



HAL
open science

Probing temporal modulation detection in white noise using intrinsic envelope fluctuations: A reverse-correlation study

Léo Varnet, Christian Lorenzi

► **To cite this version:**

Léo Varnet, Christian Lorenzi. Probing temporal modulation detection in white noise using intrinsic envelope fluctuations: A reverse-correlation study. Doctoral. ASA P&P Technical Committee, Online, France. 2022. hal-04526748

HAL Id: hal-04526748

<https://hal.science/hal-04526748>

Submitted on 2 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Probing temporal modulation detection in white noise using intrinsic envelope fluctuations: A reverse-correlation study

Léo Varnet¹, Christian Lorenzi¹

¹LSP, École Normale Supérieure de Paris

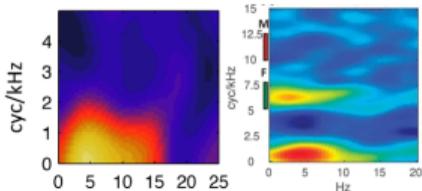
ASA P&P Technical Committee, 05/04/2022

Auditory revcorr studies (full diagram on <https://dbao.leo-varnet.fr/>)

high level

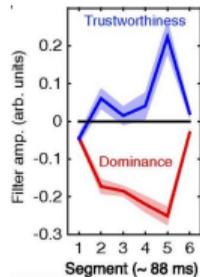
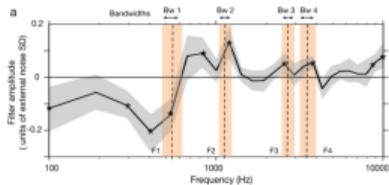
sentence recognition

[Venezia et al., 2016, 2019]



paralinguistics

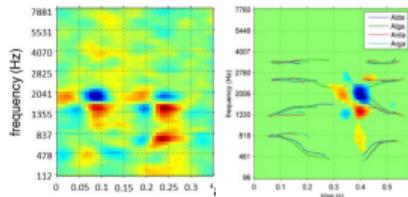
[Ponsot et al., 2018a, 2018b]



low level

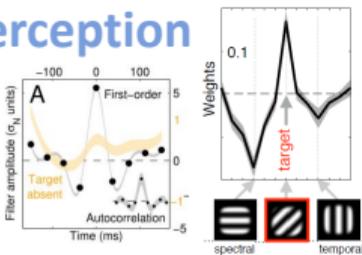
phoneme categorization

[Varnet et al., 2013, 2015]



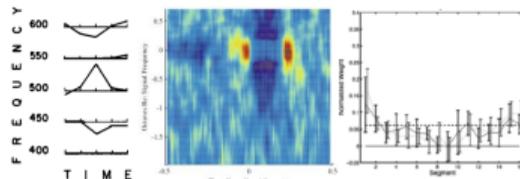
modulation perception

[Ponsot et al., 2020; Joosten & Neri, 2012]



pure-tone detection & loudness perception

[Ahumada & Lovell, 1971; Shub & Richards, 2009; Ponsot et al. 2013]

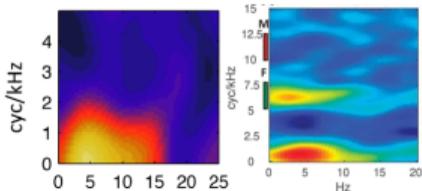


Auditory revcorr studies (full diagram on <https://dbao.leo-varnet.fr/>)

high level

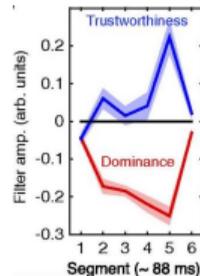
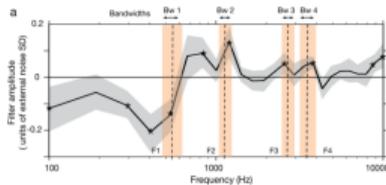
sentence recognition

[Venezia et al., 2016, 2019]



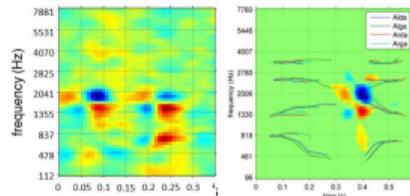
paralinguistics

[Ponsot et al., 2018a, 2018b]



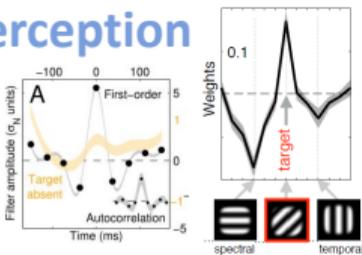
phoneme categorization

[Varnet et al., 2013, 2015]



modulation perception

[Ponsot et al., 2020; Joosten & Neri, 2012]

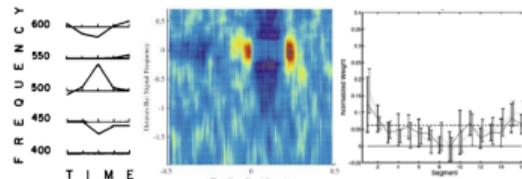


**YOU ARE
HERE**



pure-tone detection & loudness perception

[Ahumada & Lovell, 1971; Shub & Richards, 2009; Ponsot et al. 2013]

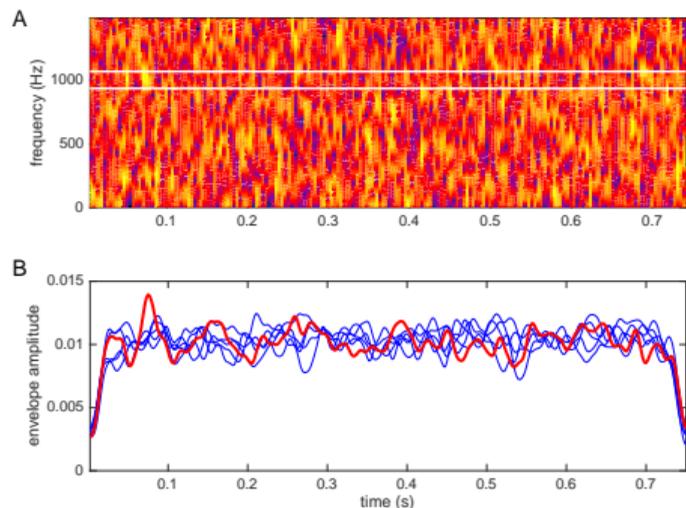


low level

Intrinsic fluctuations and revcorr

- Part of the deleterious effect of a steady noise on AM perception is due to the confusion of useful modulations in the signal with **intrinsic envelope fluctuations** arising from the filtering of noise into cochlear critical bands [*Dau et al., 1997, 1999*]
- Revcorr techniques are particularly suitable for exploring the effect of noise on perception

Using the revcorr method to **probe internal representations of AM** by relating random envelope fluctuations of masking noise with the response of listeners on a trial-by-trial basis



Envelopes for 6 realizations of a white noise, bandpass-filtered between 935 and 1068 Hz

Materials & methods

Signals:

