KU LEUVEN



Making Sense of Sound

ESAT Inaugural Lecture 2014

Prof. dr. ir. Toon van Waterschoot

Faculty of Engineering Technology
ESAT – Department of Electrical Engineering
KU Leuven, Belgium

Academic Career









1979: born in Lier, Belgium

1987: started playing music & being fascinated by sound

1996: decided to study Electrical Engineering
(for specializing in audio signal processing)
2001: graduated as Master in Electrical Engineering
2002: started teaching at Antwerp Maritime Academy
2003: started PhD at KU Leuven (Prof. Marc Moonen)
2009: graduated as PhD in Electrical Engineering
2010: moved to TU Delft as postdoc (Prof. Geert Leus)
2011: returned to KU Leuven as FWO postdoc
(Prof. Marc Moonen & Prof. Moritz Diehl)
2013: appointed as Assistant Professor at KU Leuven

KU LEUVEN

Affiliations



- **KU Leuven, Belgium:** Assistant Professor
  - **STADIUS** Center for Dynamical Systems, Signal Processing, and Data Analytics
  - **AdvISe** – Advanced Integrated Sensing Lab (KU Leuven Campus Geel)
- **TU Delft, The Netherlands:** Visiting Researcher
  - Circuits & Systems Group
- **University of Lugano, Switzerland:** Visiting Lecturer

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Advanced Integrated Sensing
AdvISe

TU Delft



KU LEUVEN

Teaching

- **KU Leuven, Faculty of Engineering Technology**
  - Control Theory
  - Digital Signal Processing
  - Databases & Web Technology (2014-2015)



- **KU Leuven, Faculty of Engineering Science**
  - Optimization (2014-2015)



- **University of Lugano, Faculty of Informatics**
  - Digital Signal Processing
    (with Prof. M. Moonen & Prof. S. Holdt Jensen)



KU LEUVEN

Research Projects

- FWO Postdoctoral Research Fellowship
  A Signal-Oriented Approach to Acoustic Signal Enhancement
- FP7 Marie Curie Initial Training Network “DREAMS”
  Dereverberation & Reverberation of Audio, Music, and Speech
- IWT O&O Project “Cochlear-IV”
  Signal Transmission Schemes for Auditory Implants
- KU Leuven Programme Financing “OPTEC”
  Optimization in Engineering Center
- IWT Strategic Basic Research Project “ALADIN”
  Adaptation & Learning for Assistive Domestic Vocal Interfaces
- IWT Strategic Basic Research Project “SINS”
  Sound Interfacing Through the Swarm

