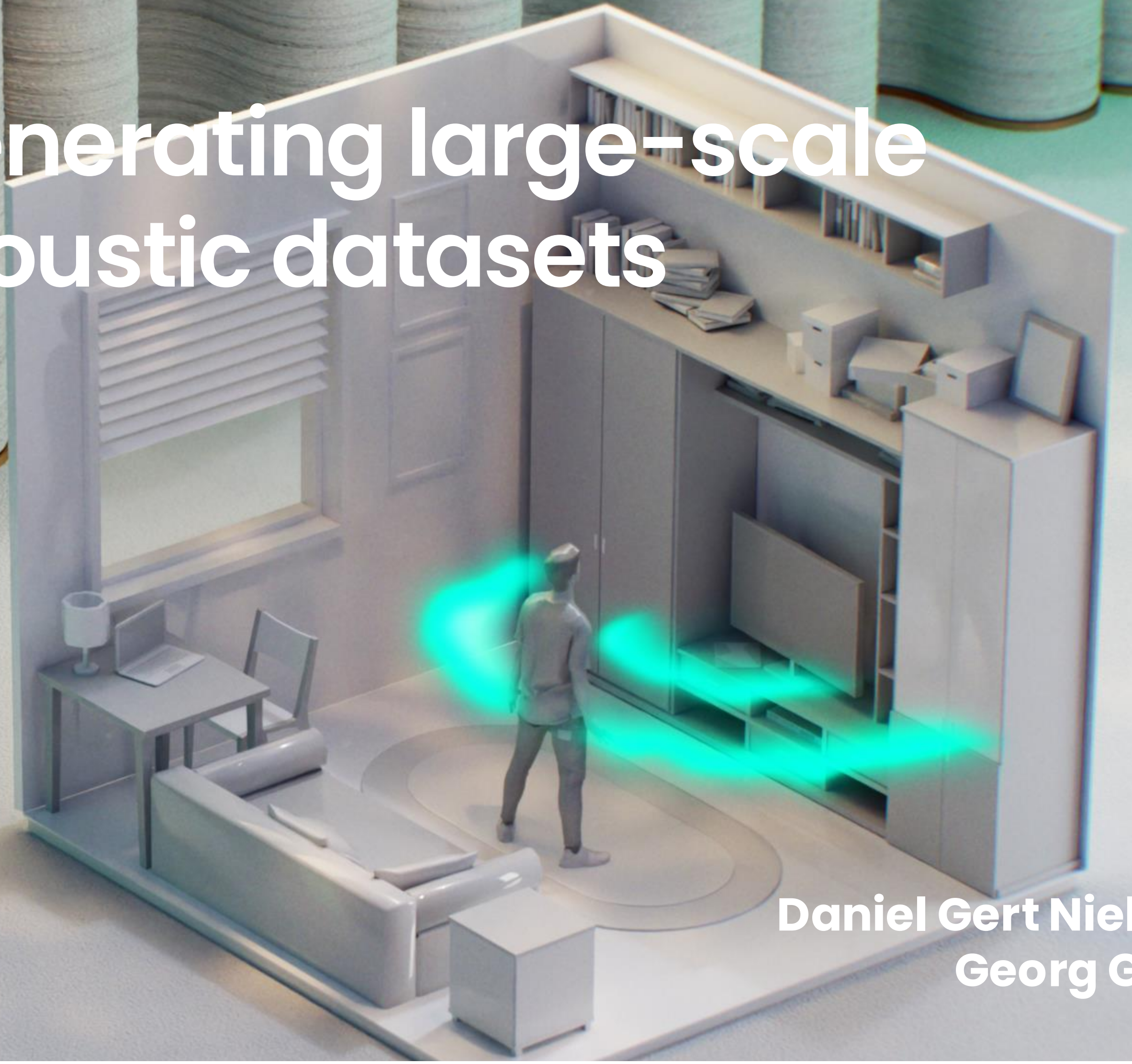
Recent advances in generating large-scale high-quality room-acoustic datasets

08 September 2025

AES International Conference on Artificial Intelligence and Machine Learning for Audio



Code examples for this tutorial
www.treble.tech/aes-sdk-tutorial

Daniel Gert Nielsen
Georg Götz

treble

Enabling a better sounding world

About us

Daniel Gert Nielsen
Senior Product Manager

Georg Götz
Senior Audio Research Engineer

Agenda

1. What improves ML models?
2. Large-scale room-acoustic data acquisition
   - The many degrees-of-freedom in room-acoustic datasets
   - Available datasets
   - Measurements vs. Simulations
   - Automating measurements
   - Recent advances in room-acoustic simulation
3. Simulation paradigms
4. Do high-quality room-acoustic datasets improve the performance of data-driven methods?
5. Room-acoustic simulations as an alternative to measurements for audio-algorithm evaluation
6. Setting up diverse room-acoustic datasets: Example workflow
7. Advanced source and receiver modeling in simulations
8. Conclusions

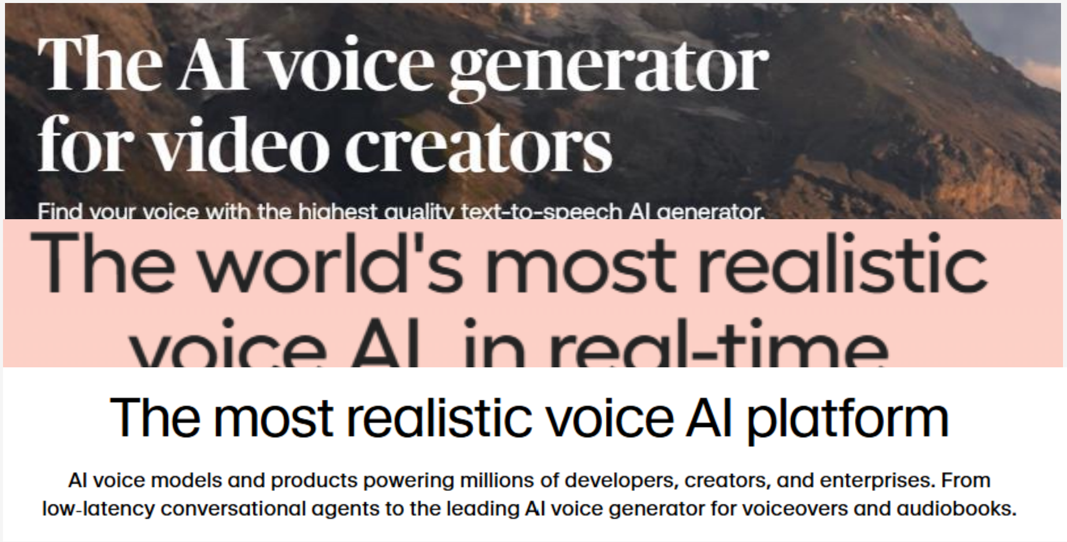
What improves ML models?
Current advancements and challenges

Trends in AI & ML for Audio

- Audio will be the main interface to technology in 5 years
  - We are seeing the building blocks being formed today

Trends in AI & ML for Audio

- Audio will be the main interface to technology in 5 years
  - We are seeing the building blocks being formed today
- Generative AI
  - Speech to text, Text to music, text to speech, voice cloning



Trends in AI & ML for Audio

- Audio will be the main interface to technology in 5 years
  - We are seeing the building blocks being formed today
- Generative AI
  - Speech to text, Text to music, text to speech, voice cloning
- Audio understanding and Analysis
  - Emotion recognition, who is talking?

Trends in AI & ML for Audio

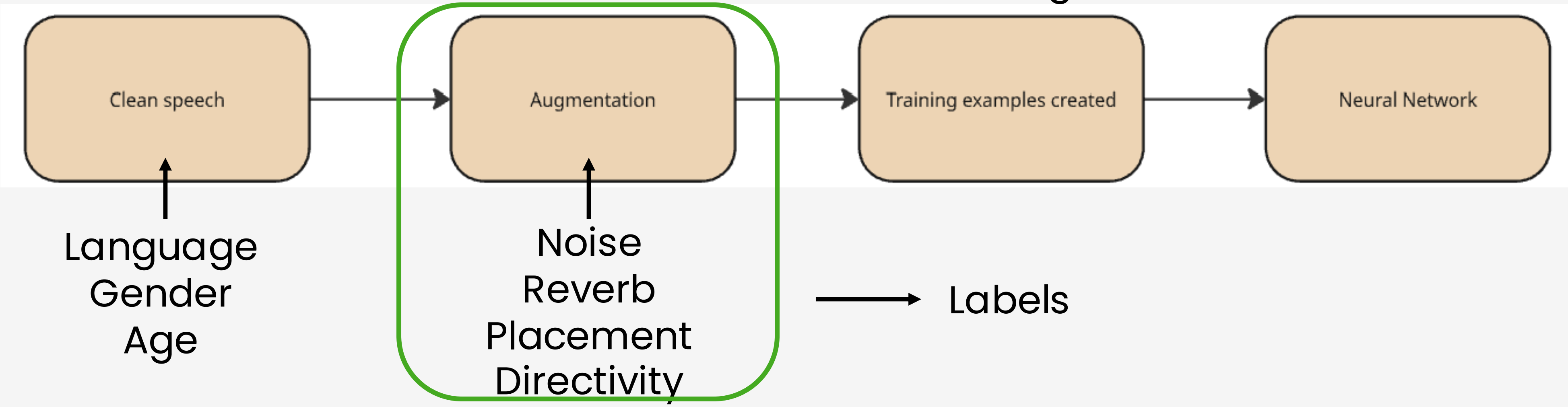
- Audio will be the main interface to technology in 5 years
  - We are seeing the building blocks being formed today
- Generative AI
  - Speech to text, Text to music, text to speech, voice cloning
- Audio understanding and Analysis
  - Emotion recognition, who is talking?
- Sound source localization, blind room estimation
- Audio Separation, Speech Enhancement and de-noising
  - The pre-requisite for the above applications

Trends in AI & ML for Audio

- Audio will be the main interface to technology in 5 years
  - We are seeing the building blocks being formed today
- Generative AI
  - Speech to text, Text to music, text to speech, voice cloning
- Audio understanding and Analysis
  - Emotion recognition, who is talking?
- Sound source localization, blind room estimation
- Audio Separation, Speech Enhancement and de-noising
  - The pre-requisite for the above applications

Challenges in AI & ML for Audio

- Data scarcity
  - Vast amount of diverse data is needed for ML algorithms

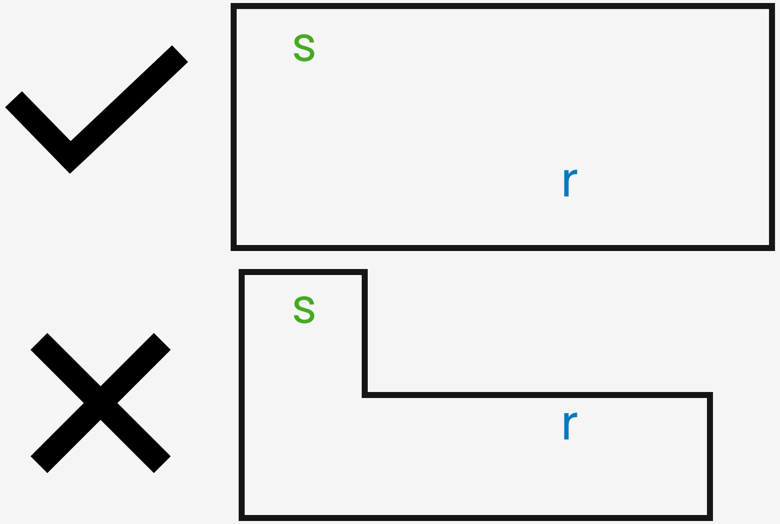


Trends in AI & ML for Audio

- Audio will be the main interface to technology in 5 years
  - We are seeing the building blocks being formed today
- Generative AI
  - Speech to text, Text to music, text to speech, voice cloning
- Audio understanding and Analysis
  - Emotion recognition, who is talking?
- Sound source localization, blind room estimation
- Audio Separation, Speech Enhancement and de-noising
  - The pre-requisite for the above applications

