PADERBORN
UNIVERSITY

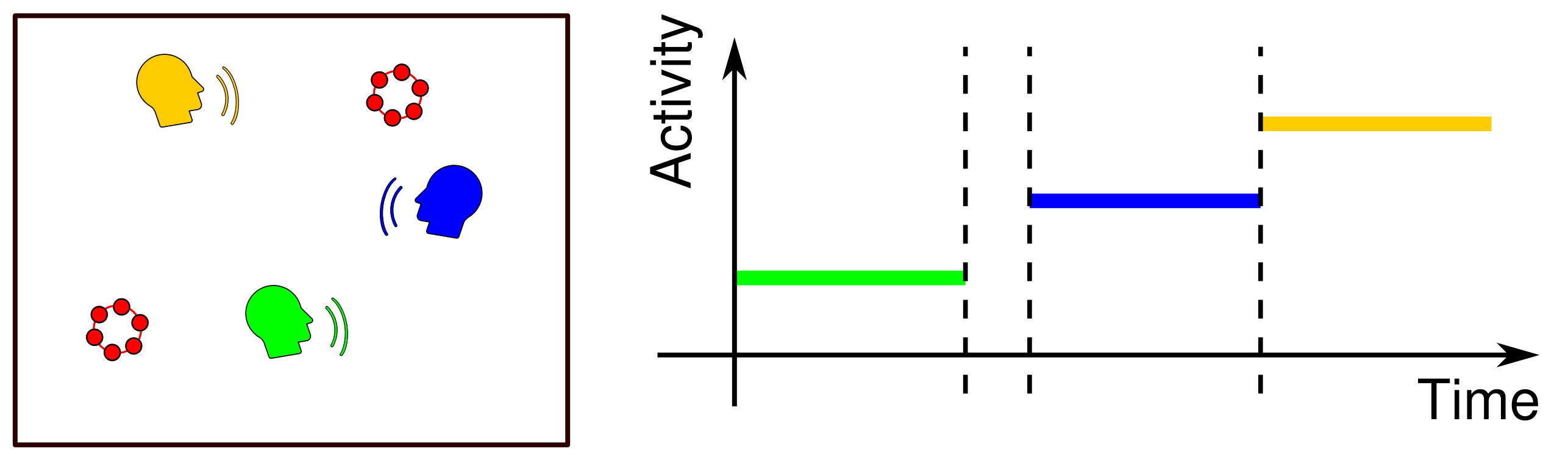
Computer Science, Electrical
Engineering and Mathematics
Communications Engineering
Prof. Dr.-Ing. Reinhold Häb-Umbach

[4381]

# On Synchronization of Wireless Acoustic Sensor Networks in the Presence of Time-varying Sampling Rate Offsets and Speaker Changes

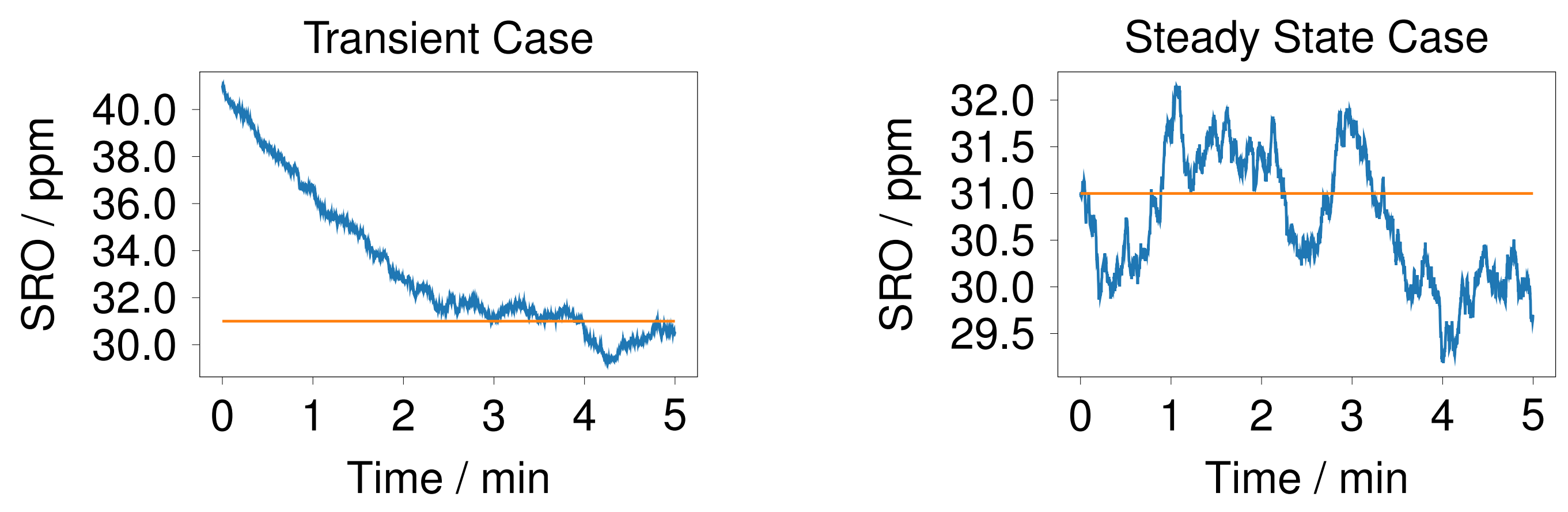
Tobias Gburrek, Joerg Schmalenstroeer, Reinhold Haeb-Umbach
{gburrek, schmalen, haeb}@nt.uni-paderborn.de

