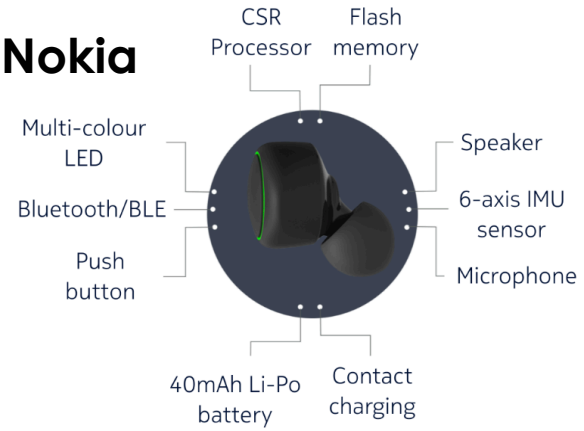
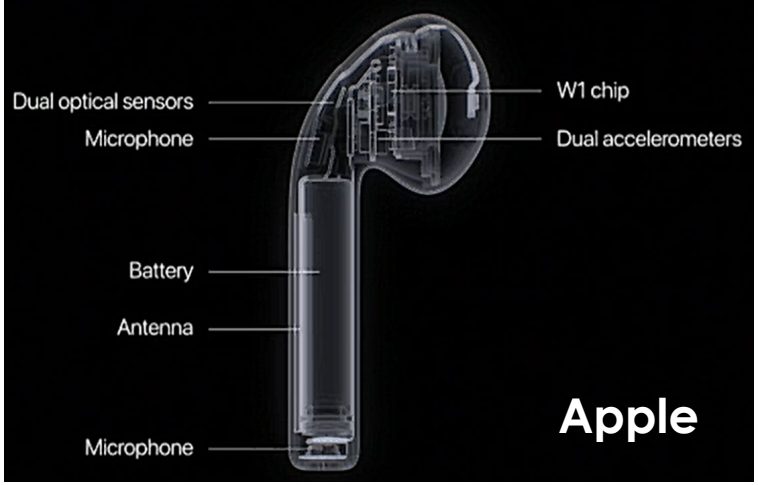
Earphones: The Next Computing Platform after Smartphones

Romit Roy Choudhury

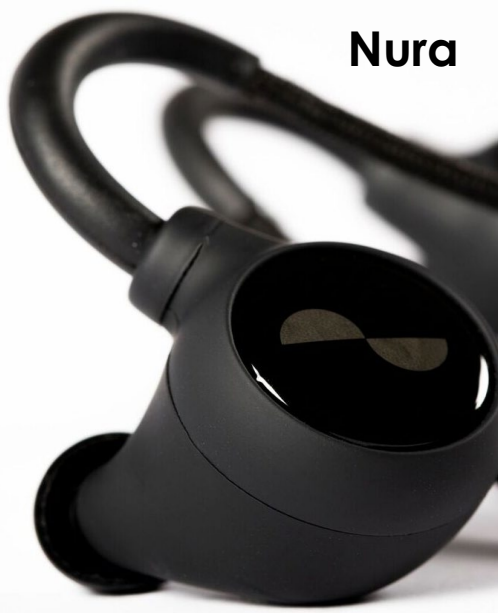
ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Nokia





Nura



Bose

Samsung

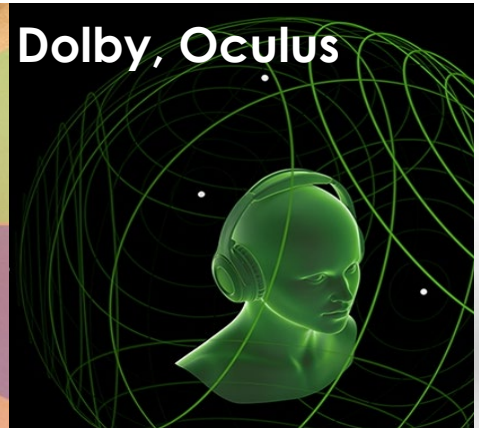
Swipe
Volume Up/Down

Tap & Hold
Launch your default Voice Command setting (Bixby, Google Voice, S Voice) Ambient Sound mode, Reject a call

Double Tap & Hold
Next playlist

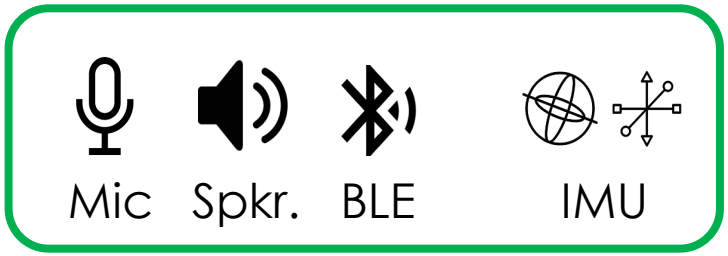
QUALCOMM

The Future of Personal Audio

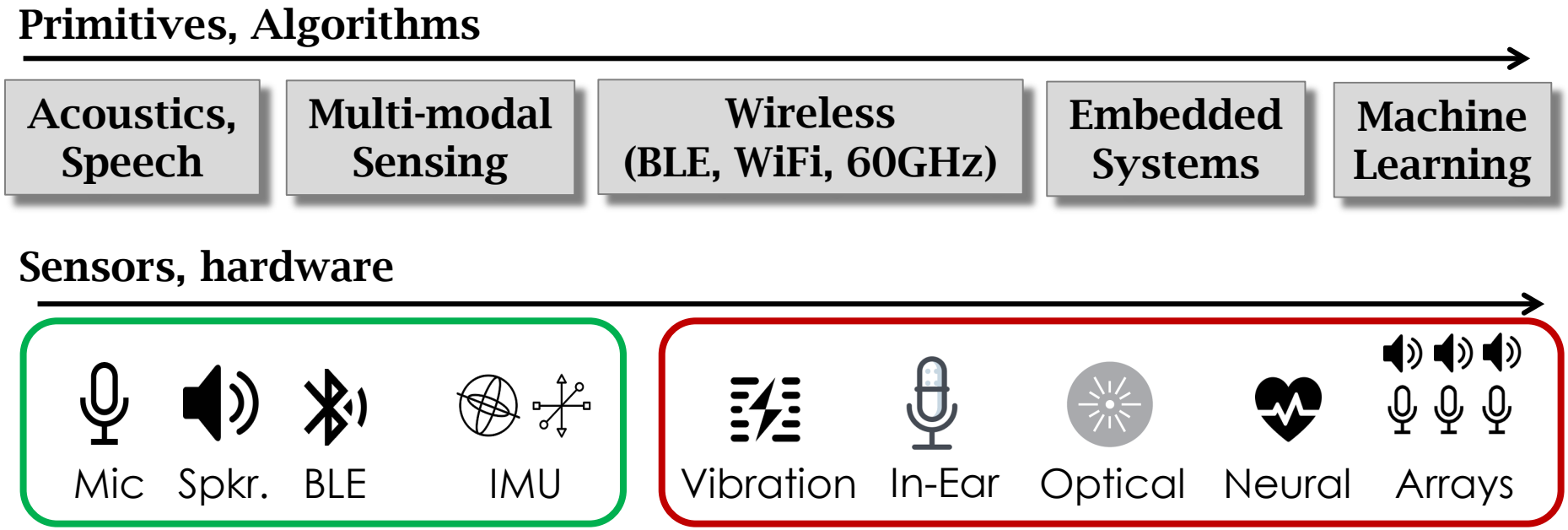


… just the tip of the iceberg

Sensors, hardware



Sensors, hardware


Primitives, Algorithms
Acoustics, Speech
Multi-modal Sensing
Wireless (BLE, WiFi, 60GHz)
Embedded Systems
Machine Learning
Sensors, hardware
Mic
Spkr.
BLE
IMU
Vibration
In-Ear
Optical
Neural
Arrays

App Store

Sound+ 3D

Acoustic

AR

indoor location

Security

Health Sense

HEALTHY EATING

Translate

Primitives, Algorithms

Acoustics, Speech | Multi-modal Sensing | Wireless (BLE, WiFi, 60GHz) | Embedded Systems | Machine Learning

Sensors, hardware

Mic | Spkr. | BLE | IMU

Vibration | In-Ear | Optical | Neural | Arrays

2005

2015

2018

2023

But why Earables? **Why now?**

But why Earables? Why now?

1 **Reading and typing, a cognitive re-focus**

Voice + hearing seamless for humans

But why Earables? Why now?

1 Reading and typing, a cognitive re-focus

Voice + hearing seamless for humans

2 Phone, watch, fitbit = quantified lower body

Earables gateway to upper body, head

But why Earables? Why now?

1 Reading and typing, a cognitive re-focus

Voice + hearing seamless for humans

2 Phone, watch, fitbit = quantified lower body

Earables gateway to upper body, head

3 Socially well accepted (unlike, Google Glass)

Takes off one important risk factor

But what are **hurdles** …

show-stoppers?

Many **Hurdles** …

Many **Hurdles** …

1 Energy

Many **Hurdles** …

1 Energy

2 Discomfort

Many **Hurdles** …

1 Energy

2 Discomfort

3 Privacy / Shy

Many **Hurdles** …

1 Energy

2 Discomfort

3 Privacy / Shy

4 Health, radiation

Many **Hurdles** …but not **Insurmountable**

1 Energy

2 Discomfort

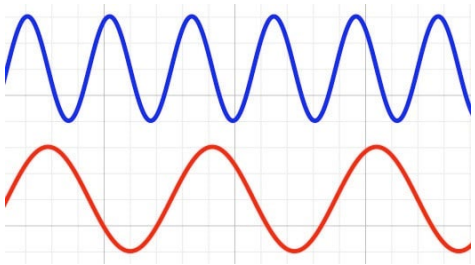
3 Privacy / Shy

4 Health, radiation

Contact Mic.