Unmod. tone vs. 4-Hz AM

Duration = 0.75 s

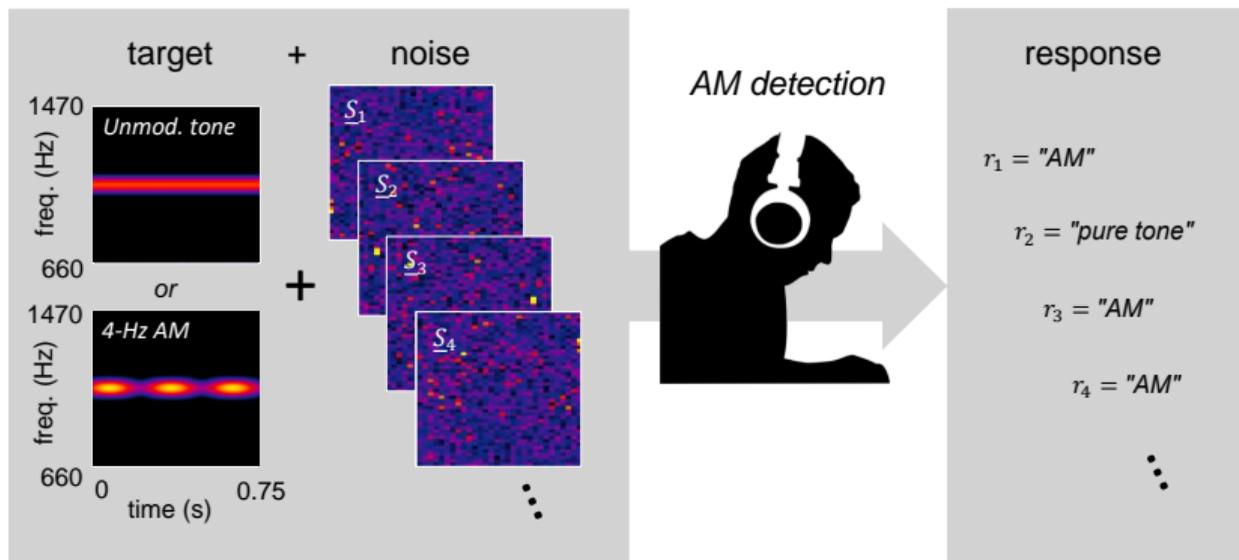
Carrier freq. = 1 kHz

Fixed AM phase

Task: yes/no detection task
in white noise (-10 dB SNR)

Modulation depth (m)
adapted from trial to trial to
target 70.7% correct
responses

Participants: 8 NH listeners
+ 1 participant with only one
NH ear



Materials & methods

Signals:

Unmod. tone vs. 4-Hz AM

Duration = 0.75 s

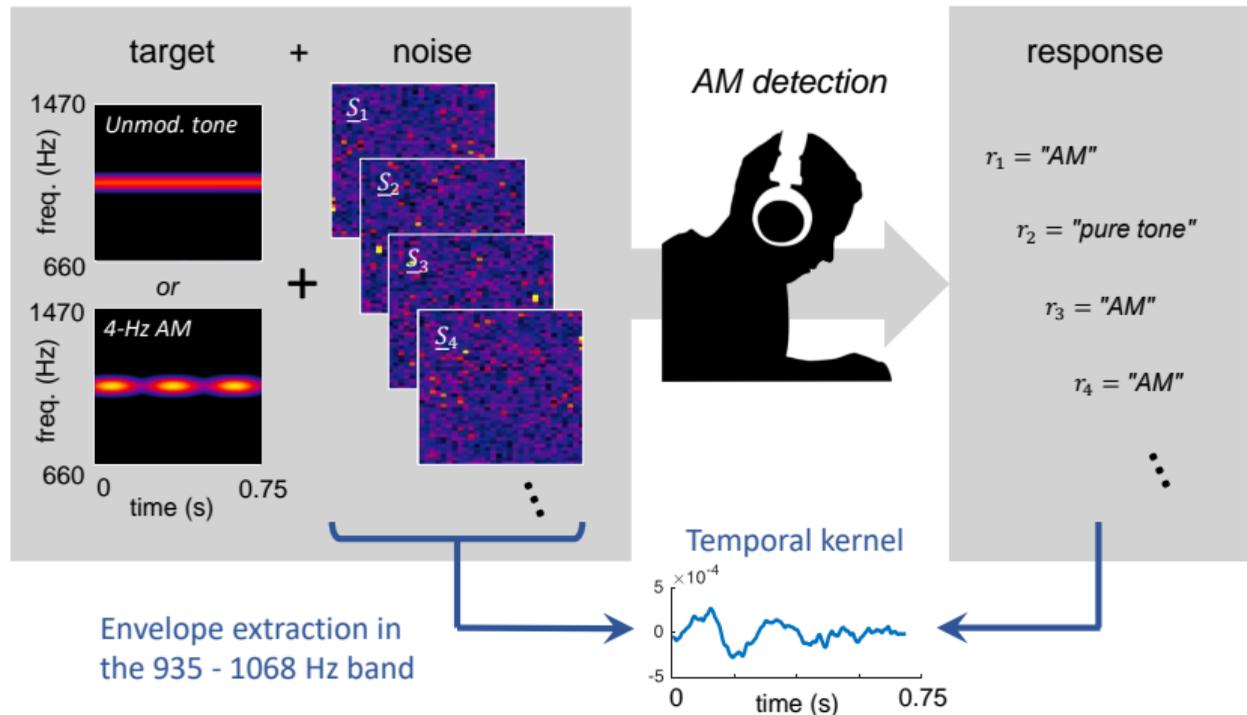
Carrier freq. = 1 kHz

Fixed AM phase

Task: yes/no detection task
in white noise (-10 dB SNR)

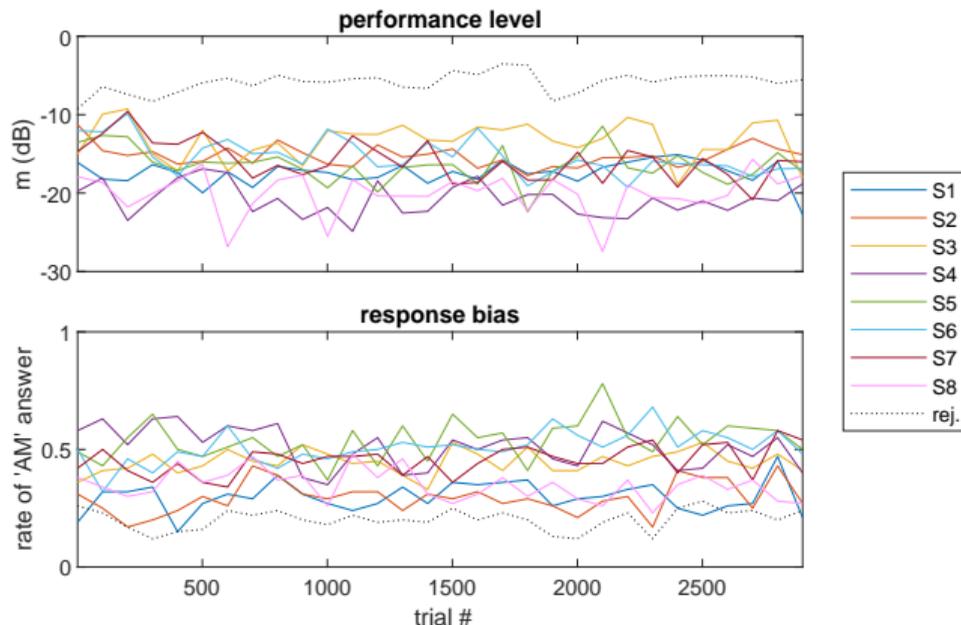
Modulation depth (m)
adapted from trial to trial to
target 70.7% correct
responses