Challenges in AI & ML for Audio

- Data scarcity
  - Vast amount of diverse data is needed
- Robustness issues



Trends in AI & ML for Audio

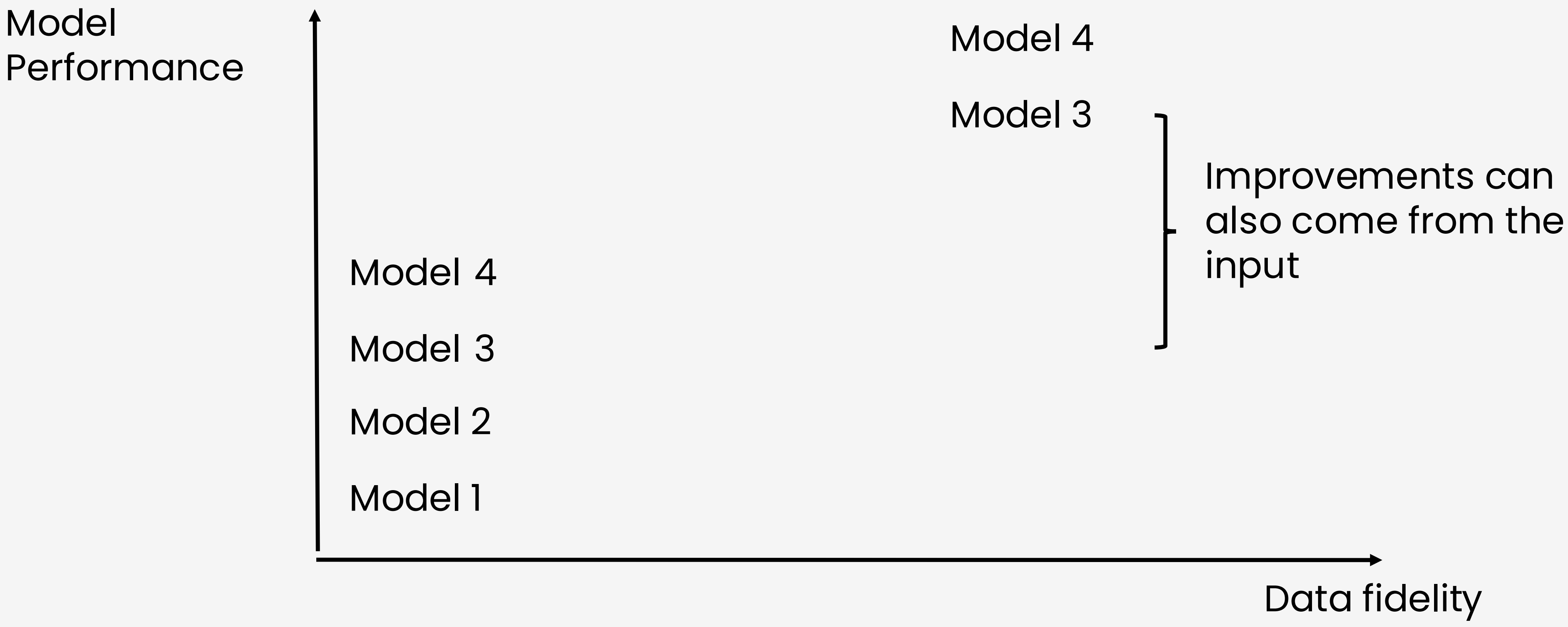
- Audio will be the main interface to technology in 5 years
  - We are seeing the building blocks being formed today
- Generative AI
  - Speech to text, Text to music, text to speech, voice cloning
- Audio understanding and Analysis
  - Emotion recognition, who is talking?
- Sound source localization, blind room estimation
- Audio Separation, Speech Enhancement and de-noising
  - The pre-requisite for the above applications

Challenges in AI & ML for Audio

- Data scarcity
  - Vast amount of diverse data is needed for ML algorithms
- Robustness issues
- Metrics are King
  - PESQ, MOS, SDR, FAD, WER, STOI,....
  - How do we perceive it?
    - Time consuming but important
- AI on the edge - low latency is a requirement for real time applications

Model improvement dimensions

- The race is on! Rapid development in ML models constantly improves performance



Large-scale room-acoustic data acquisition

Challenges and current trends

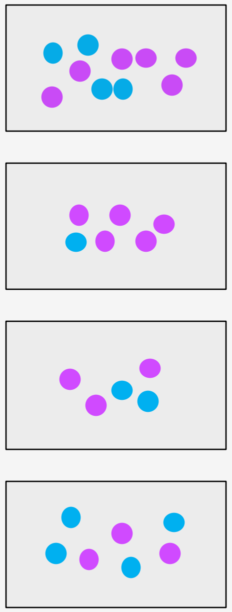
Room-acoustic datasets should be diverse

DISENTANGLING THE VARIOUS DEGREES OF FREEDOM

Sound source
Speech signal

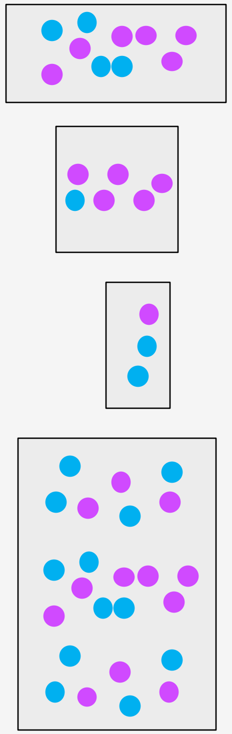
Receiver

Different source-receiver configurations

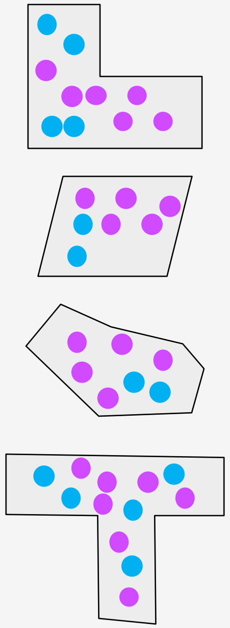


Also: directivities

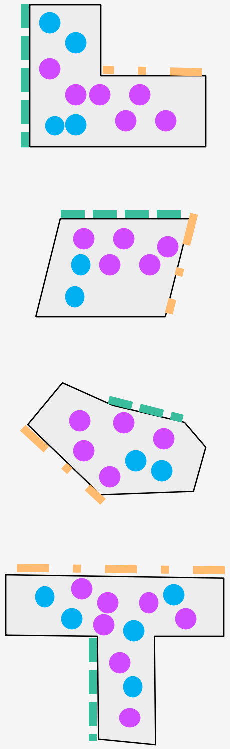
Different room volumes



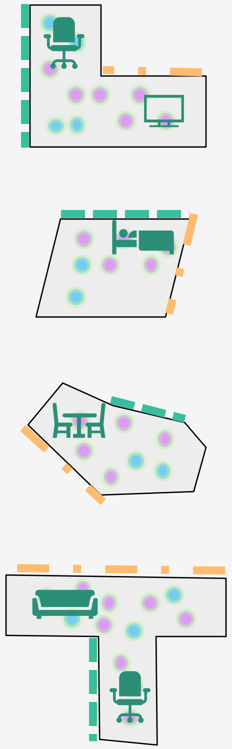
Different room shapes



Different absorption



Different scattering objects



Room-acoustic datasets should be diverse

What effects do cultural/aesthetic differences have on room acoustics or device placement? Should we account for them in room-acoustic datasets?

“German living room”

https://app.dropinblog.com/uploaded/blogs/34241141/files/German_Accessories.png

“Japanese living room”

https://tiletoria.co.za/wp-content/uploads/2025/05/Japanese-interior-design-living-room.jpg

Realistic and diverse training datasets improve the ML model performance

Table 2: *Localization results on three real test sets achieved by the SRP-PHAT baseline and by the supervised model [34] trained using various simulation modes. Mean angular errors (MAE) are displayed with their 95% confidence interval. Bold numbers indicate the best system in each column and the systems statistically equivalent to it. Statistical significance was assessed using McNemar's test for the Recall metric and 95% confidence intervals over angular error differences for the MAE metric.*

Real Test Sets →	**VoiceHome-2** [30]		**DIRHA** [31]		**STARS22** [32]	
Methods	↑ Recall	↓ MAE (°)	↑ Recall	↓ MAE (°)	↑ Recall	↓ MAE (°)
SRP-PHAT	70%	9.9 ± 1.5	61%	15.0 ± 2.3	45%	14.9 ± 0.6
Naive Training	78%	7.6 ± 1.2	77%	8.4 ± 1.4	57%	12.9 ± 0.6
Advanced Training	**85%**	$\mathbf{5.8 \pm 0.8}$	**84%**	$\mathbf{6.3 \pm 1.0}$	61%	$\mathbf{11.4 \pm 0.5}$
Ablation study						
without wall realism	83%	6.2 ± 0.8	81%	7.5 ± 1.4	59%	12.1 ± 0.6
without source realism	82%	7.1 ± 1.1	80%	7.8 ± 1.2	**63%**	$\mathbf{11.4 \pm 0.6}$
without receiver realism	N/A	N/A	78%	8.3 ± 1.5	53%	13.4 ± 0.6

P. Srivastava, A. Deleforge, A. Politis, and E. Vincent, "How to (Virtually) Train Your Speaker Localizer," *INTERSPEECH 2023*, pp. 1204–1208, 2023

More evidence later in this talk!

What datasets are already publicly available?

Dataset	Acquisition	Rooms	Acoustic conditions per room	Scattering objects	Source-receiver configurations	Directivity
dEchorate (*Di Carlo et al.*)	Measurement	1 (Shoebox)	11	❌	6 arrays with 5 mics each 6 sources	Mono mics Directional loudspeaker
Arni (*Prawda et al.*)	Measurement	1 (Shoebox)	5312	❌	5 mics 1 source	Mono mics Omni loudspeaker
Arni – 6DoF Spatial (*McKenzie et al.*)	Measurement	1 (Shoebox)	5	❌	7 mics 3 sources	HOA mic arrays Directional loudspeaker
BUT ReverbDB (*Szöke et al.*)	Measurement	8 (Various)	1	✅	31 mics 5 sources	Mono mics Directional loudspeaker
BIRD (*Grondin et al.*)	Simulation (Image-source model)	12,500 (Shoebox)	1	❌	2 mics 4 sources	Mono receivers Omni sources
Motus (*Götz et al.*)	Measurement	1 (Shoebox)	830	✅	1 mic 4 sources	HOA mic array Directional loudspeaker
ACE Challenge (*Eaton et al.*)	Measurement	7 (Various)	1	✅	2 mics 1 source	Several mic arrays, also HOA Directional loudspeaker
MIT IR Survey (*Traer and McDermott*)	Measurement	271	1	✅	2 mics 1 source	Mono mics Directional loudspeaker
GWA (*Tang et al.*)	Simulation (Hybrid)	18,900 (Various)	1	✅	Receiver grid Source grid	Mono receivers Omni sources
Loudspeaker array (*Erbes et al.*)	Measurement	1	4	❌	1 mic (80 head rotations) 64 sources	Mono mic and dummy head Directional loudspeakers
ADREAM (*Winter et al.*)	Measurement	1	1	✅	4 mics with 78 head rotations + 20 mics with single head rotation 4 sources	Dummy head
MeshRIR (*Koyama et al.*)	Measurement	2	1	❌	Mic grid (robot), 4,400 positions 33 sources	Mono mics Directional loudspeaker

… and many more!

Simulations or measurements?