Hollow

Whisper Zone



Low Frequency

We are building a

software library

for Earable Computing

With particular interest in the more challenging problems …

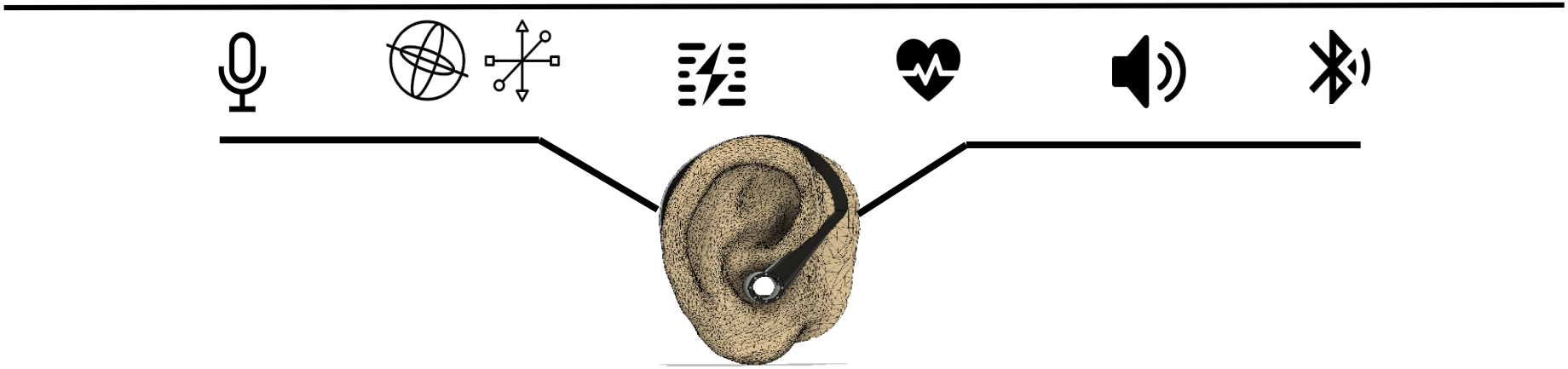
We are building a

software library

for Earable Computing

- Voice assistants
- Spatial acoustics
- Motion tracking
- Beamforming to Speech
- Noise Cancellation
- Face/Jaw Activity (eating, drinking)
- In-Ear Anatomy (health, security)

Sensor and Hardware Layer



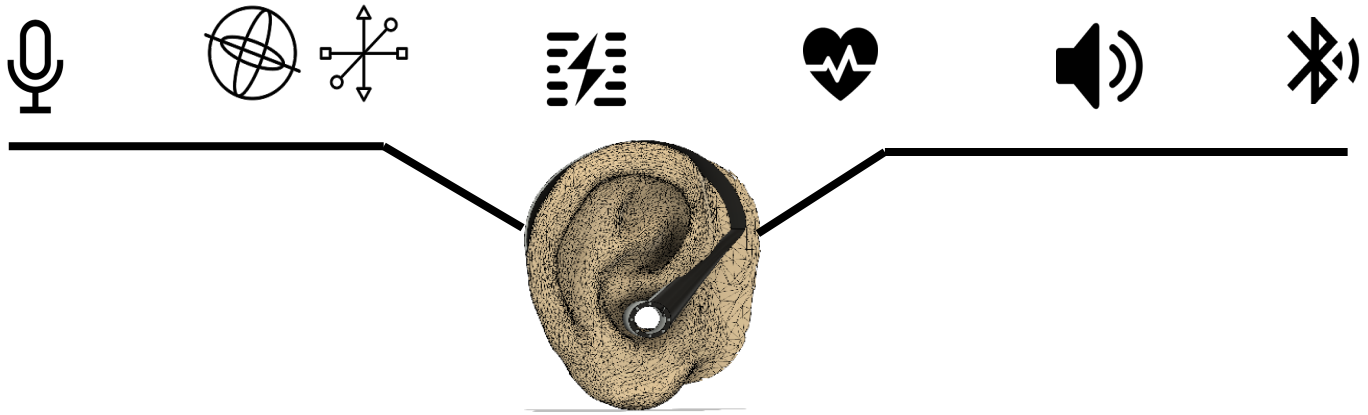
We are building a

software library

for Earable Computing

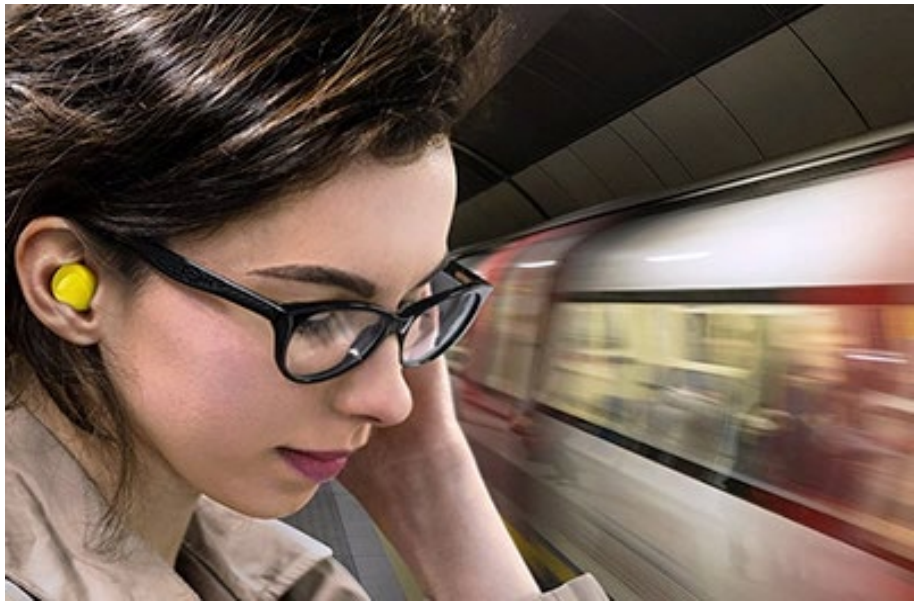
- **Voice assistants**
- Spatial acoustics
- Motion tracking
- Beamforming to Speech
- Noise Cancellation
- Face/Jaw Activity (eating, drinking)
- In-Ear Anatomy (health, security)

Sensor and Hardware Layer



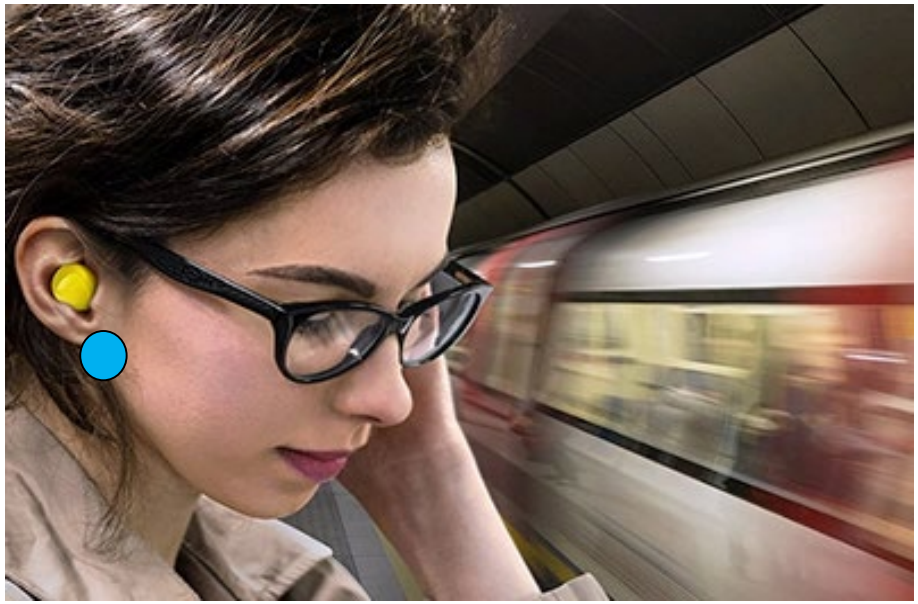
Voice assistants

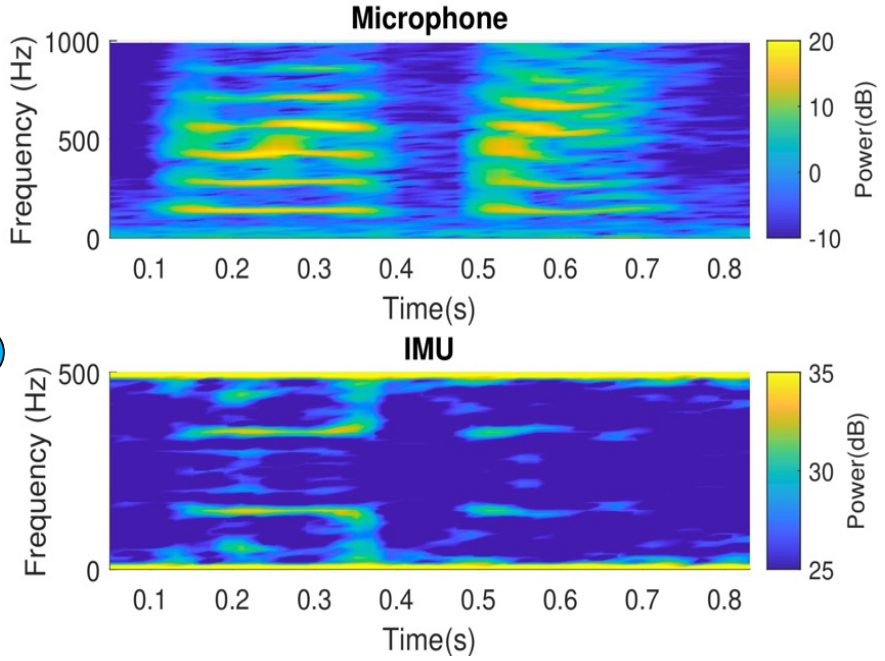
Low SNR Speech Recognition **(whisper)**



Voice assistants

Low SNR Speech Recognition **(whisper)**





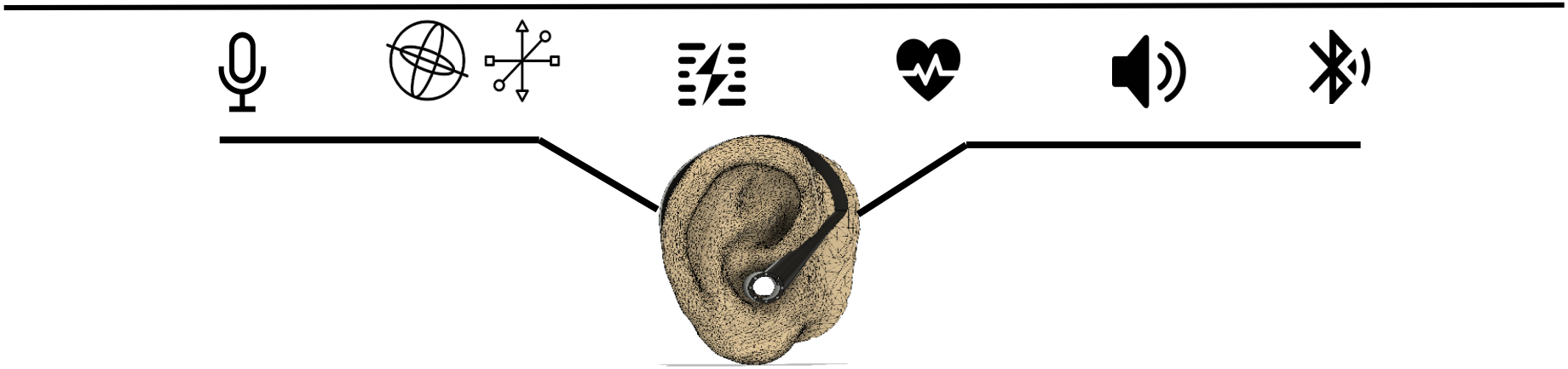
$$h_{air}(t) * V(t) + N_{backgr}(t)$$

$$\left(f_{alias}\left(h_{solid}(t) * V(t)\right)\right) + n_{thermal}(t)$$

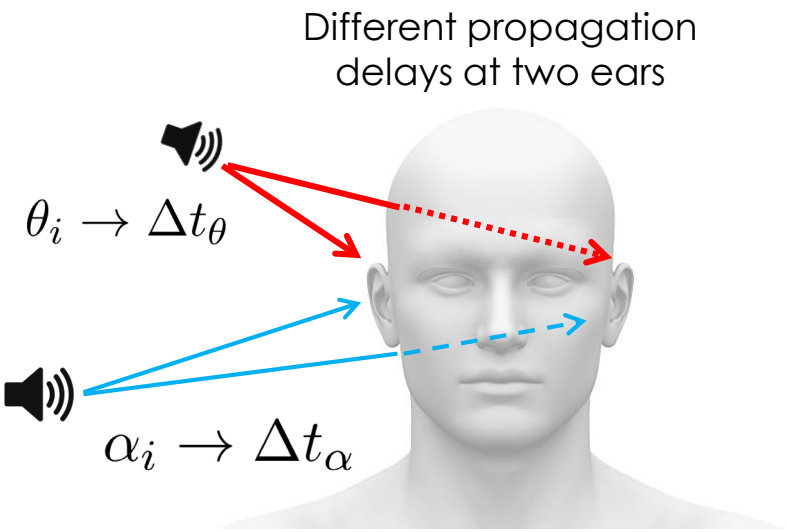
→ **Whisper Decoder** → $\hat{V}(t)$ → ASR