Participants: 8 NH listeners
+ 1 participant with only one
NH ear



Performance over the course of the experiment

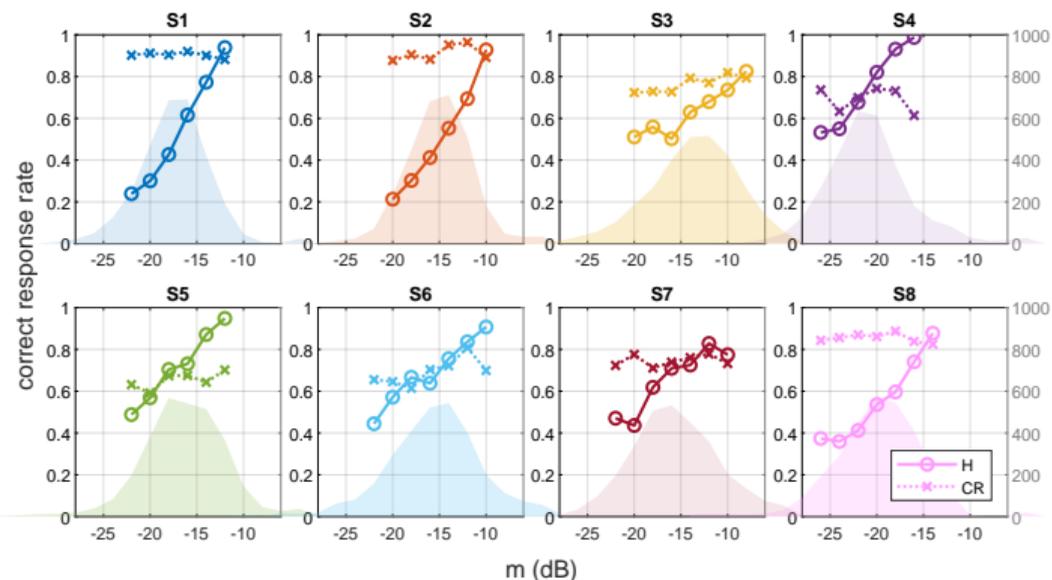
- 3.000 trials per participant (\approx 3h including breaks)
- Stable behavior in terms of performance and bias
- High performance level (mean $m \approx -16$ dB, at -10 dB SNR)
- Most participants are unbiased!



Performance as a function of modulation depth

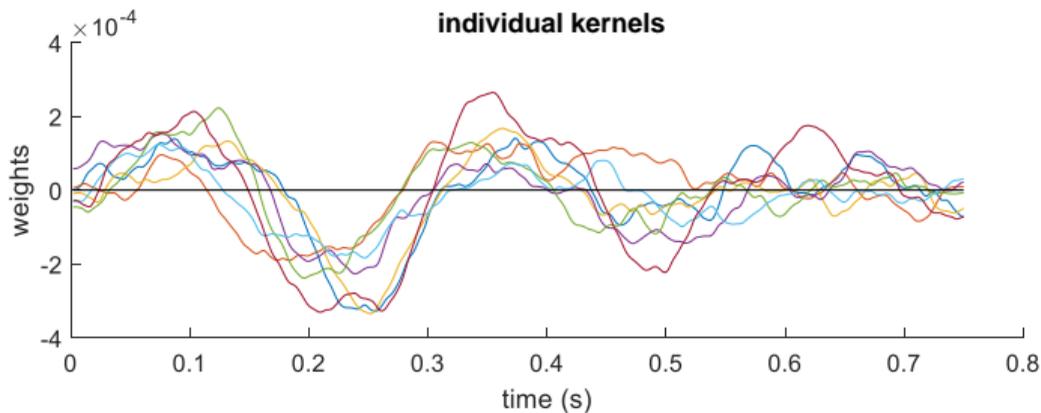
Performance in target-present trials (percentage of Hits) strongly depends on modulation depth.

Performance in target-absent trials (percentage of Correct Rejections) does not.



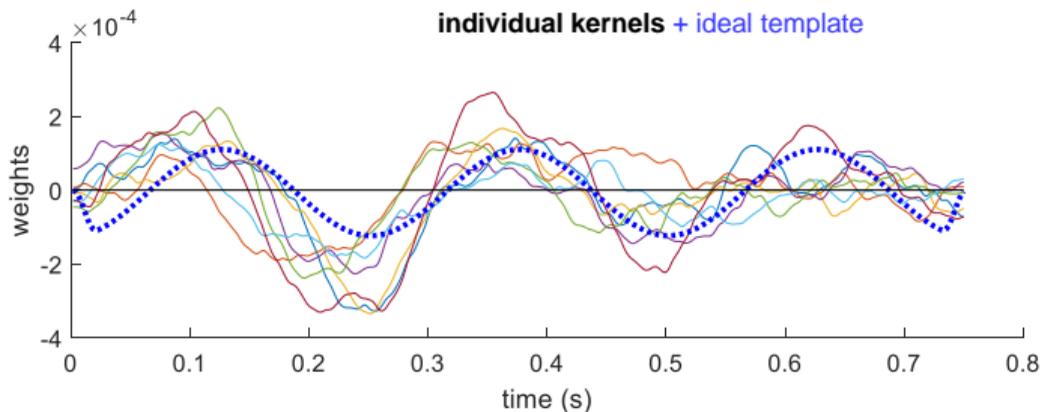
Temporal kernel (= pattern of intrinsic fluctuations of the noise consistently associated with an 'AM' response)

- Similar pattern for all 8 participants
- **Strong 4-Hz component**



Temporal kernel (= pattern of intrinsic fluctuations of the noise consistently associated with an 'AM' response)

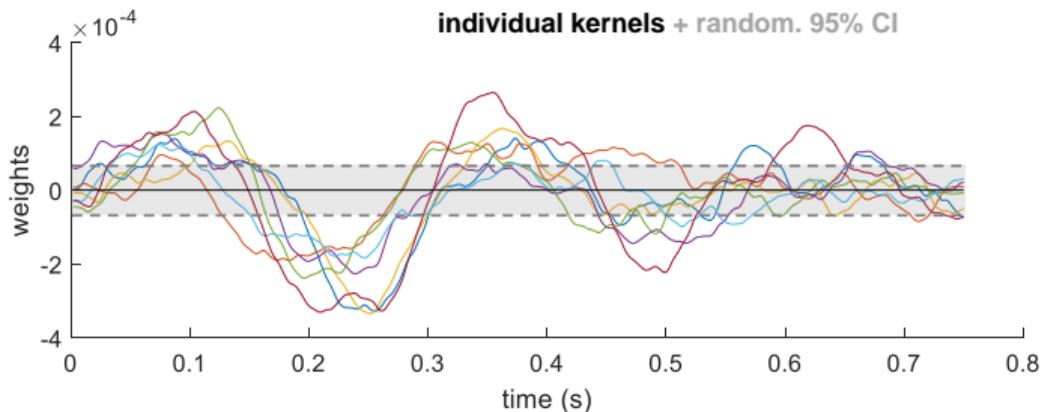
- Similar pattern for all 8 participants
- **Strong 4-Hz component**, very close to the **AM target**



