A Comparative Study of Auditory Attention Decoding Algorithms

Simon Geirnaert, Servaas vandecappelle,
Tom Francart, Alexander Bertrand

KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

expORL

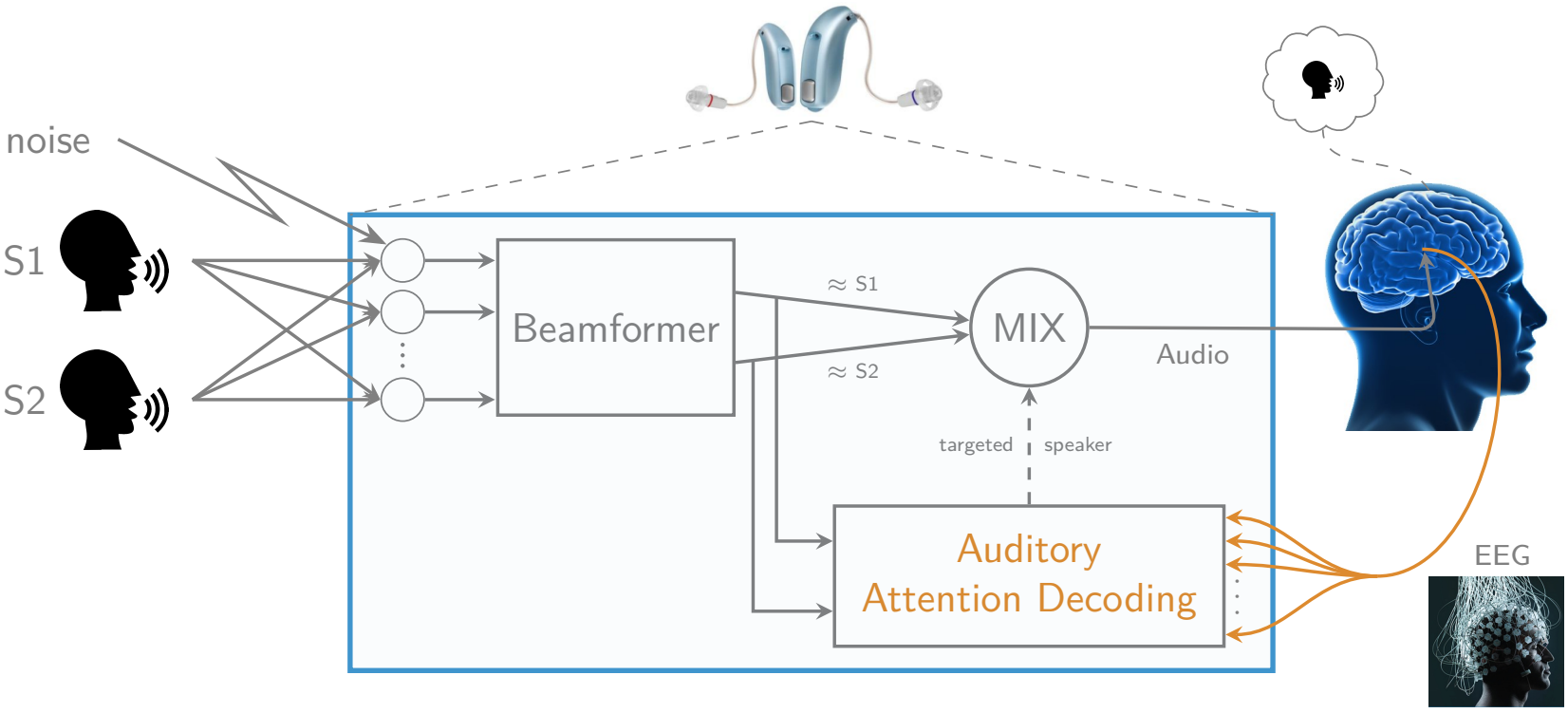
AESoP

17/09/2019

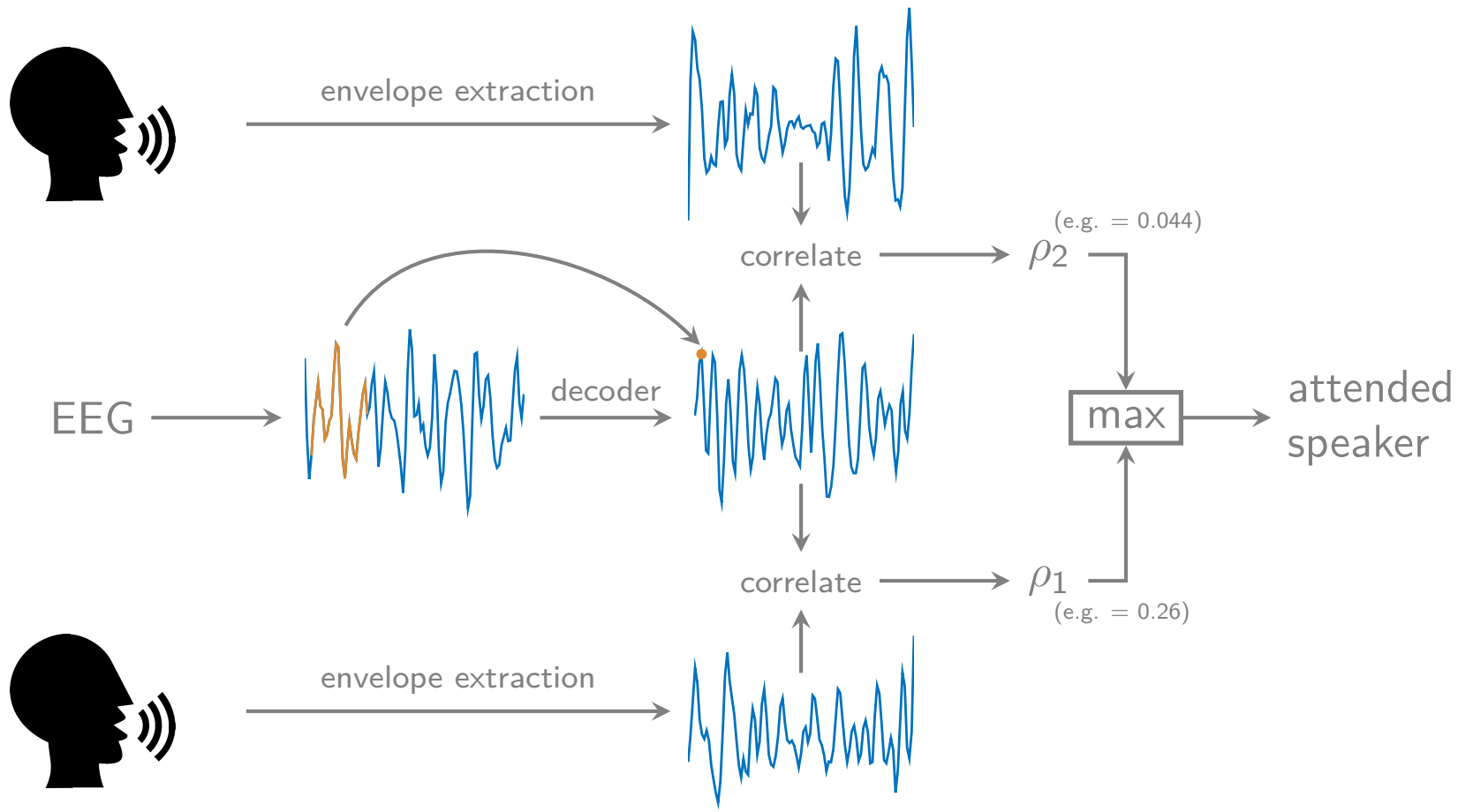
fwo Research Foundation Flanders
Opening new horizons

erc European Research Council

The concept of neuro-steered hearing prostheses



A backward decoding framework for AAD: stimulus reconstruction



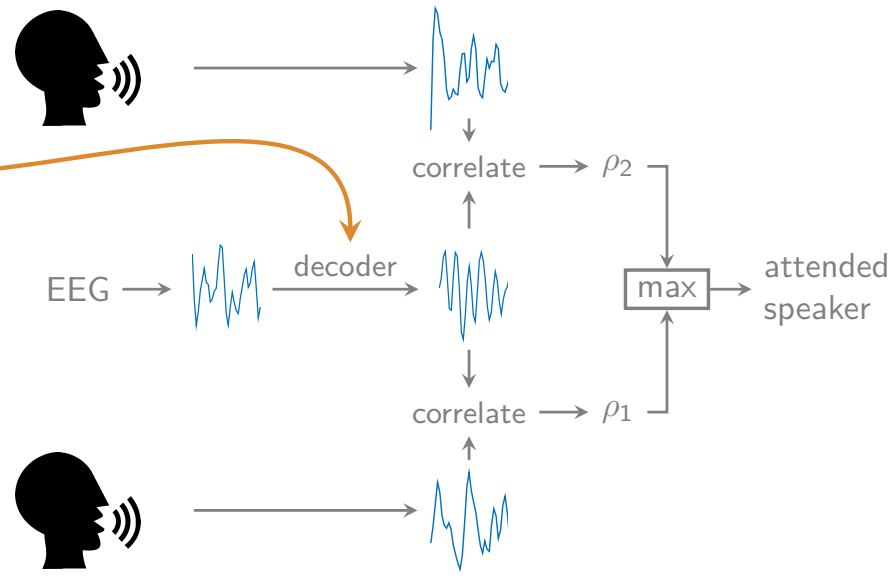
AAD = auditory attention decoding

A backward decoding framework for AAD: stimulus reconstruction

Training

- Lots of data to fit the decoder (~minutes)

Testing



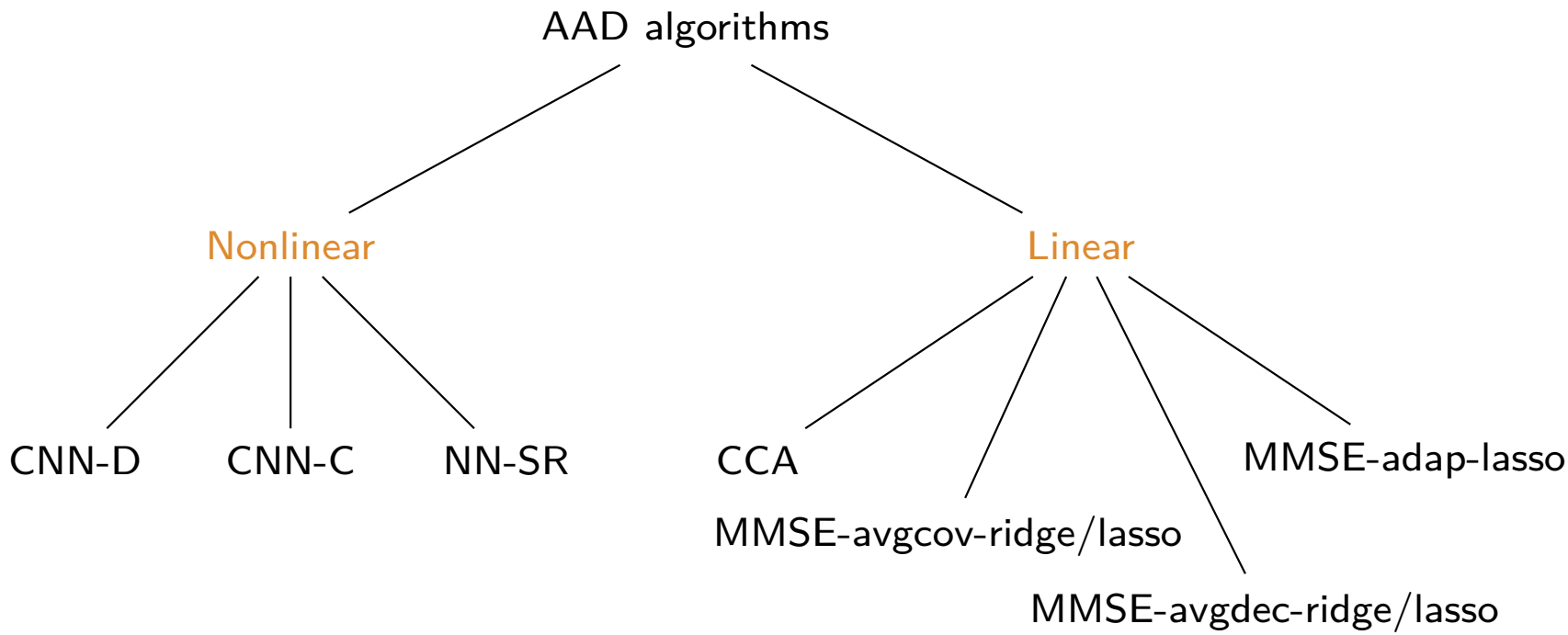
- Small decision window (~seconds)