- Voice assistants
- Spatial acoustics
- Motion tracking
- Beamforming to Speech
- Noise Cancellation
- Face/Jaw Activity (eating, drinking)
- In-Ear Anatomy (health, security)

Sensor and Hardware Layer

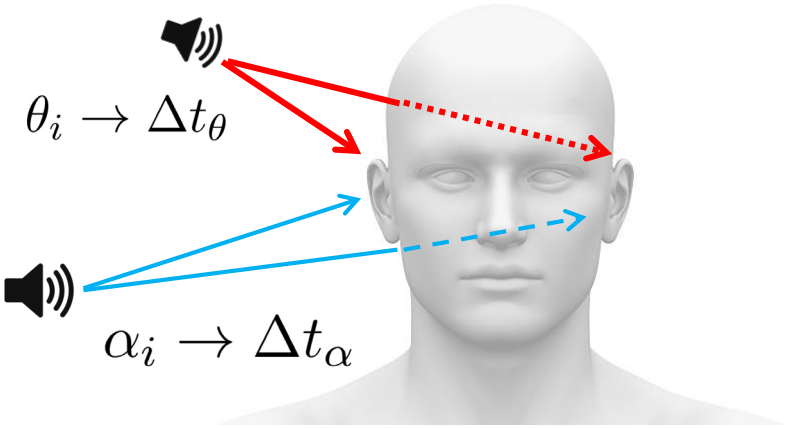


Spatial Acoustics



Spatial Acoustics

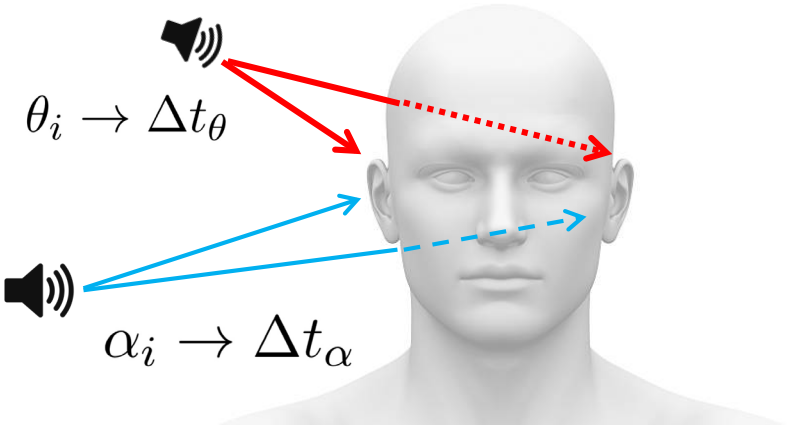
Different propagation delays at two ears



Inject **artificial delays** as if sounds coming from a specific direction

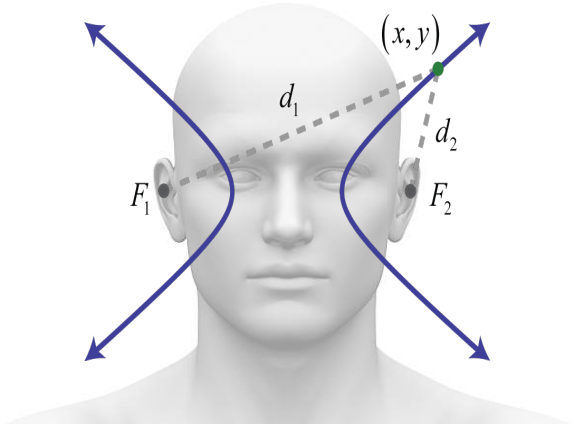
Spatial Acoustics

Different propagation delays at two ears



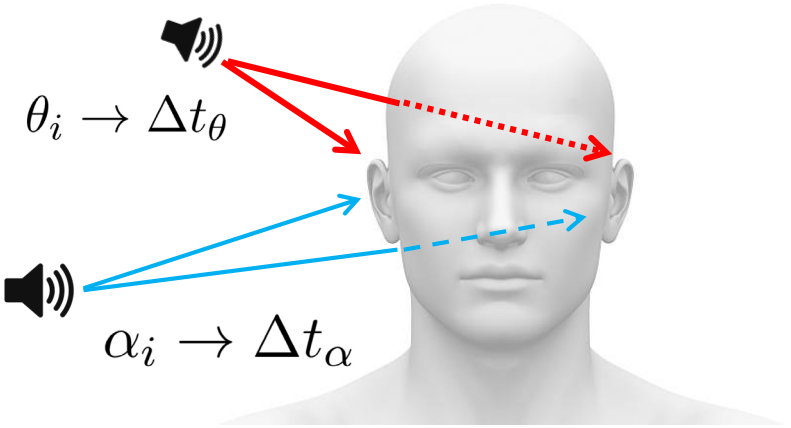
Inject **artificial delays** as if sounds coming from a specific direction

But delays are not unique along **hyperbola**



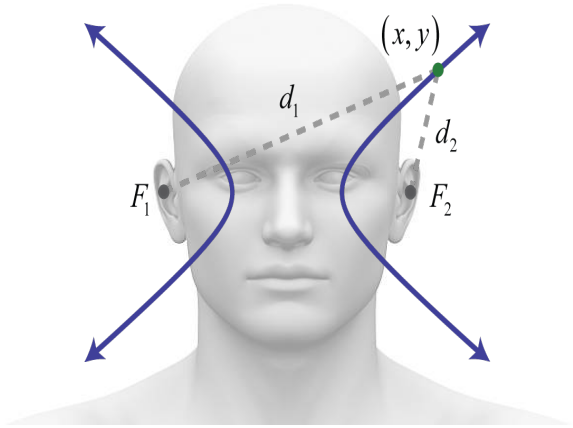
Spatial Acoustics

Different propagation delays at two ears

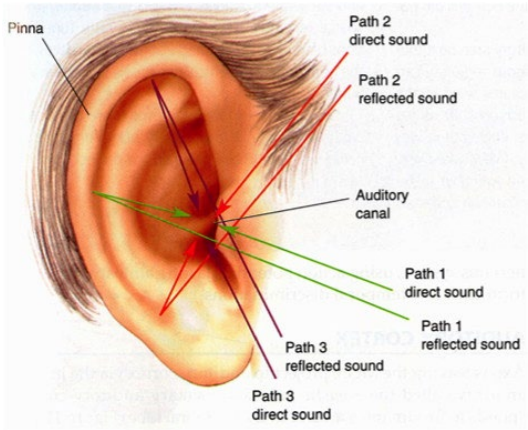


Inject **artificial delays** as if sounds coming from a specific direction

But delays are not unique along **hyperbola**

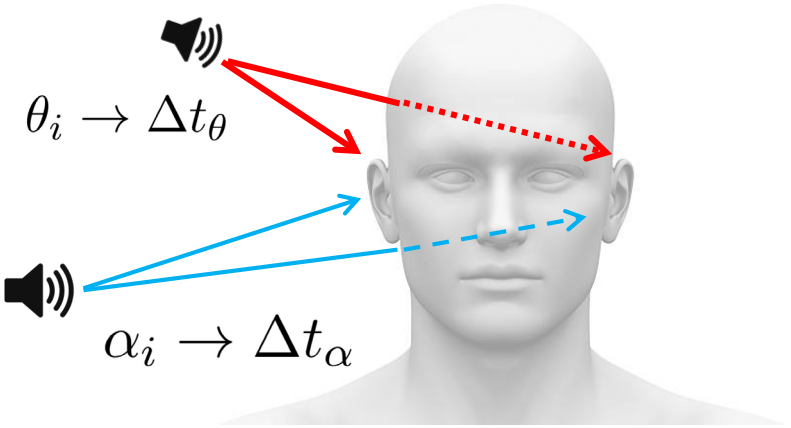


Personal Transfer Function



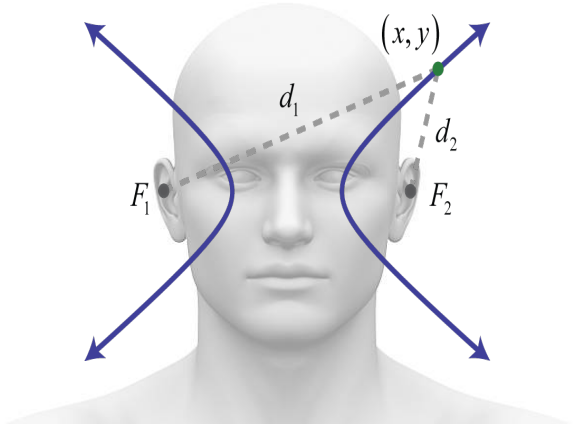
Spatial Acoustics

Different propagation delays at two ears

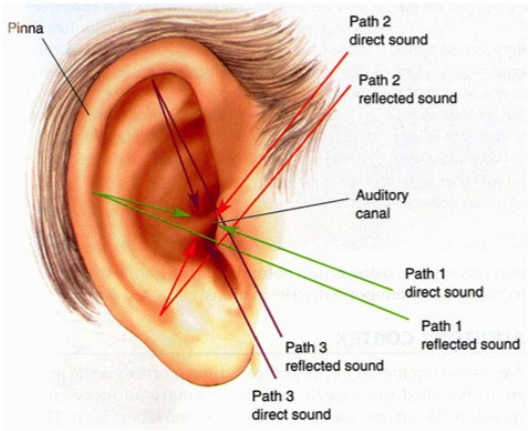


Inject **artificial delays** as if sounds coming from a specific direction

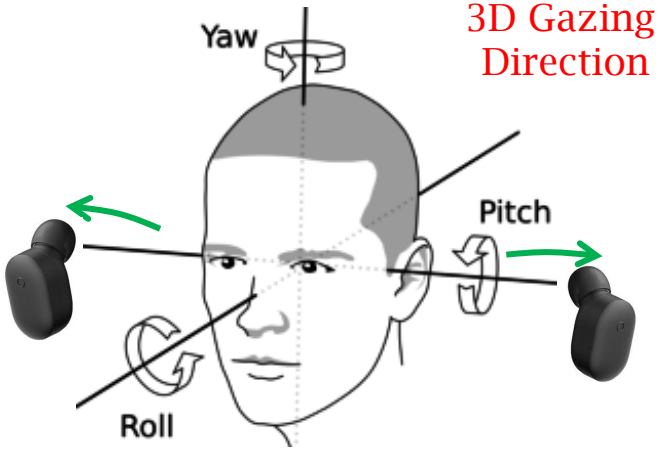
But delays are not unique along **hyperbola**