⇒ When the noise intrinsic envelope shows a strong 4-Hz component, it is more likely to be confused with the AM target to be detected

Temporal kernel (= pattern of intrinsic fluctuations of the noise consistently associated with an 'AM' response)

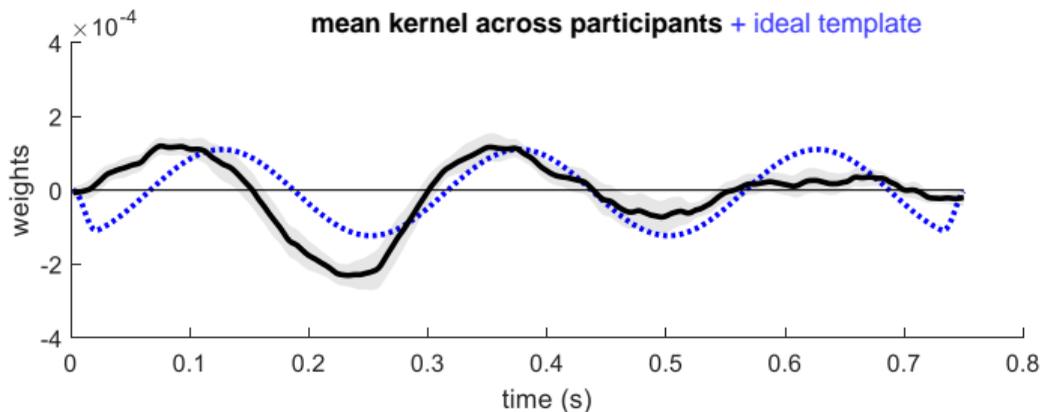
- Similar pattern for all 8 participants
- **Strong 4-Hz component**, very close to the **AM target**



⇒ When the noise intrinsic envelope shows a strong 4-Hz component, it is more likely to be confused with the AM target to be detected

Temporal kernel (= pattern of intrinsic fluctuations of the noise consistently associated with an 'AM' response)

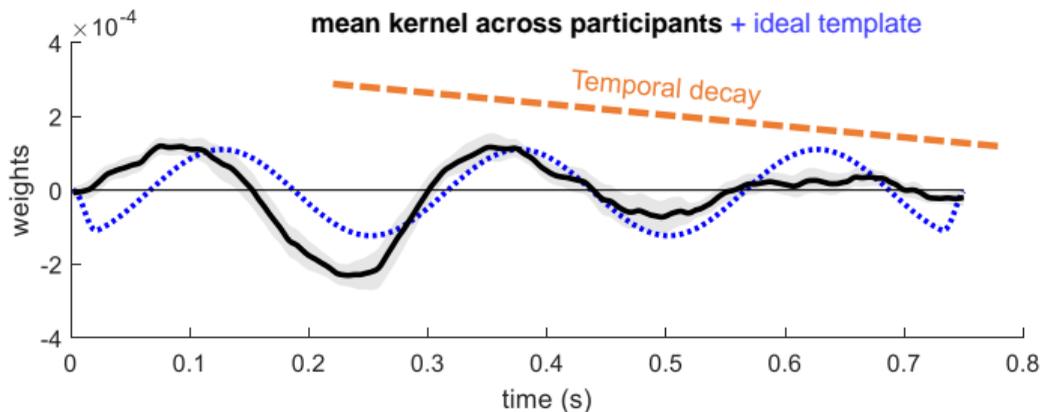
- Similar pattern for all 8 participants
- **Strong 4-Hz component**, very close to the **AM target**



⇒ When the noise intrinsic envelope shows a strong 4-Hz component, it is more likely to be confused with the AM target to be detected

Temporal kernel (= pattern of intrinsic fluctuations of the noise consistently associated with an 'AM' response)

- Similar pattern for all 8 participants
- **Strong 4-Hz component**, very close to the **AM target**

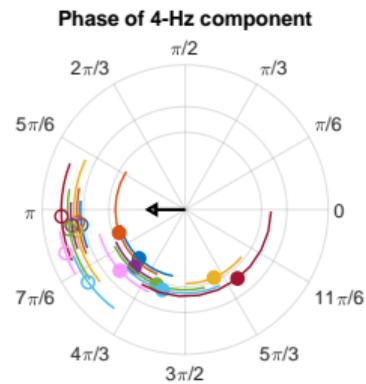
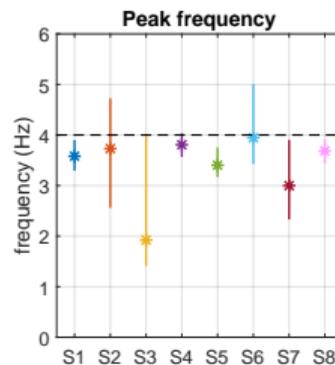
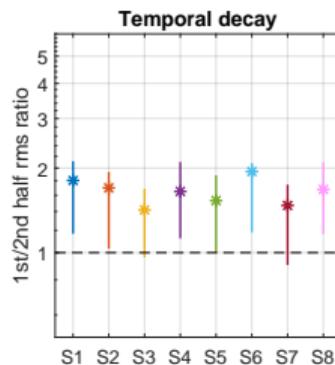


⇒ When the noise intrinsic envelope shows a strong 4-Hz component, it is more likely to be confused with the AM target to be detected

Differences with the ideal template inform us about the underlying auditory processes

Kernel metrics

Deviations to the ideal template were quantified and assessed at the individual level using 3 metrics derived from the kernels (+ 95% bootstrap intervals).



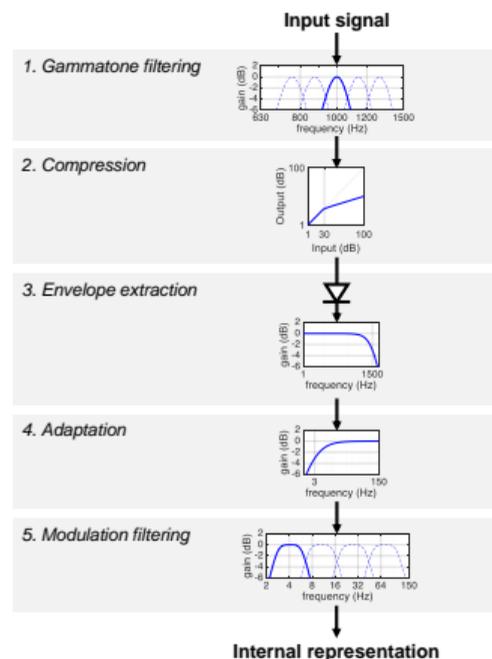
Identifying the source of these suboptimalities will require a model of the human auditory system...