ACCURACY, EFFORT, AND SCALABILITY

Simulations

- Exactly reproducible
- High scalability
- High diversity
- High flexibility
- Prototyping devices/rooms before building them

- Tradeoff: Simulation speed and cost vs. accuracy
- Input data uncertainty

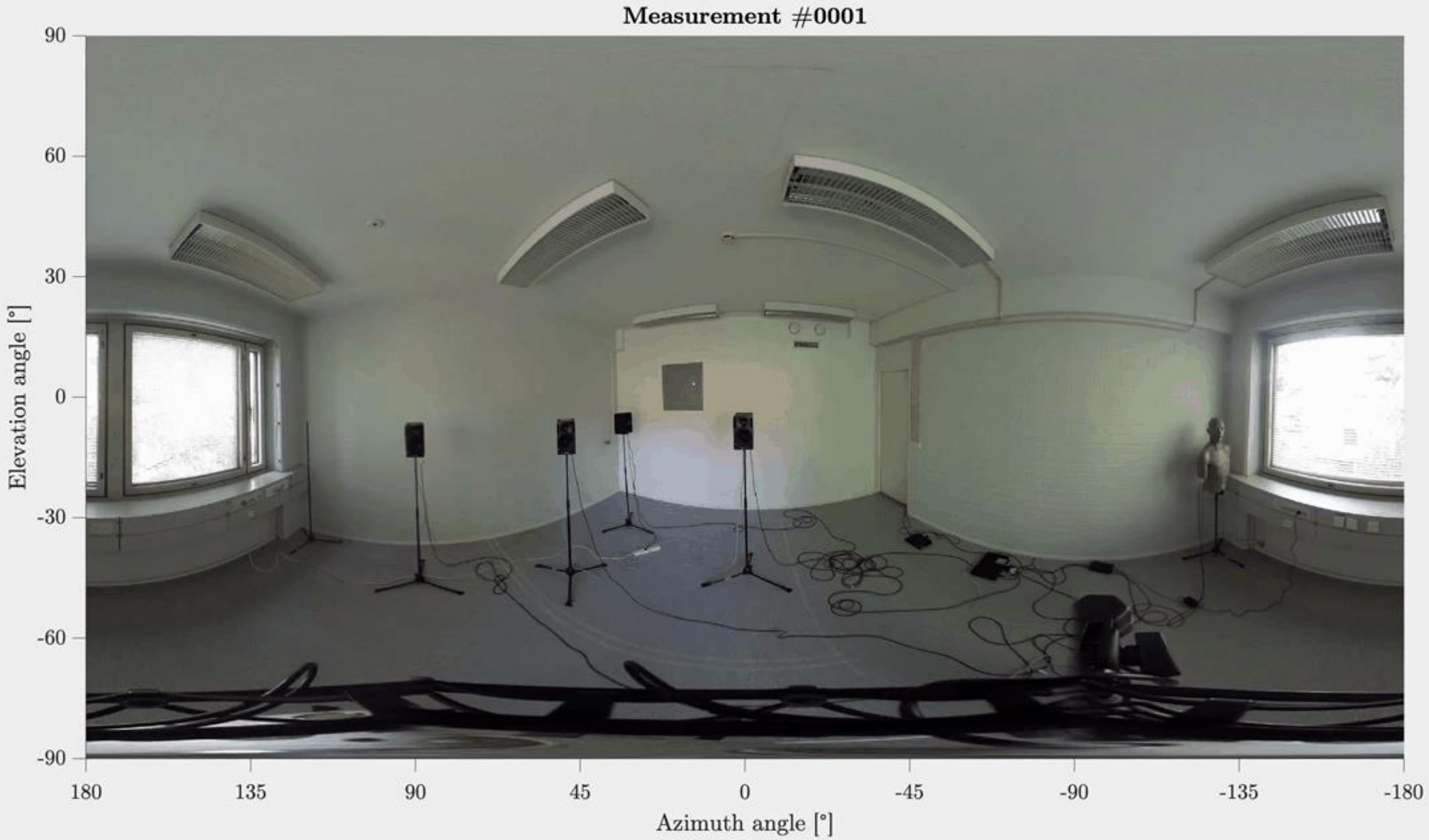
Measurements

- Capturing an environment "as it is"

- Time-consuming, tedious, labor-intensive, but automation is possible
- Limited diversity
- Require measurement equipment (possibly expensive)
- Difficult to reproduce exactly (setups, measurement procedure, human error)

- Low scalability
- Inherently noisy

Example: A summer of 830 measurements

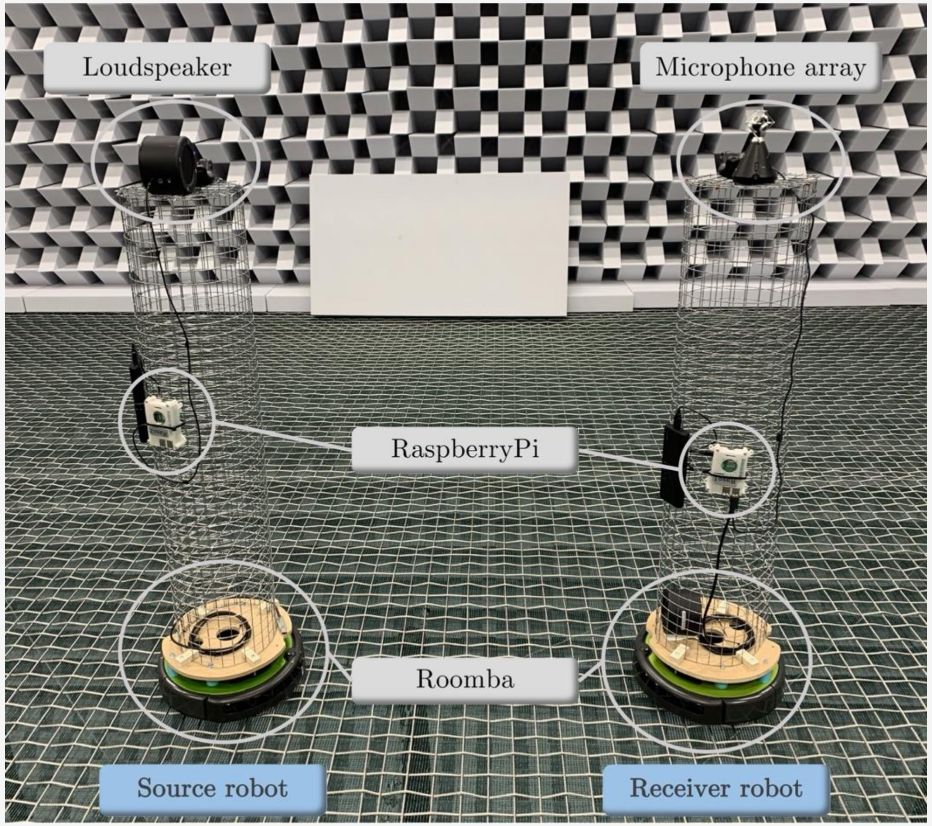


G. Götz, S. J. Schlecht, and V. Pulkki, "A dataset of higher-order Ambisonic room impulse responses and 3D models measured in a room with varying furniture," *2021 Immersive 3D Audio Archit. Automot. I3DA*, 2021

Example: ~~A summer of 830 measurements~~
One afternoon and 30 lines of code

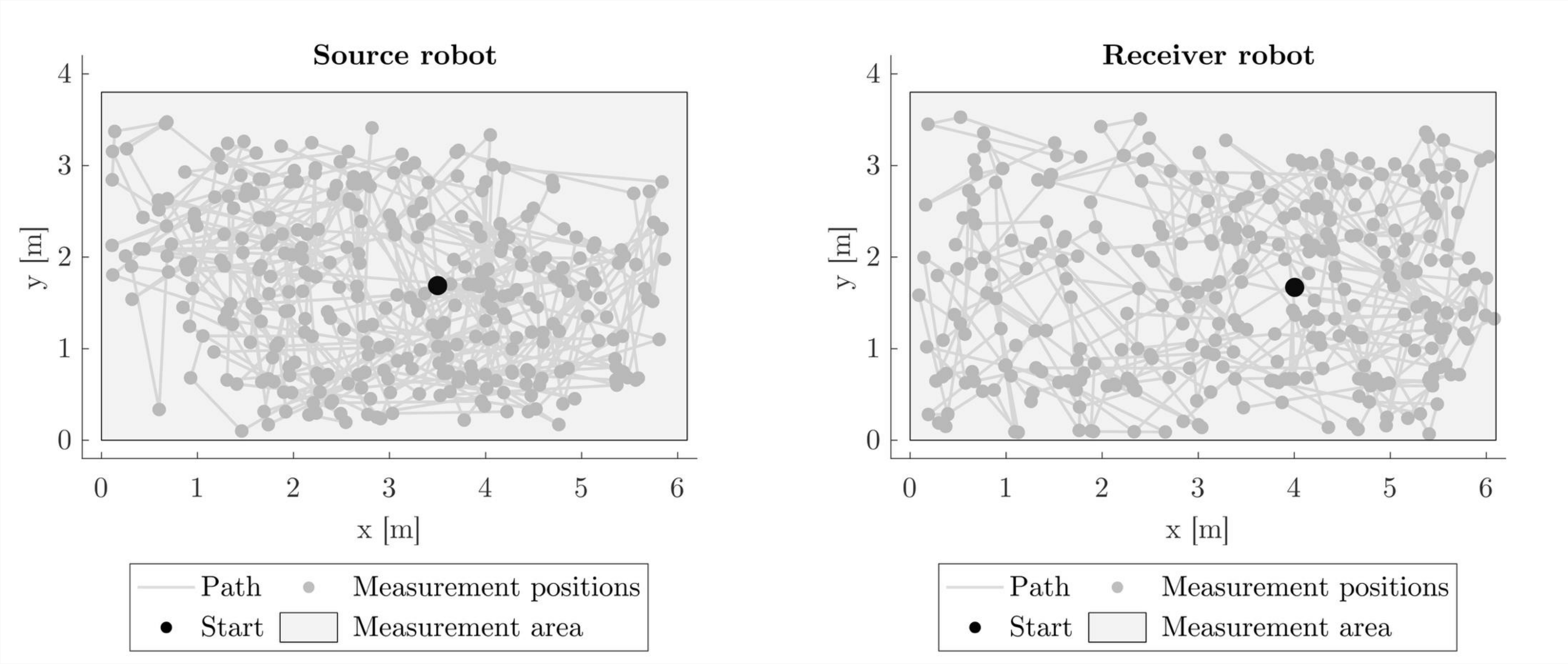
```python
# When randomly placing geometry components, rotation_settings are specified as a ComponentAnglePool.
# A ComponentAnglePool specifies which orientations a geometry component can have.
# If no rotation_settings is given, the component will be placed with a random orientation.
whole_angles = treble.ComponentAnglePool([0, 90, 180, 270])
half_angles = treble.ComponentAnglePool([0, 45, 90, 135, 180, 225, 270, 315])

# Define the objects and quantities that we want to place in the rooms.
placements = [
    # We want 2 tables, and they can be any table from the library. preferred_count is the max count of the component that will be placed. This does not guarantee that this many objects will appear, for example if space is limited.
    treble.GeometryComponentPlacement(
        components=tsdk.geometry_component_library.query(group="table"),
        preferred_count=2,
        rotation_settings=whole_angles,
        min_dist_from_objects=0.25,
        min_dist_from_walls=0.25,
    ),
    # 4 chairs.
    treble.GeometryComponentPlacement(
        components=tsdk.geometry_component_library.query(group="chair"),
        preferred_count=2,
        rotation_settings=half_angles,
        min_dist_from_objects=0.25,
        min_dist_from_walls=0.25,
    ),
    # 1 sofa.
    treble.GeometryComponentPlacement(
        components=tsdk.geometry_component_library.query(group="sofa"),
        preferred_count=1,
        rotation_settings=whole_angles,
        min_dist_from_objects=0.25,
        min_dist_from_walls=0.25,
    ),
    # 2 big boxes.
    treble.GeometryComponentPlacement(
        components=treble.GeometryComponentGenerator.create_box(
            treble.BoundingBox.from_points(-0.5, 0.5, -0.5, 0.5, 0, 1)
        ),
        preferred_count=2,
        # You can specify minimum distance from other components and walls.
        min_dist_from_objects=0.5,
        min_dist_from_walls=0.5,
    ),
    # 5 small boxes.
    treble.GeometryComponentPlacement(
        components=treble.GeometryComponentGenerator.create_box(
            treble.BoundingBox.from_points(-0.1, 0.1, -0.1, 0.1, 0, 0.2)
        ),
        preferred_count=5,
        min_dist_from_objects=0.02,
    ),
]

for room_type in generated_room_defs:
    print(f"==== Populating {room_type} rooms with objects. ====")
    for room_idx, room in enumerate(tqdm(generated_room_defs[room_type])):
        # Populate with objects
        room.populate_with_geometry_components(
            components=placements, selection_algorithm=treble.ComponentSelectionAlgorithm.random
        )
```

