



Exploiting non-negative matrix factorization for binaural sound source localization in the presence of directional interference

Ingvi Örnolfsson*, Torsten Dau*, Ning Ma[†] and Tobias May*

**Department of Health Technology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*

[†]Department of Computer Science, University of Sheffield, Sheffield S1 4DP, United Kingdom

Why study sound localization?



Beamforming for hearing aids



Social robots



Auditory scene analysis

Machine listening system

Task to solve:

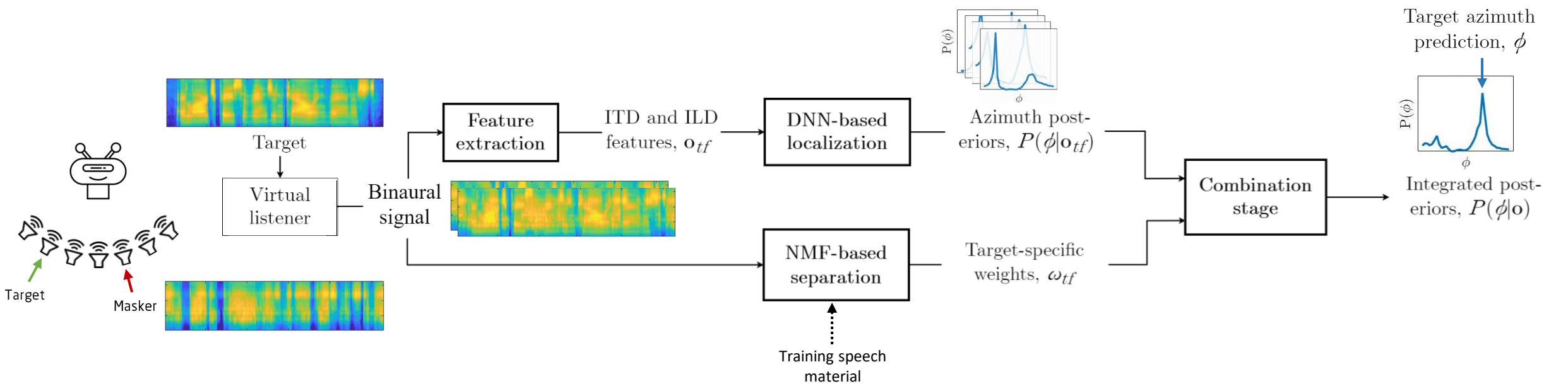
- ◆ Predict azimuth angle ϕ of a target speech source
- ◆ An additional masking source will be present at a different angle

Problem:

- ◆ The directional masking source provides competing ITD and ILD cues

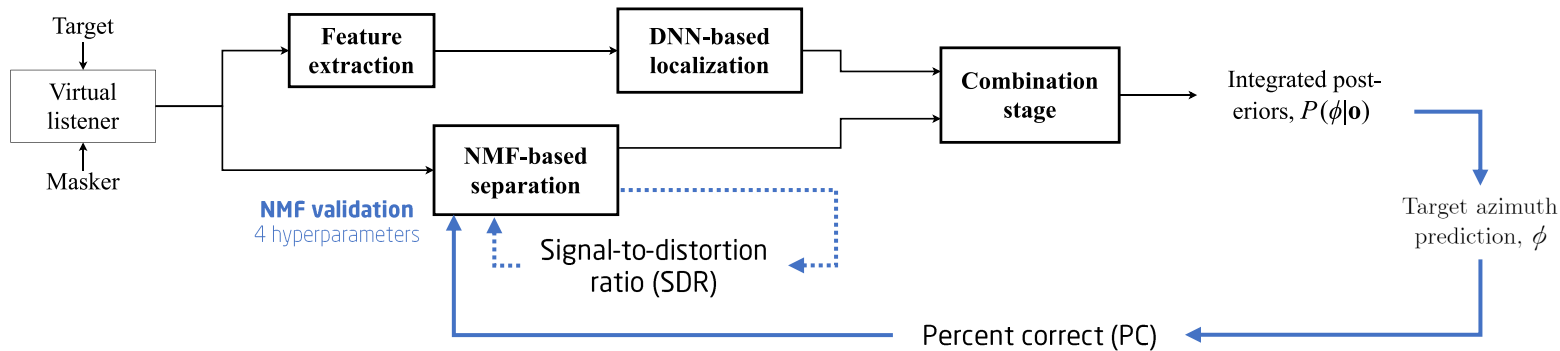
Solution:

- ◆ Introduce a separation stage that indicates where speech is dominating the mixture

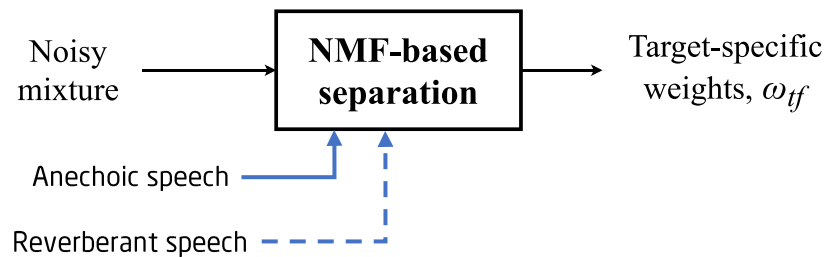


Research questions

1. Which performance metric should be used for NMF validation?



2. What material should be used for NMF training?



3. Can NMF beat a learning-free approach?

- ◆ NMF-based separation will be compared the APAB algorithm

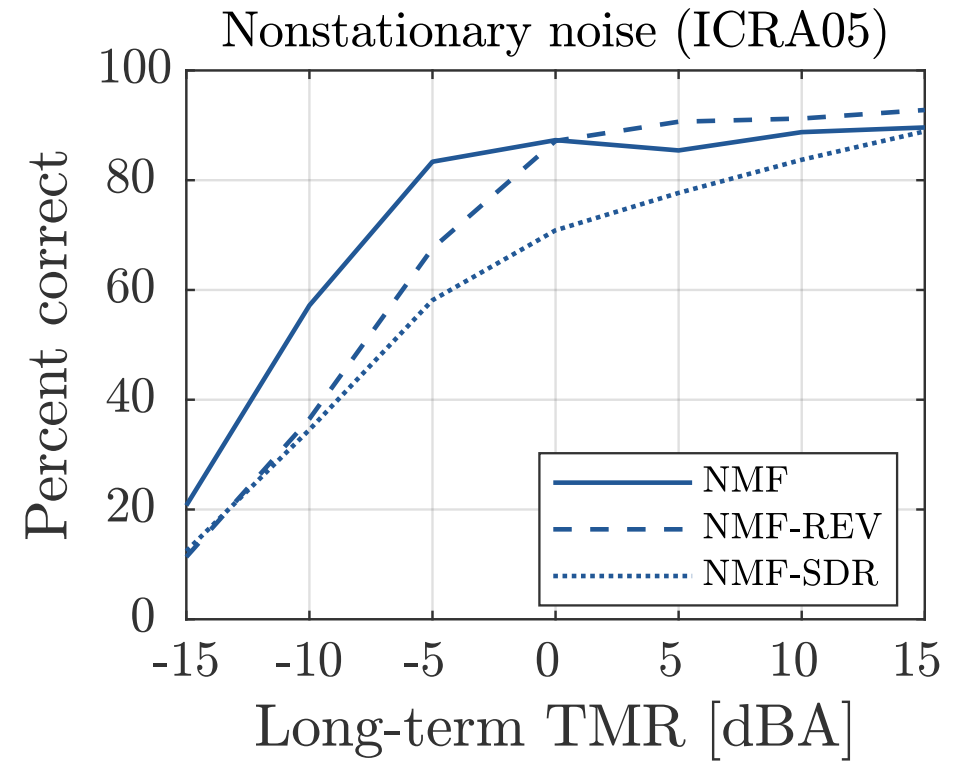
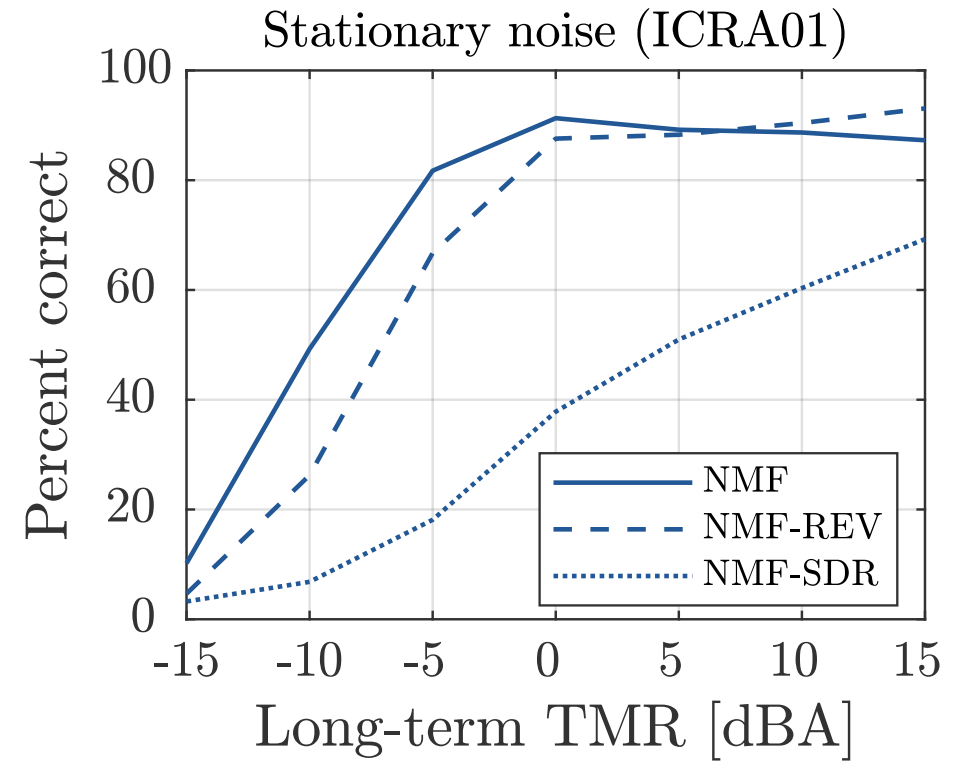
Summary of algorithms