Benchmarking of different AAD algorithms

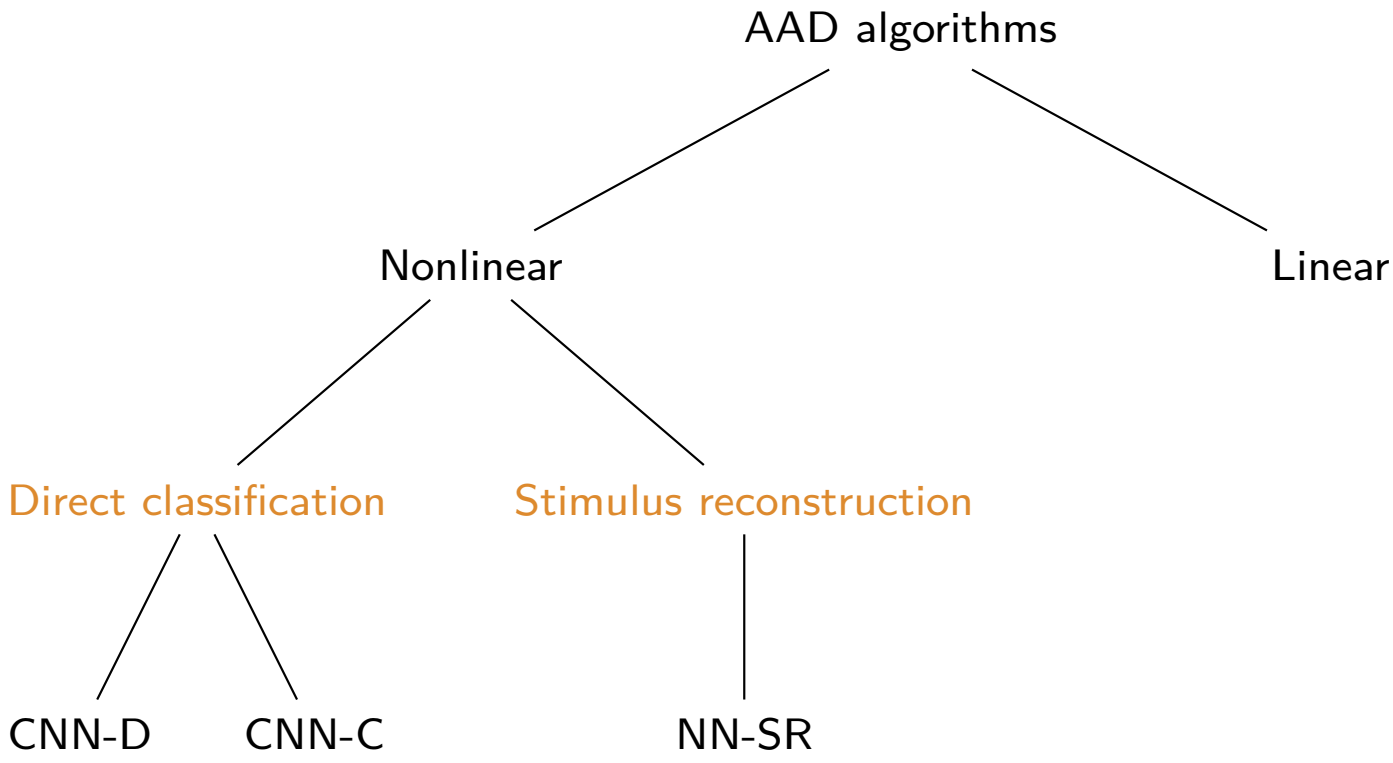
Paper	Code/Method
O'Sullivan et al., 2014	MMSE-avgdec-ridge
Biesmans et al., 2017	MMSE-avgcov-ridge
Alickovic et al., 2019	MMSE-avgdec-lasso
Alickovic et al., 2019	MMSE-avgcov-lasso
Miran et al., 2018	MMSE-adap-lasso
de Cheveigné et al., 2018	CCA
de Taillez et al., 2017	NN-SR
Deckers et al., 2018 (**P30**)	CNN-D
Ciccarelli et al., 2018	CNN-C

The implementations of the algorithms have been validated by the authors of the corresponding papers

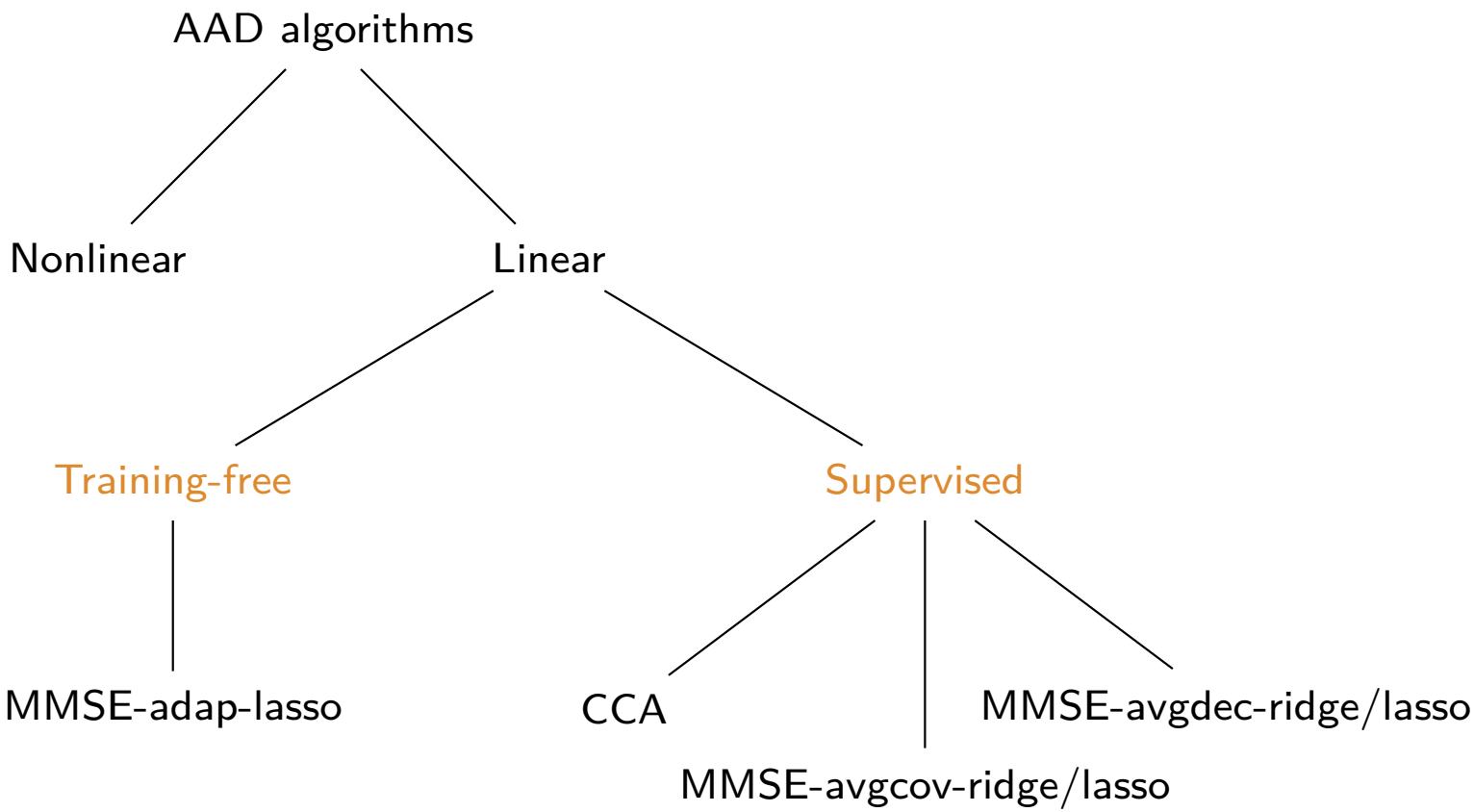
Benchmarking of different AAD algorithms



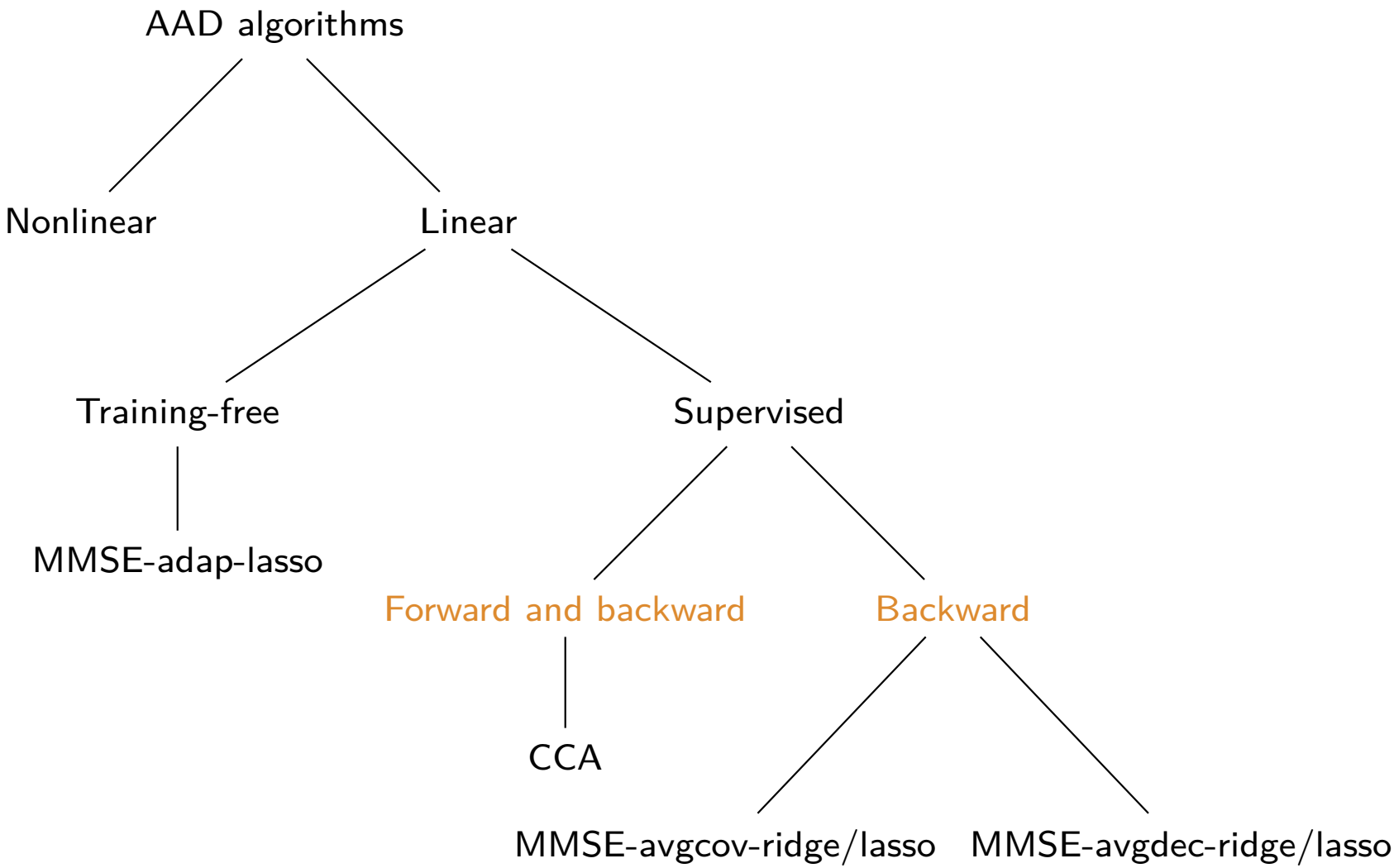
Benchmarking of different AAD algorithms