Example: Automating room-acoustic measurements with robots



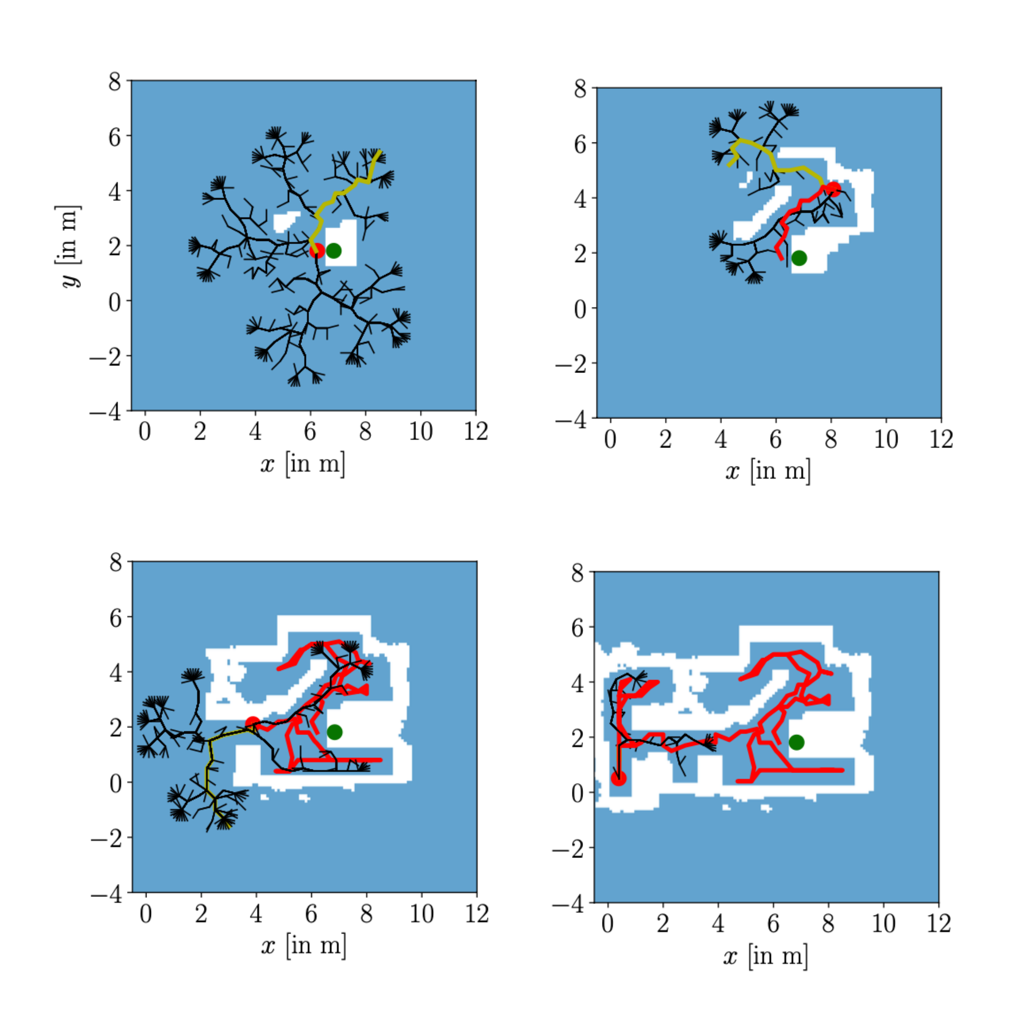
G. Götz, A. M. Ornelas, S. J. Schlecht, and V. Pulkki, "Autonomous Robot Twin System for Room Acoustic Measurements," *J. Audio Eng. Soc.*, vol. 69, no. 4, pp. 261–272, 2021

A random walk measurement strategy covers a shoebox room without leaving out any large areas



G. Götz, A. M. Ornelas, S. J. Schlecht, and V. Pulkki, “Autonomous Robot Twin System for Room Acoustic Measurements,” *J. Audio Eng. Soc.*, vol. 69, no. 4, pp. 261–272, 2021

Advanced robot systems enable advanced measurement strategies in complex environments



G. Götz, I. Ananthabhotla, S. V. A. Garí, and P. Calamia, "Autonomous Room Acoustic Measurements using Rapidly-Exploring Random Trees and Gaussian Processes," in Proc. 10th Conv. European Acoustics Association (Forum Acusticum). Torino, Italy, 2023.

Remaining challenges in automated measurements

- Accurate SLAM (simultaneous localization and mapping)
- Robust automated measurement procedures and automatic detection of artifacts
- Truly(!) flexible robots:
    - Size that fits "everywhere"
    - Moving furniture or objects?
    - What about height?

https://www.rc-zeppelin.com/slike/1.5%20m/1.5-m-RC-Blimp.jpg

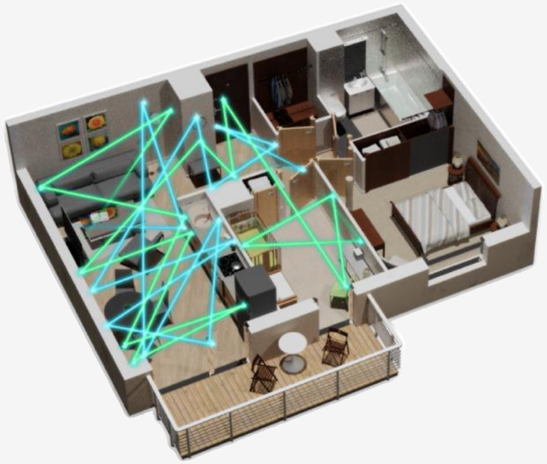
Simulation paradigms
A comparison of acoustic simulation methods

The two simulation paradigms

- In traditional room acoustic simulation there are two main methods
  - Geometrical Acoustics
    - Approximates acoustic waves as rays – works decently at high frequencies
  - Wave-based methods
    - Solves the wave equation – needs a lot of compute power at high frequencies
- In terms of data generation for audio Geometrical acoustics are by far the most popular one, as it is used in OS such as PRA, gpuRIR, and Habets ISM

Numerical ("wave-based")

Directly solving the wave equation: inherently capturing wave phenomena like diffraction, phase and scattering.

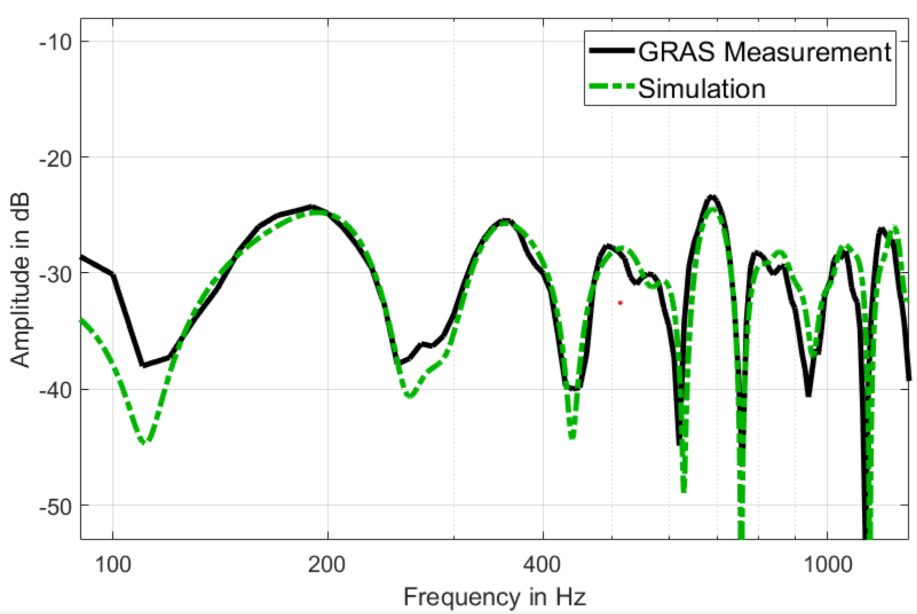


Geometrical acoustics

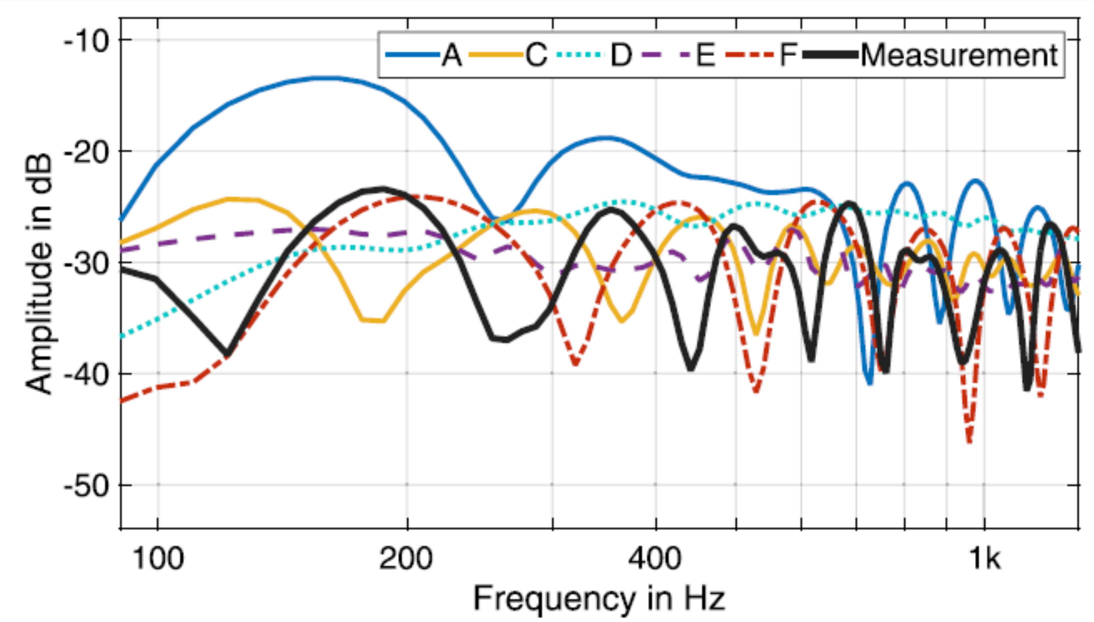
High-frequency approximation, lacking wave-nature of sound. Computationally efficient.

Reflection off a complex surface

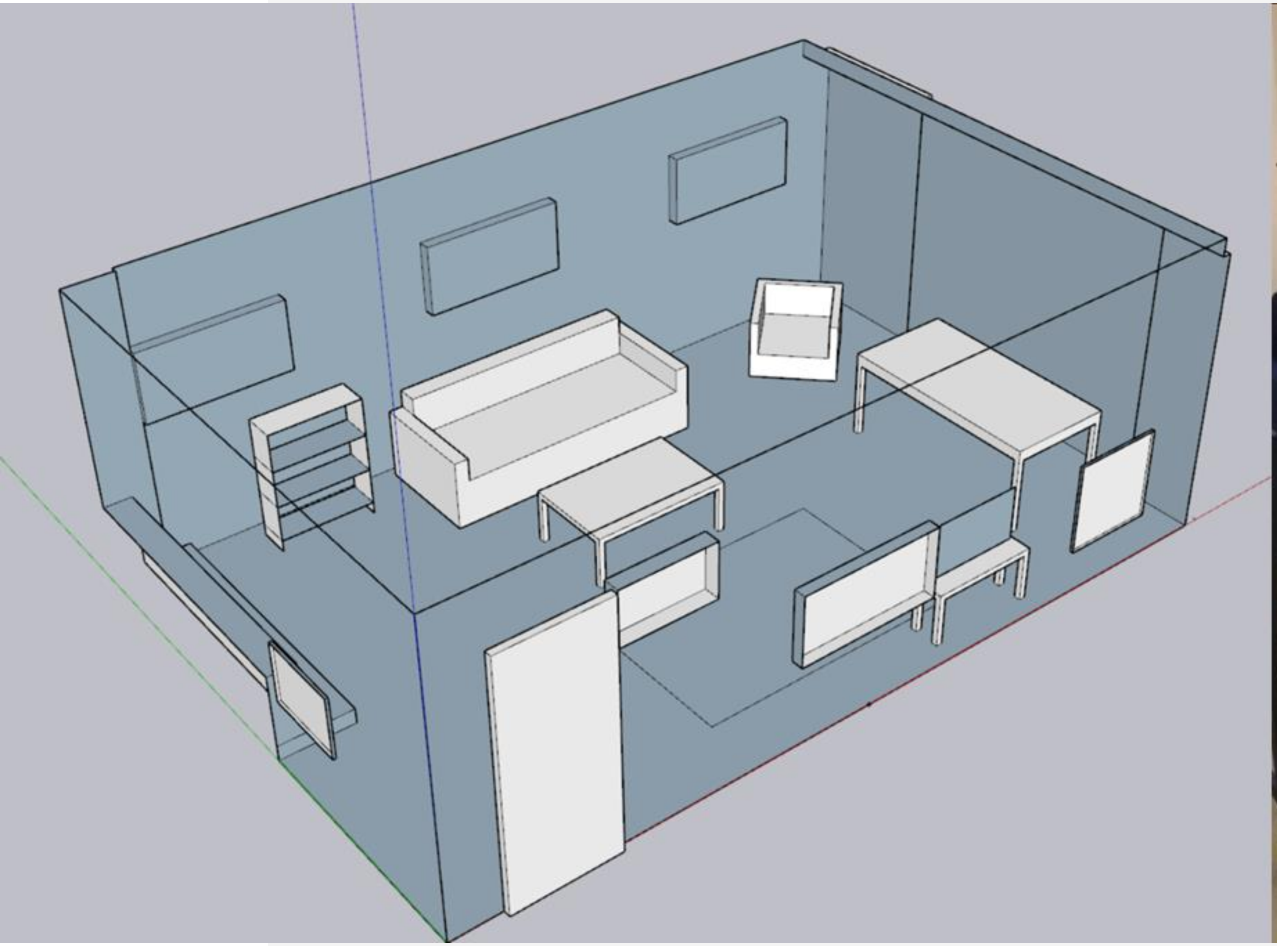
Numerical ("wave-based")