fwo

DREAMS

Cochlear™

OPTEC
K.U.Leuven Optimization in Engineering Center

aladi

INS

KU LEUVEN

Research Team



Enzo
De Sena

Postdoc
DREAMS



Niccoló
Antonello

PhD Student
DREAMS



Pablo
Peso

PhD Student
Nuance/Imperial
DREAMS



Giuliano
Bernardi

PhD Student
Cochlear-IV



Gert
Dekkers

PhD Student
ALADIN/SINS



Ante
Jukić

PhD Student
Univ. of Oldenburg
DREAMS



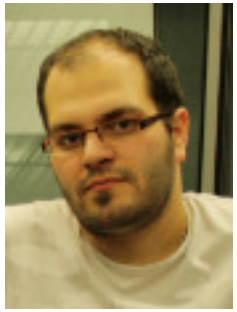
Giacomo
Vairetti

PhD Student
DREAMS



Mina
Shehata

PhD Student
ALADIN/SINS



Neo
Kaplanis

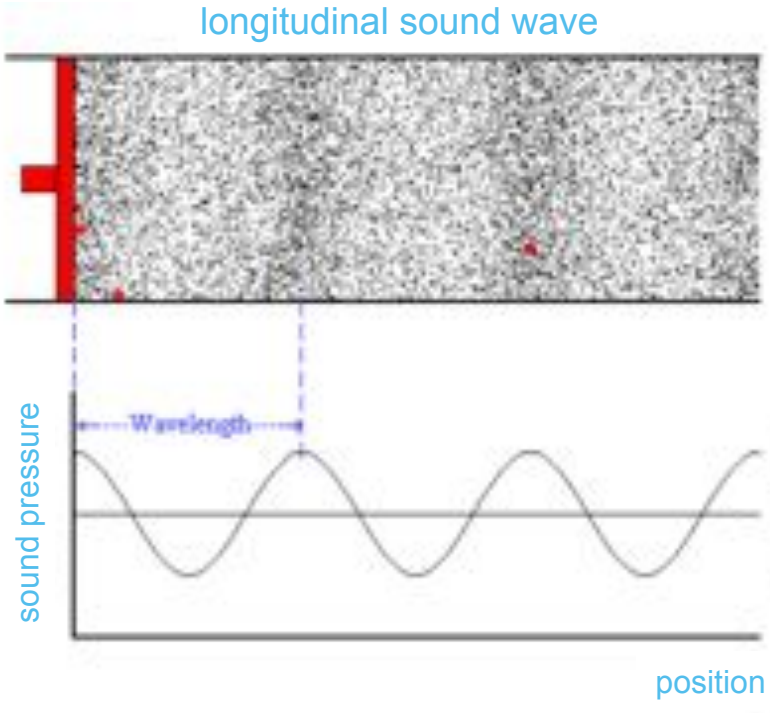
PhD Student
B&O/Aalborg Univ.
DREAMS

KU LEUVEN

Sound and Electrical Engineering

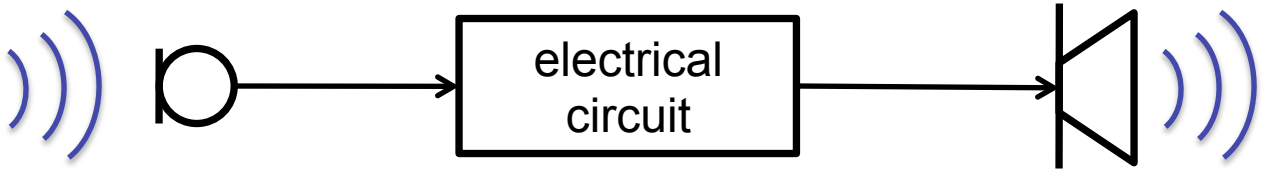
- Sound = pressure wave propagating in medium (air)
- Pressure wave can be converted into electrical voltage and vice versa
  - microphone
  - loudspeaker





© Institute for Sound and Vibration Research, Southampton University

- Audio signal processing:

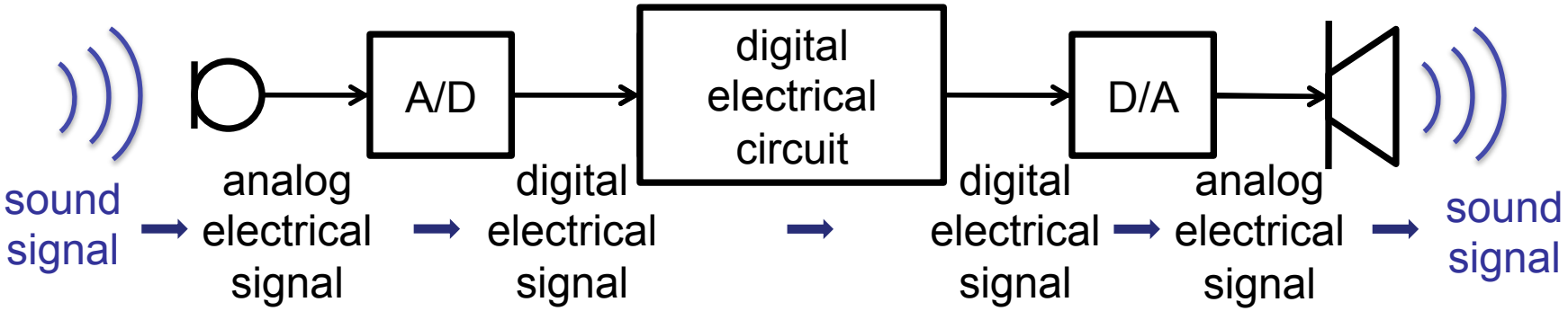


sound signal → electrical signal → electrical signal → sound signal

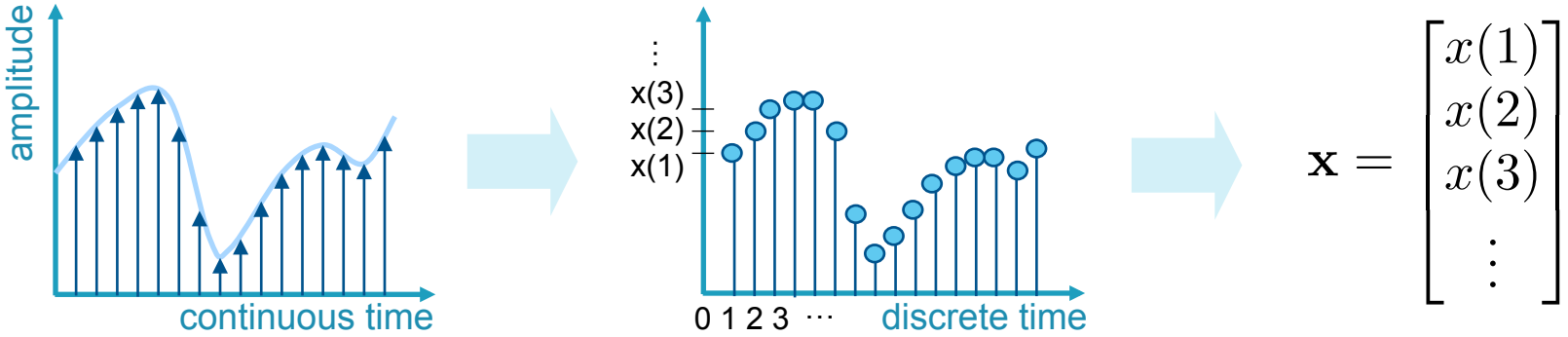
KU LEUVEN

Digital Signal Processing

- Digital audio signal processing:



  - analog-to-digital (A/D) conversion: sampling + quantization

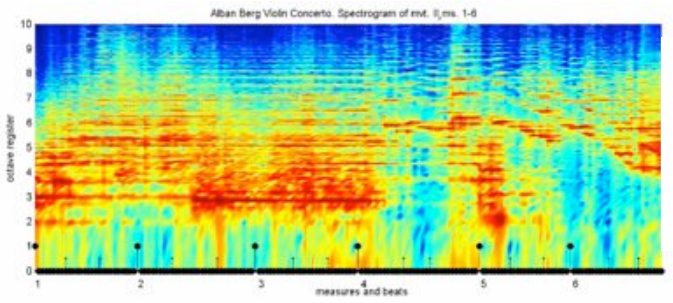


  - algorithm = analysis/manipulation of sound by means of mathematical operations (+−×÷) & matrix algebra

Challenges in Digital Audio Processing

- Modeling:
  - sound = natural/physical phenomenon
  - no exact/unique signal models available
- Processing:
  - sound = broadband signal
  - high sampling frequency
  - high computational processing load
- Evaluation:
  - what really matters is: “How does it sound?”
  - subjective measure that is hard to objectify
  - listening tests: time-consuming & expensive



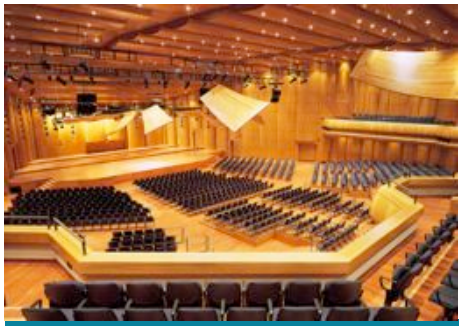


© Diego Banuelos



© IELTS

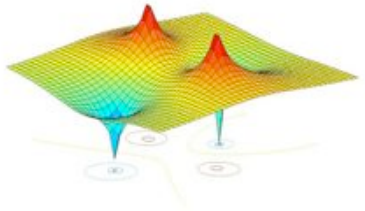
KU LEUVEN



Room Acoustics



Psycho-acoustics



Numerical Optimization



Signal processing

Making Sense of Sound



Machine Learning

Audio research is by definition multi-disciplinary research!

KU LEUVEN

Making Sense of Sound

- How to improve the perceived quality of sound?

  Audio signal enhancement

- How to model the interaction between sound and environment?

  Acoustic modeling

- How to extract information from sound?

  Audio analysis

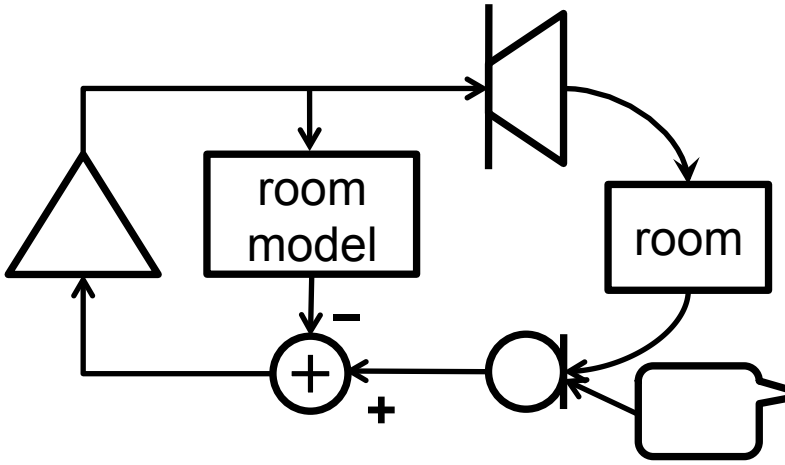
- How to create an immersive and healthy sound experience?

  Audio reproduction

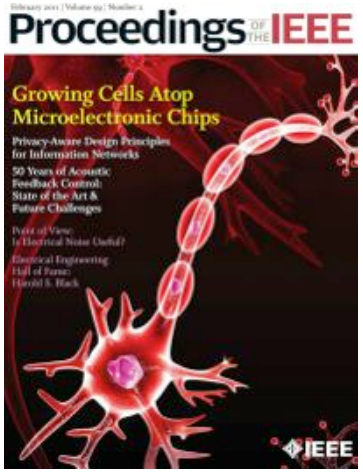


Acoustic feedback control

- Acoustic feedback control:
  - acoustic feedback from loudspeaker to microphone → howling
  - adaptive feedback canceller estimates acoustic room model & removes acoustic feedback without distorting source signal
  - structuring 5 decades of acoustic feedback research