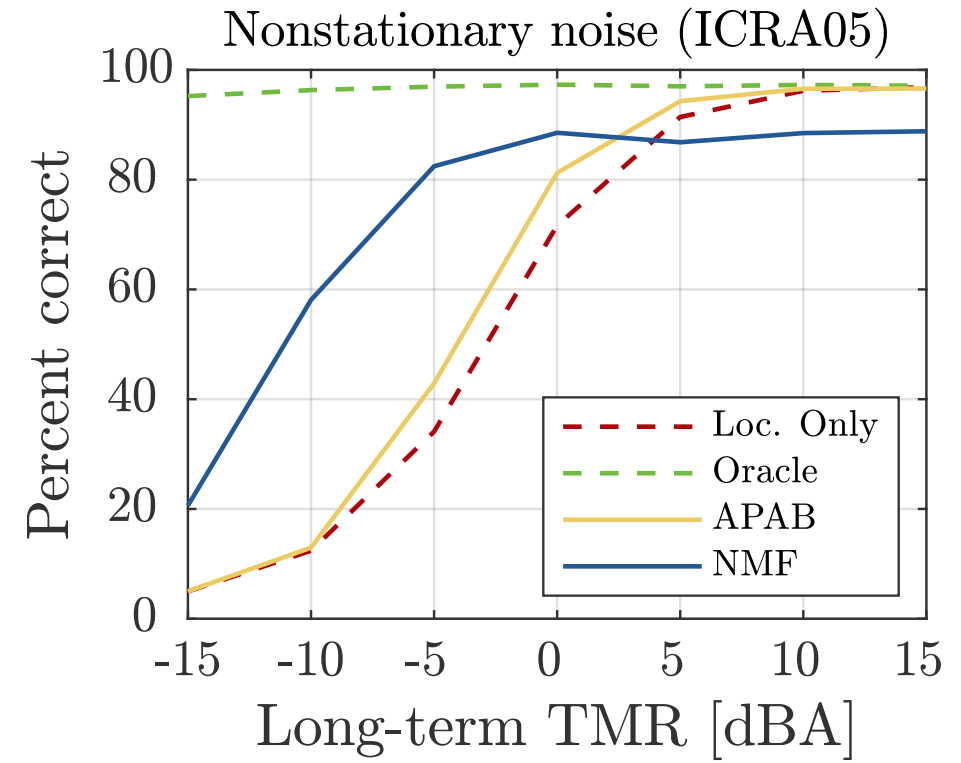
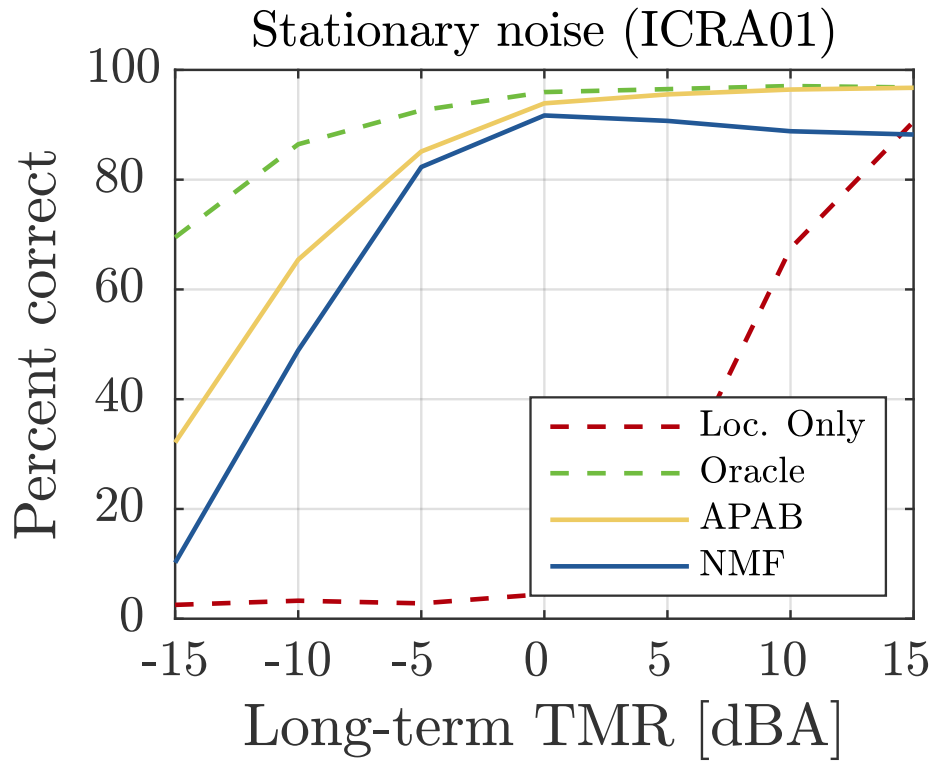
<i>NMF variants</i>	<i>Validation metric</i>	<i>Training material</i>	<i>Legend</i>
NMF	PC	Anechoic	
NMF-SDR	SDR	Anechoic	
NMF-REV	PC	Reverberant	
<i>Baselines</i>		<i>Description</i>	
Loc. Only	Separation stage is excluded.		
Oracle	Separation is based on oracle information about source activity.		
APAB	Separation is obtained via a variance-based noise reduction technique.		