## Scenario



- Independent analog-digital converters (ADCs) ⇒ Sampling rate offset (SRO) & Sampling time offset (STO)
- Previous works: Constant SRO and fixed source positions
- Here: Time-varying SRO and source position changes

## Time-varying SRO



- Transient case: Temperature changes, . . .
- Steady state fluctuations: Supply voltage changes, . . .
- Simulation via an Ornstein-Uhlenbeck process:

$$\varepsilon[\ell]=\varepsilon[\ell-1]+\theta\cdot(\mu_\infty-\varepsilon[\ell-1])+x_\varepsilon[\ell],\ \ x_\varepsilon[\ell]\sim\mathcal{N}(0;\sigma^2_{\mathrm{OU}})$$

## Coherence Drift

- Complex-conjugated product of consecutive coherence functions: $P_\Gamma(\ell,k)=\Gamma_{12}(\ell,k)\cdot\Gamma^*_{12}(\ell-\ell_d,k)$
- Phase of coherence prodruct $\angle P_\Gamma(\ell,k)\propto\bar{\varepsilon}_{12}[\ell]$ if
  - Source position is the same in both frames
  - SRO is close to zero

## Dynamic Weighted Average Coherence Drift (DWACD)

- Adaptation of online WACD [1]:
  - $P_{\mathrm{WACD}}(\ell,k)=\alpha\cdot P_{\mathrm{WACD}}(\ell-1,k)+(1-\alpha)\cdot P_\Gamma(\ell,k)$
  - Reduce frame length used for coherence estimation
  - Reduce temporal distance $\ell_d$
  - Estimate SRO from $P_{\mathrm{WACD}}(\ell,k)$:

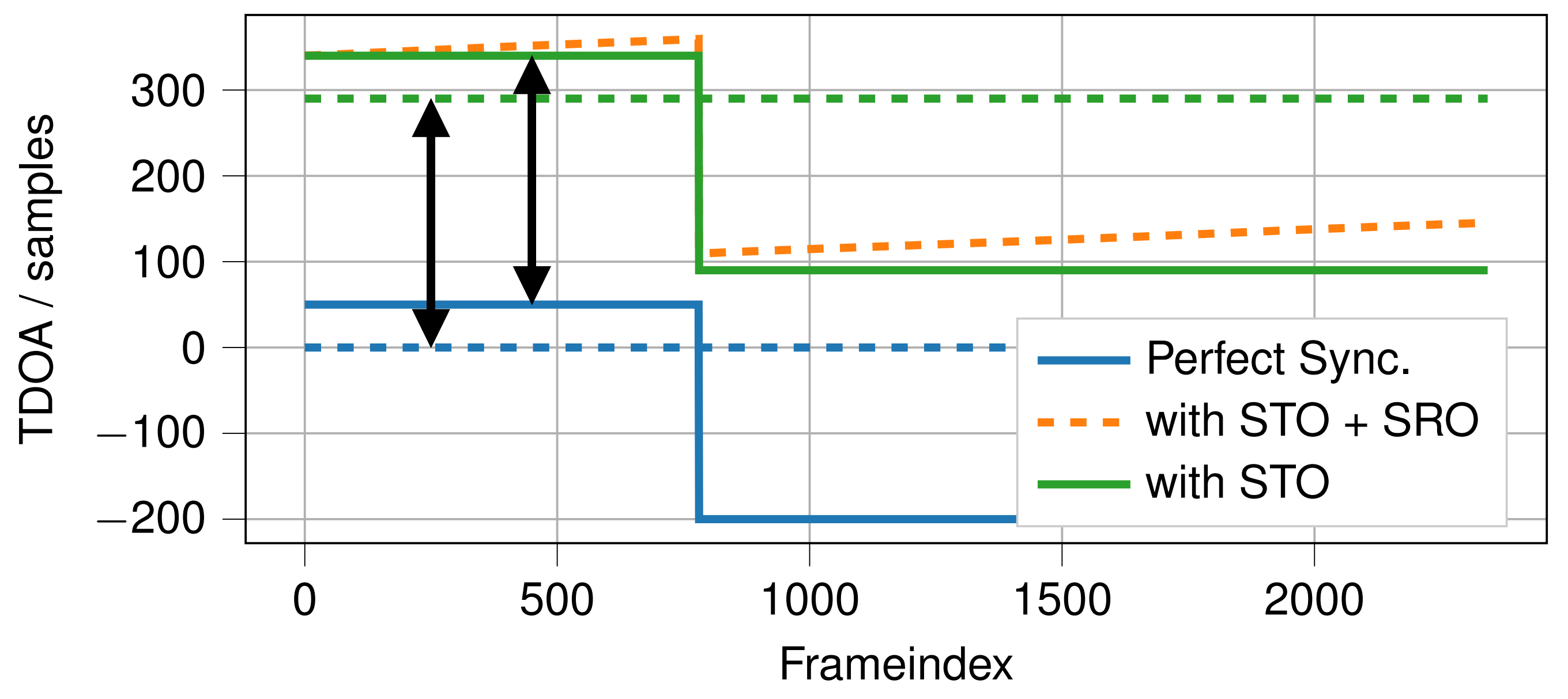
$$P_{\mathrm{WACD}}(\ell,k)\approx W(\ell,k)\cdot\exp\left(j\frac{2\pi k}{N}\ell_d B_s\bar{\varepsilon}_{12}[\ell]\right)$$

$$\widehat{\varepsilon}_{12}[\ell]=-\frac{1}{\ell_d B_s}\cdot\underset{\lambda}{\operatorname{argmax}}\,|p_{\mathrm{WACD}}(\ell,\lambda)|$$

$$\text{with } p_{\mathrm{WACD}}(\ell,\lambda)=\mathrm{IFFT}\{P_{\mathrm{WACD}}(\ell,k)\}$$

  - Use current SRO estimate to resample the next signal frame for coherence estimation

## STO Estimation



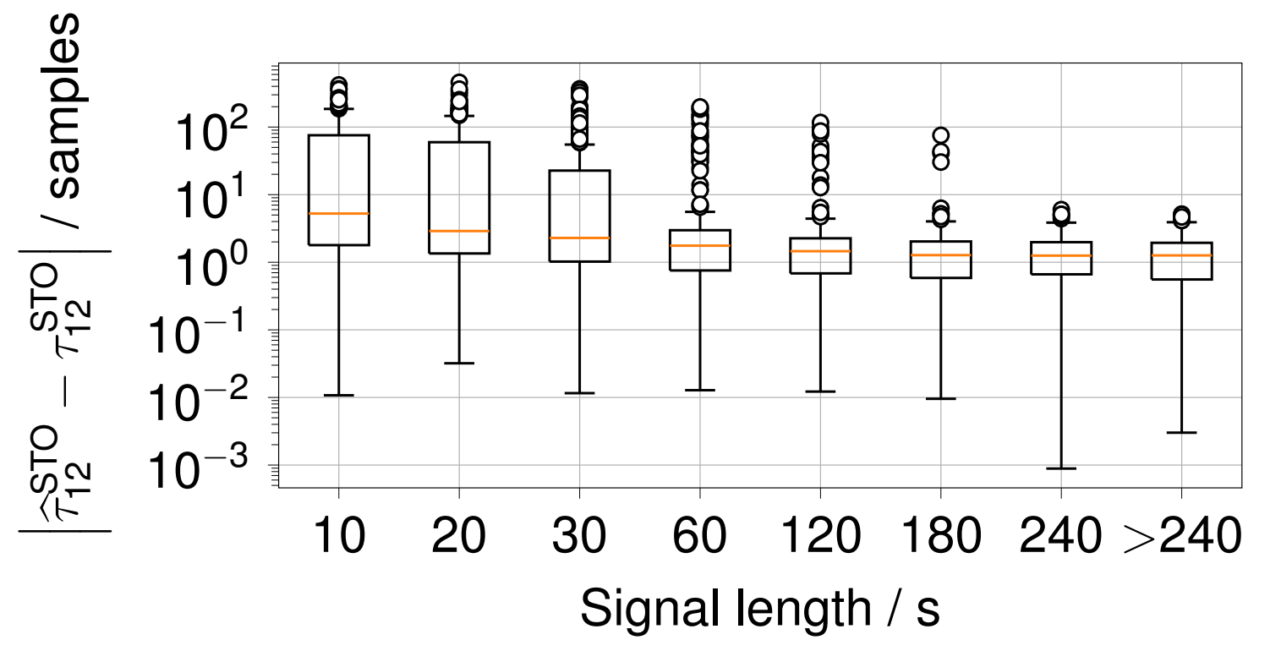
- Goal: Compensate for STO such that synchronized signals correctly represent the time of flight (TOF) from source position to microphone position
- Principle: Align time difference of arrival (TDOA) with time difference of flight (TDOF)
  - TDOA: $\tau_{12}[\ell]=(d_2[\ell]-d_1[\ell])/c\cdot f_s-(T_2-T_1)\cdot f_s$
  - Estimated source-microphone distances [2] → TDOF
- STO estimation via a least squares problem