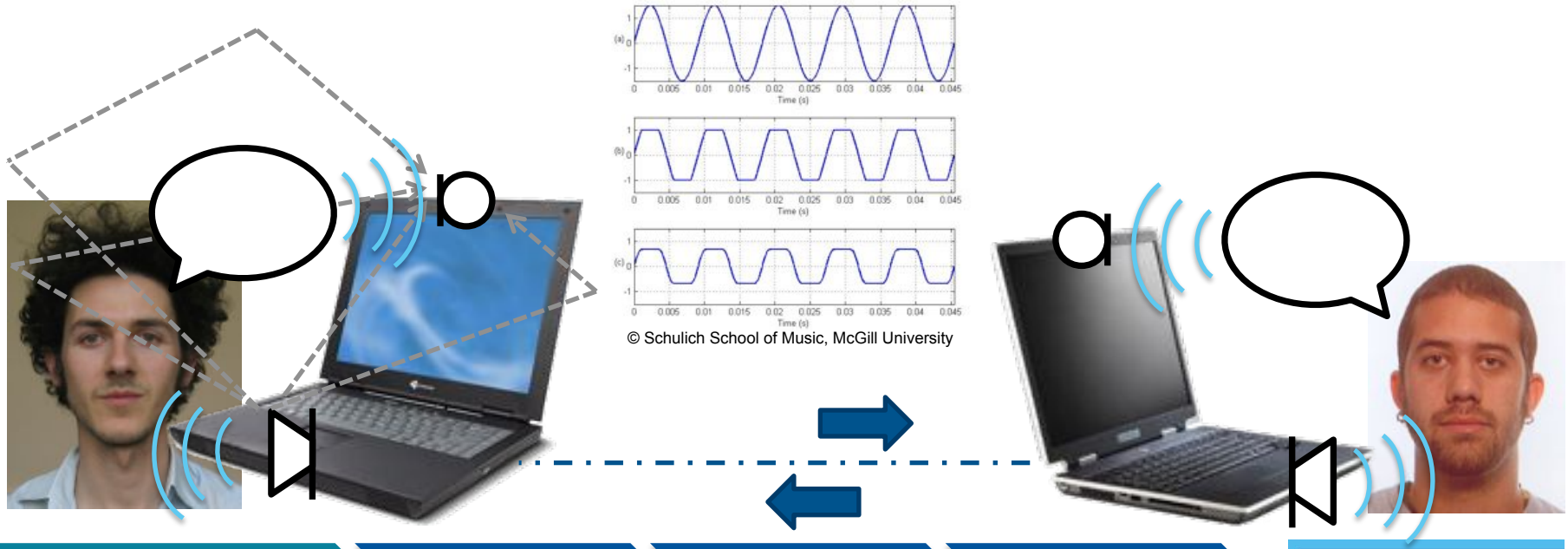






Acoustic echo cancellation

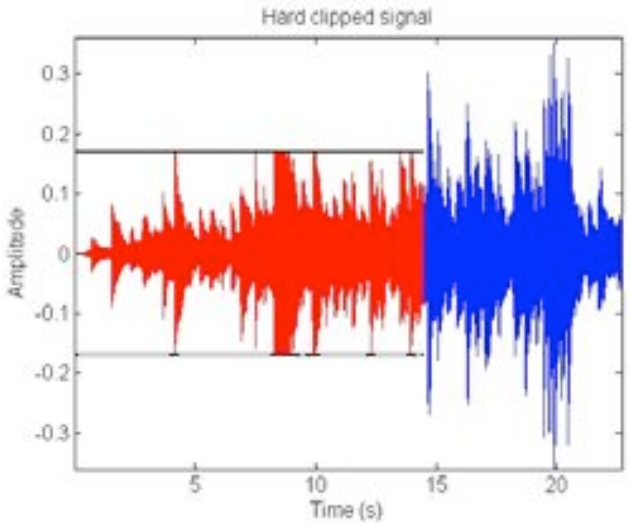
- Acoustic echo cancellation
  - echo removal by adaptive estimation of acoustic room model
  - increased double-talk robustness (when both speakers active)
  - joint estimation of room model & nonlinear loudspeaker model