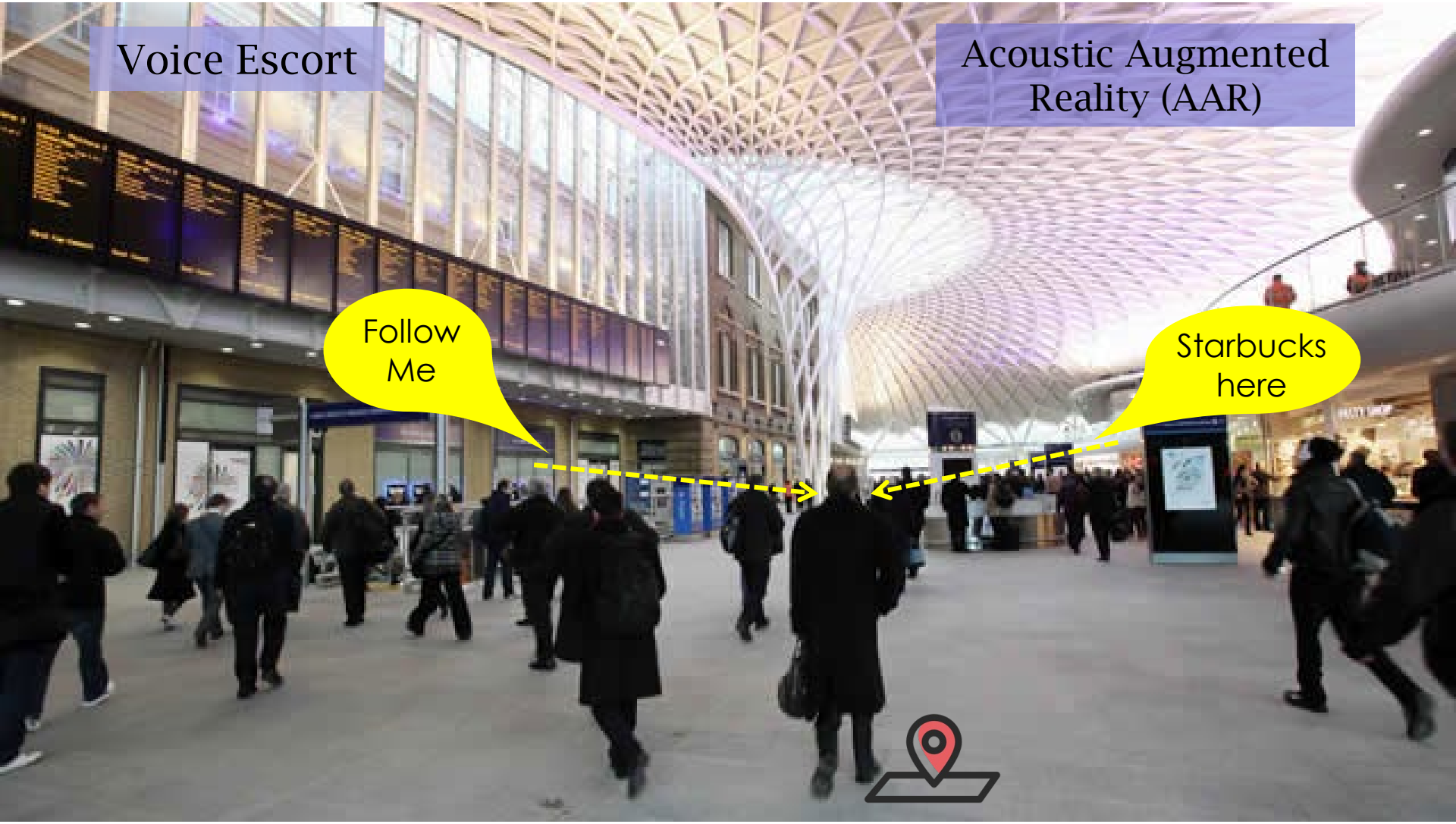
Personal Transfer Function



3D Gazing Direction



Spatial Acoustics



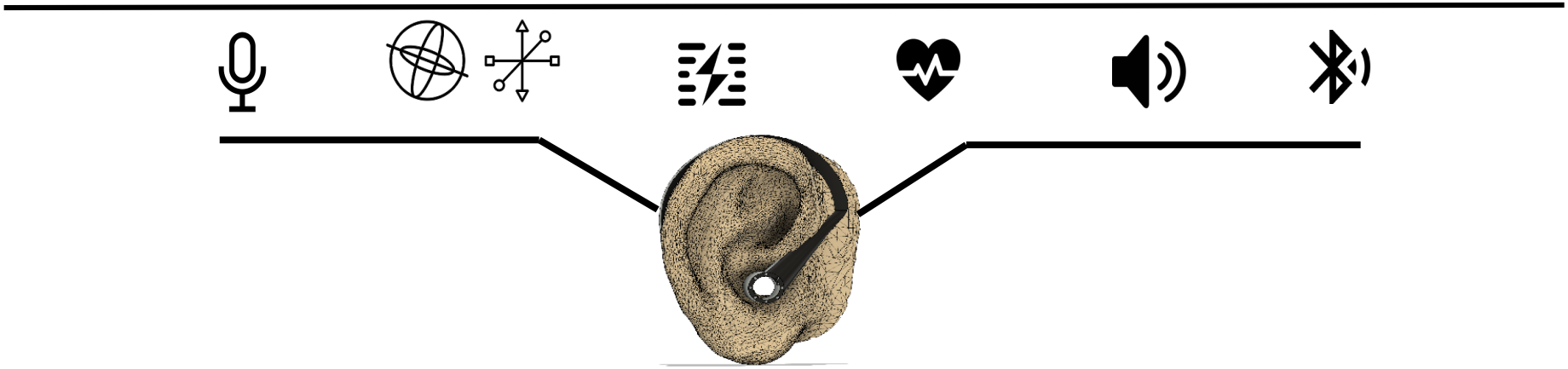
We are building a

software library

for Earable Computing

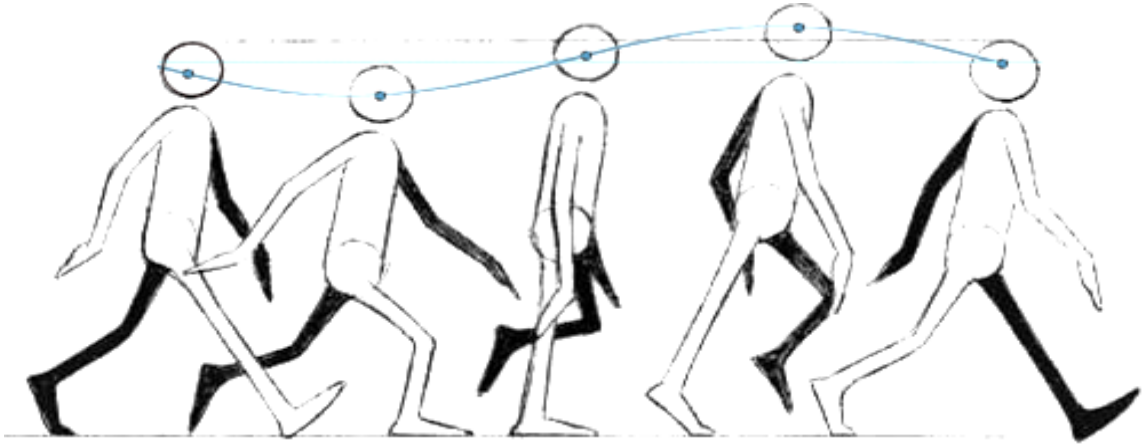
- Voice assistants
- Spatial acoustics
- Motion tracking
- Beamforming to Speech
- Noise Cancellation
- Face/Jaw Activity (eating, drinking)
- In-Ear Anatomy (health, security)