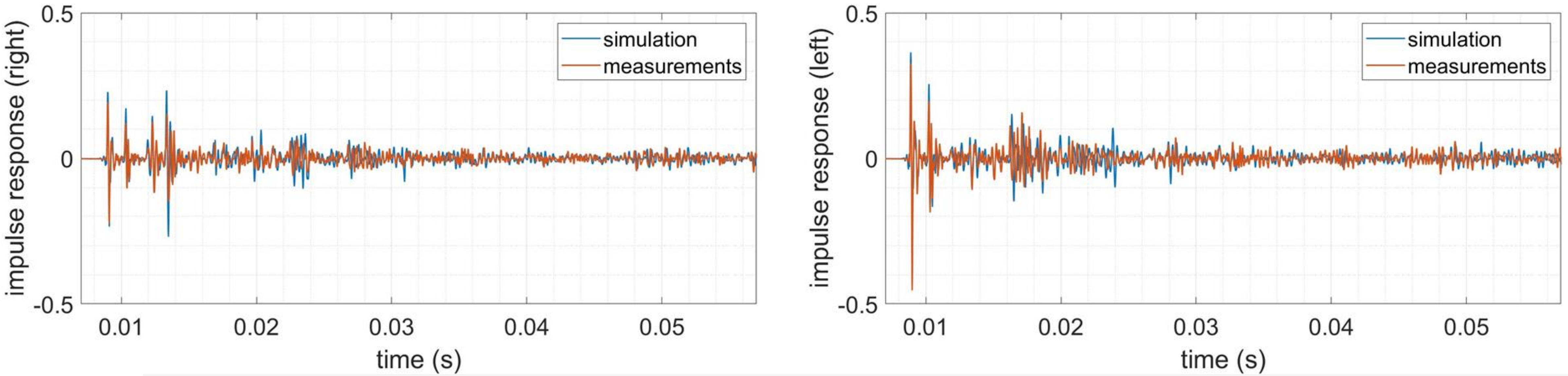
Geometrical acoustics



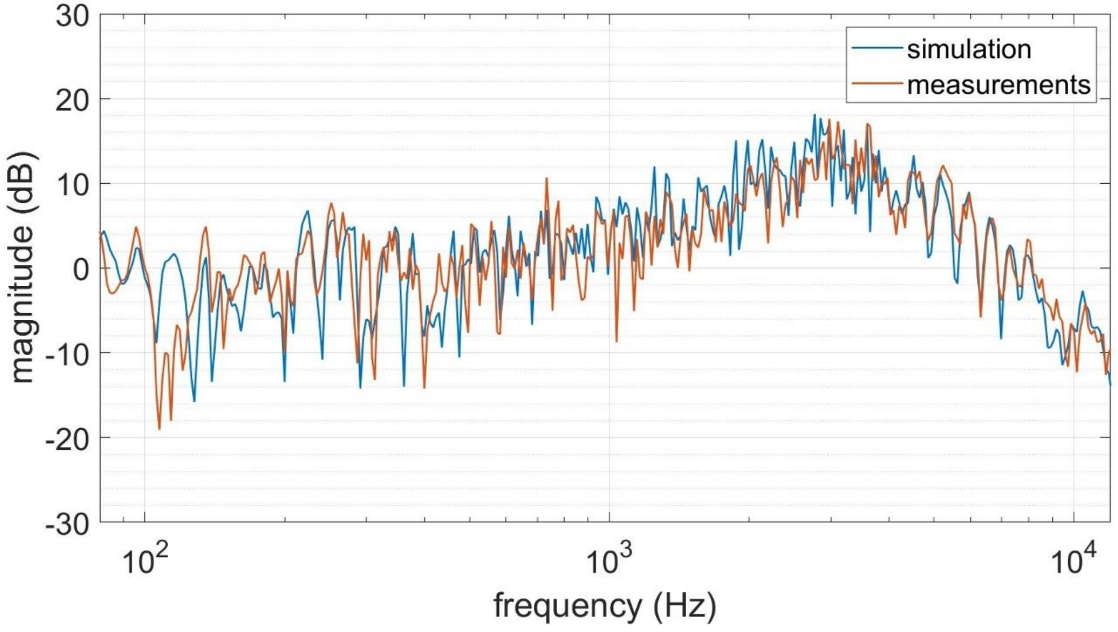
Complex acoustic conditions

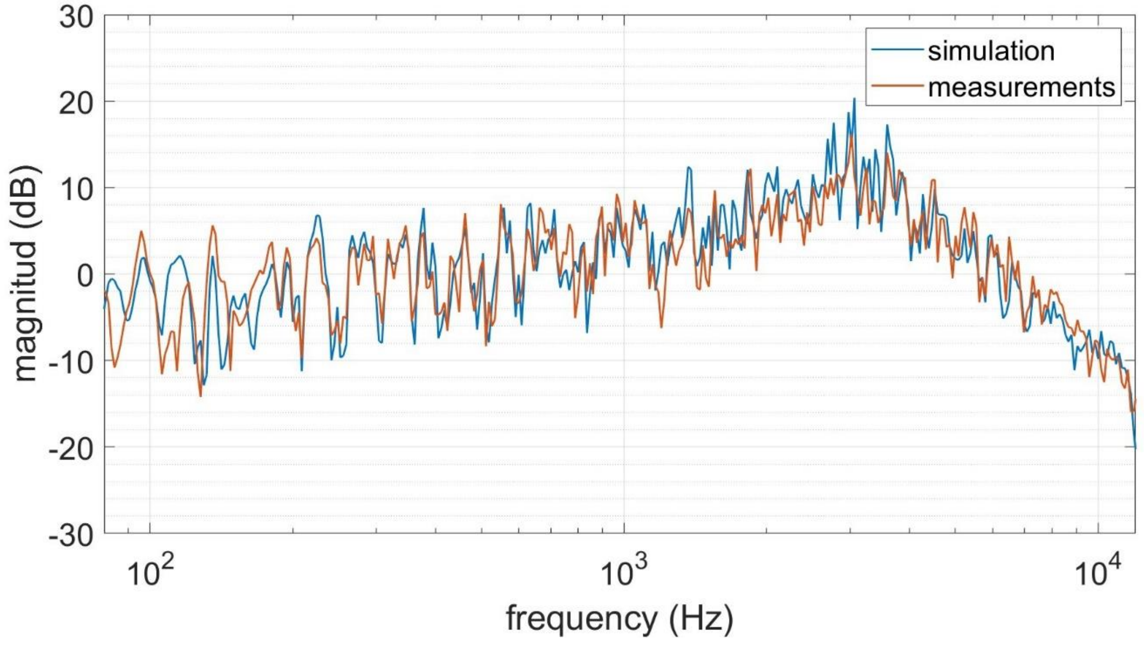


Binaural IRs



Binaural IR results



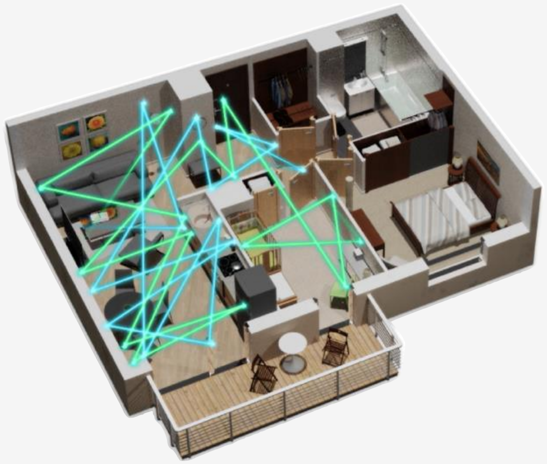


The two simulation paradigms

- Wave-based simulations are expensive at high frequencies
- Works well for small datasets
    - E.g. for evaluation tasks
- A hybrid approach is more suitable for large dataset generation

Numerical ("wave-based")

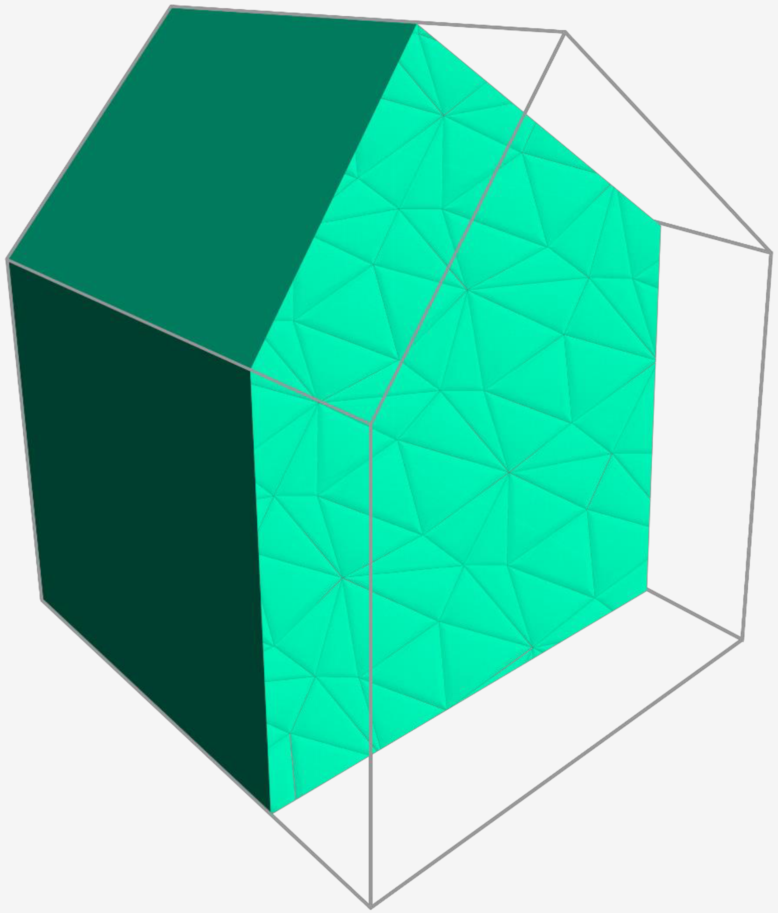
Directly solving the wave equation: inherently capturing wave phenomena like diffraction, phase and scattering.