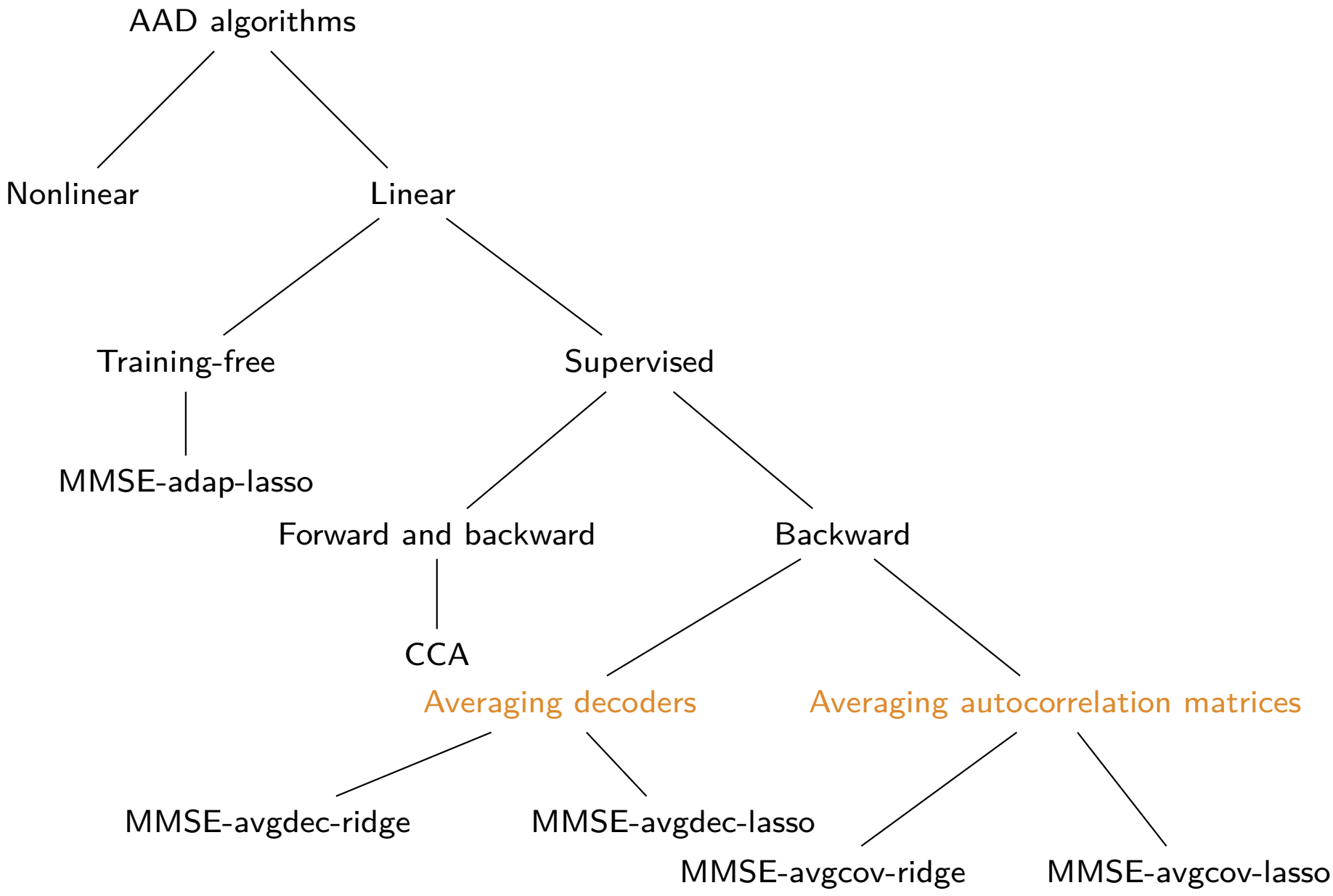
Benchmarking of different AAD algorithms



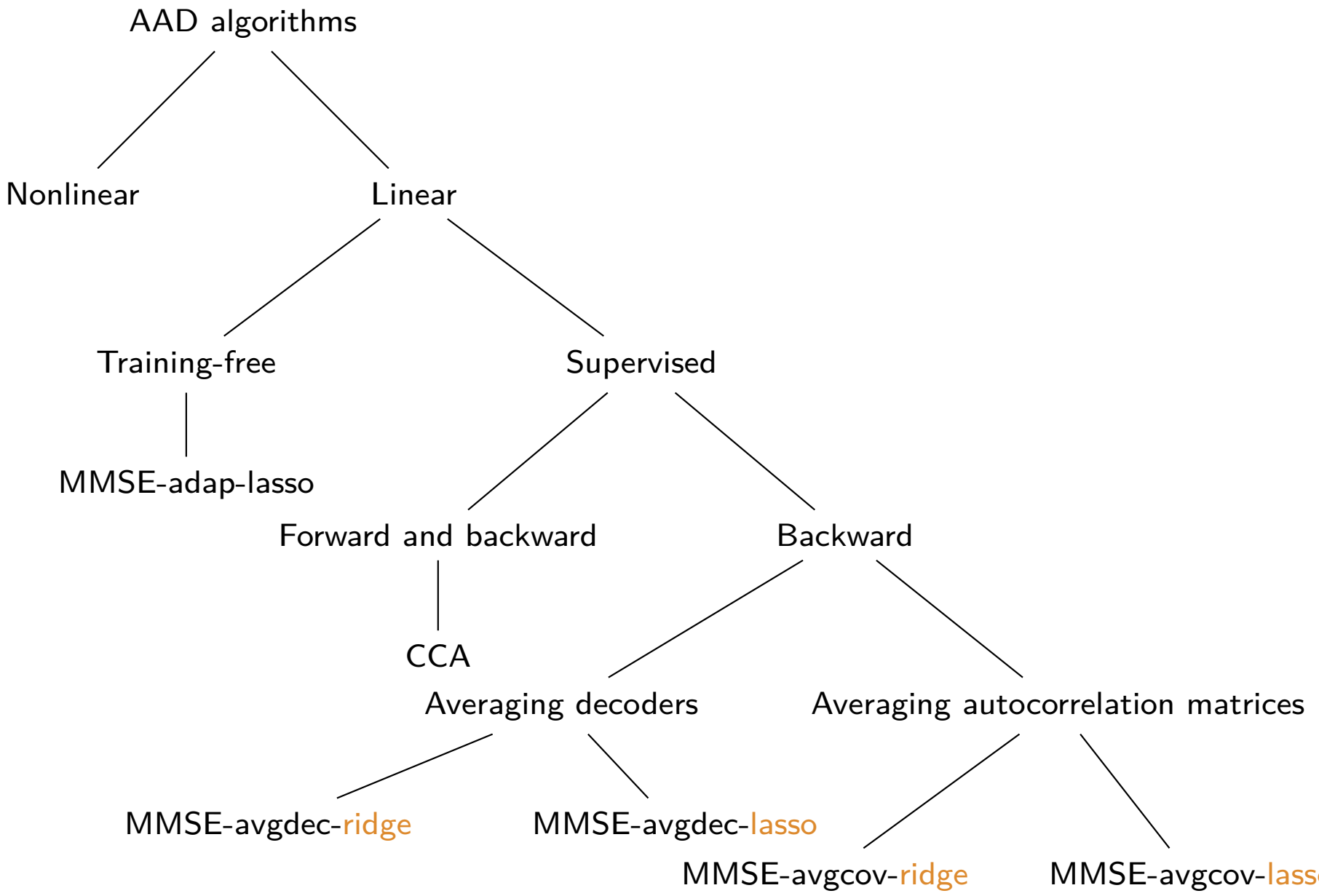
Benchmarking of different AAD algorithms



Benchmarking of different AAD algorithms



Benchmarking of different AAD algorithms



Take-home messages

I

Deep learning methods *can* achieve very high performances, but suffer from computational complexity and high variability

II

Within linear methods, a combination of backward and forward modeling works best (CCA)

III

Average covariance matrices, rather than decoders

The setup: data, testing procedure

Two independent datasets are used:

AADKUL-2015	Fuglsang-2018
	

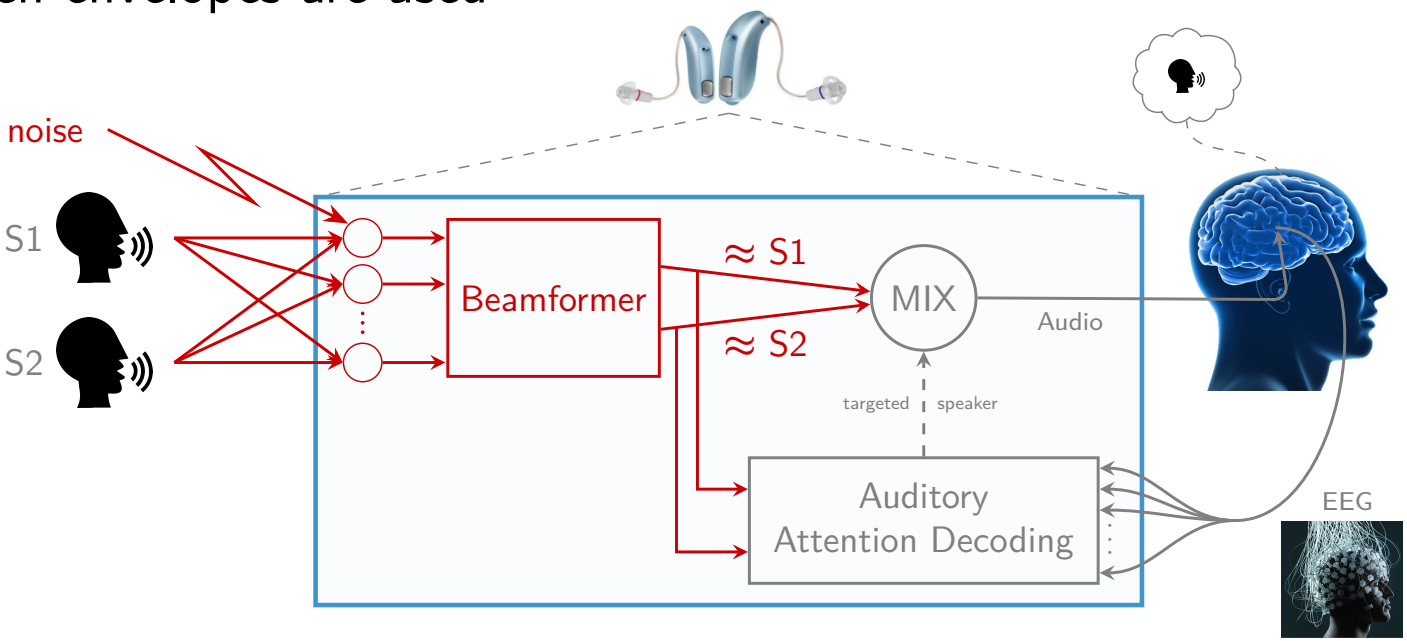
All performances and hyperparameters are cross-validated

The setup: data, testing procedure

Two independent datasets are used:

AADKUL-2015	Fuglsang-2018

Clean speech envelopes are used

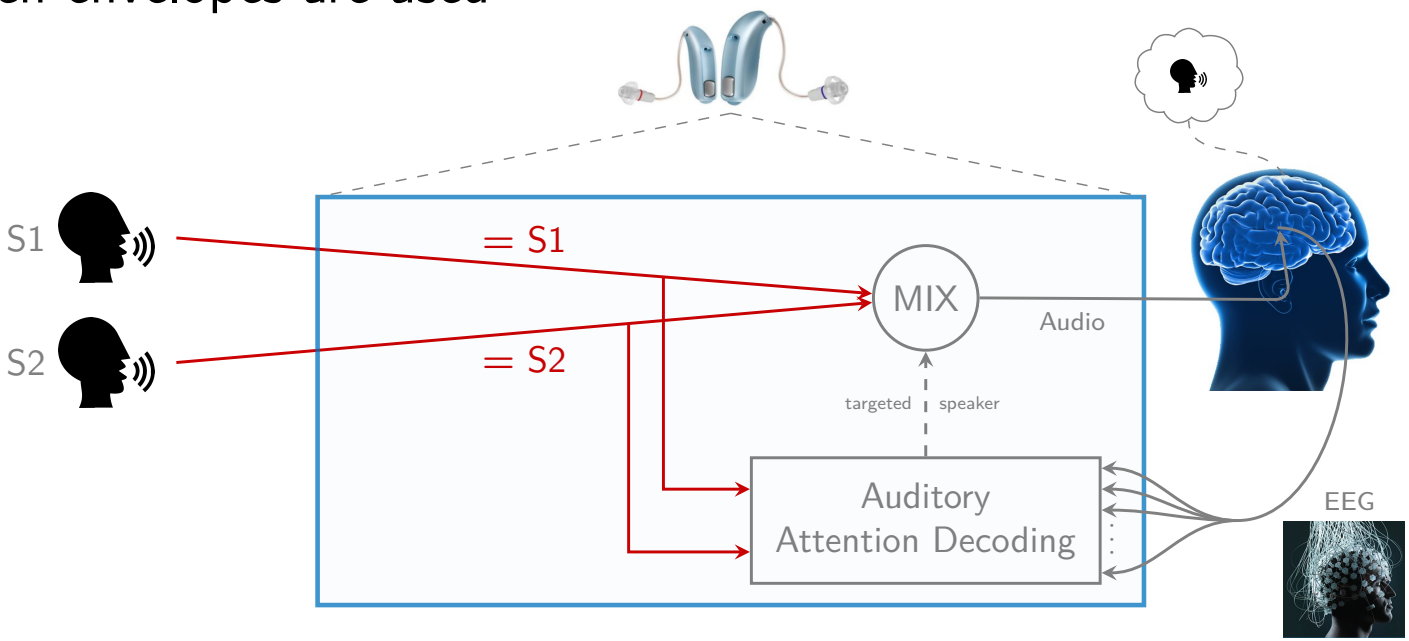


The setup: data, testing procedure

Two independent datasets are used:

AADKUL-2015	Fuglsang-2018

Clean speech envelopes are used



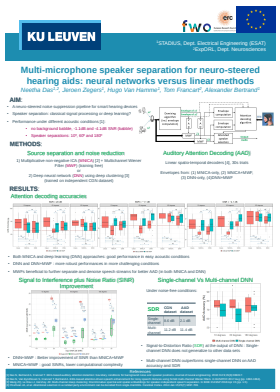
The setup: data, testing procedure

Two independent datasets are used:

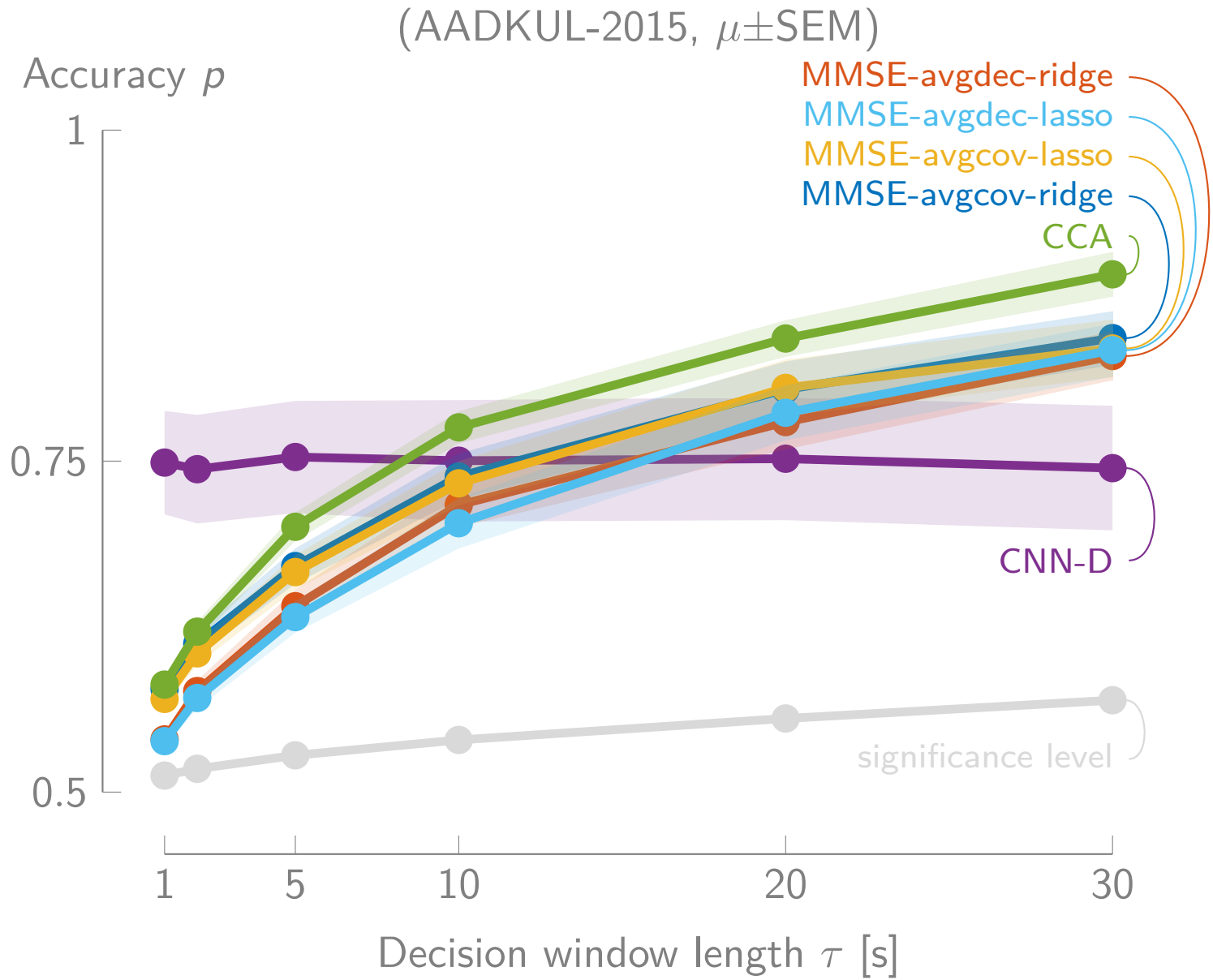
AADKUL-2015	Fuglsang-2018

Clean speech envelopes are used, but covered in . . .

- (Van Eyndhoven et al., 2016), (Han et al., 2019), (Aroudi et al., 2019), . . .
- Poster *Multi-microphone speaker separation for neuro-steered hearing aids: neural networks versus linear methods* of Neetha Das (**P24**)

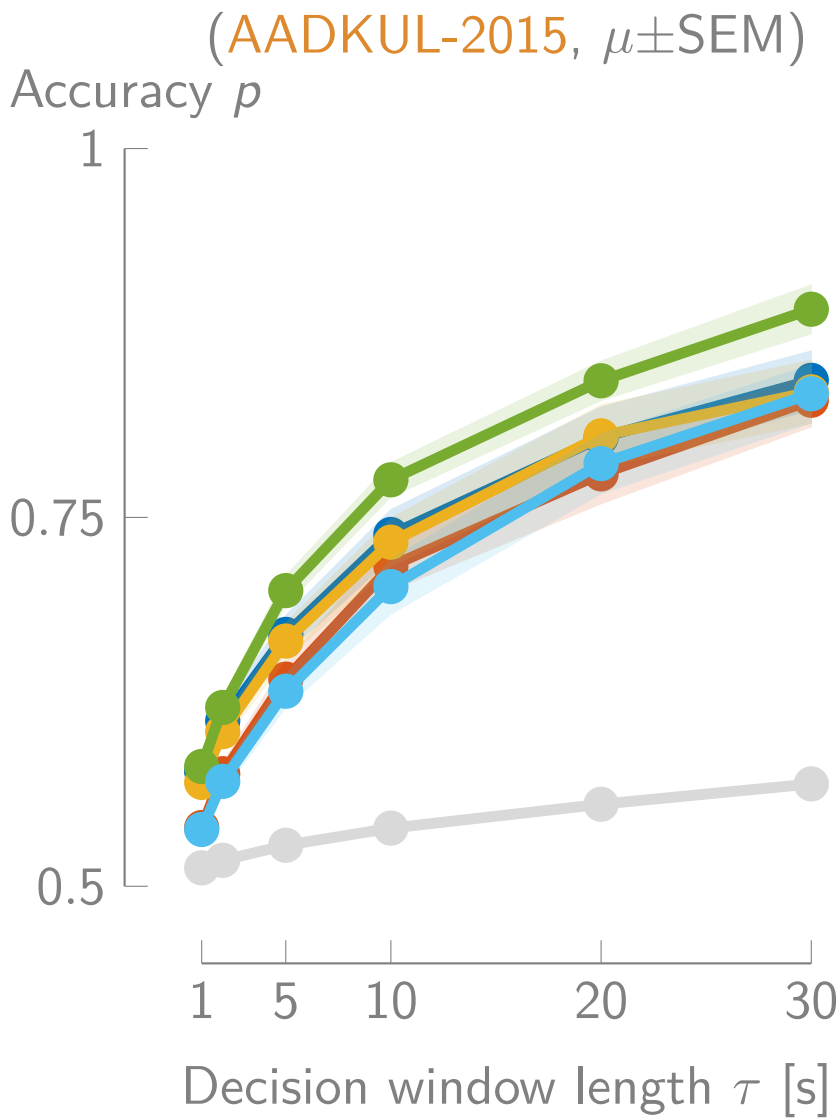


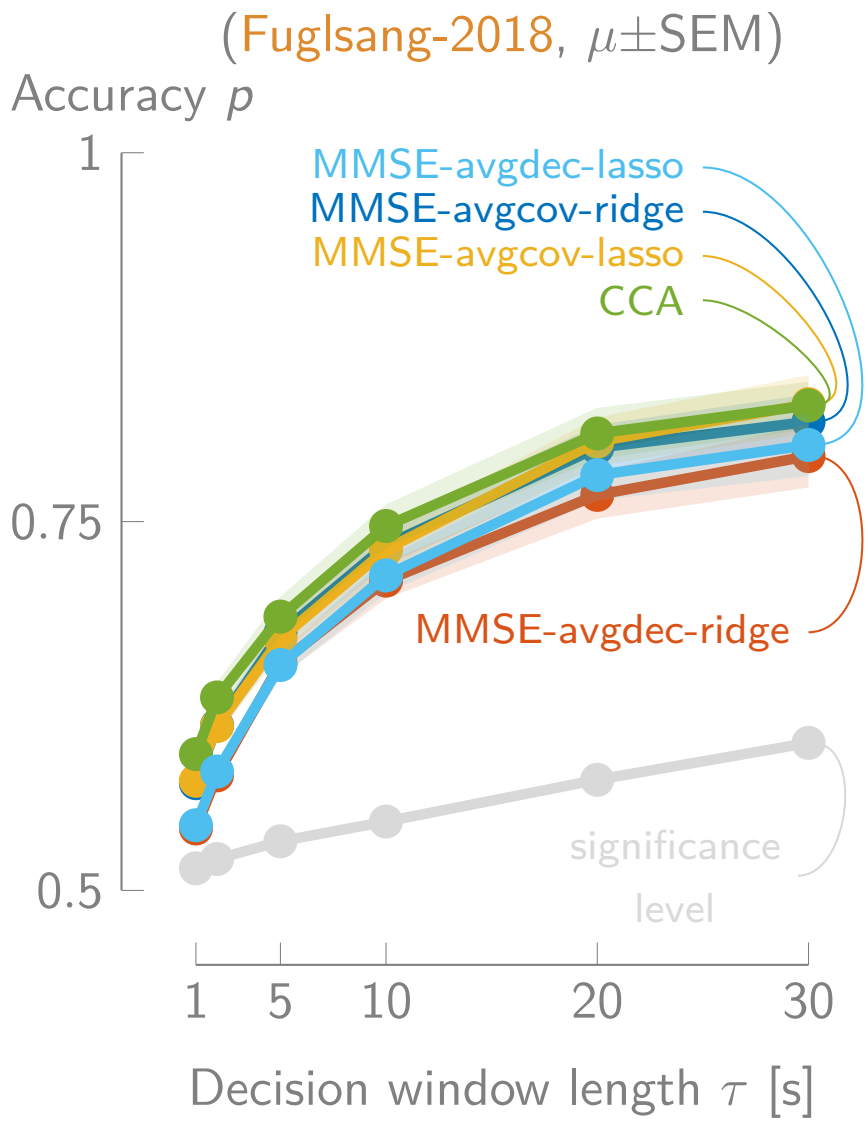
The $p(\tau)$-performance curves: a visual analysis



Disclaimer: these are preliminary results (some algorithms still missing)

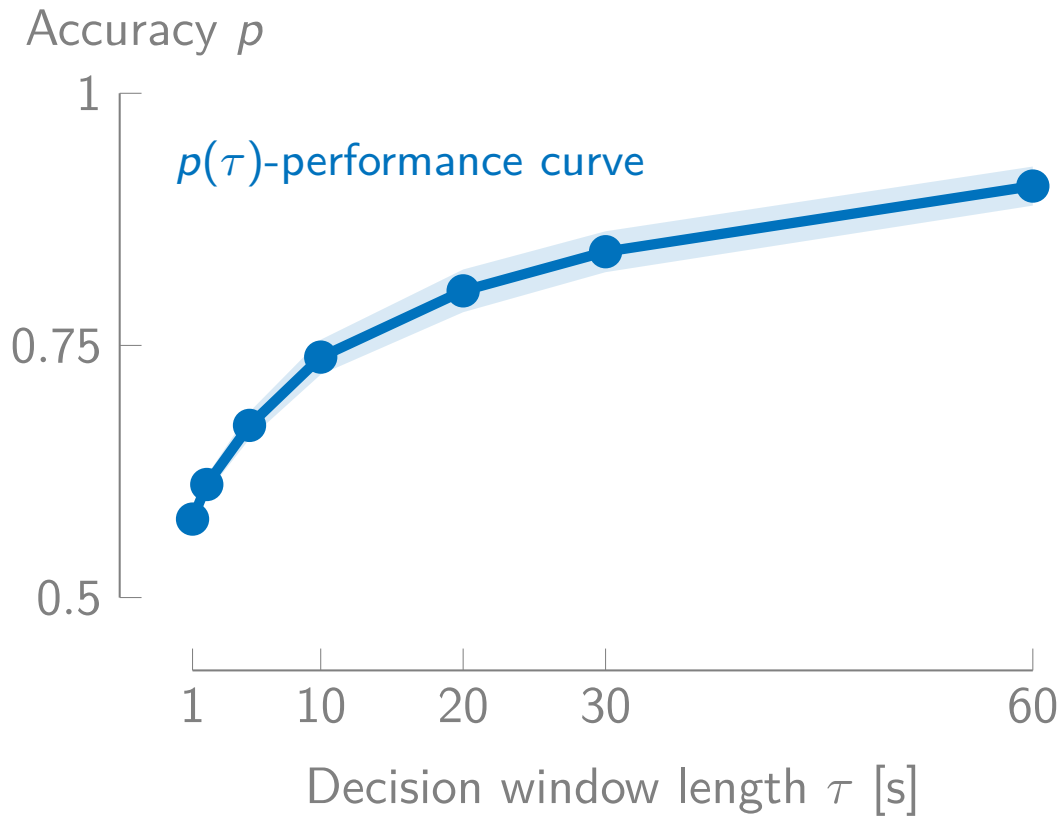
The $p(\tau)$-performance curves: a visual analysis





Disclaimer: these are preliminary results (some algorithms still missing)

A new performance metric: the MESD



Which point is most practical for an attention-tracking gain control system?