Sensor and Hardware Layer



Motion tracking

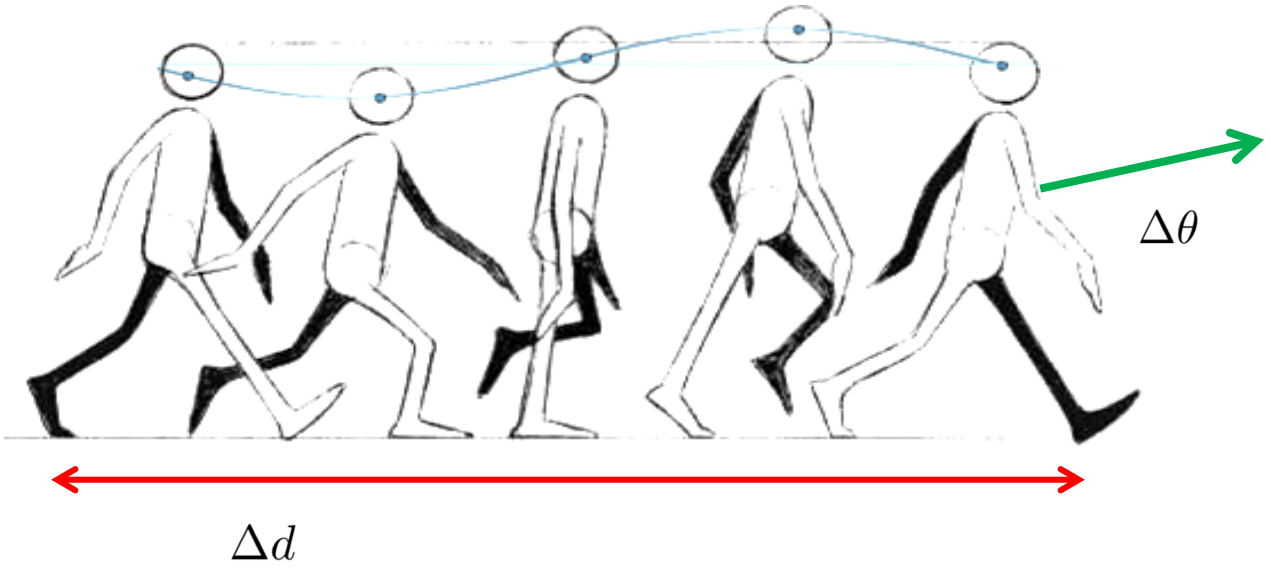
Indoor localization



Motion tracking

Indoor localization

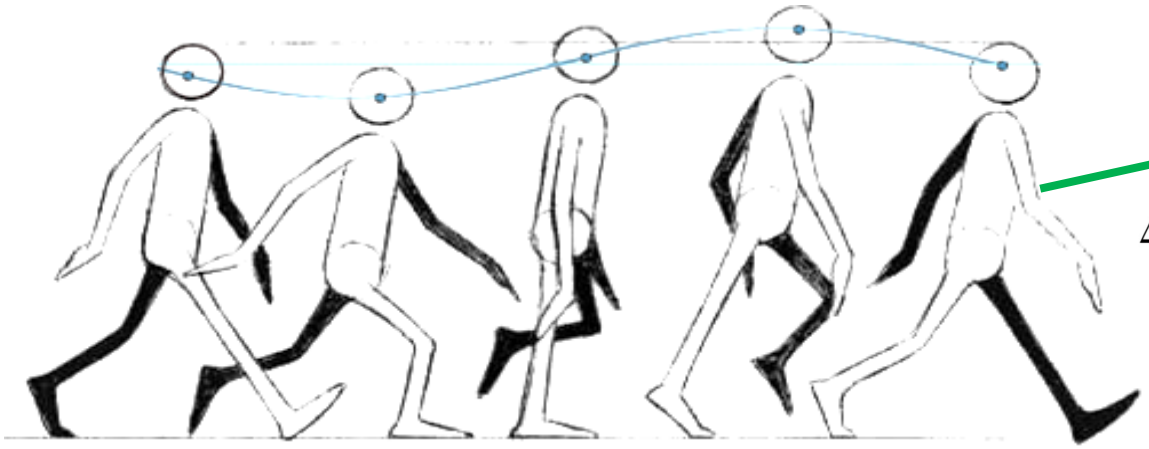
$$Location(L_t) = L_0 + \sum_{steps} f(\Delta d, \Delta\theta)$$



Motion tracking

Indoor localization

$$Location(L_t) = L_0 + \sum_{steps} f(\Delta d, \Delta\theta)$$



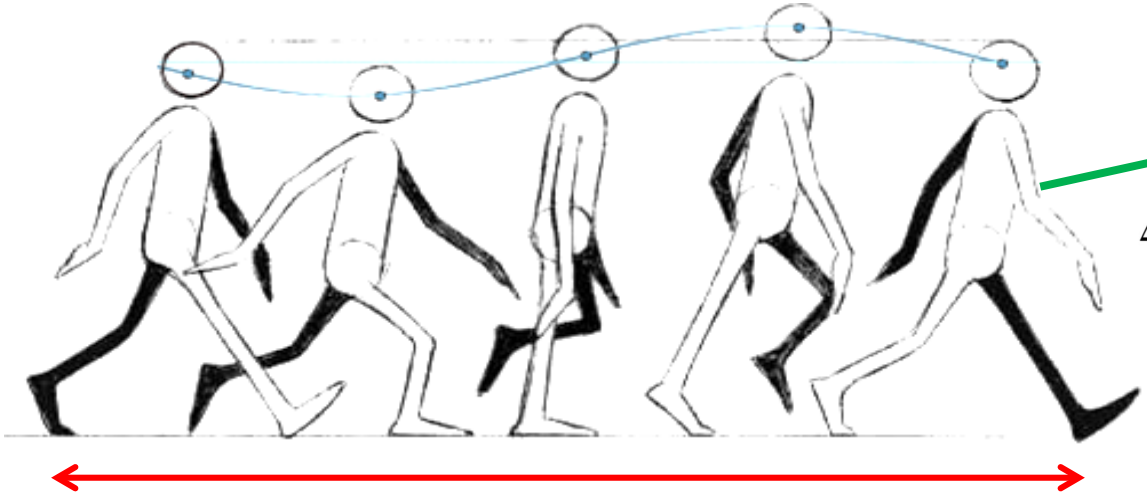
$\Delta\theta = f\left(gyroscope, compass\right)$

$\Delta d = f\left(accelerometer, compass\right)$

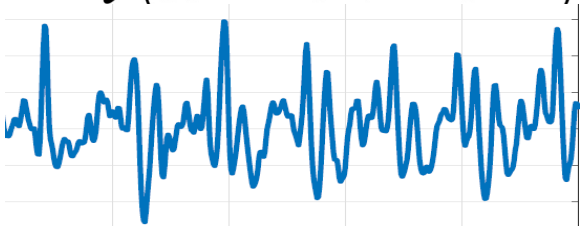
Motion tracking

Indoor localization

$$Location(L_t) = L_0 + \sum_{steps} f(\Delta d, \Delta\theta)$$



$$\Delta\theta = f\left(gyroscope, compass\right)$$

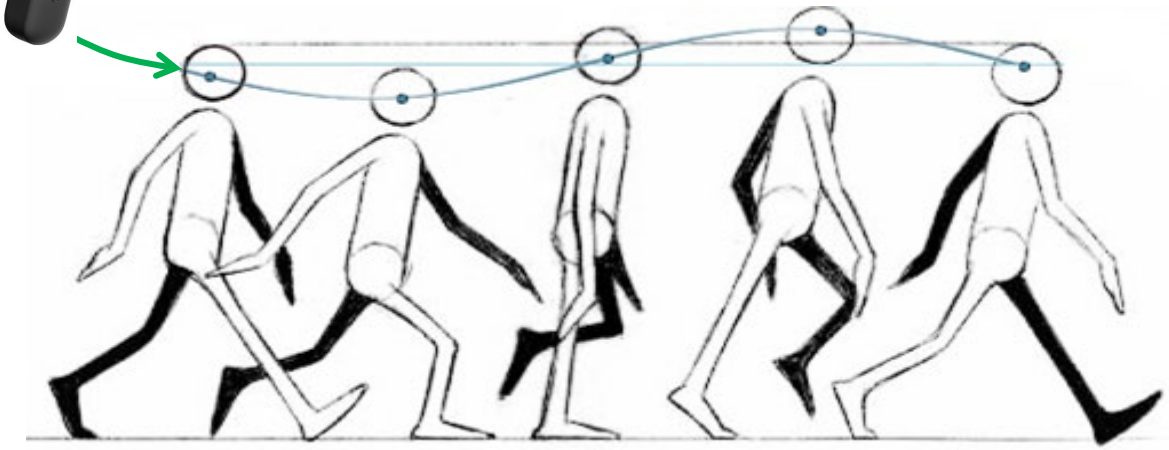


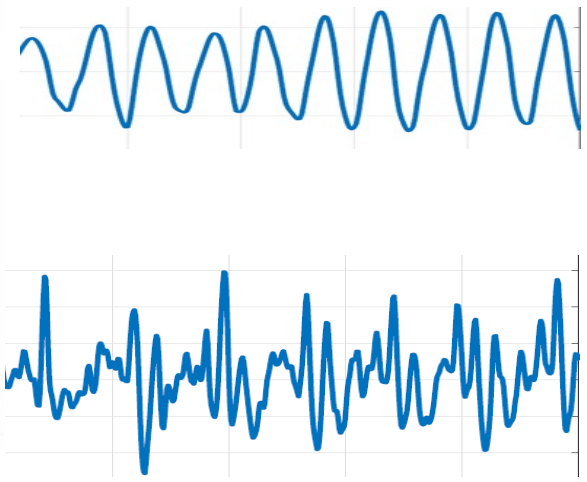
$$\Delta d = f\left(accelerometer, compass\right)$$