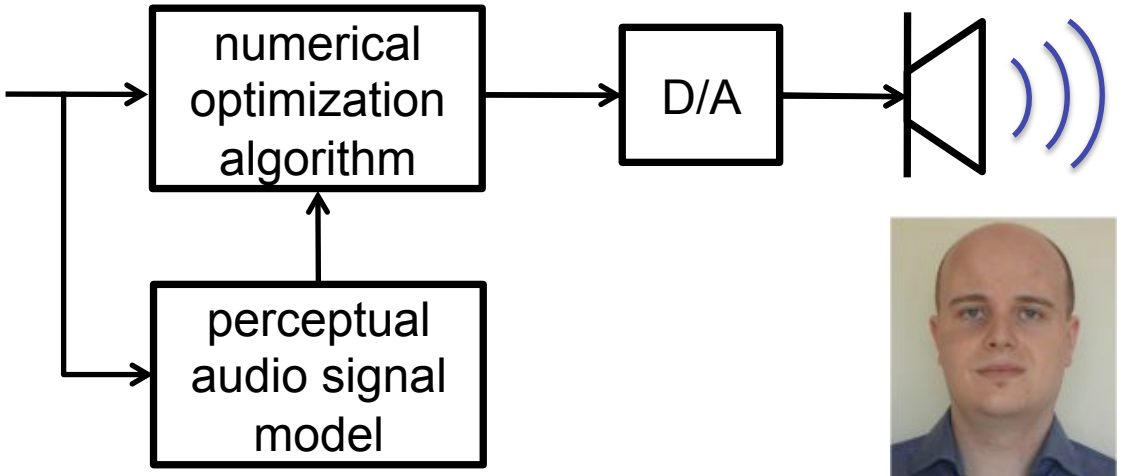


© Schulich School of Music, McGill University

Clipping & loudspeaker compensation

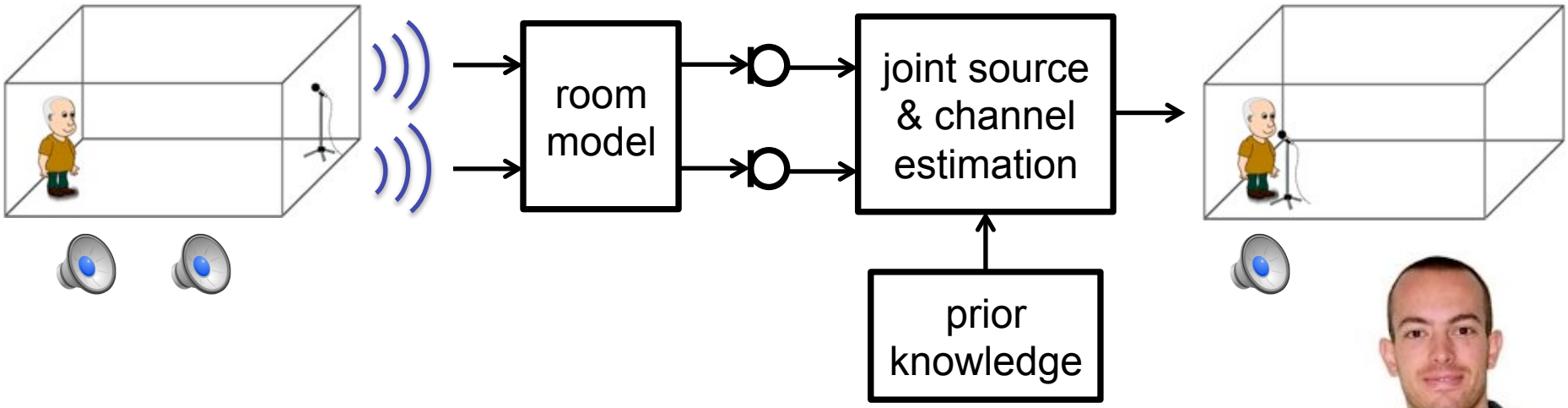
- Clipping & nonlinear loudspeaker distortion
  - analog audio equipment (amplifiers, loudspeakers) as well as digital signal representations have limited dynamic range
  - hard or soft clipping precompensation → signal distortion
  - perceptually optimal clipping & loudspeaker precompensation





Dereverberation

- Multi-channel blind dereverberation
  - far-talk microphones record reverberant instead of “dry” sound
  - dereverberation = removing effect of room acoustics
  - joint source & channel estimation = underdetermined problem

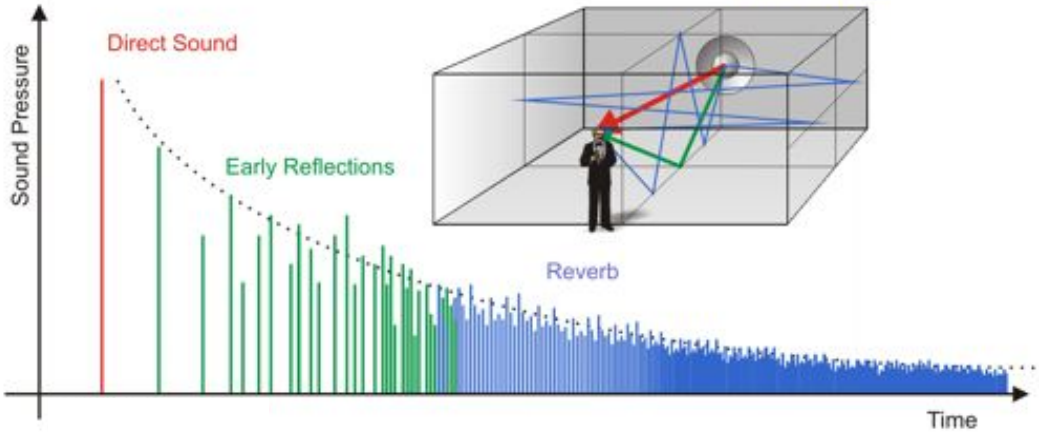


  - prior knowledge: speech model, room decay, …

Room modeling

- Room impulse response (RIR)

