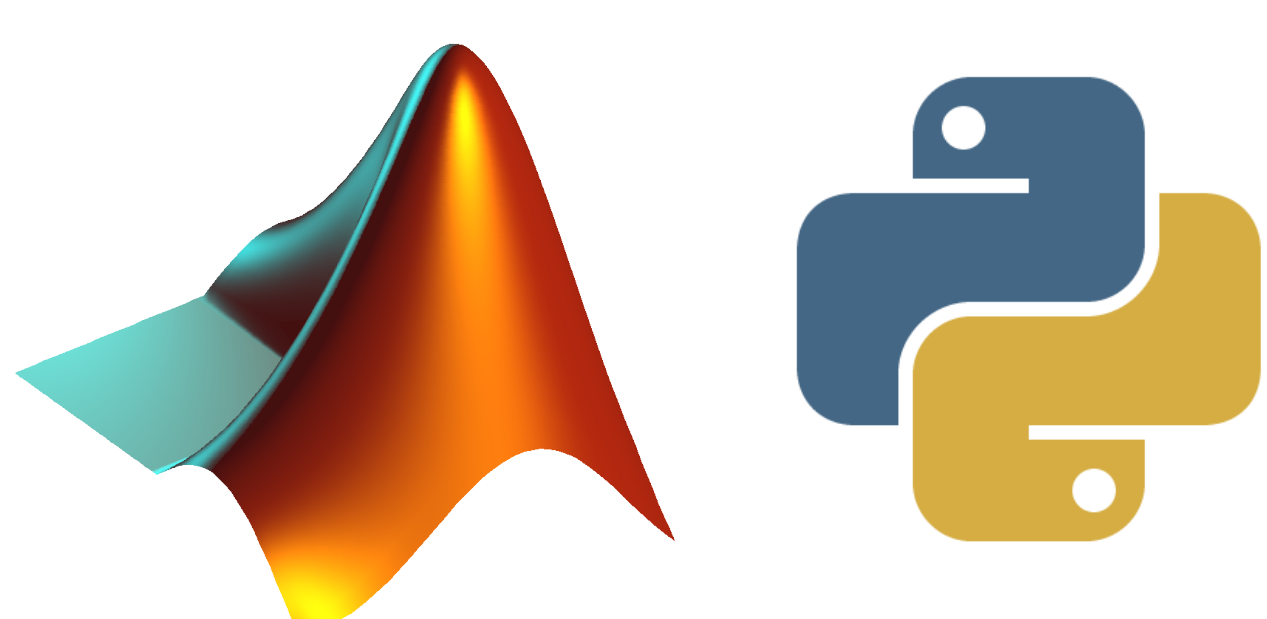
Three C++ core executables



- Representation**
mixture covariance matrix $\hat{\mathbf{R}}_{\mathbf{x},fn}$ is Hermitian → stored in binary format
- Parameter estimation**
model structure is hierarchical → stored in XML format
- Wiener filtering**

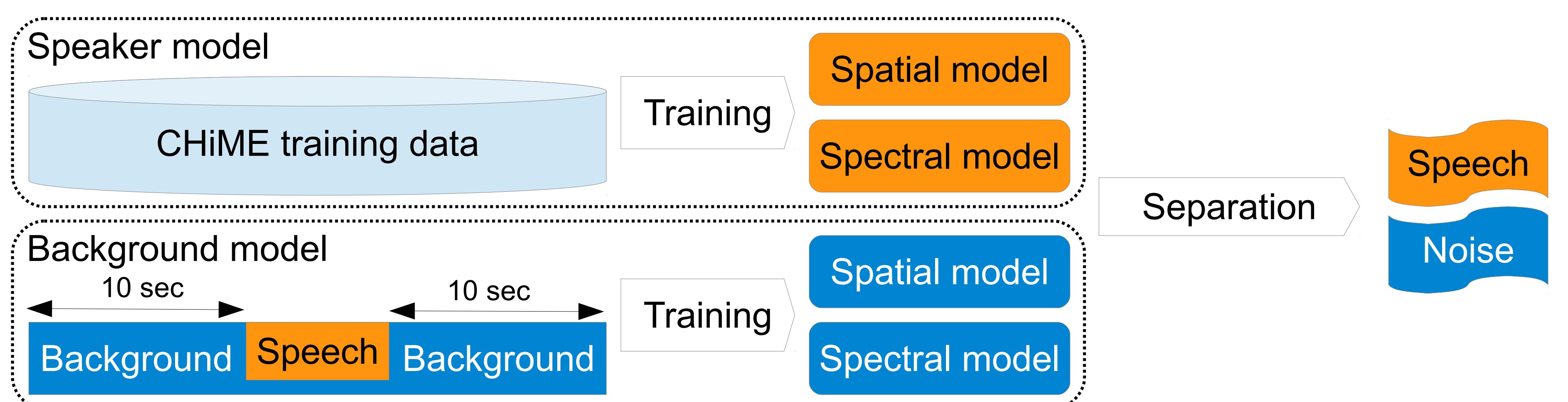
User scripts

- Used to initialize parameters and to glue executables together
- Available in MATLAB and Python



EXAMPLE USES

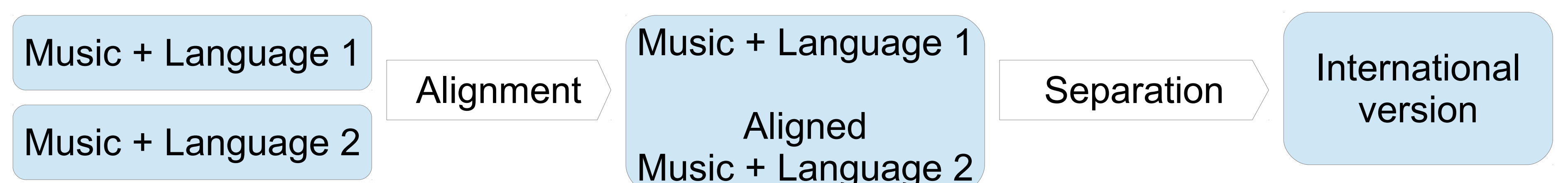
Track 1 of the 2nd CHiME Challenge



Play with:

- Speaker spectral model (NMF, none, excitation-filter)
- Separation settings (iterations, NMF states)

Movie soundtrack



Play with spatial model (rank-1 or full rank)

CODE

<http://bass-db.gforge.inria.fr/fasst/>

- GNU/Linux, OSX, Windows
- License → Q Public License Version 1.0
- We ♥ feedback!

REFERENCES

- A. Ozerov, E. Vincent, and F. Bimbot, "A General Flexible Framework for the Handling of Prior Information in Audio Source Separation" in *IEEE Transactions on Audio, Speech, and Language Processing*, 2012.
- E. Vincent, et al., "The Second 'CHiME' Speech Separation and Recognition Challenge: An overview of challenge systems and outcomes" in *Proc. ASRU*, 2013.