Motion tracking

Indoor localization

$$Location(L_t) = L_0 + \sum_{steps} f(\Delta d, \Delta\theta)$$

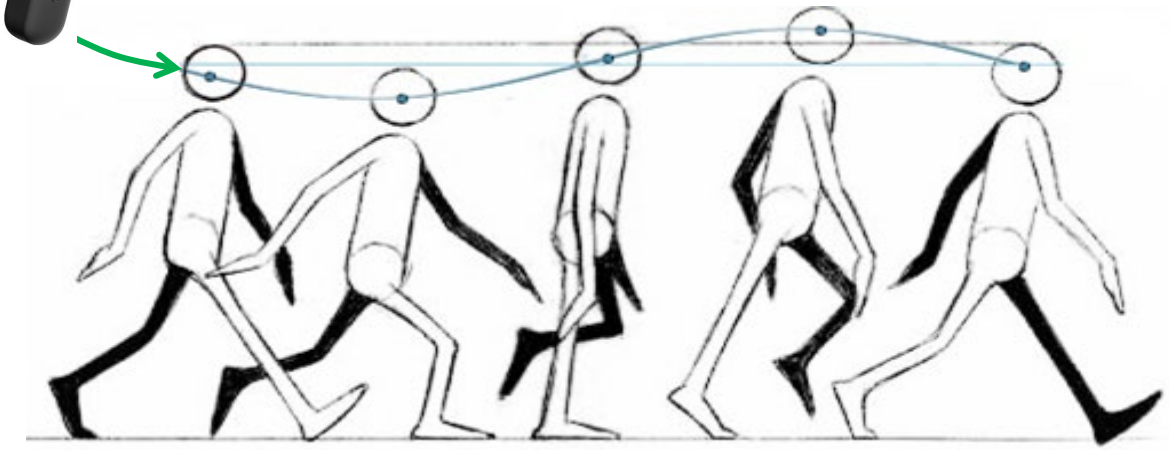


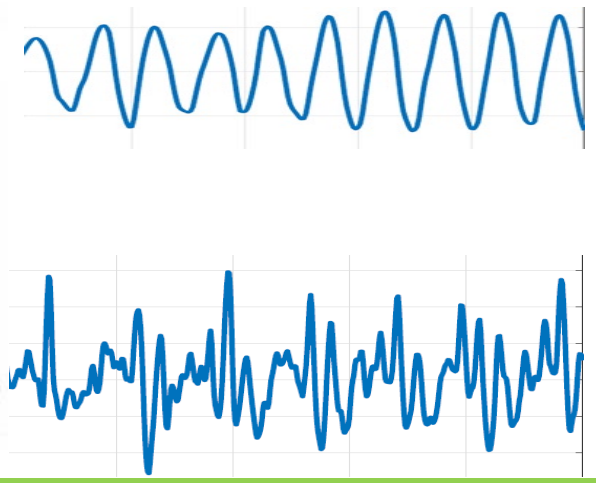


Motion tracking

Indoor localization

$$Location(L_t) = L_0 + \sum_{steps} f(\Delta d, \Delta\theta)$$



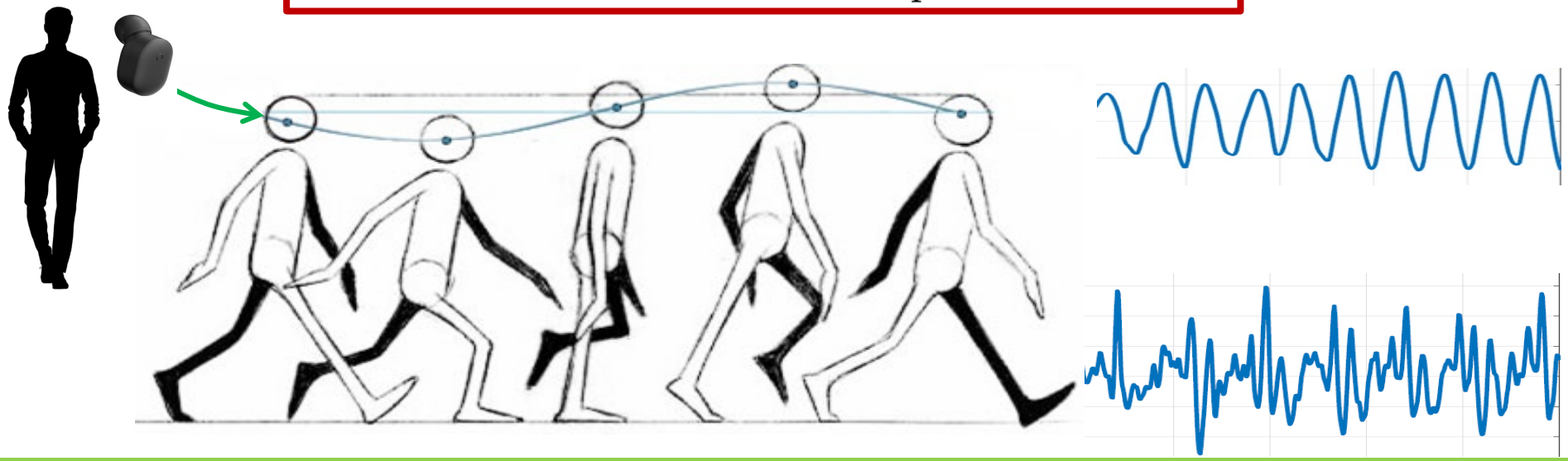


Body/skeleton is a natural low pass filter

Motion tracking

Indoor localization

$$Location(L_t) = L_0 + \sum_{steps} f(\Delta d, \Delta\theta)$$



Body/skeleton is a natural low pass filter

Lower body IMU (smartphone) →
Head IMU (earphone) →
Kinematic Motion models →

Location Tracking → $\langle Loc(t), Gaze(t) \rangle$

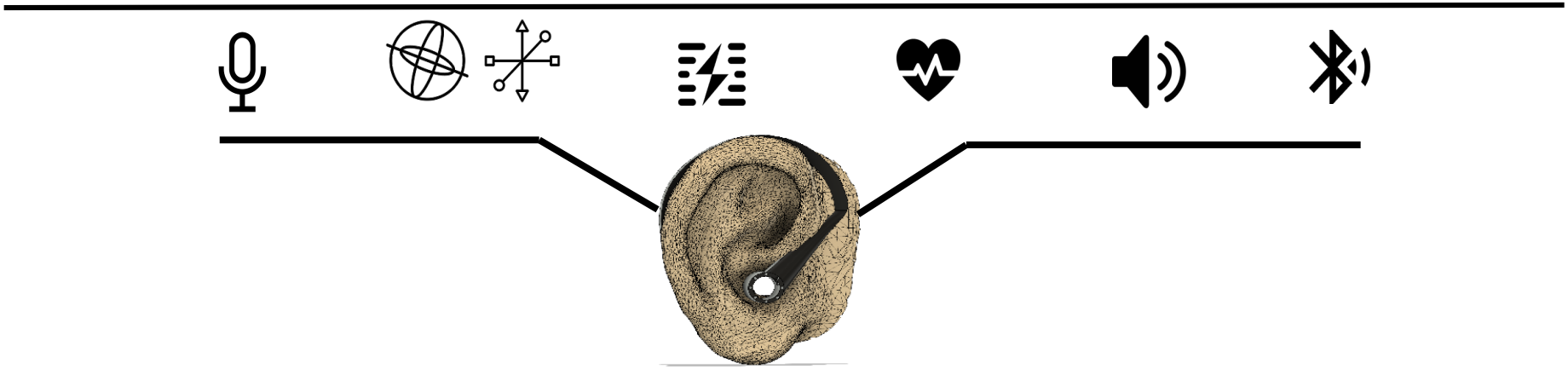
We are building a

software library

for Earable Computing

- Voice assistants
- Spatial acoustics
- Motion tracking
- **Beamforming to Speech**
- Noise Cancellation
- Face/Jaw Activity (eating, drinking)
- In-Ear Anatomy (health, security)

Sensor and Hardware Layer

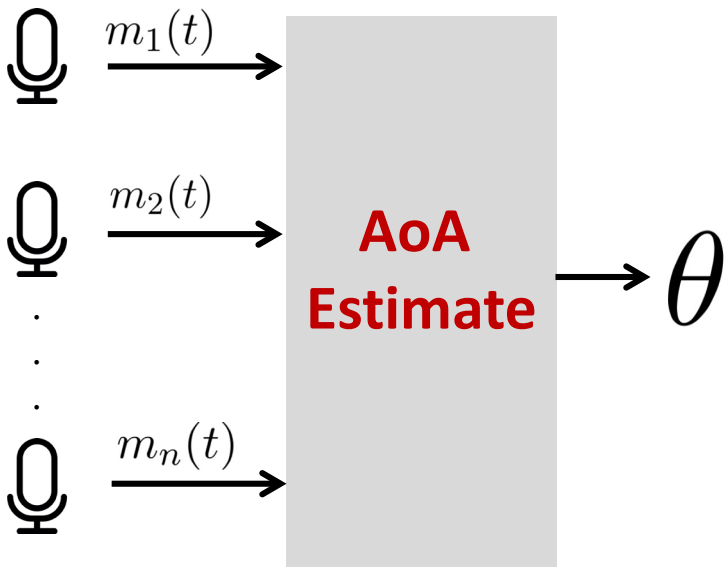


Beamforming to Speech

Estimate Angle of Arrival (AoA)