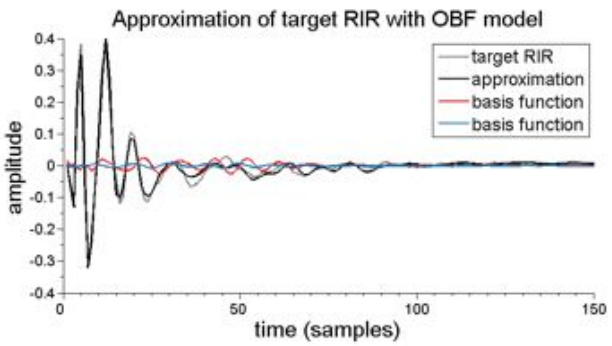


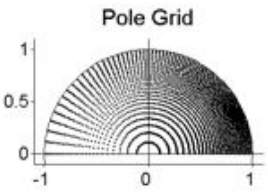


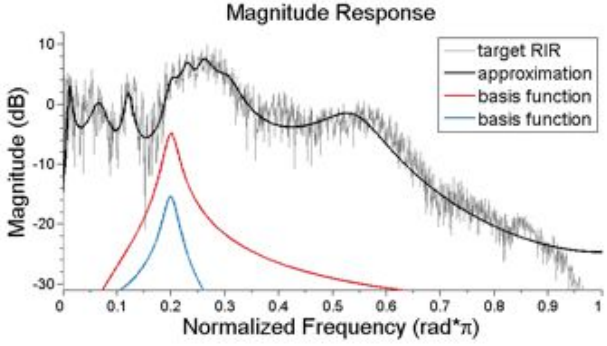
- Problems
  - room impulse responses are very long (10^3–10^4 coefficients)
  - truncation results in poor spectral accuracy
  - high spatial variability
  - inefficient representation of perceptual attributes

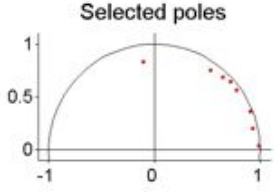
Room modeling: Orthonormal basis functions

- Room modeling using orthonormal basis functions (OBFs)
  - RIR: linear combination of many unit impulse basis functions
  - OBF model: linear combination of few “long” basis functions







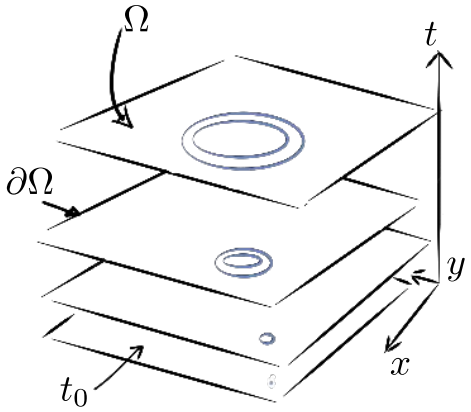


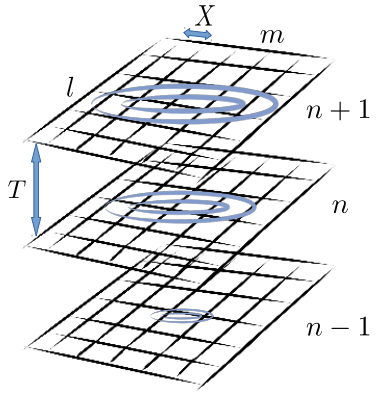
- Scalable OBF model building algorithm:
  - dictionary of OBFs from “overcomplete” pole grid
  - select poles using sparsity-promoting criterion

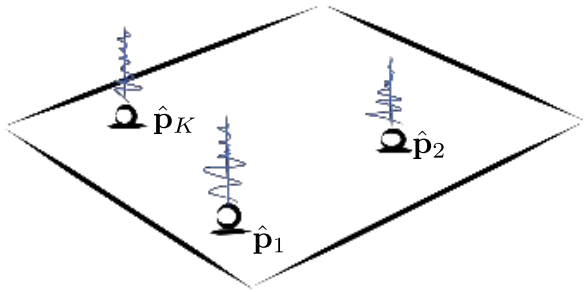


Room modeling: Numerical acoustics

- Room modeling using numerical acoustics
  - modeling of complete space-time acoustic wave field
  - space-time sampling acoustic wave equation (FEM/FDTD)





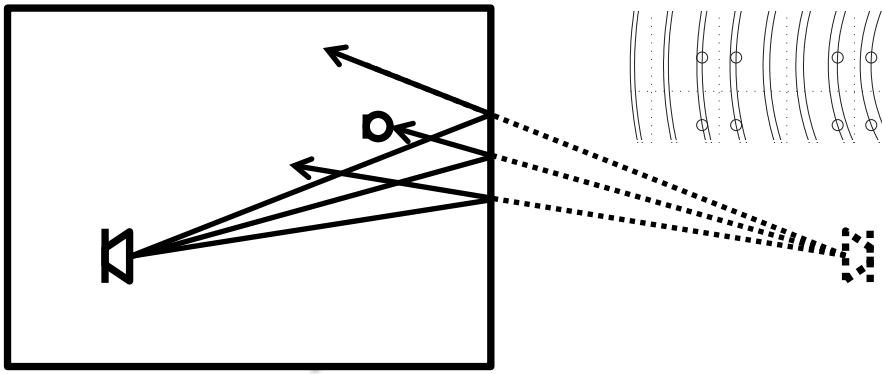


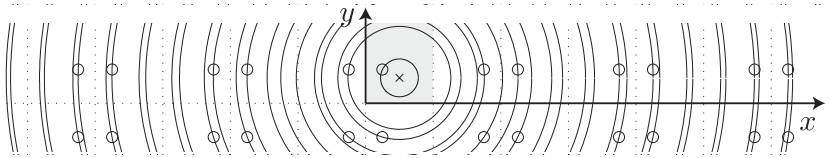
- Wave field sensing
  - use microphone measurements in numerical model
  - source & field estimation using sparse optimization