## Experiments

| | $\varepsilon\neq$ const. | Multi-Pos. | Silence | Method | avg. RMSE of SRO / ppm | avg. RMSE of SRO shift / samples | max. RMSE of SRO shift / samples |
|---|---|---|---|---|---|---|---|
| Setup-1 | | | | Online WACD | 0.21 | 0.14 | 0.50 |
| | | | | DXCP-PhaT | 0.15 | 0.36 | 0.68 |
| | | | | DWACD | 0.40 | 0.15 | 0.50 |
| Setup-2 | ✓ | | | Online WACD | 0.63 | 0.73 | 2.09 |
| | ✓ | | | DXCP-PhaT | 0.66 | 0.97 | 2.73 |
| | ✓ | | | DWACD | 0.51 | 0.27 | 1.04 |
| Setup-3 | ✓ | ✓ | ✓ | Online WACD | 2.98 | 6.04 | 21.00 |
| | ✓ | ✓ | ✓ | DXCP-PhaT | 28.96 | 21.84 | 161.54 |
| | ✓ | ✓ | ✓ | DXCP-PhaT$_8$ | 1.31 | 2.70 | 7.76 |
| | ✓ | ✓ | ✓ | DWACD | 0.57 | 0.32 | 1.20 |
| Setup-4 | ✓ | ✓ | | Online WACD | 2.80 | 3.25 | 10.96 |
| | ✓ | ✓ | | DXCP-PhaT | 22.42 | 16.61 | 160.49 |
| | ✓ | ✓ | | DXCP-PhaT$_8$ | 1.28 | 2.81 | 6.93 |
| | ✓ | ✓ | | DWACD | 0.64 | 0.32 | 1.10 |

Average RMSE of the SRO in ppm

| $\sigma_{\mathrm{SRO}}$ / ppm | Online WACD | DXCP-PhaT | DWACD |
|---|---|---|---|
| 0 - 1 | 0.55 | 0.54 | 0.49 |
| 1 - 2 | 0.60 | 0.61 | 0.51 |
| 2 - 3 | 0.63 | 0.69 | 0.51 |
| 3 - 4 | 0.71 | 0.80 | 0.54 |



## Summary

- Time-varying SROs and source position changes
- SRO estimation via DWACD method
  - Previously proposed online SRO estimators fail to properly handle dynamic scenarios
  - Adaptation of our online WACD method to dynamic scenario
- STO estimation method
  - Synchronization which reflects the physical TDOFs
  - Usage of source-microphone distance estimates as support information

## Open Source Toolbox

`https://github.com/fgnt/paderwasn`

[1] A. Chinaev, G. Enzner, T. Gburrek, and J. Schmalenstroeer, “Online Estimation of Sampling Rate Offsets in Wireless Acoustic Sensor Networks with Packet Loss,” in Proc. EUSIPCO, 2021

[2] T. Gburrek, J. Schmalenstroeer, and R. Haeb-Umbach, “On source-microphone distance estimation using convolutional recurrent neural networks,” in Proc. ITG-Conference, 2021.

ICAASP 2022, Singapore

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project 282835863