Modulation Filterbank Model

The **Modulation Filterbank** model: a simple, widely-accepted front-end model of the auditory system. [e.g., *Dau et al., 1997; Ewert & Dau, 2000; King et al., 2019*]



Auditory Modeling Toolbox
 (<https://www.amtoolbox.org/>)
 [Majdak et al., 2021; Osses et al., 2021]



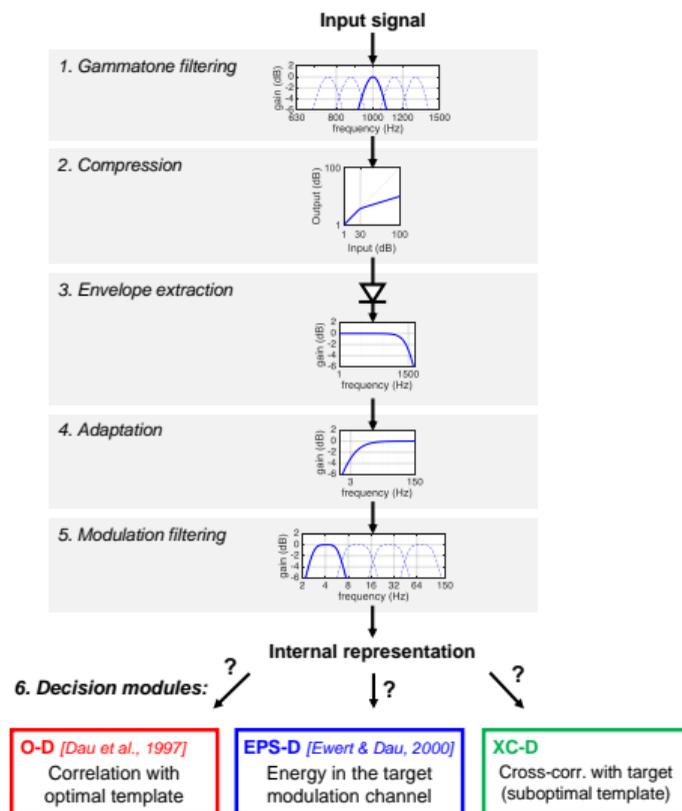
Modulation Filterbank Model

The **Modulation Filterbank** model: a simple, widely-accepted front-end model of the auditory system. [e.g., *Dau et al., 1997*; *Ewert & Dau, 2000*; *King et al., 2019*]

However, no general agreement on how the auditory system converts this internal representation into a decision statistics.

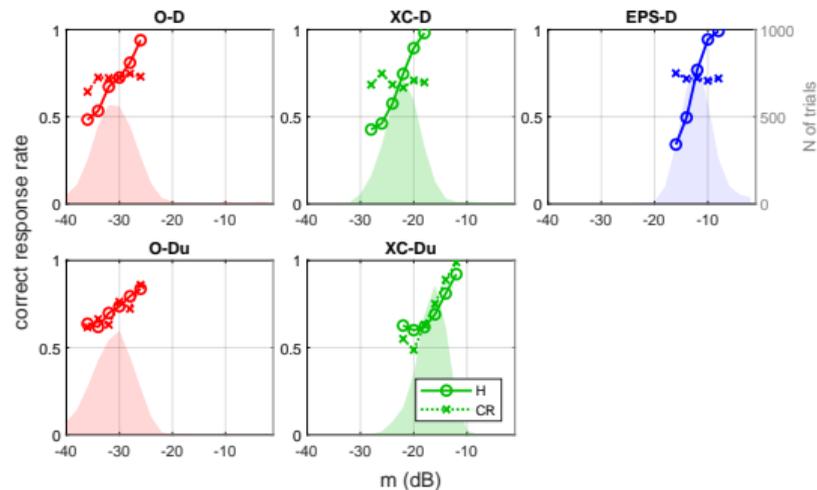


Auditory Modeling Toolbox
(<https://www.amtoolbox.org/>)
[*Majdak et al., 2021*; *Osses et al., 2021*]



Performance as a function of modulation depth

O-D and XC-D base their decision on the comparison with a template of the target. Is this template updated throughout the experiment or not?

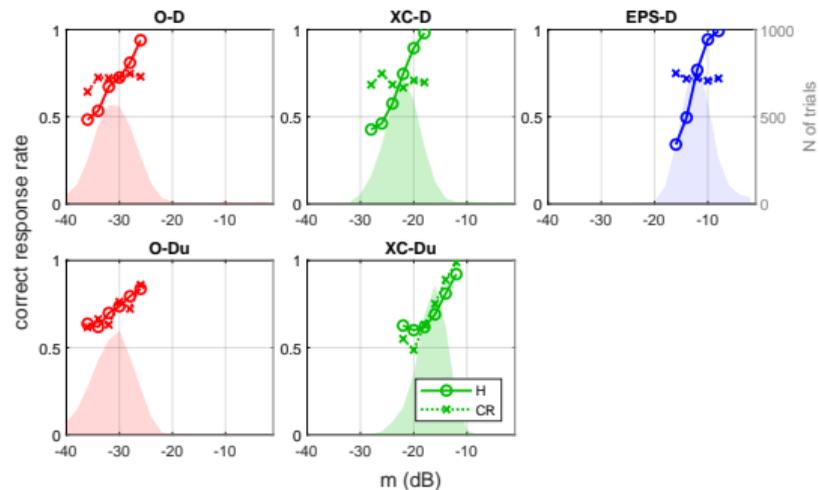


Performance as a function of modulation depth

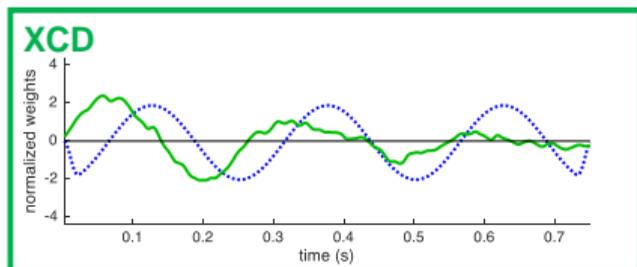
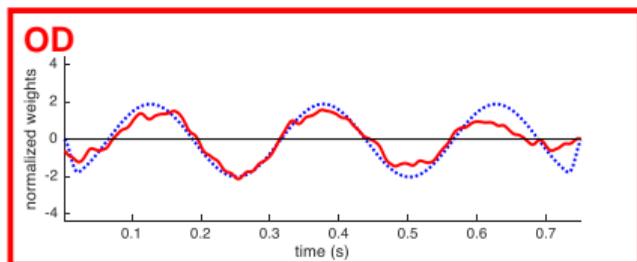
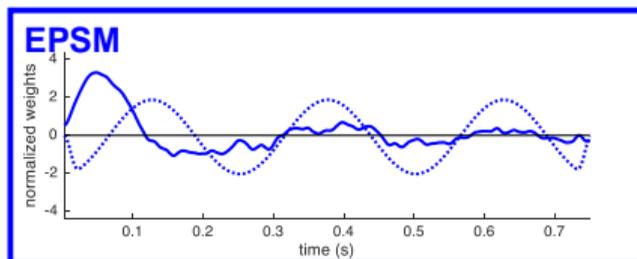
O-D and XC-D base their decision on the comparison with a template of the target. Is this template updated throughout the experiment or not?