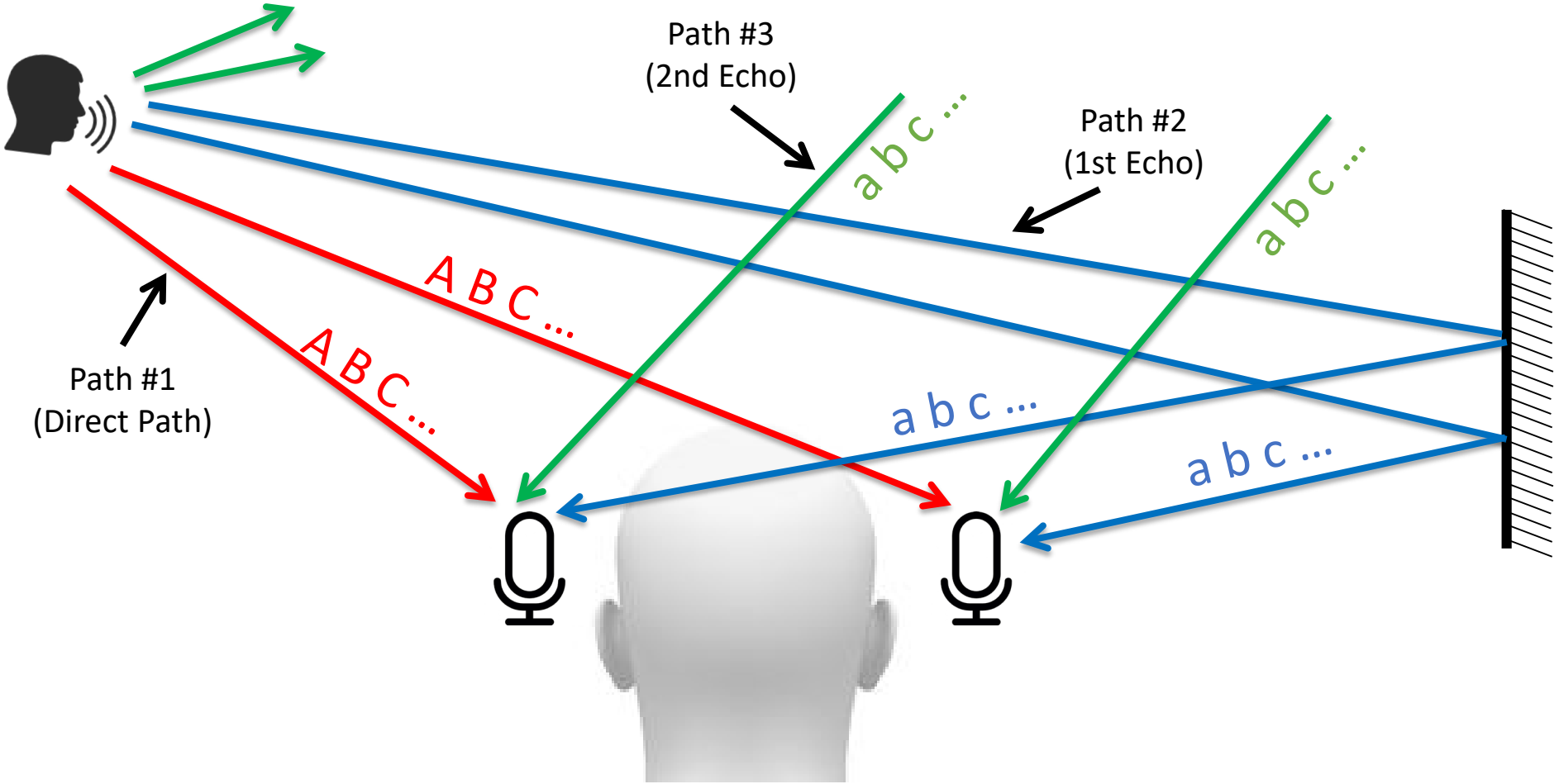
Beamforming to Speech

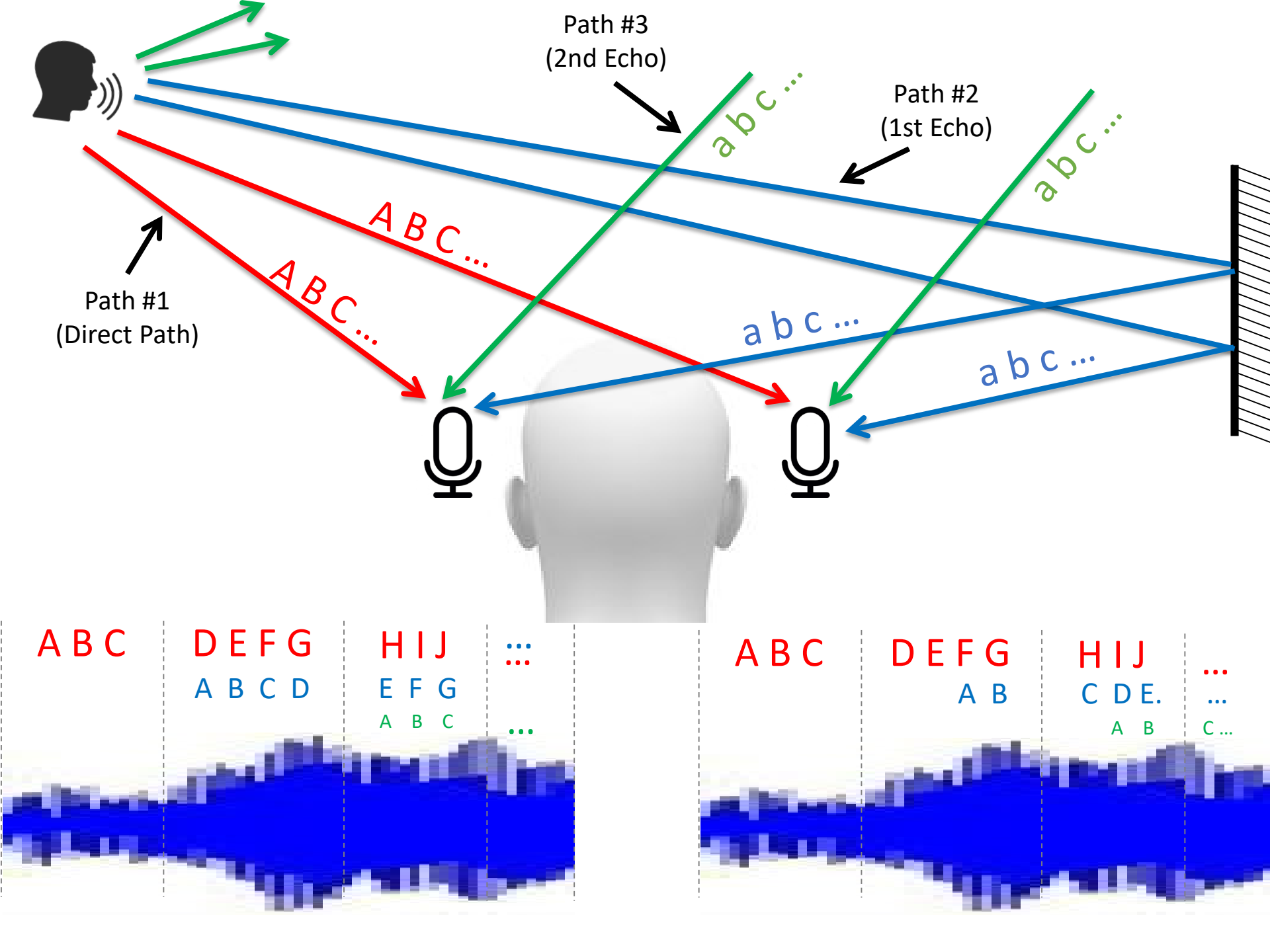
Estimate Angle of Arrival (AoA)



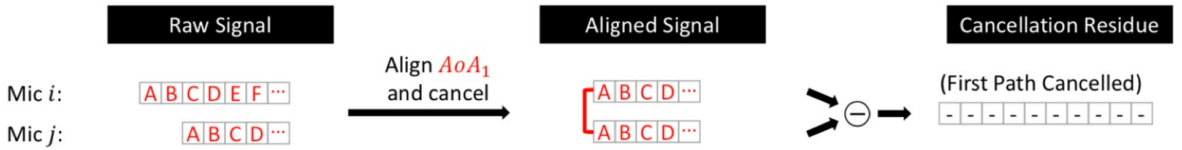
Beamforming to Speech

Estimate Angle of Arrival (AoA)

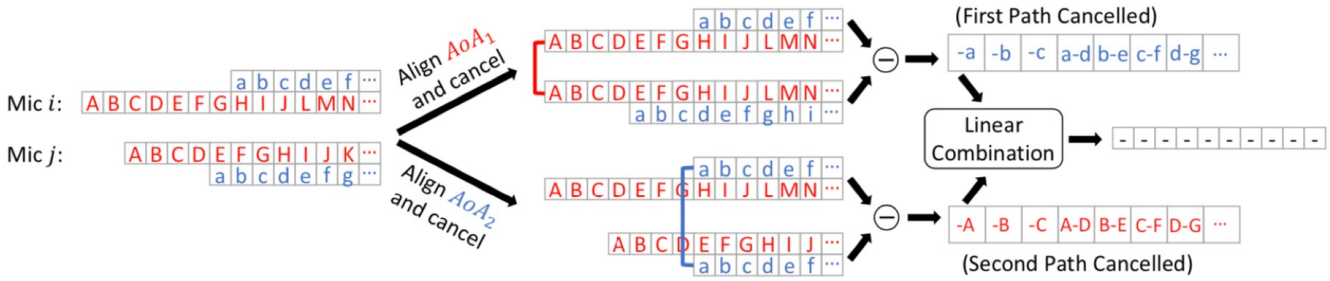



Path #3
(2nd Echo)
Path #2
(1st Echo)
Path #1
(Direct Path)
A B C …
A B C …
a b c …
a b c …
a b c …
a b c …
A B C
D E F G
A B C D
H I J
E F G
A B C
…
…
A B C
D E F G
A B
H I J
C D E.
A B
…
…
C …

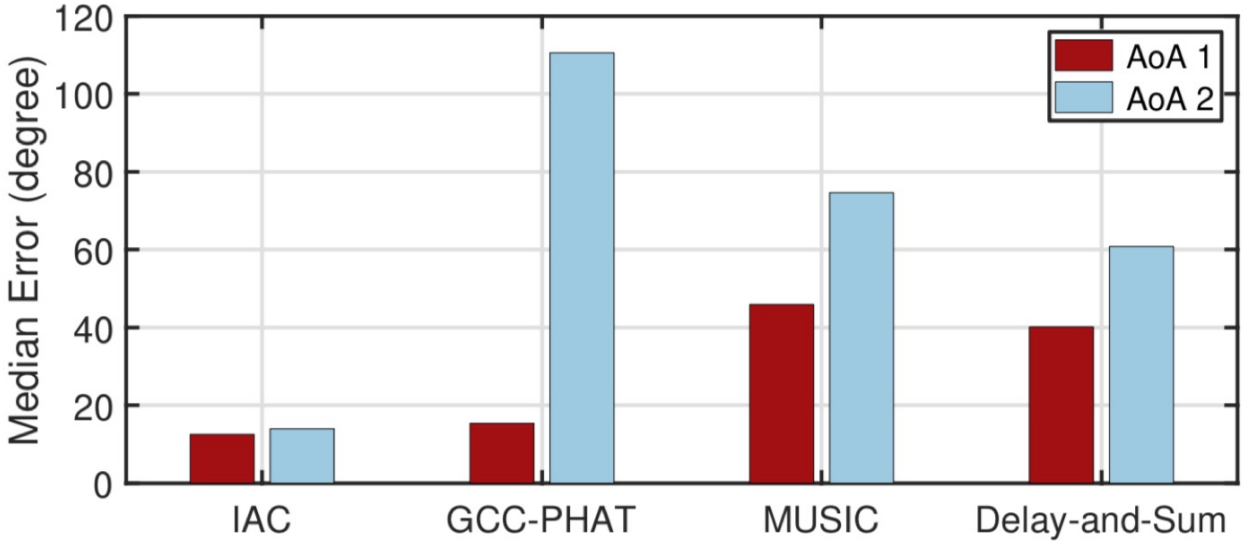
Iterative Align and Cancel (IAC) Algorithm



(a) Solving AoA_1 for the first path. After aligning with the correct AoA, the aligned signals will cancel each other.



THEOREM 3.1 (IAC AOA DECODING). *For a given pair of microphones, the k residue vectors from aligning and canceling each of the k AoAs are linearly dependent.*



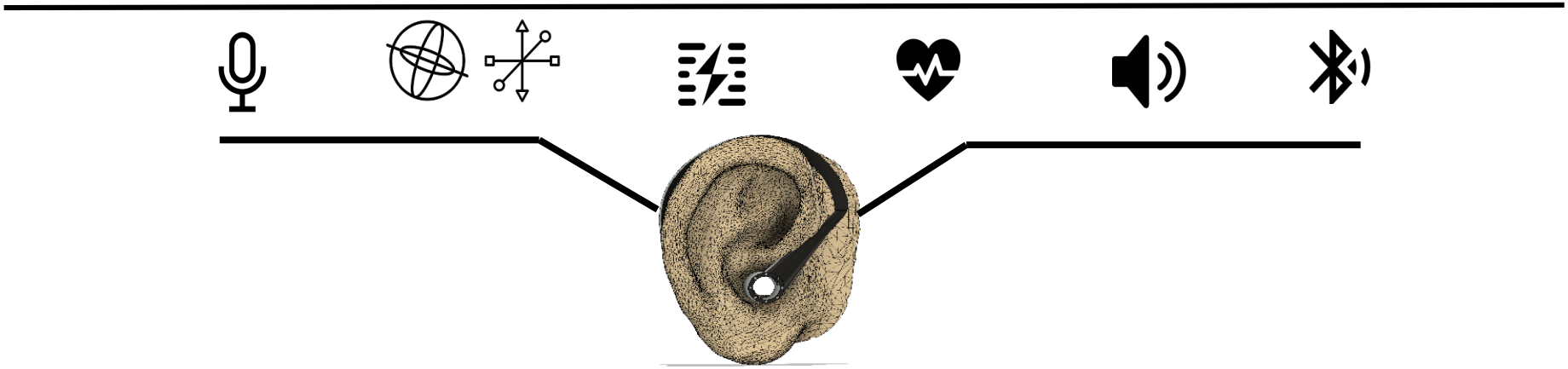
We are building a

software library

for Earable Computing

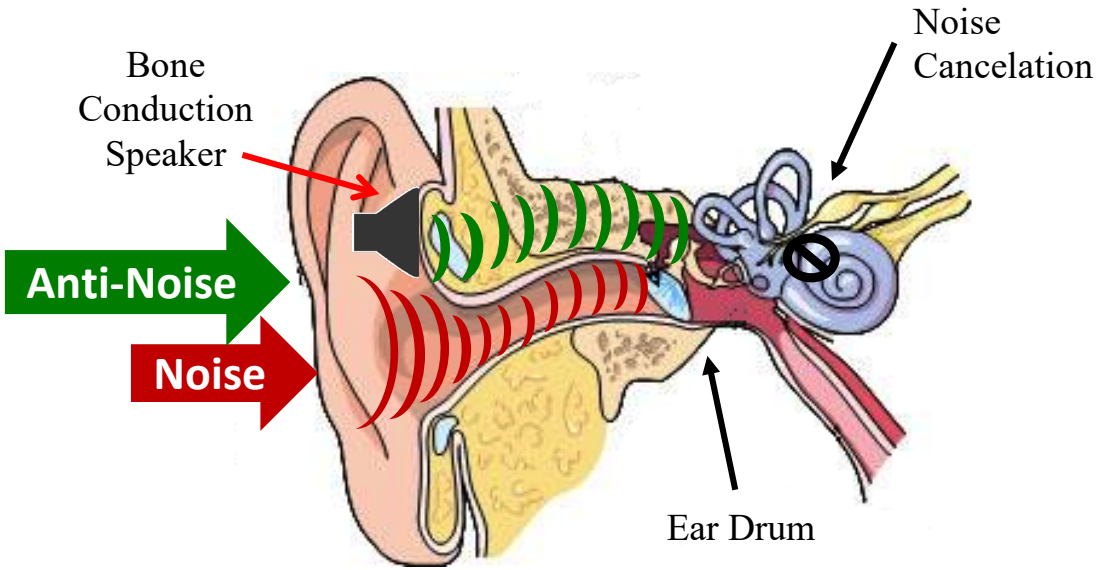
- Voice assistants
- Spatial acoustics
- Motion tracking
- Beamforming to Speech
- Noise Cancellation
- Face/Jaw Activity (eating, drinking)
- In-Ear Anatomy (health, security)