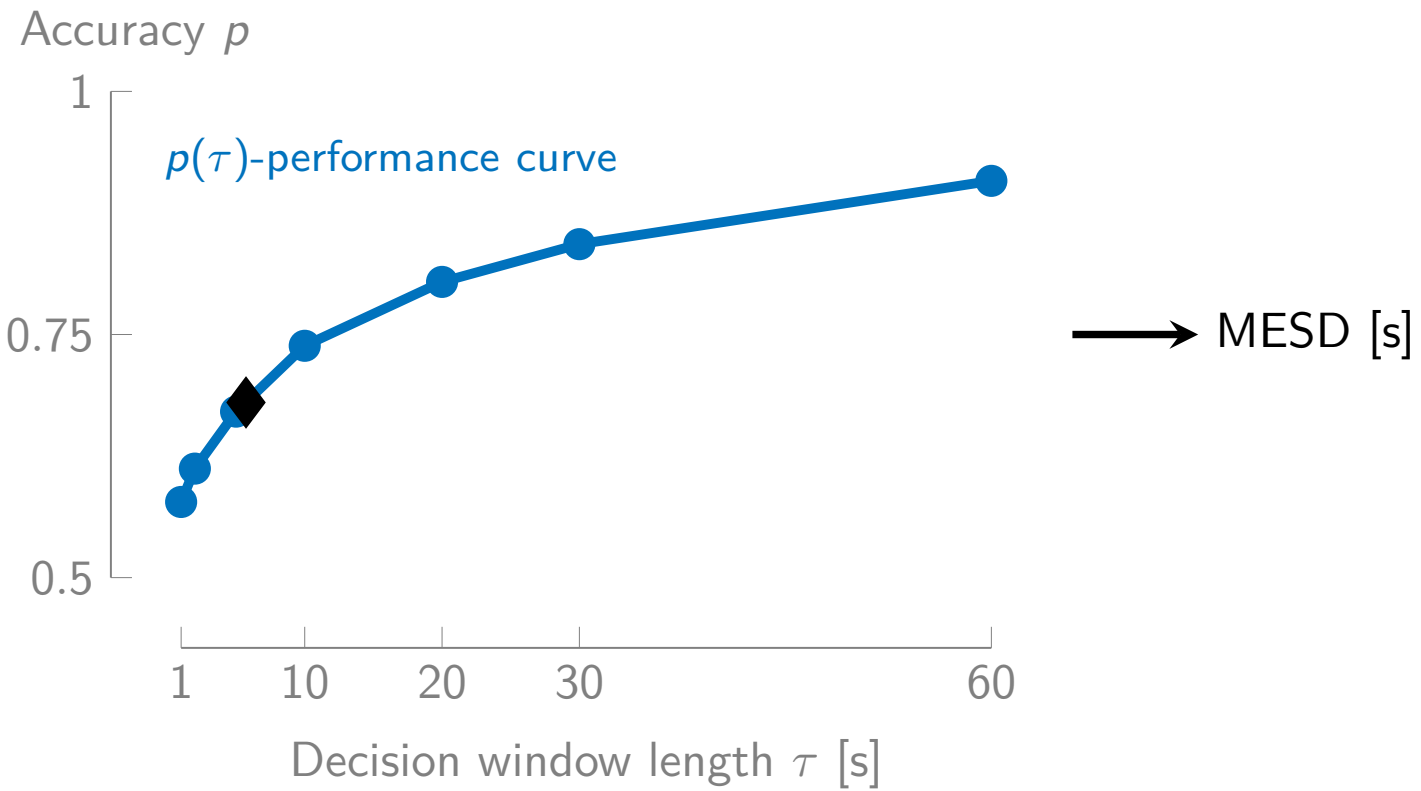
- Fast but inaccurate?
- Slow but accurate?

A new performance metric: the MESD

Which point is most practical for an attention-tracking gain control system?

- Fast but inaccurate?
- Slow but accurate?

Optimizing the attention-tracking gain control system leads to the minimal expected switch duration (MESD) performance metric



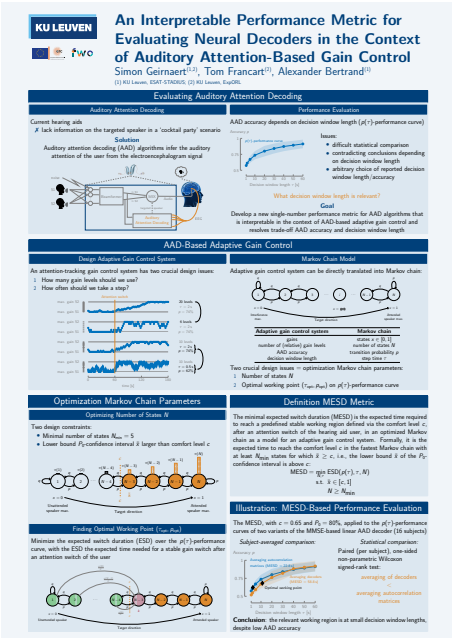
A new performance metric: the MESD

Which point is most practical for an attention-tracking gain control system?

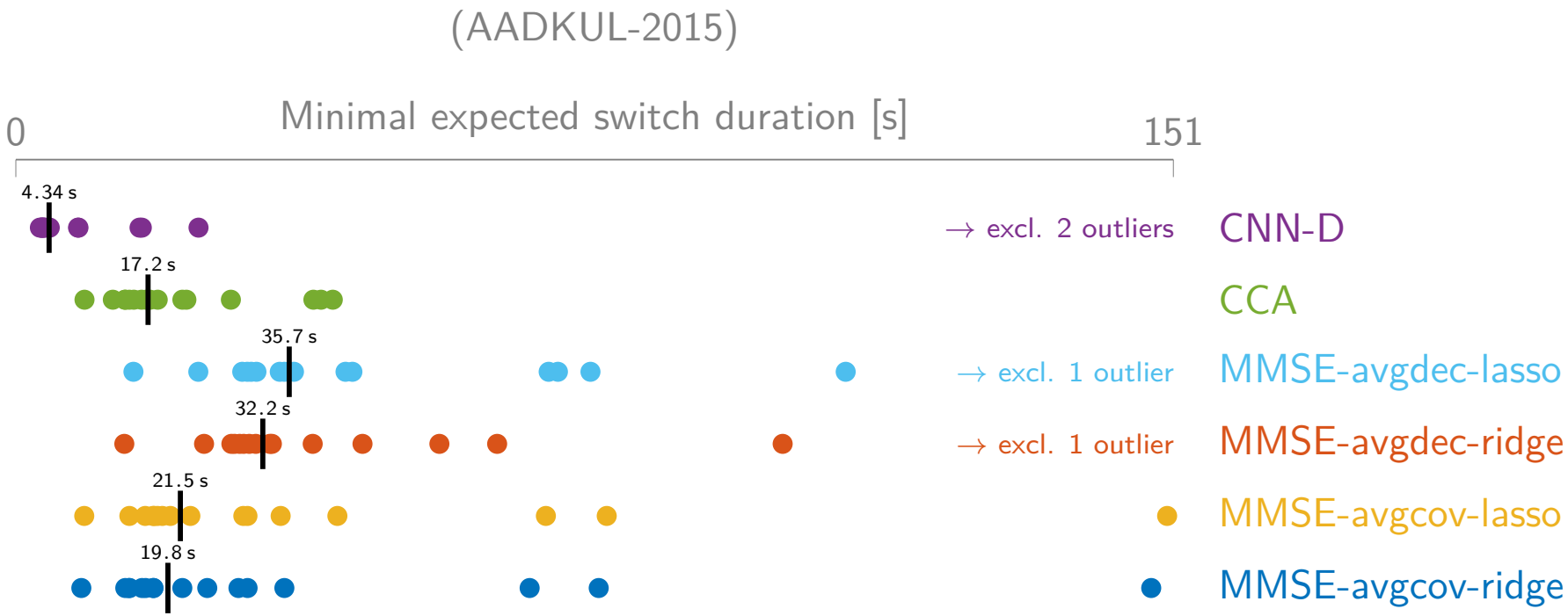
- Fast but inaccurate?
- Slow but accurate?

Optimizing the attention-tracking gain control system leads to the minimal expected switch duration (MESD) performance metric

*Find out more at poster **P32**!*

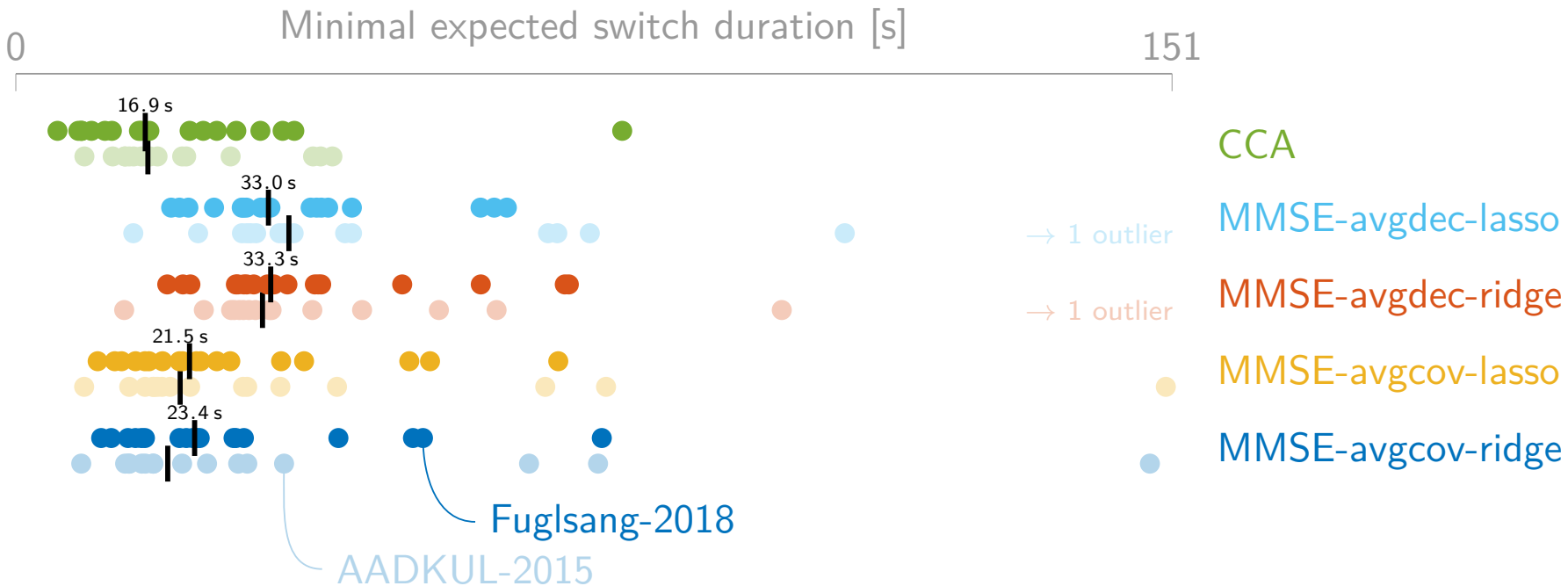


Comparing the per-subject MESDs



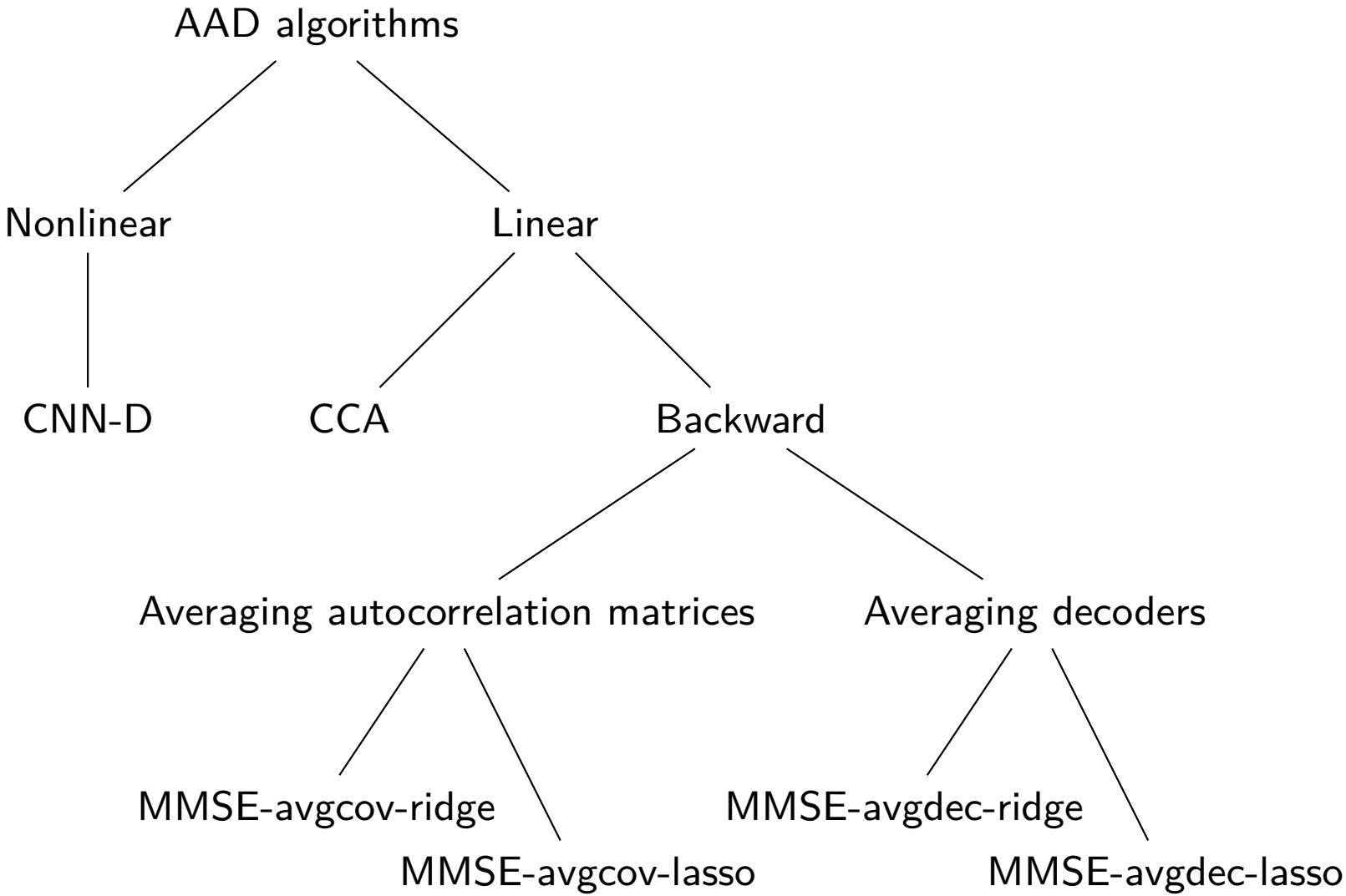
Disclaimer: these are preliminary results (some algorithms still missing)