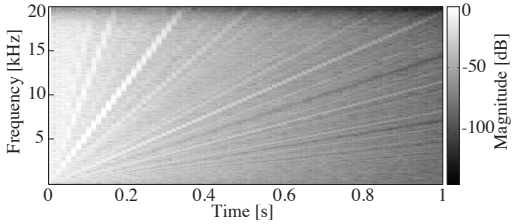


Room modeling: Image method

- Room modeling using image method (1979)
  - room boundary reflections modeled by image sources
- Perceptual performance of image method
  - despite 2000+ citations, perceptual evaluation is lacking
  - time alignment of image sources → “flanging” effect
  - solution: small random displacement of image sources









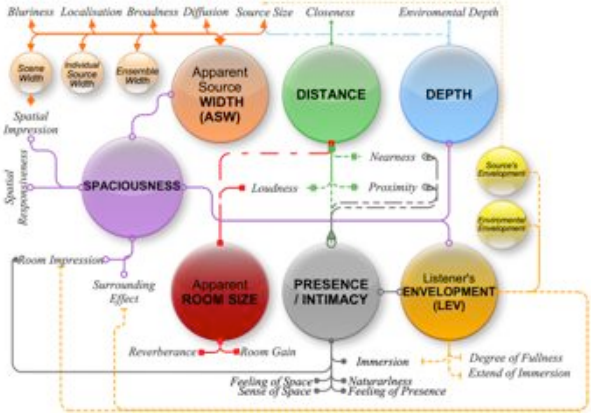
Room modeling: Perceptual models

- Perceptual attributes of room acoustics
  - broad literature on concert halls, few results for small rooms

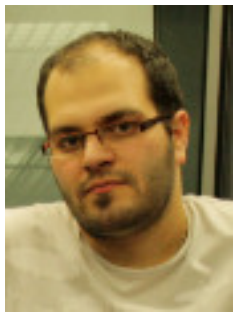






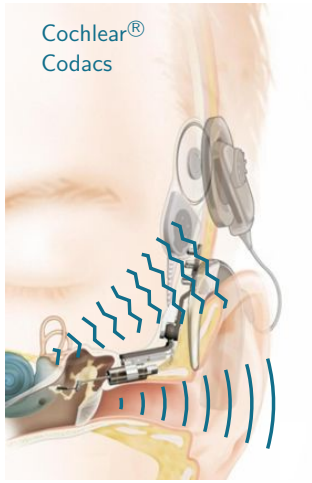


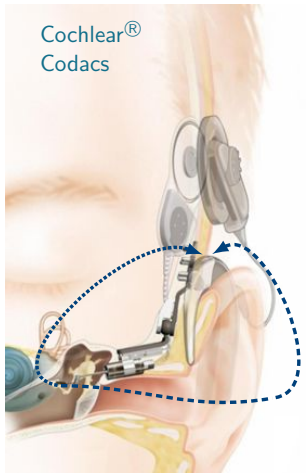
- Perceptual models for small room acoustics
  - identification & classification of perceptual attributes
  - measurement & simulation of small room acoustics
  - subjective listening tests link physics and perception

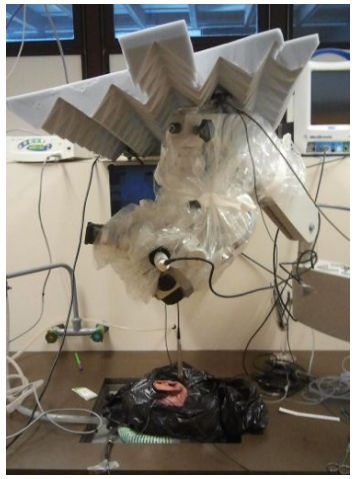


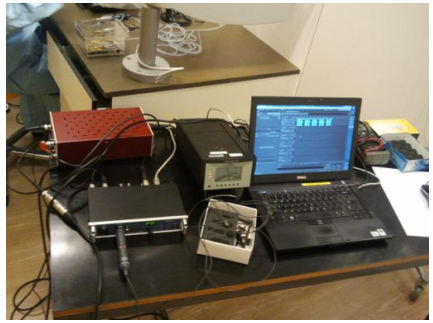
Ear modeling

- Ear modeling
  - acoustic/mechanic sound propagation from inner to outer ear
  - cause of acoustic feedback in acoustic hearing implants