	<i>Validation</i>	<i>Test</i>
Targets	TIMIT-TRAIN	TIMIT-TEST
Maskers	NatNoises-VAL	NatNoises-TEST & ICRA
Rooms	Surrey anechoic	All five Surrey rooms
Azimuths	-90° to 90°, 5° steps	
Long-term TMRs	-15dBA to 15dBA, 5dBA steps	
Evaluation metric	PC or SDR, depending on algorithm	PC

SDR: See reference [11] in the paper*

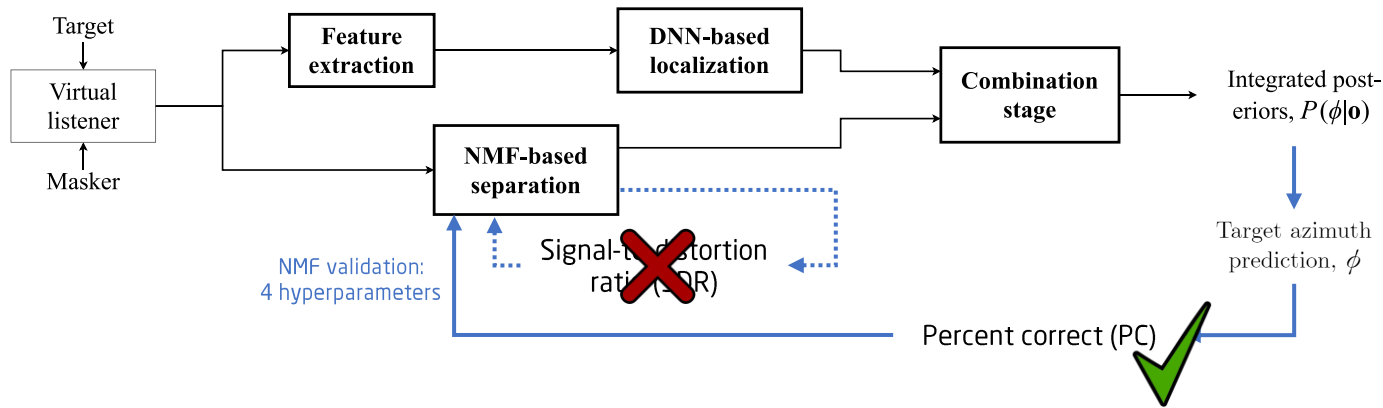
PC: 1 if correct azimuth, 0 otherwise



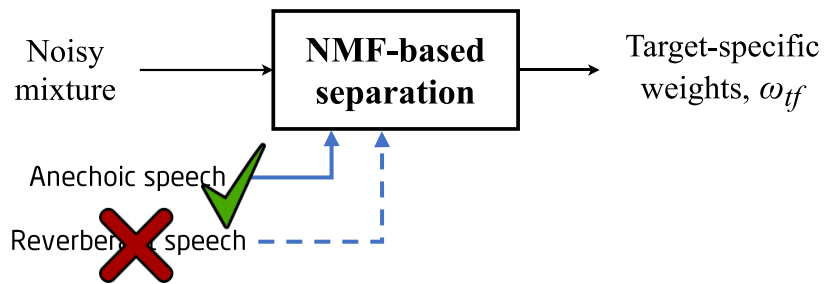


Conclusion - research questions revisited

1. Which performance metric should be used for NMF validation?

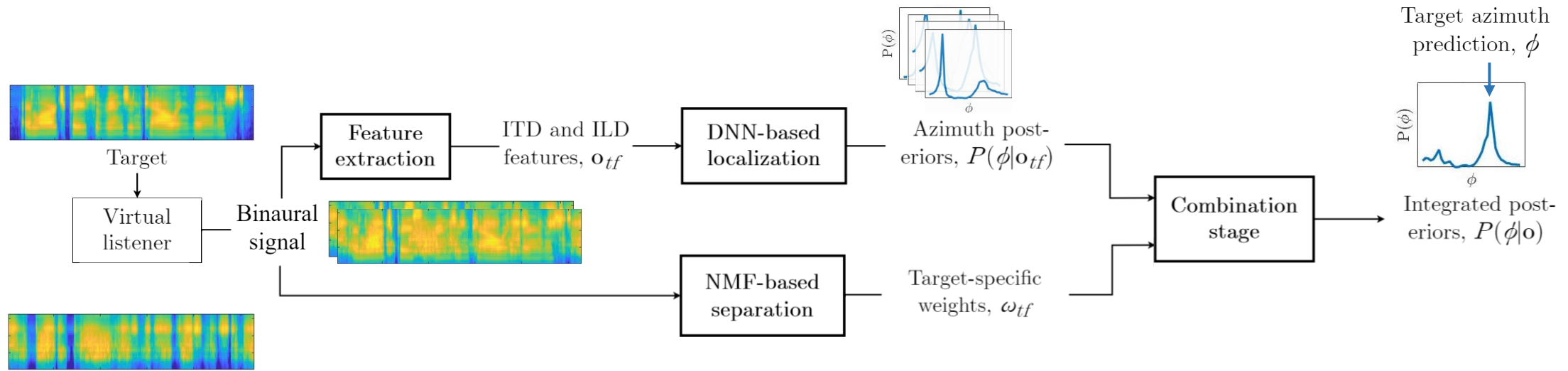


2. What material should be used for NMF training?



3. Can NMF beat a learning-free approach?

- ◆ In nonstationary noise, yes!
- ◆ In stationary noise, NMF is slightly worse than the APAB algorithm.



- ◆ Study further variations on the NMF
- ◆ Different source separation strategies: DNN, ICA, etc.
- ◆ What about a different combination stage?
- ◆ Use human performance as a baseline

Thank you for your time!

Please join us for our poster session:

AUD-8: Audio and Speech Source Separation 4: Multi-Channel Source Separation

Wednesday, 9 June from 13:00 to 13:45 in Eastern Daylight Time