Comparing the per-subject MESDs



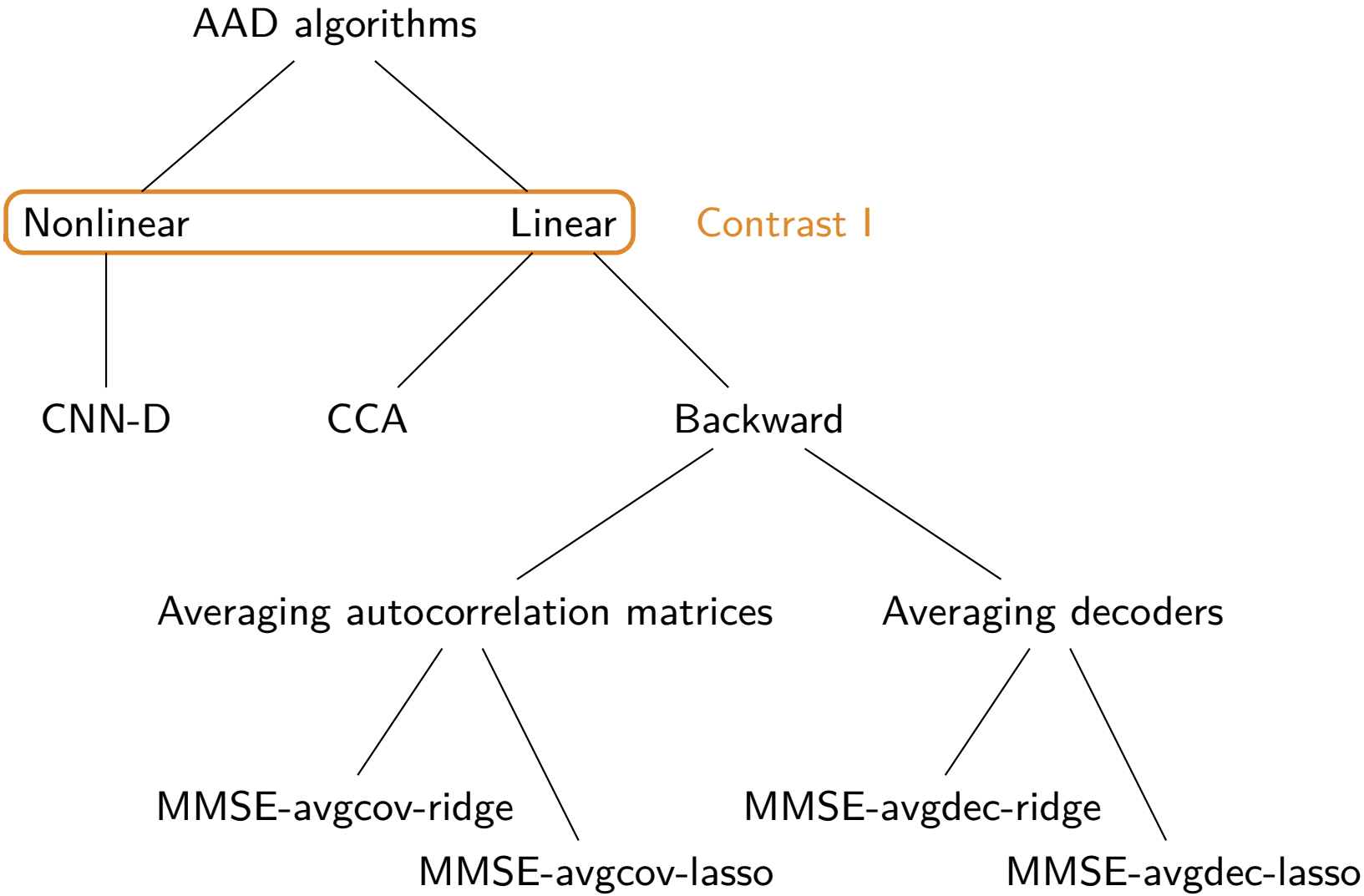
Disclaimer: these are preliminary results (some algorithms still missing)

Statistical analysis on MESD: linear mixed-effects model



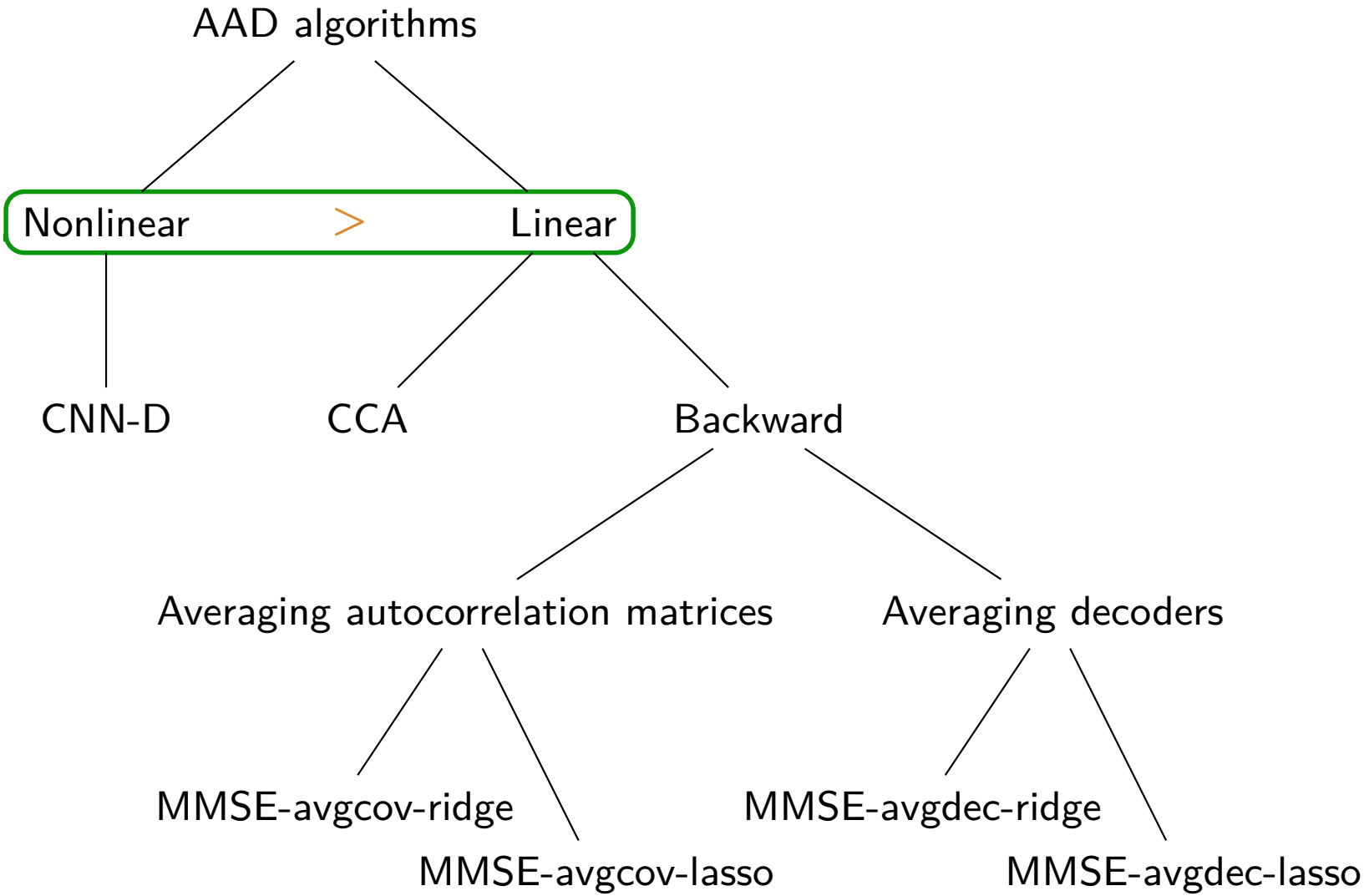
Only on AADKUL-2015 for the moment (two outlying subjects removed)
Disclaimer: these are preliminary results (some algorithms still missing)

Statistical analysis on MESD: linear mixed-effects model



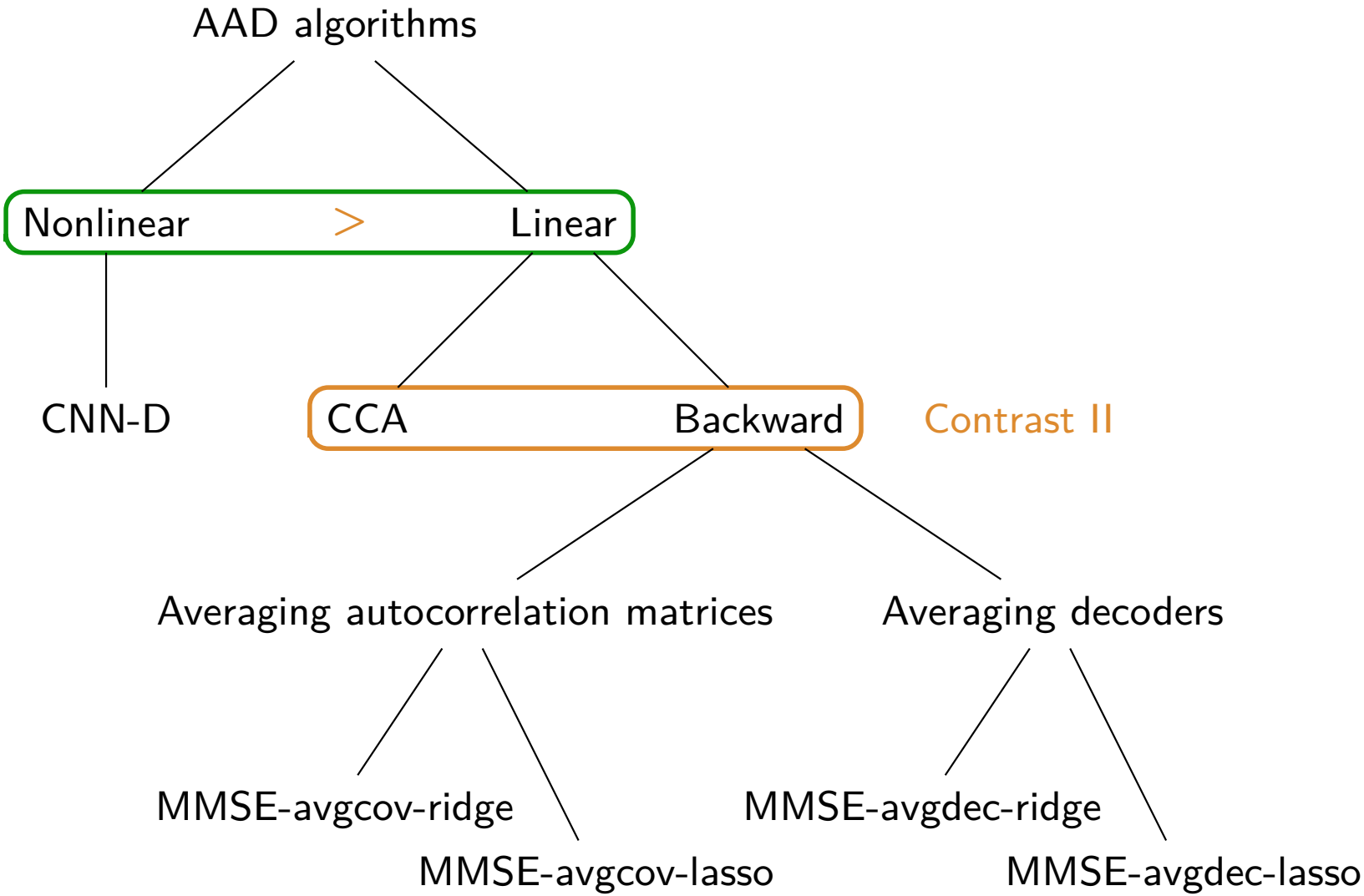
Only on AADKUL-2015 for the moment (two outlying subjects removed)
Disclaimer: these are preliminary results (some algorithms still missing)