Only non-updating models can account for the fact that the performance in target-absent trials does not depend on the performance in target-present trials.

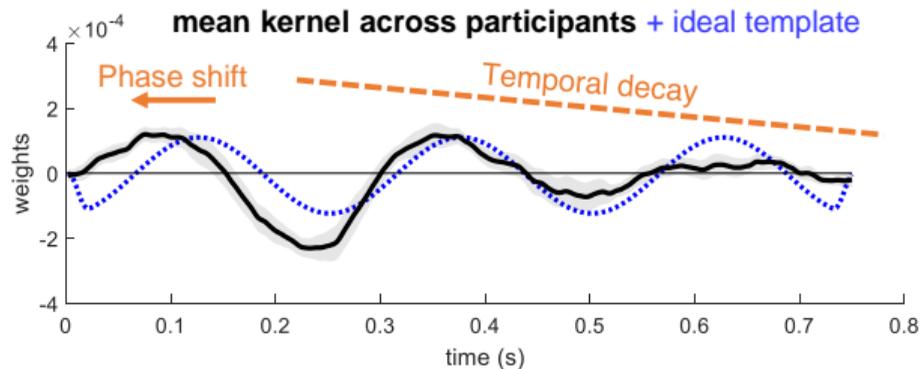
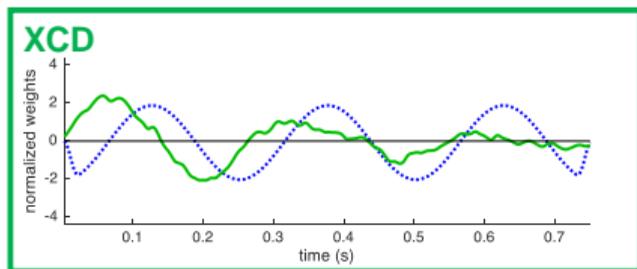
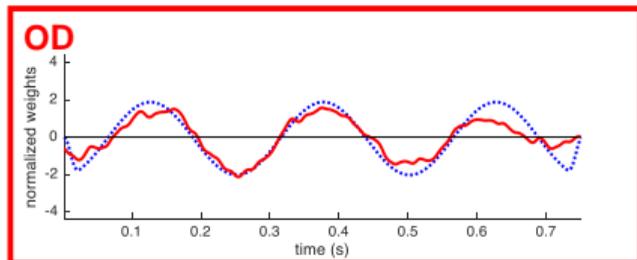
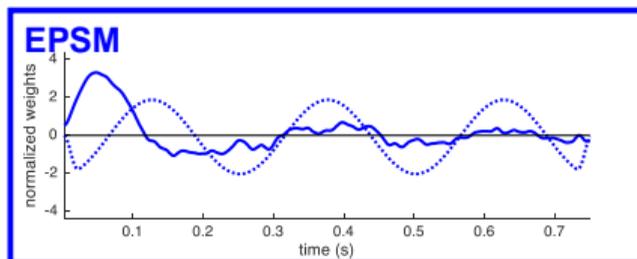
⇒ supports the idea that human participants used a fixed decision rule over the duration of the experiment.



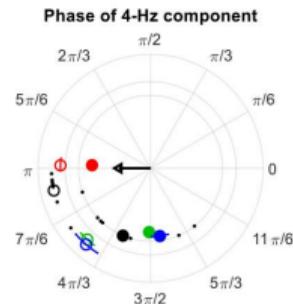
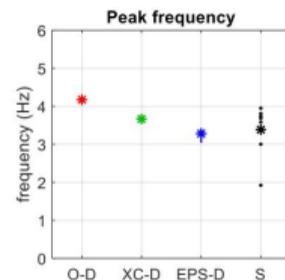
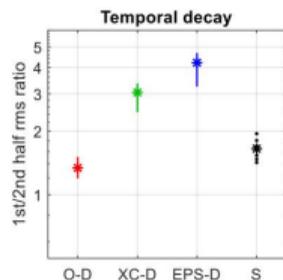
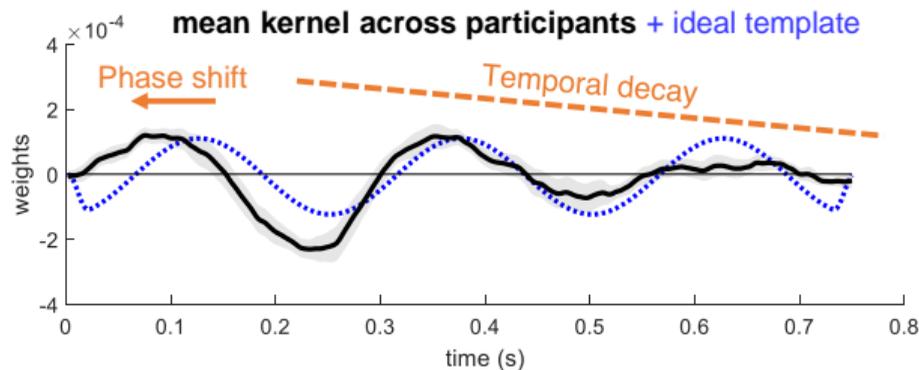
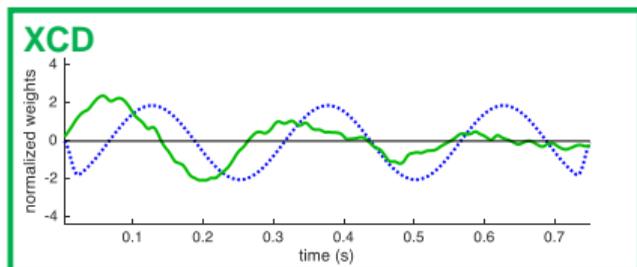
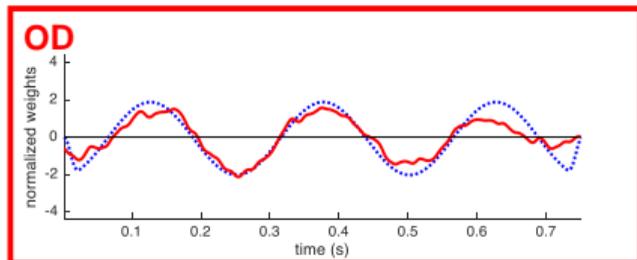
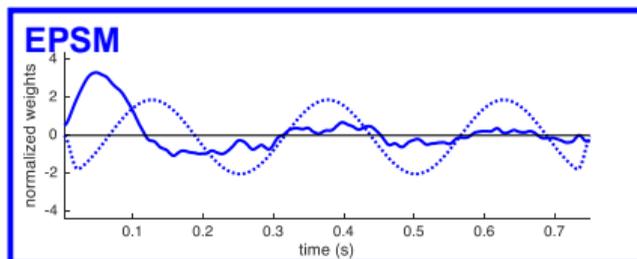
Simulated kernels