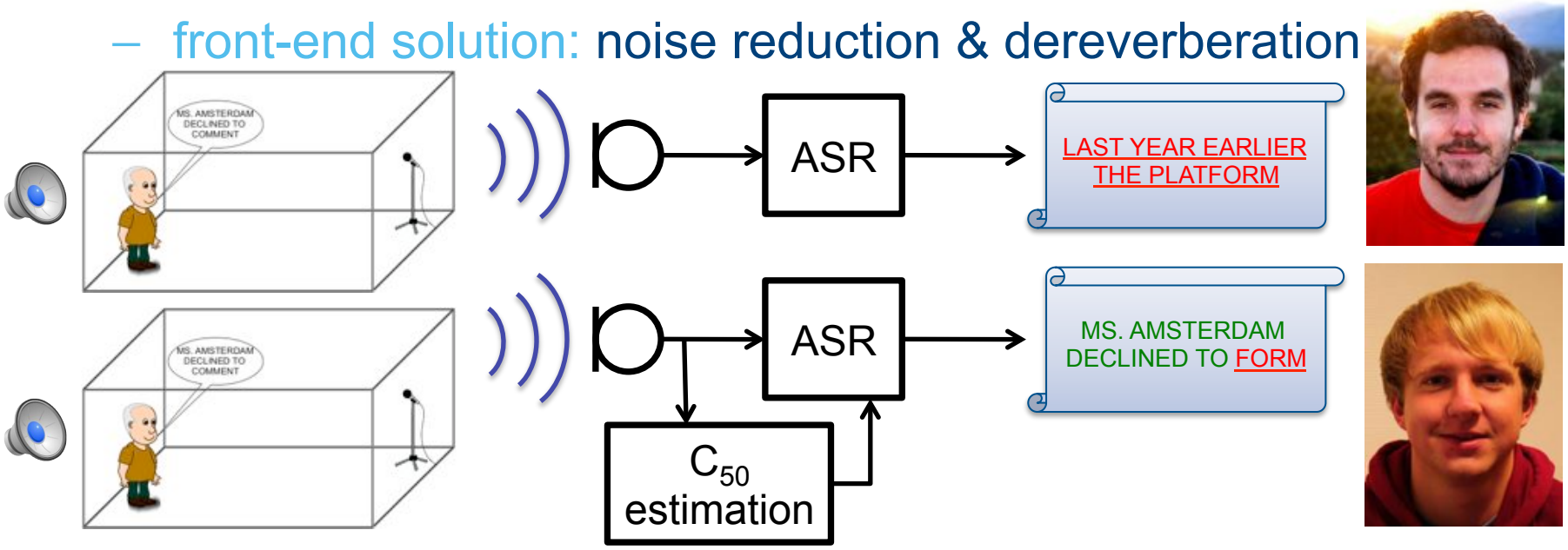


- How to characterize inner-to-outer-ear response?
  - in-vitro measurements on cadaver heads



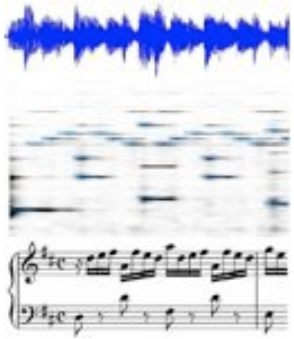
Robust speech recognition

- Automatic speech recognition (ASR)
  - applications: speech-to-text, voice control, surveillance, …
  - standard ASR fails in background noise & reverberation
  - back-end solution: C_{50}-based acoustic model selection
  - front-end solution: noise reduction & dereverberation



Acoustic event detection

- Acoustic event detection
  - acoustic event: change in physical condition of sound source
  - applications: music transcription, acoustic surveillance, …
- Event detection using sparse signal representations
  - two-class dictionaries with steady-state and transient atoms
  - criteria to search for most efficient (sparse) representation
  - detection based on relative importance of two classes



© Centre for Digital Music, Queen Mary University of London



© SST, Inc.



© Why Cry



Model-based audio analysis

Audio analysis **Applications**

- Hearing aids
- Music technology
- Bio-monitoring
- Personal audio devices
- Ambient assisted living
- Music recommendation
- Acoustic surveillance
- Multimedia archiving









Audio analysis **Problems**

- Modeling
- Estimation
- Detection
- Classification
- Separation
- Optimization

Horizon 2020 Marie Skłodowska-Curie Actions ITN proposal

AUDIOTRACKS

Training, Research, and Career Development in Audio Analysis

Acoustic signal enhancement | Acoustic modeling | Audio analysis | Audio reproduction

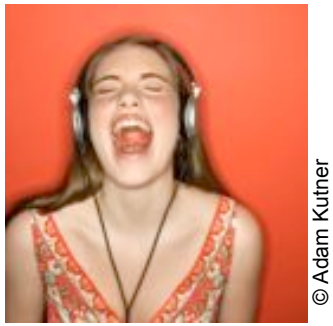
KU LEUVEN

Audio reproduction

- Immersive sound experience
  - reproduction of arbitrary sound fields
  - limiting number of loudspeakers
- Healthy sound experience
  - “binge listening”, earphones misuse
  - limiting sound pressure levels without compromising sound experience
- Multi-disciplinary approach
  - psychoacoustics
  - electro- & room acoustics
  - numerical optimization


© Fraunhofer IDMT


© Vilamundo


© Adam Kutner

Thank you for listening!

KU LEUVEN