Statistical analysis on MESD: linear mixed-effects model



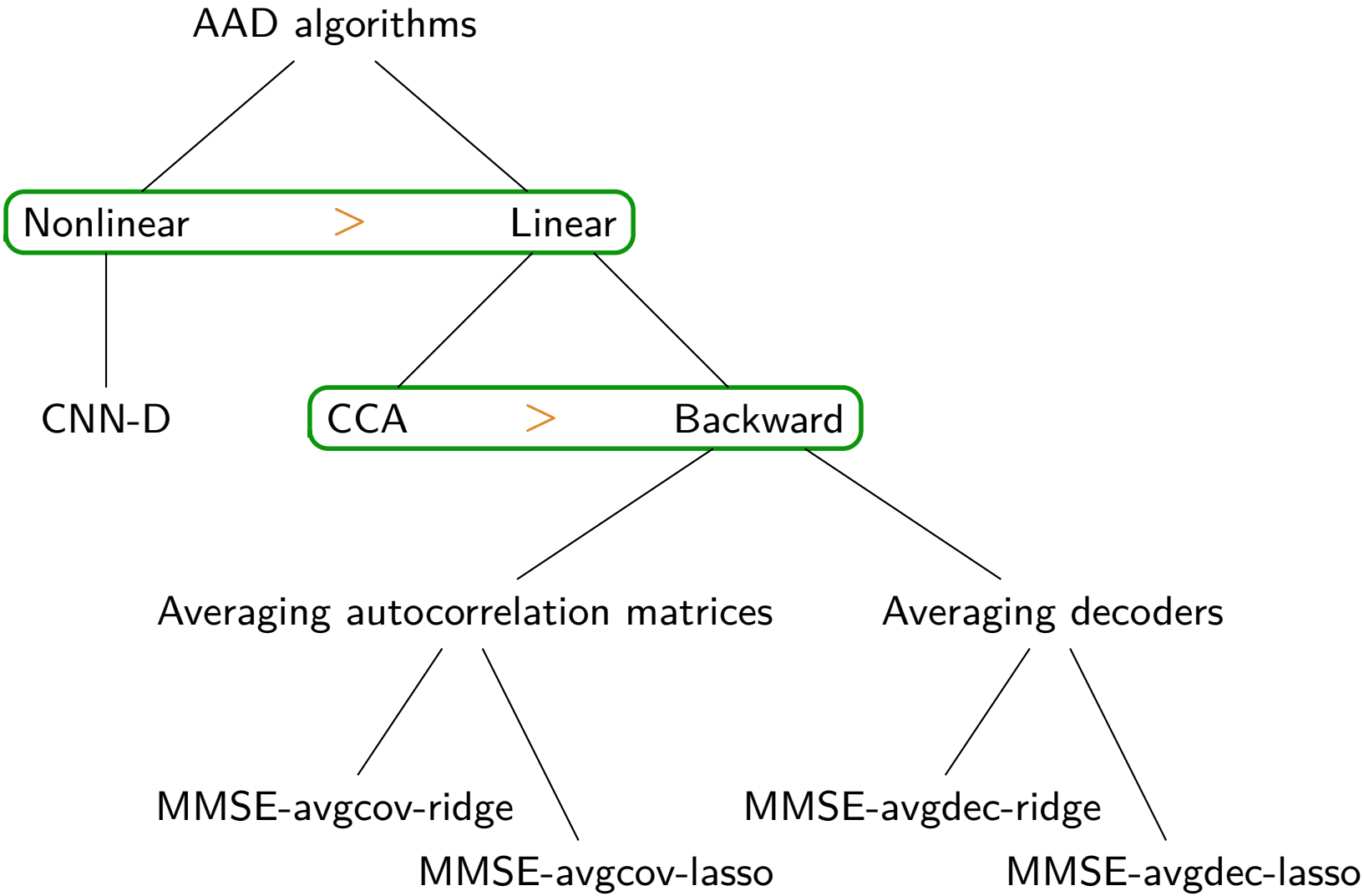
Only on AADKUL-2015 for the moment (two outlying subjects removed)
Disclaimer: these are preliminary results (some algorithms still missing)

Statistical analysis on MESD: linear mixed-effects model



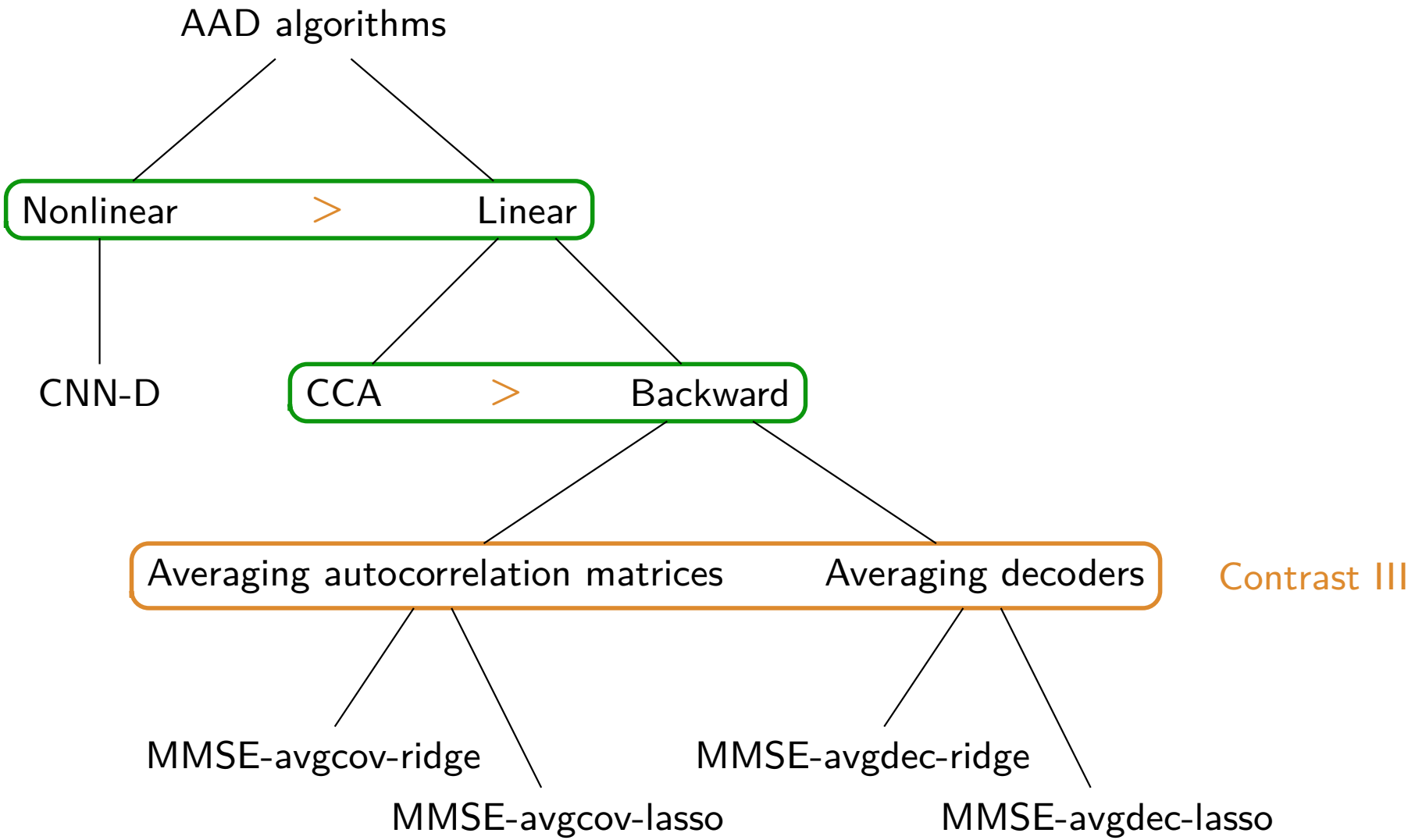
Only on AADKUL-2015 for the moment (two outlying subjects removed)
Disclaimer: these are preliminary results (some algorithms still missing)

Statistical analysis on MESD: linear mixed-effects model



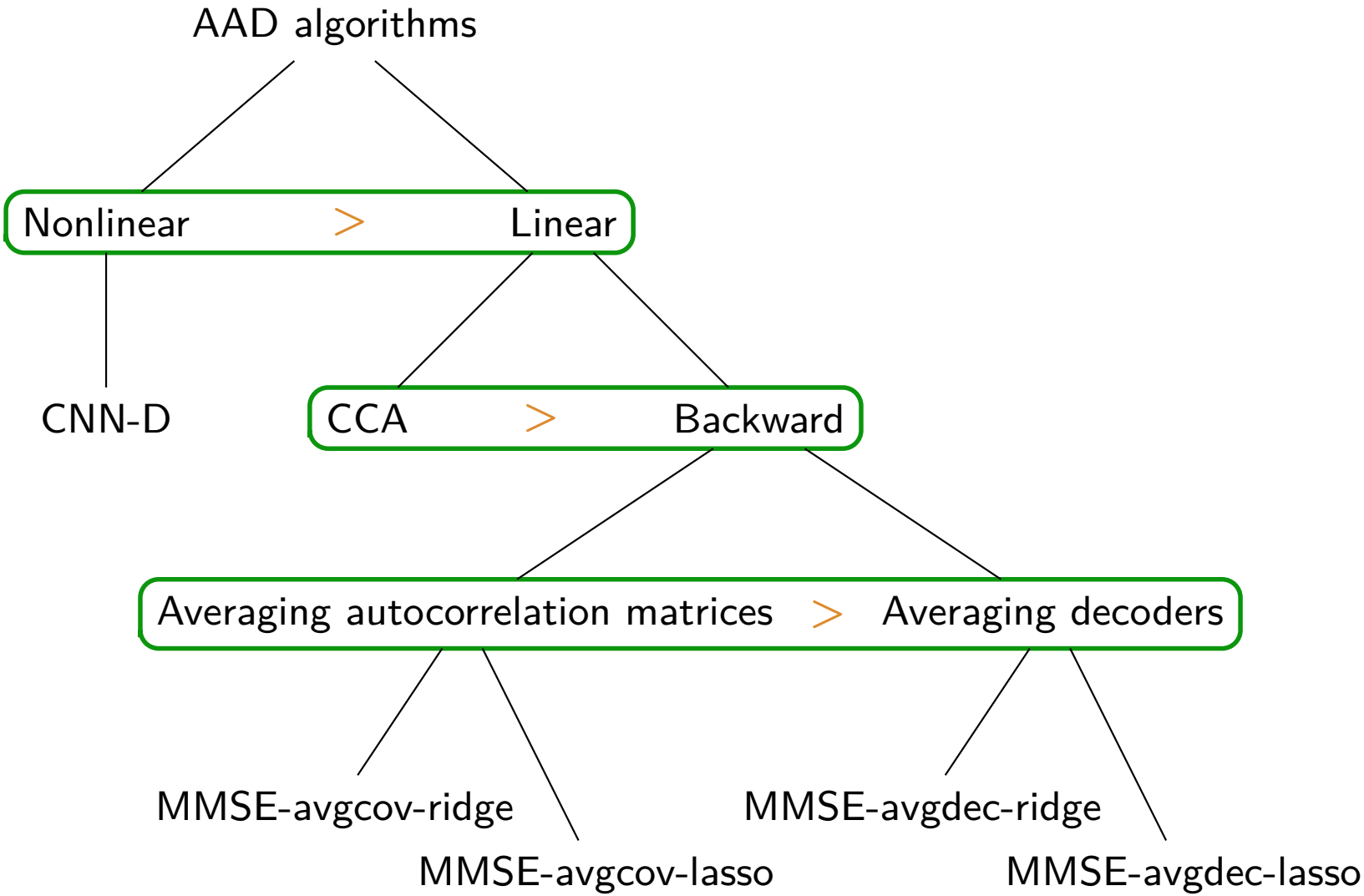
Only on AADKUL-2015 for the moment (two outlying subjects removed)
Disclaimer: these are preliminary results (some algorithms still missing)