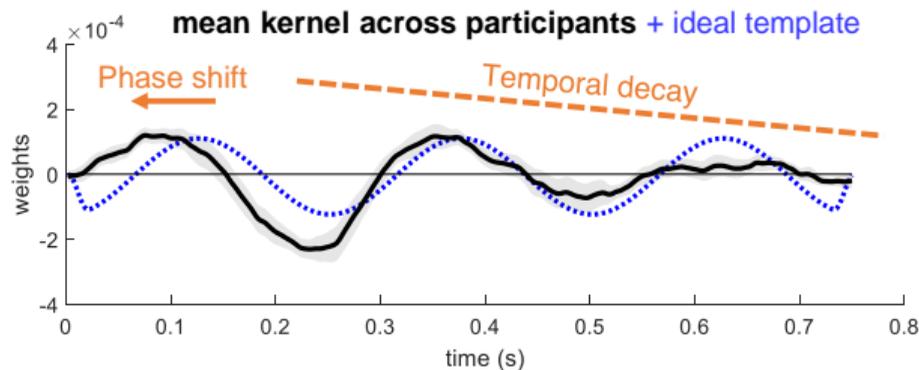
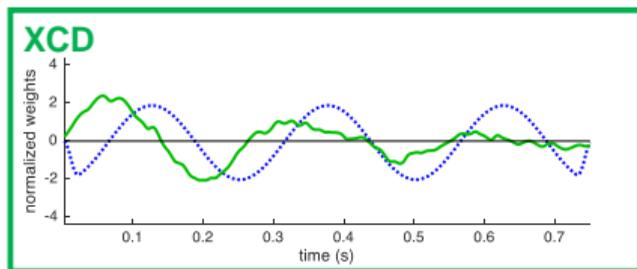
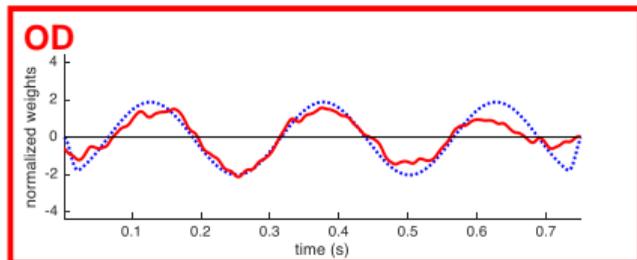
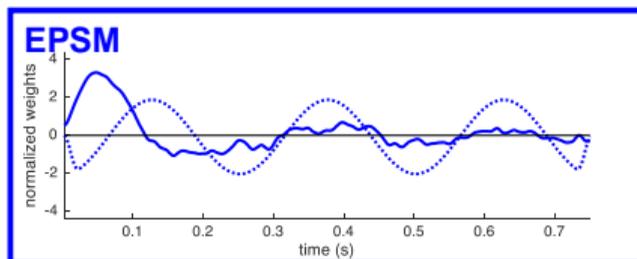
Simulated kernels



Simulated kernels



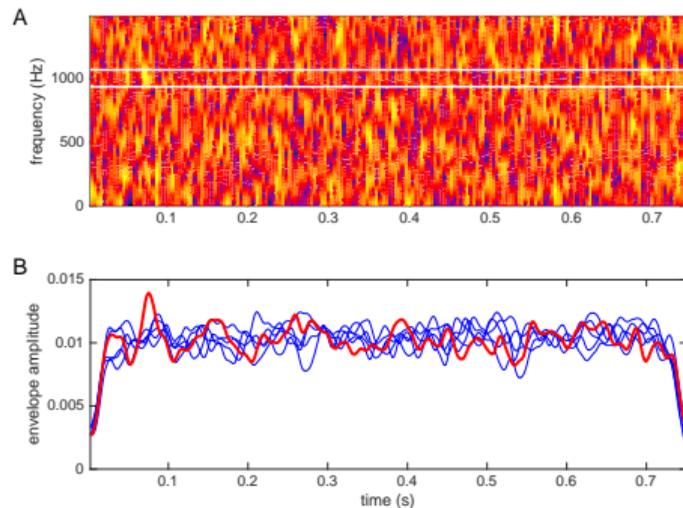
Simulated kernels

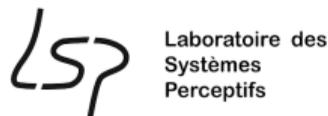


Only **XCD** can reproduce the characteristics of the human kernel, suggesting that the auditory system uses a suboptimal template-matching strategy for this task

Conclusions

- Direct illustration of the **effect of intrinsic fluctuations on perception**
 - The **auditory revcorr** technique can be used to **probe internal representations and decision strategies** in real and simulated listeners
- Listeners are able to encode a 4-Hz AM target with its phase, but they use a **suboptimal detection strategy**