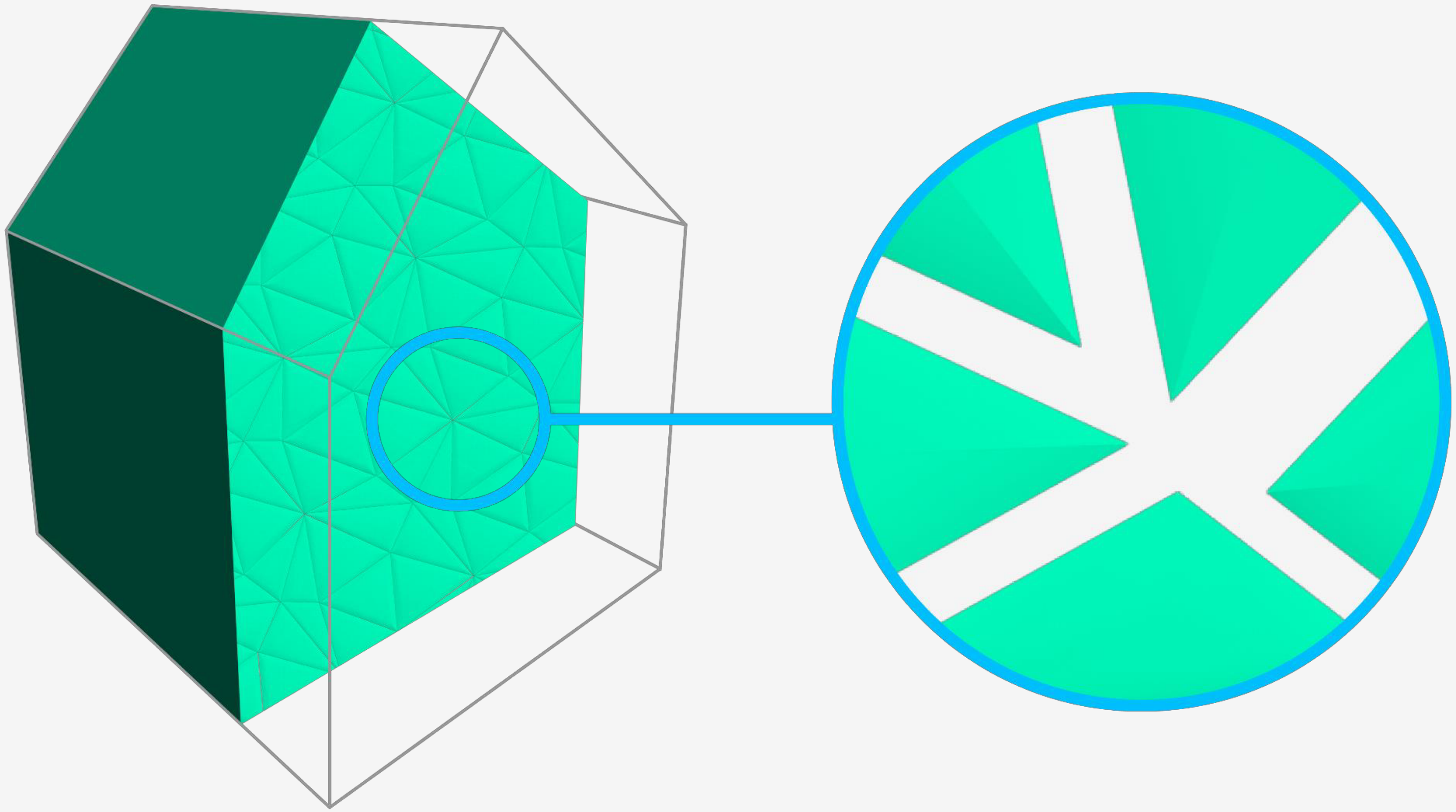
Geometrical acoustics

High-frequency approximation, lacking wave-nature of sound. Computationally efficient.

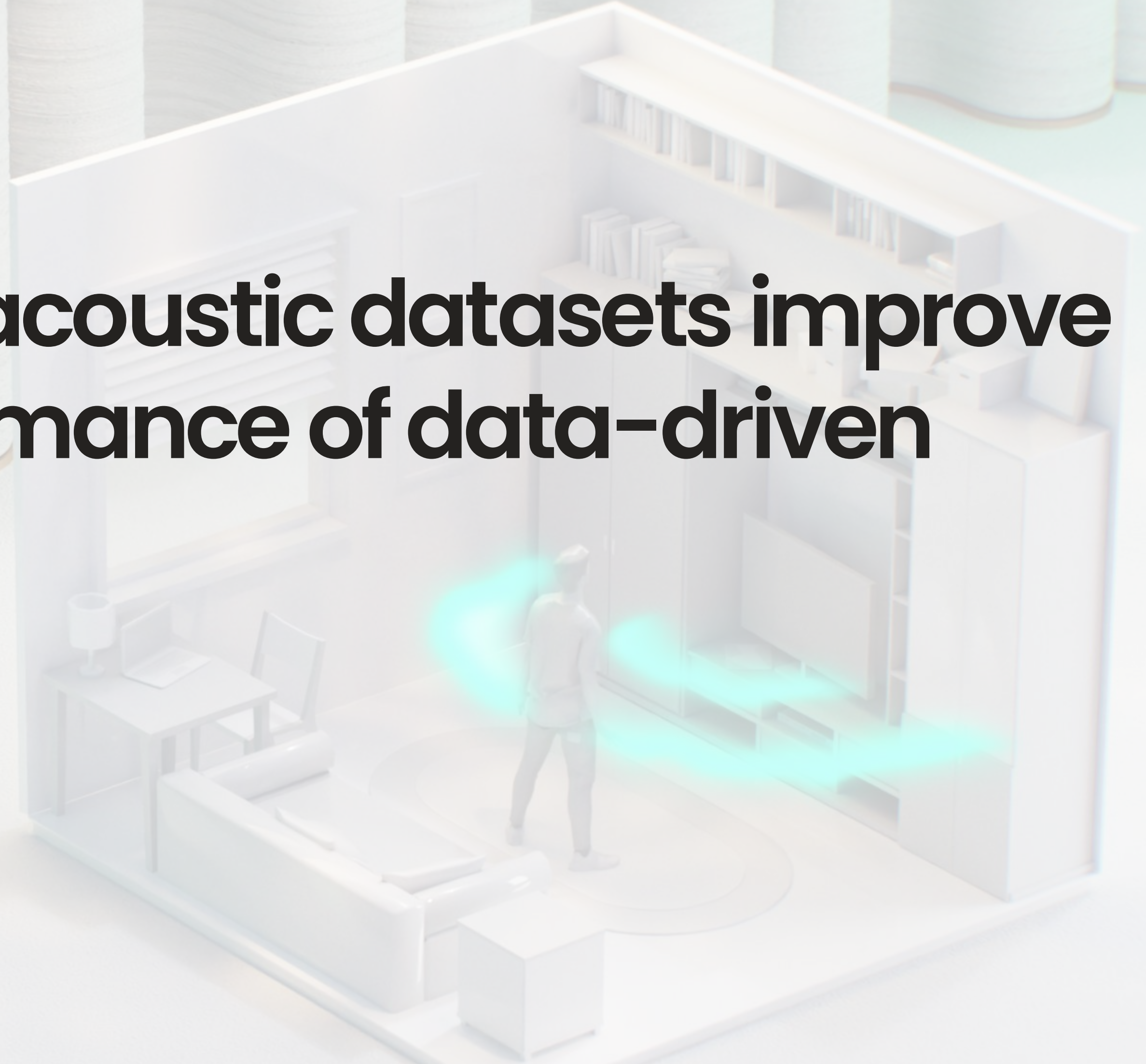
Massively parallel wave-based modeling



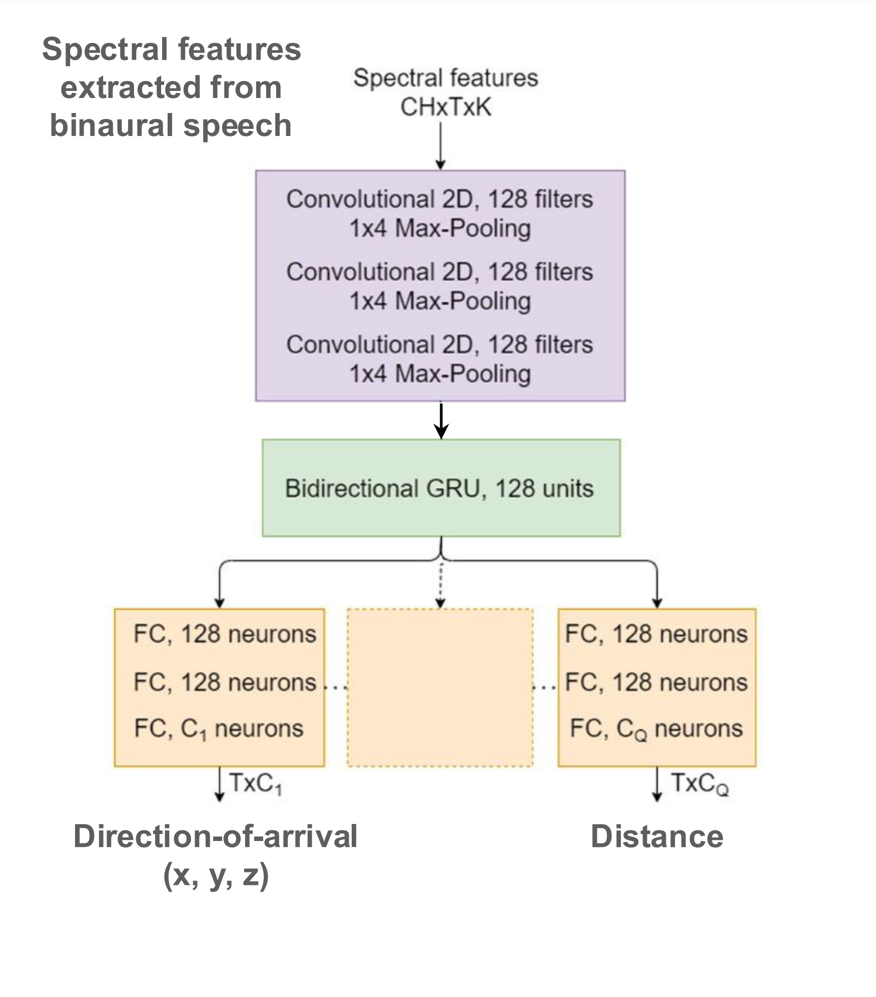
Massively parallel wave-based modeling



Do high-quality room-acoustic datasets improve the downstream performance of data-driven methods?



We train a binaural sound source distance estimation and localization model



Krause, D. A., García-Barrios, G., Politis, A., Mesaros, A. & Krause, D. A. "Binaural Sound Source Distance Estimation and Localization for a Moving Listener." *IEEE/ACM Trans. Audio, Speech Lang. Process.* **32**, 996–1011 (2024).

We compare the performance of different training datasets

- **ISM:**
  - Dataset provided by authors
  - Image-source model (custom implementation)
  - 2500 shoeboxes
  - Non-uniform absorption

- **Hybrid:**
  - Numerical wave-based simulation (DG-FEM) until 700 Hz
  - Ray radiosity + image-source model above 700 Hz
  - Implemented in Treble SDK
  - 2500 shoeboxes
  - Non-uniform absorption

- **PRA:**
  - Image-source model (50th order)
  - Implemented in PyRoomAcoustics
  - 2500 shoeboxes
  - Non-uniform absorption
  - Exactly replicating the room and source-receiver setup from Hybrid

We evaluate the performance of the trained models on a measured evaluation dataset

- Combination of three datasets containing measured BRIRs
  - Loudspeaker Array

    V. Erbes, M. Geier, S. Weinzierl, S. Spors (2015): Database of single-channel and binaural room impulse responses of a 64-channel loudspeaker array. Proc. of the 138th Int. AES Convention, Warsaw, Poland.

  - ADREAM

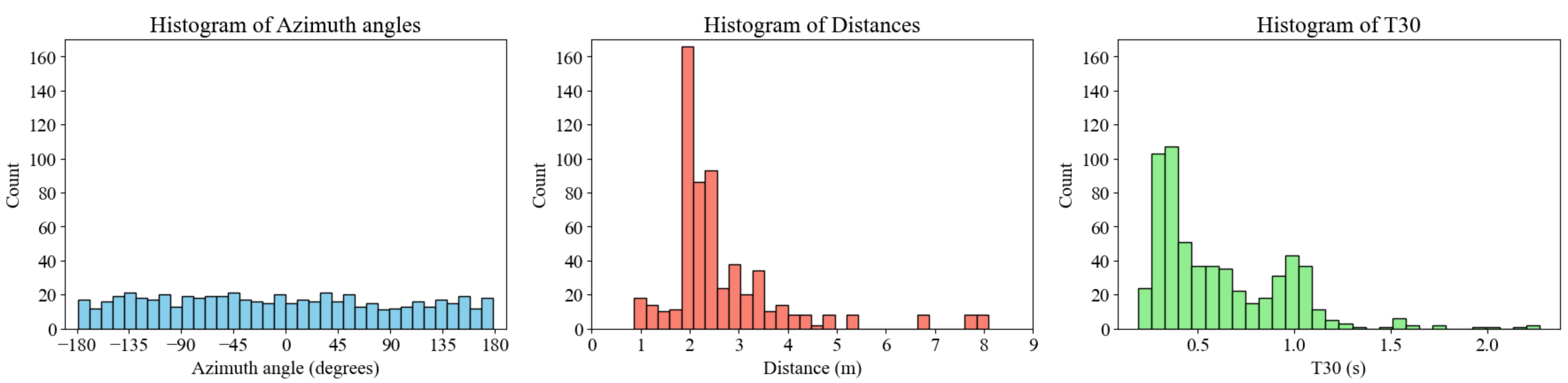
    F. Winter, H. Wierstorf, A. Podlubne, T. Forgue, J. Manhès, M. Herrb, S. Spors, A. Raake, and P. Danès, "Database of binaural room impulse responses of an apartment-like environment," Proc. of 140th Aud. Eng. Soc. Conv., Paris, 2016