Statistical analysis on MESD: linear mixed-effects model



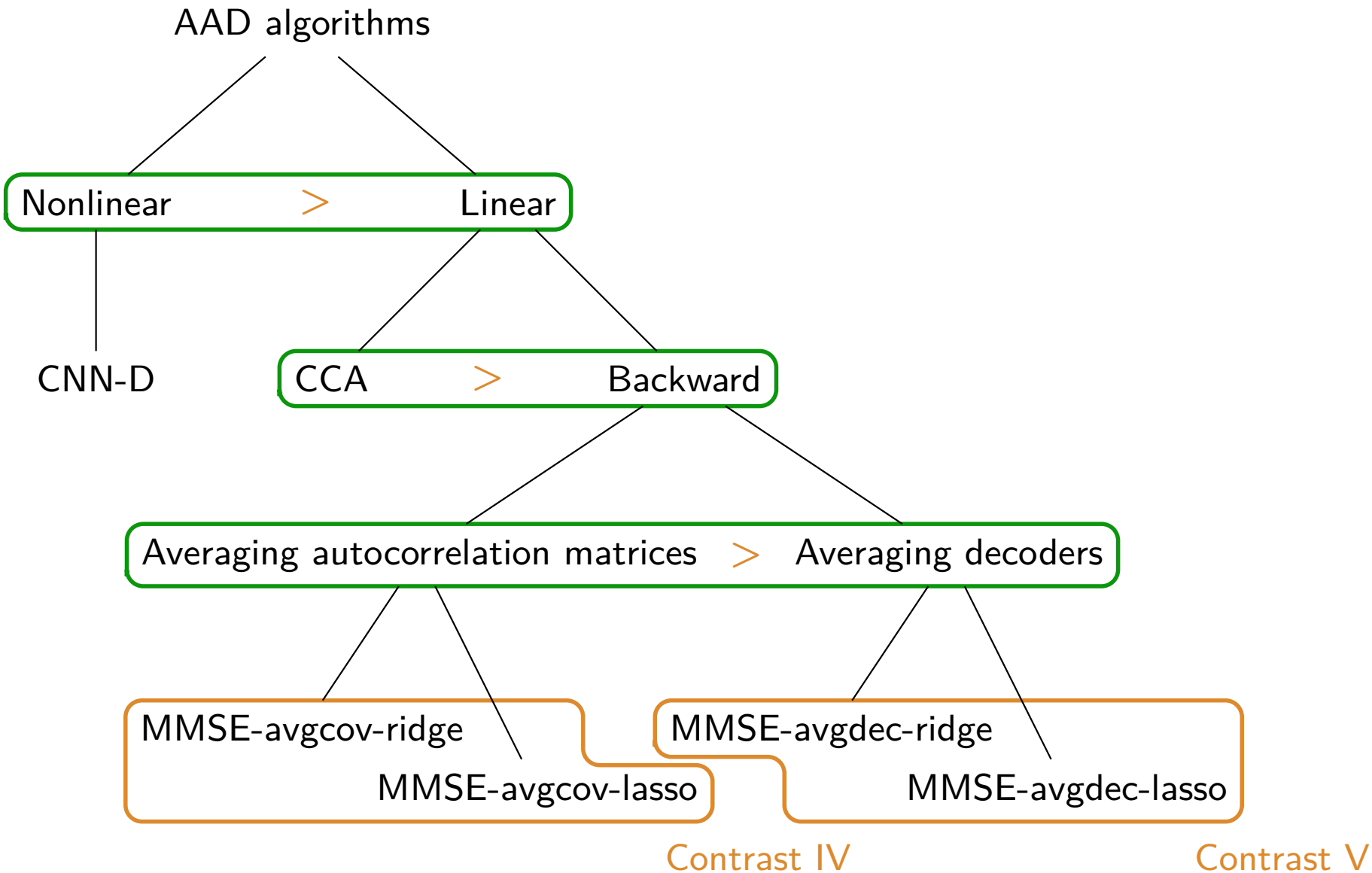
Only on AADKUL-2015 for the moment (two outlying subjects removed)
Disclaimer: these are preliminary results (some algorithms still missing)

Statistical analysis on MESD: linear mixed-effects model



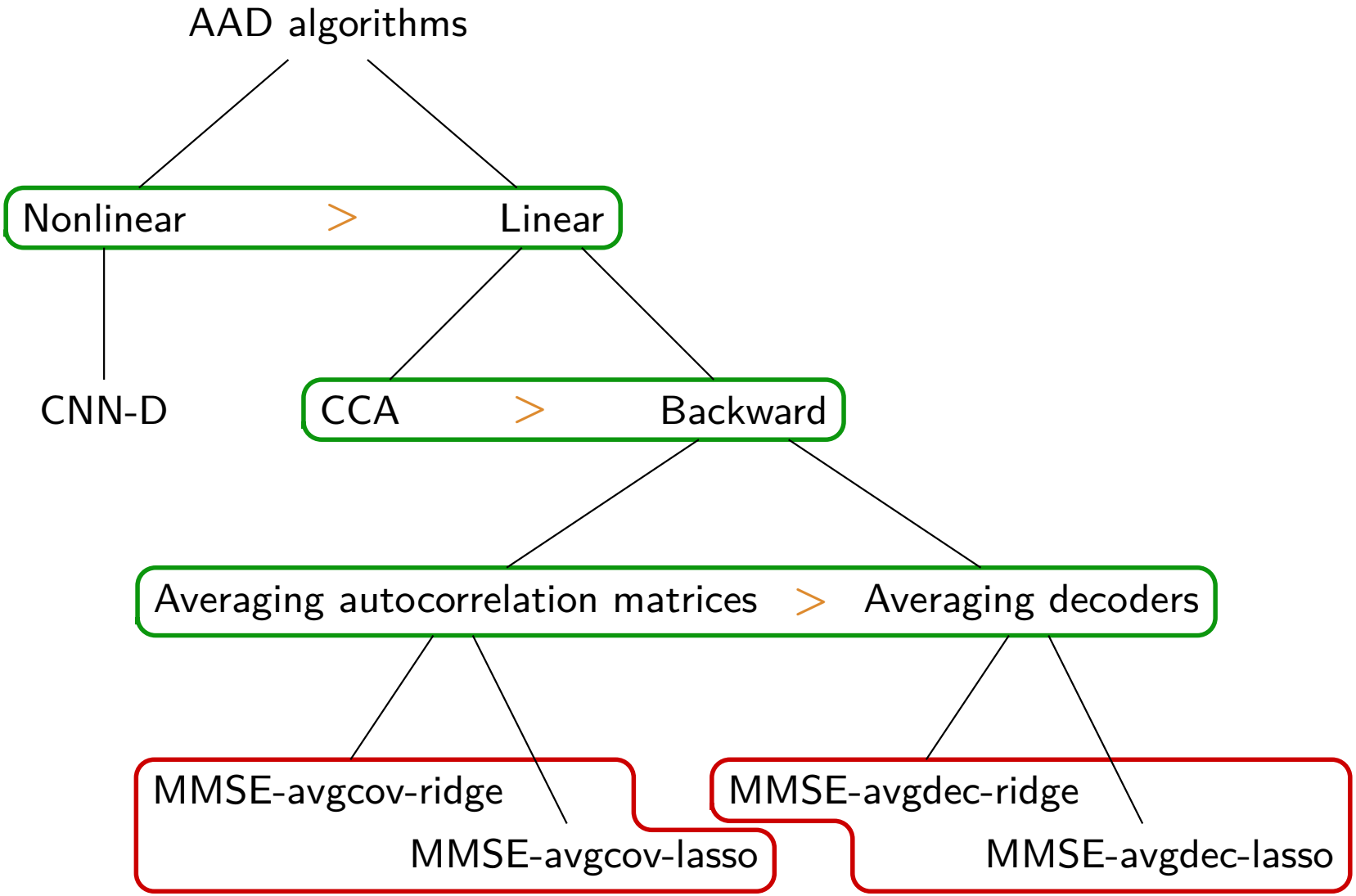
Only on AADKUL-2015 for the moment (two outlying subjects removed)
Disclaimer: these are preliminary results (some algorithms still missing)

Statistical analysis on MESD: linear mixed-effects model



Only on AADKUL-2015 for the moment (two outlying subjects removed)
Disclaimer: these are preliminary results (some algorithms still missing)

Statistical analysis on MESD: linear mixed-effects model



Only on AADKUL-2015 for the moment (two outlying subjects removed)
Disclaimer: these are preliminary results (some algorithms still missing)

Take-home messages

I

Deep learning methods *can* achieve very high performances, but suffer from computational complexity and high variability

II

Within linear methods, a combination of backward and forward modeling works best (CCA)

III

Average covariance matrices, rather than decoders

Attention is vitality. It connects you with others. It makes you eager. Stay eager.

– Susan Sontag

Interested in more? Questions?

simon.geirnaert@esat.kuleuven.be

fwo Research Foundation Flanders
Opening new horizons

erc European Research Council
Established by the European Commission

STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics

KU LEUVEN

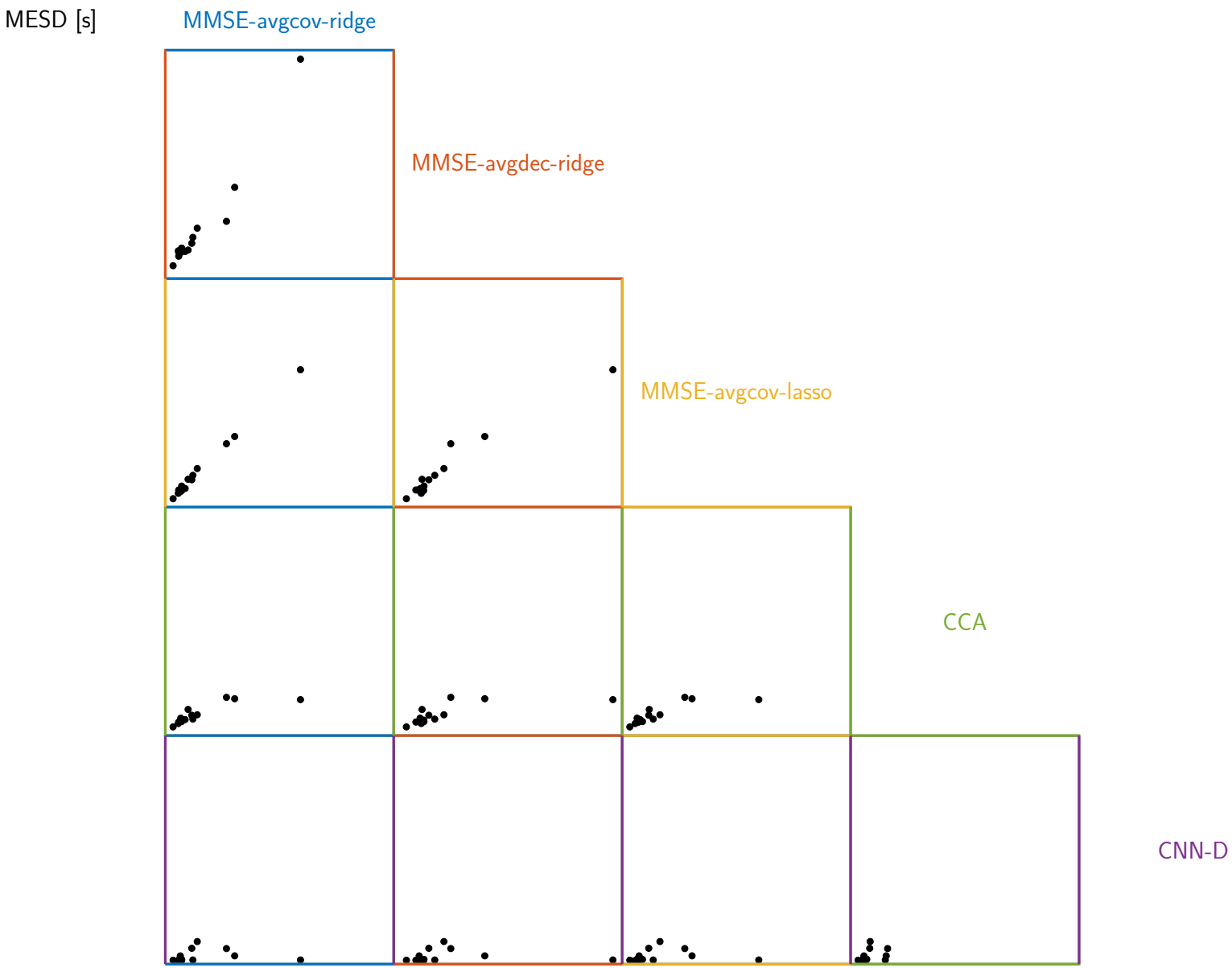
KU LEUVEN expORL

The setup: data, testing procedure

Two independent datasets are used:

AADKUL-2015	Fuglsang-2018
• 16 subjects	• 18 subjects
• 72 min of data per subject	• 50 min of data per subject
• 64-channel Biosemi EEG system	• 64-channel Biosemi EEG system
• Dry and HRTF-filtered stimuli ($-90°$, $+90°$)	• HRTF-filtered stimuli ($-60°$, $+60°$)
	• $\neq$ acoustic room conditions

Comparing the per-subject MESDs: correlations



Disclaimer: these are preliminary results (some algorithms still missing)