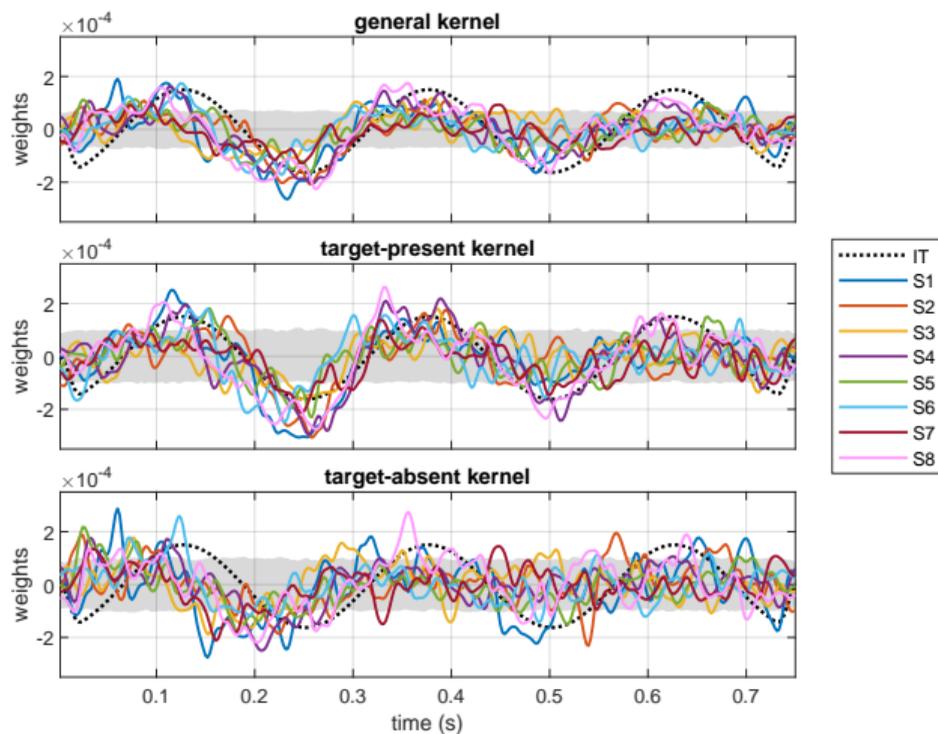




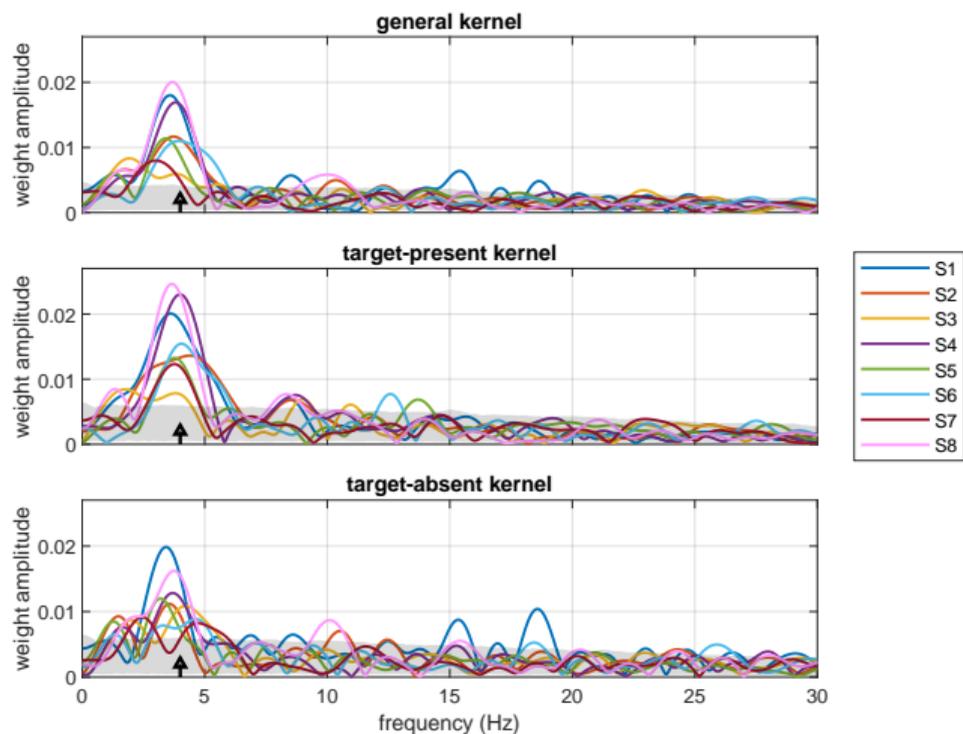
Thank you for your attention!

and thanks to **Alejandro Osses & Christian Lorenzi**

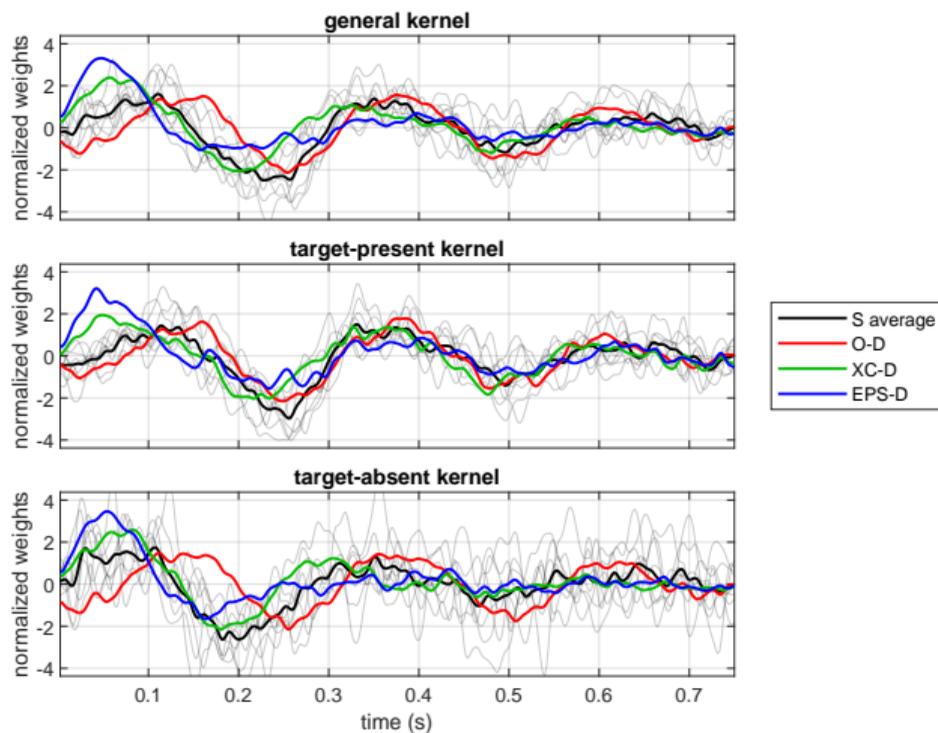
Annex 1: target-present and target-absent kernels



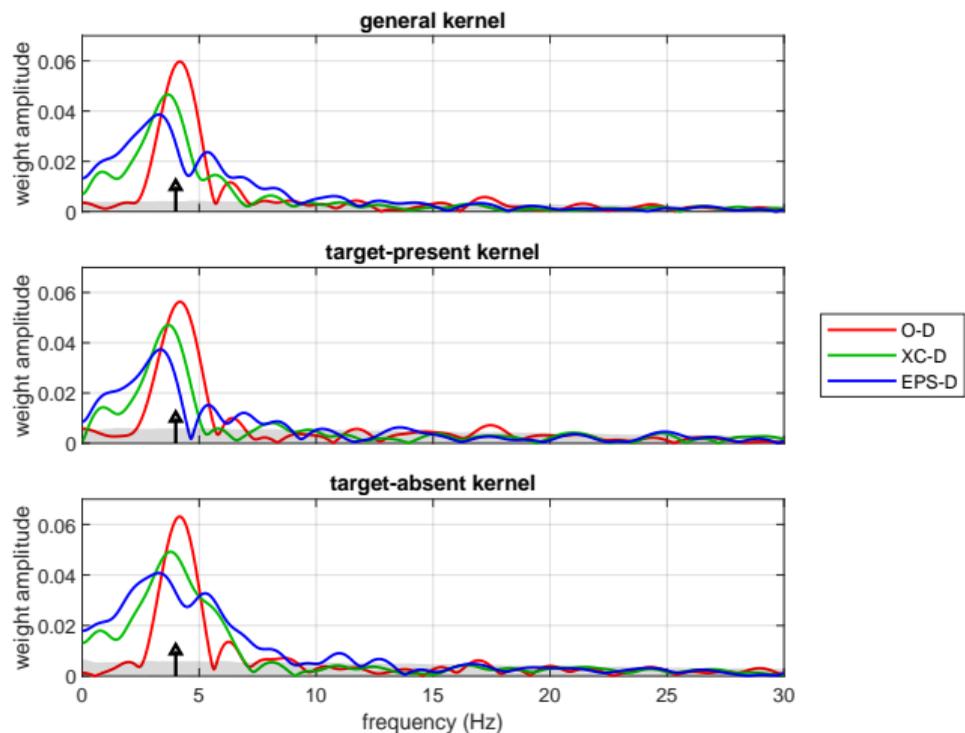
Annex 2: Fourier kernels



Annex 3: simulated kernels



Annex 4: simulated Fourier kernels



Annex 5: simulated kernels with padded stimuli