  - Measurements from Treble's furnished lab

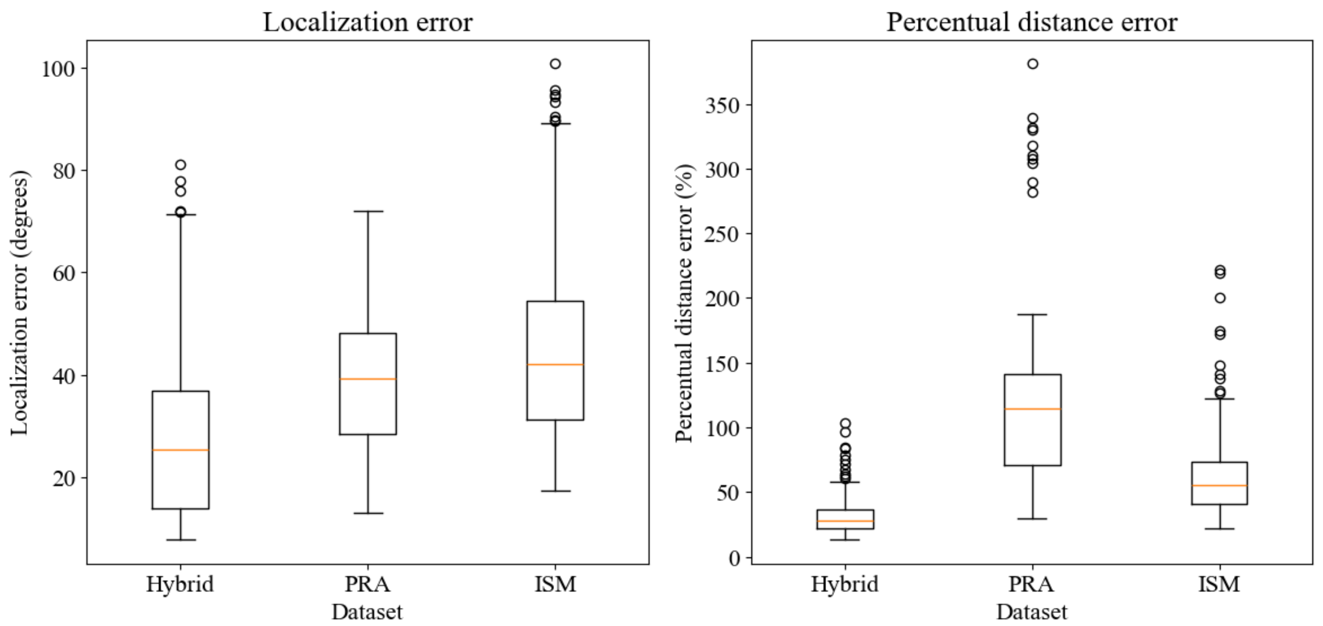
    Custom dataset recorded in the Treble Lab, including furniture

- Approximately 600 BRIRs in total

The evaluation dataset consists of measurements and covers various acoustic conditions



The performance of the model improves when training on high-quality simulations



Training dataset	Localization error (°) ↓		Percentual distance estimation error (%) ↓	
	Median	90th percentile	Median	90th percentile
Hybrid	**25.37**	**49.08**	**27.755**	**48.940**
PRA	39.20	55.58	114.583	164.163
ISM	42.07	68.51	55.309	97.881

Task-specific considerations

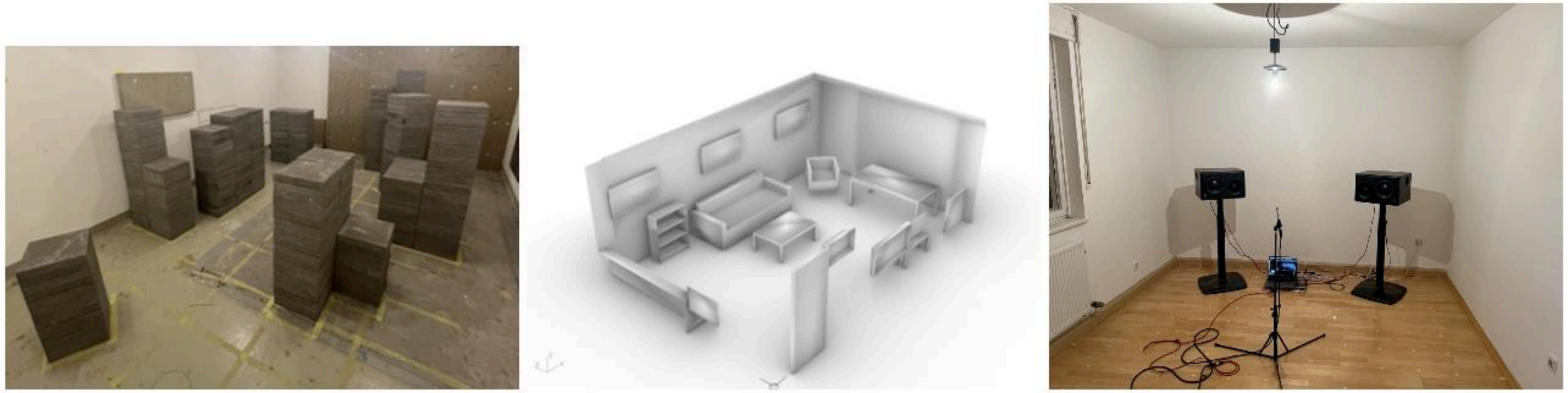
- Binaural localization model is just one example we tried out: improved performance also in other tasks
- Some tasks benefit more from wave-based data than others
- Comprehensive study following soon!
- Let's discuss if you have a specific use case in mind!

Room-acoustic simulations as an alternative to measurements for audio-algorithm evaluation

G. Götz, D. G. Nielsen, S. Guðjónsson, F. Pind, "Room-acoustic simulations as an alternative to measurements for audio-algorithm evaluation," arXiv:2509.05175

We replicate four measured datasets using different simulation paradigms

Dataset	Room	Description	Source-receiver configurations
Bricks	Room 1: Lab room, $80\,\mathrm{m}^3$ ($7\,\mathrm{m} \times 4.5\,\mathrm{m} \times 2.5\,\mathrm{m}$)	Including piles of $40\,\mathrm{cm} \times 40\,\mathrm{cm}$ bricks at multiple locations and 100 mm thick stonewool absorbers on the walls. Average reverberation time: 0.6 s	20 RIRs (10 receivers $\times$ 2 sources) Receivers: GRAS 1/2" free-field microphones Sources: Avantone MixCubes Active
Furniture	Room 1: Lab room, $80\,\mathrm{m}^3$ ($7\,\mathrm{m} \times 4.5\,\mathrm{m} \times 2.5\,\mathrm{m}$)	Including typical living room furniture and 100 mm thick stonewool absorbers on the walls. Average reverberation time 0.5 s	6 RIRs (3 receivers $\times$ 2 sources) Receivers: GRAS 1/2" free-field microphones Sources: Avantone MixCubes Active
Variable absorption	Room 1: Lab room, $80\,\mathrm{m}^3$ ($7\,\mathrm{m} \times 4.5\,\mathrm{m} \times 2.5\,\mathrm{m}$)	Including different stonewool absorber configurations on the walls, resulting in three acoustic conditions. Average reverberation time, condition 1: 0.3 s Average reverberation time, condition 2: 0.6 s Average reverberation time, condition 3: 0.9 s	12 RIRs (2 receivers $\times$ 2 sources $\times$ 3 conditions) Receivers: GRAS 1/2" free-field microphones Sources: Avantone MixCubes Active
Studio	Room 2: Studio, $38\,\mathrm{m}^3$ ($4.8\,\mathrm{m} \times 3.2\,\mathrm{m} \times 2.5\,\mathrm{m}$)	Empty studio room in a historic old building. Solid exterior walls, lightweight construction interior walls, laminate floor. Average reverberation time: 1.43 s	34 RIRs (17 receivers $\times$ 2 sources) Receivers: NTI MA220 Class 1 microphones Sources: ATC SCM25 A MK2



(a) Lab room with bricks. **(b)** Lab room with furniture. **(c)** Studio.

We replicate four measured datasets using different simulation paradigms

Name	Simulation details
DG-FEM	Wave-based: Discontinuous Galerkin finite-element method • Implemented with the Treble SDK • Surfaces modelled with complex acoustic impedances • Directional source modelling: boundary velocity source • Upper simulation frequency: 7 kHz
GA-RR	Geometrical Acoustics: Ray radiosity • Implemented with the Treble SDK • Surfaces modelled with complex acoustic impedances • Directional source modelling: fitted pattern from boundary velocity source
GA-RT	Geometrical Acoustics: Ray tracing • Implemented with PyRoomAcoustics • Surfaces modelled with energy absorption coefficients • Directional source modelling: cardioid directivity pattern

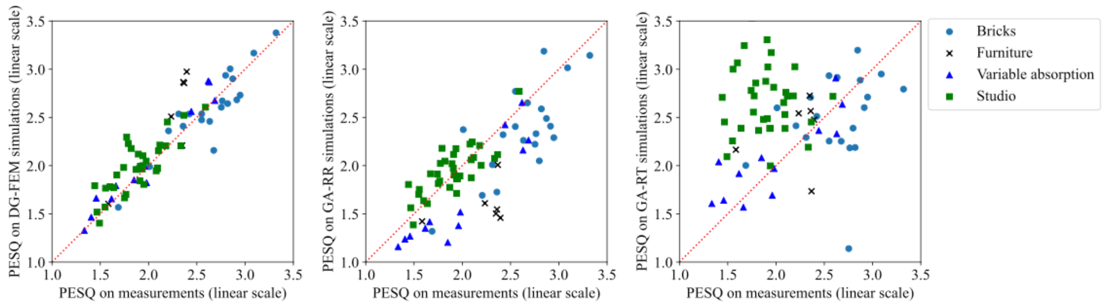
We use these evaluation datasets to evaluate the performance of three audio signal processing and ML algorithms

- **WPE-based dereverberation:**
  Traditional signal processing
  T. Yoshioka and T. Nakatani, “Generalization of multi-channel linear prediction methods for blind MIMO impulse response shortening,” IEEE Trans. Audio Speech Lang. Process., vol. 20, no. 10, pp. 2707–2720, 2012.

- **Neural network for speaker-distance estimation (SDE):**
  Pre-trained checkpoint, trained with measured data
  M. Neri, A. Politis, D. A. Krause, M. Carli, and T. Virtanen, “Speaker distance estimation in enclosures from single-channel audio,” IEEE/ACM Trans. Audio Speech Lang. Process., vol. 32, pp. 2242–2254, 2024.

- **Diffusion-based neural network for speech dereverberation:**
  Pre-trained checkpoint provided by authors
  J. Richter, S. Welker, J.-M. Lemercier, B. Lay, and T. Gerkmann, “Speech enhancement and dereverberation with diffusion-based generative models,” IEEE/ACM Trans. Audio Speech Lang. Process., vol. 31, pp. 2351–2364, 2023.

DG-FEM simulations provide the same insights into algorithm performance as measurements



(a) WPE-based dereverberation algorithm

DG-FEM simulations provide the same insights into algorithm performance as measurements

	SDE (ML)		Dereverberation (ML)						Dereverberation (DSP)					
	Distance est. error		PESQ		ESTOI		SI-SDR		PESQ		ESTOI		SI-SDR	
	$\rho \uparrow$	RMSE $\downarrow$	$\rho \uparrow$	RMSE $\downarrow$	$\rho \uparrow$	RMSE $\downarrow$	$\rho \uparrow$	RMSE $\downarrow$	$\rho \uparrow$	RMSE $\downarrow$	$\rho \uparrow$	RMSE $\downarrow$	$\rho \uparrow$	RMSE $\downarrow$
DG-FEM	**0.76**	**0.16**	**0.92**	**0.22**	**0.91**	**0.05**	**0.75**	**3.09**	**0.91**	**0.21**	**0.90**	**0.03**	**0.73**	**2.35**
GA-RR	0.59	0.24	0.68	0.51	0.70	0.11	0.61	3.68	0.77	0.35	0.76	0.06	0.57	2.86
GA-RT	0.51	0.25	0.28	0.81	0.45	0.17	0.23	5.71	0.14	0.70	0.56	0.07	0.10	4.19

Setting up diverse room-acoustic datasets
Example workflow

Code examples

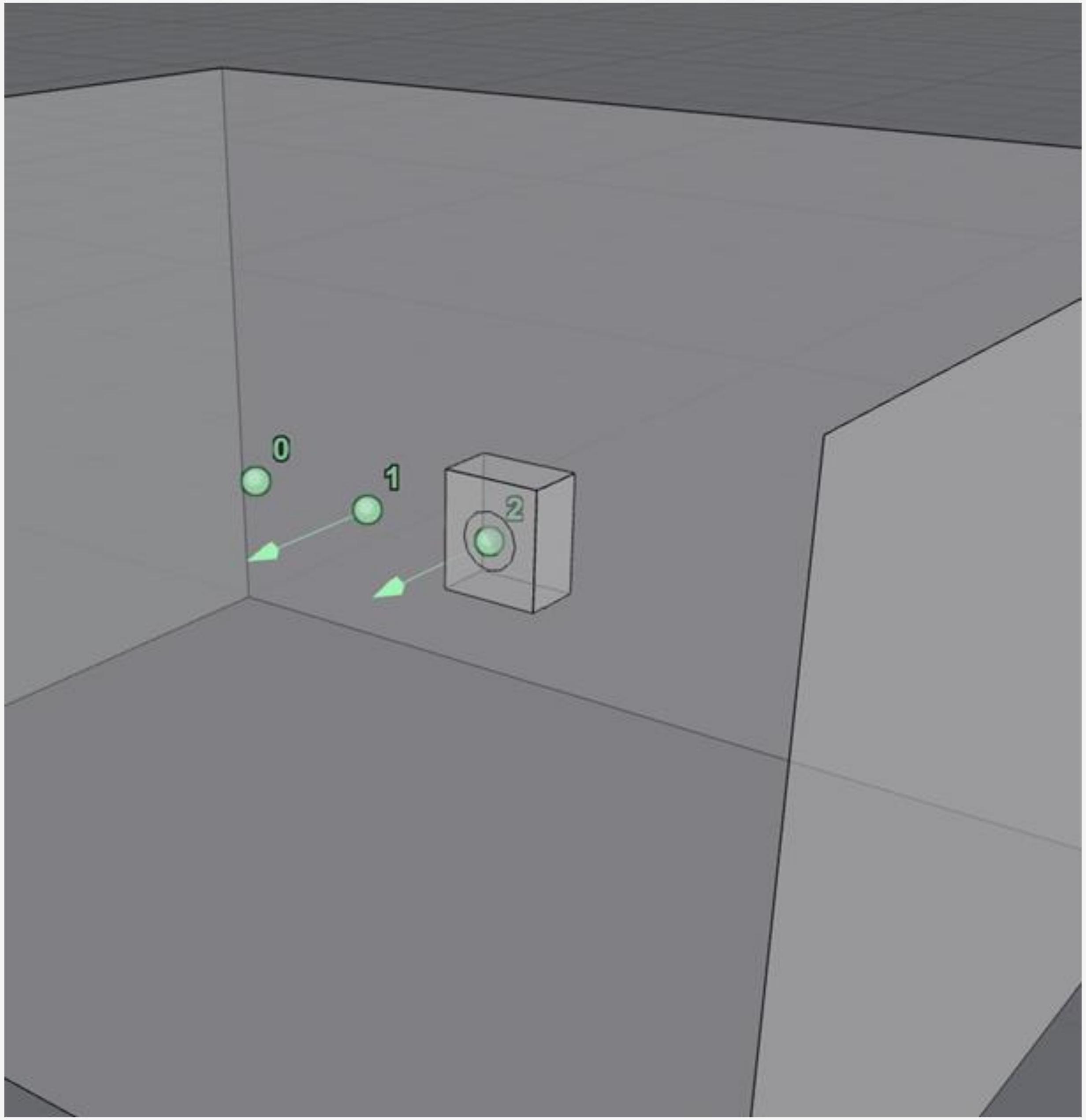
www.treble.tech/
aes-sdk-tutorial

Source and receiver modeling

Modeling sources as they are in real-life

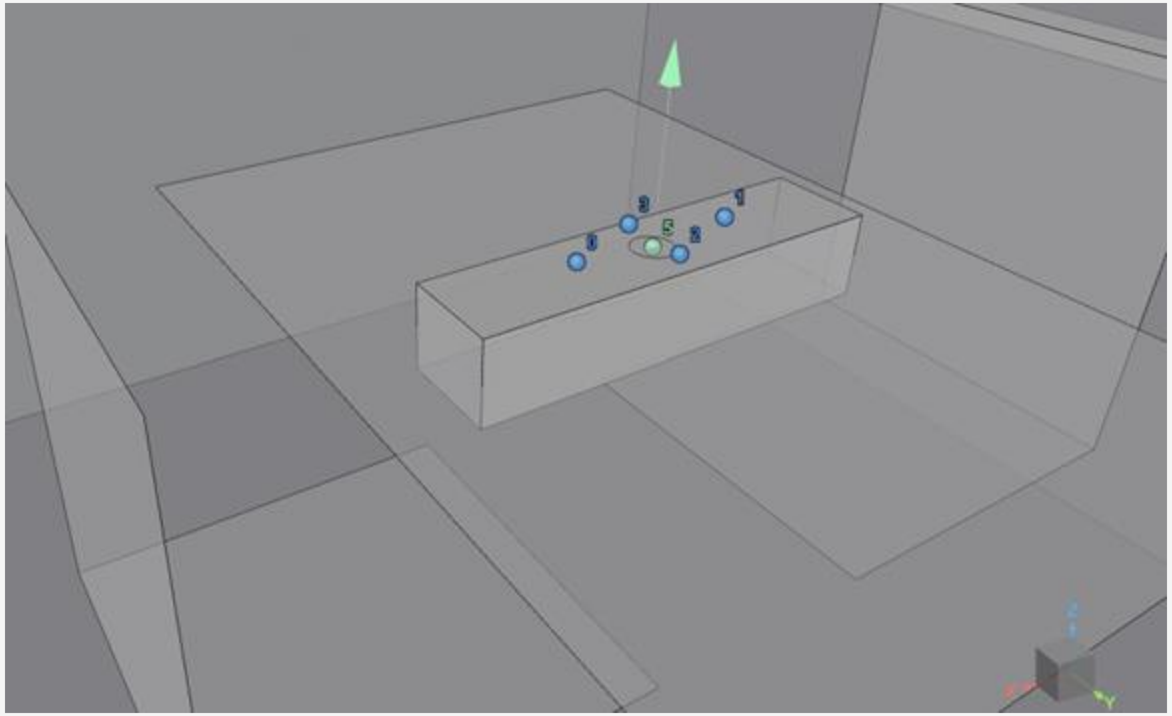
Introduction to sources

- Point sources
  - Omni-directional
  - Directional Point Sources
- Boundary Velocity sources
  - Prescribed to an existing layer
  - Injected submodel



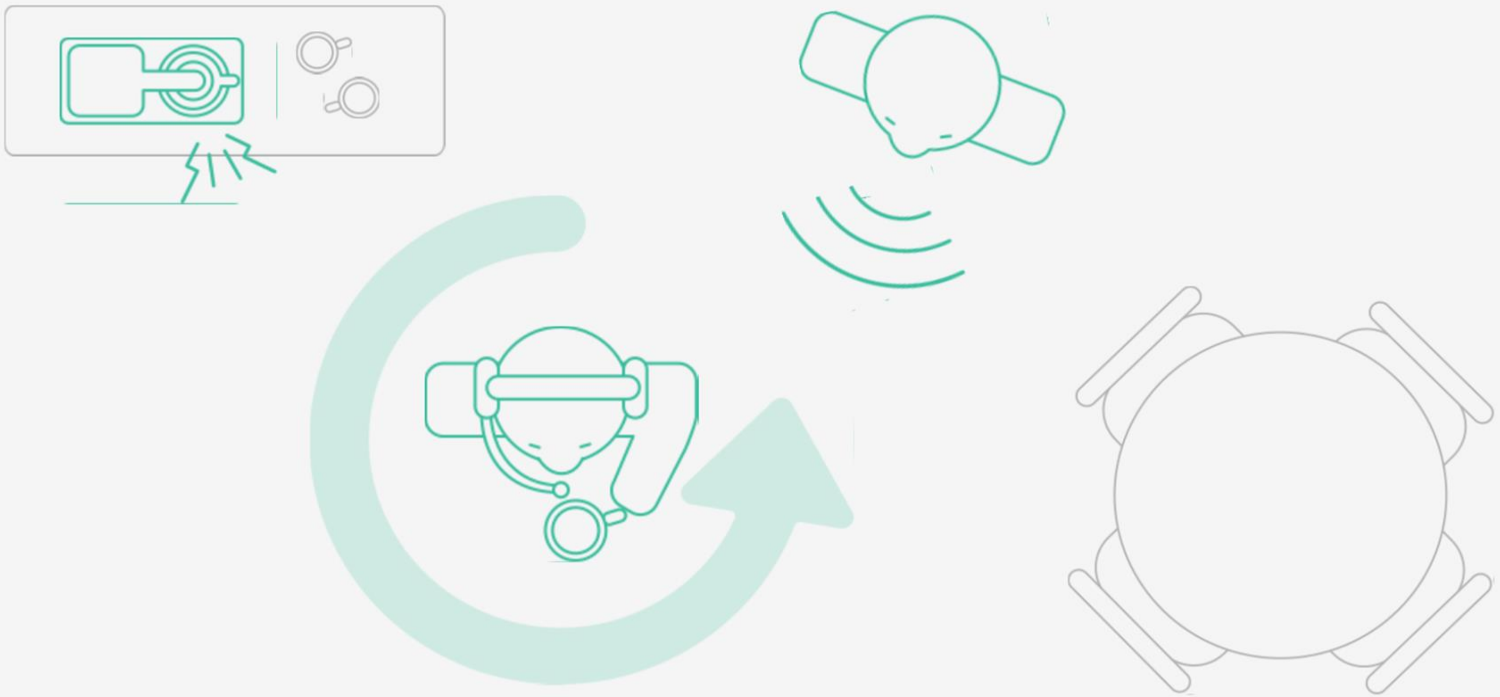
Boundary Velocity Sources

- A boundary velocity source can be arbitrarily close to objects and surfaces
  - Addresses shortcomings of point sources
- Precise simulations of both near and far-field
  - The effect of the source geometry is captured well
  - Enables on-board microphones close to the device
  - High frequency directivity patterns are captured well



Receiver Modeling

- Mono receivers
- Spatial receivers
  - Up to 32nd order ambisonics
  - Enables vast post-processing capabilities
  - Perceptual evaluation



Data augmentation and device-specific IRs

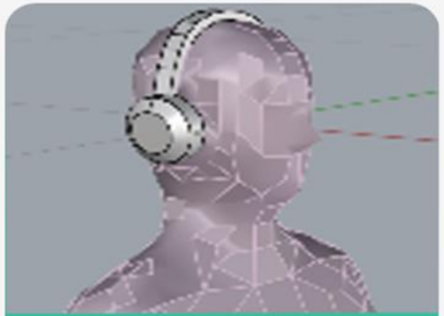
Static data

Initialize Simulation

Spatial RIRs

Augment spatial RIRs with DRTF/HRTF

Load DRTF/HRTF from TSDK library



Import device CAD and simulate DRTF

Import a measured DRTF/HRTF

Device Render

Rotate or change device

Demo

Directive sources and receivers in PyRoomAcoustics

```
import pyroomacoustics as pra

room = pra.ShoeBox(
    p=[5, 3, 3],
    materials=pra.Material(energy_absorption),
    fs=16000,
    max_order=40,
)

# add a cardioid source
dir = pra.directivities.Cardioid(DirectionVector(azimuth=-65, colatitude=90) , gain=1.0)
room.add_source([3.75, 2.13, 1.41], directivity=dir)
```

- Only ISM can be used with directive sources/receivers; ray tracer not available
- Directivity is baked into the simulation → complete re-run required when new orientation is desired
- No Ambisonics output available

Conclusions

treble

Enabling a better sounding world

Main takeaways from today

- There are two dimensions to downstream performance
  - ML model improvements
  - Input data
- There are two major components for data sets
  - Size
  - Quality
- The quality of the data is linked to how well it represents the real world
- Significant improvements to downstream performance can be gained by using higher quality data
- Measurements can be replaced by simulations for model evaluation

If interest: hands-on exercises after the tutorial

Tutorial material and SDK access:

www.treble.tech/aes-sdk-tutorial

treble

Enabling a better sounding world