Sensor and Hardware Layer



Noise Cancellation

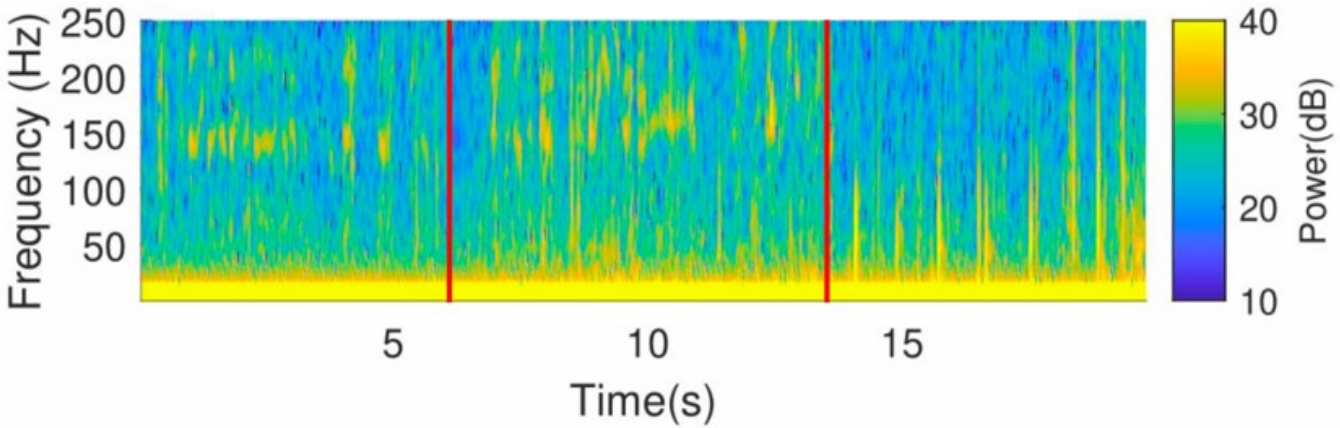
Cancellation with **Hollow** earbuds

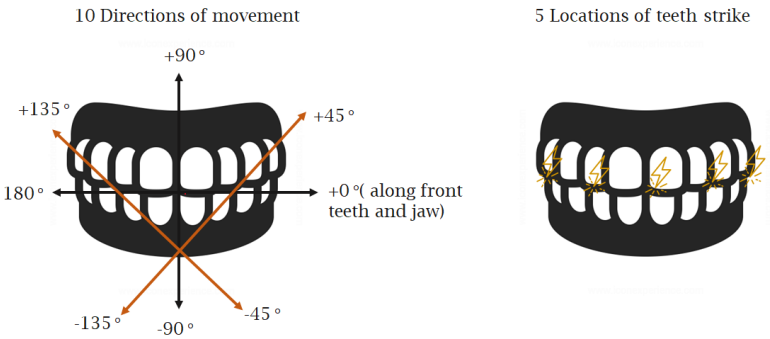


Face/jaw activity

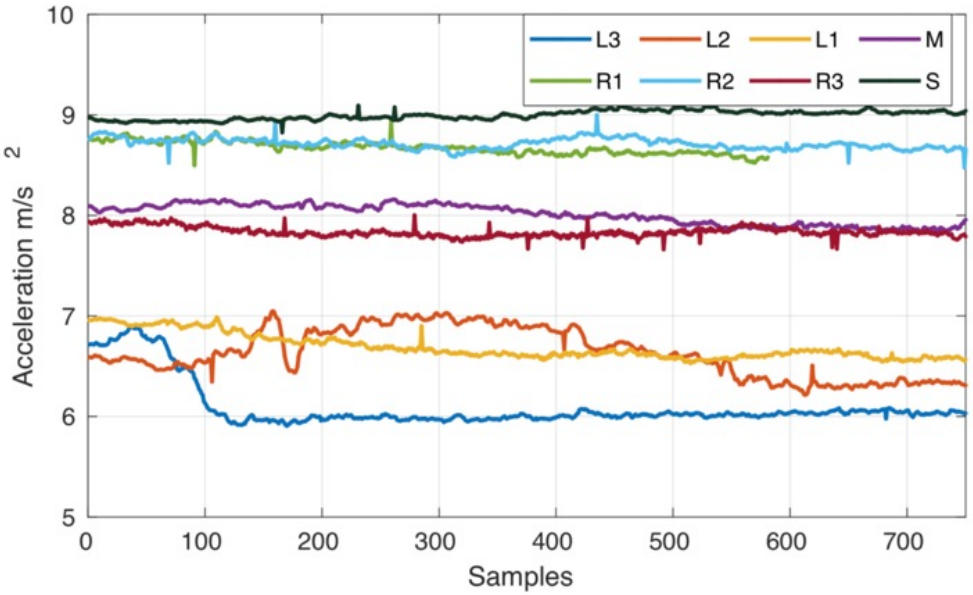
Activities = Eating, Drinking, Brushing

Different eating, drinking, tooth brushing activity



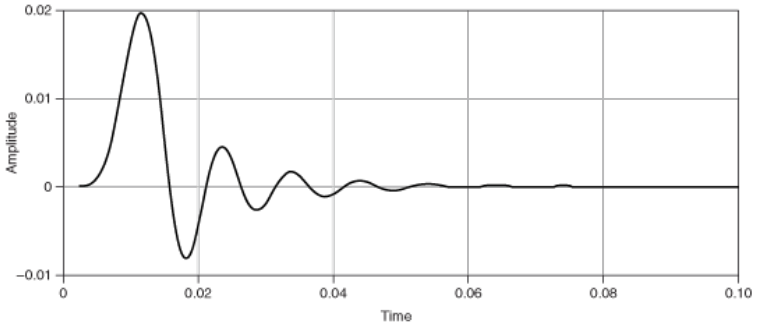


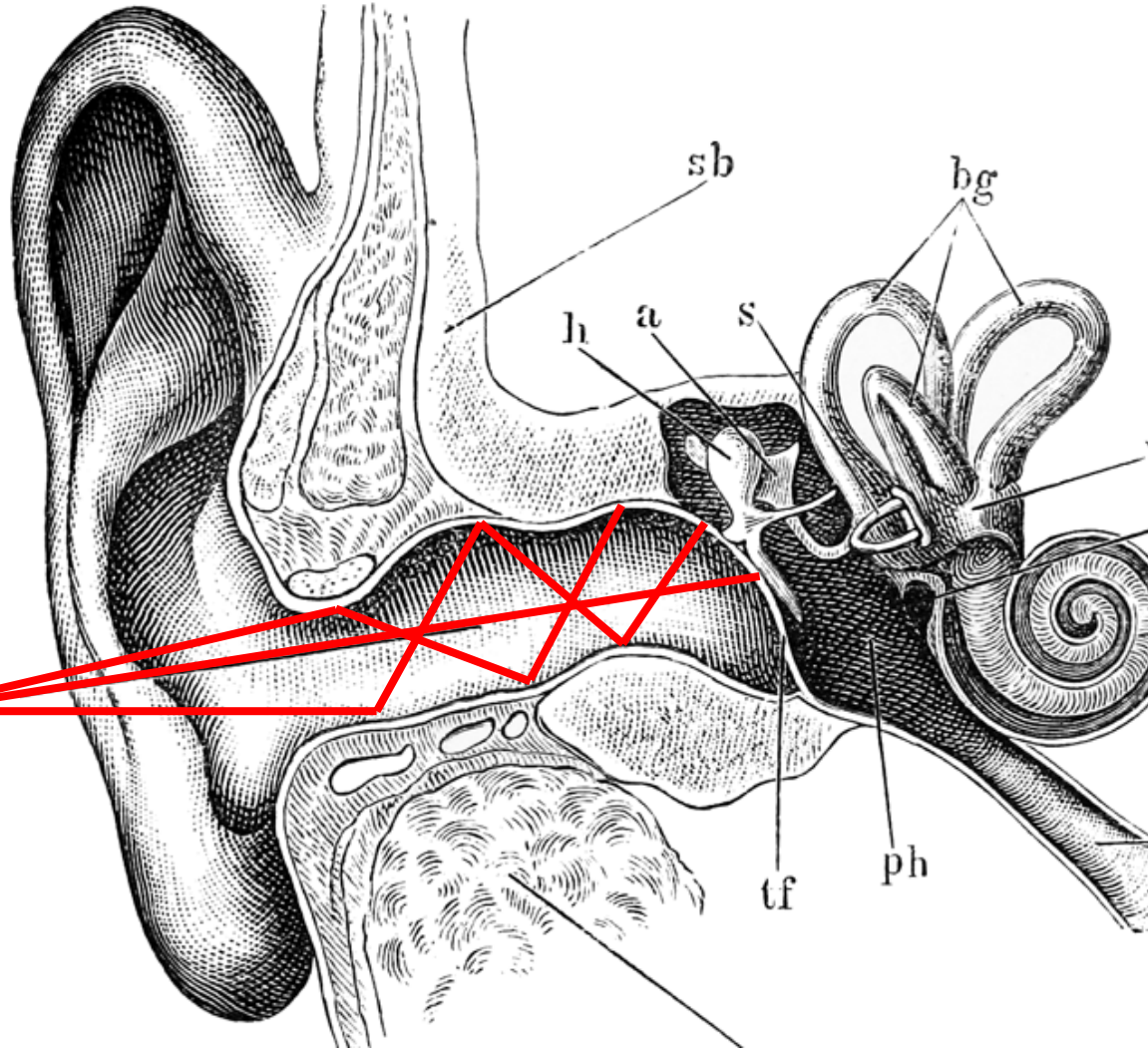
Note: in all these movements, it's the lower jaw which moves



In-ear health sensing

Ear Impulse Response (EIR)





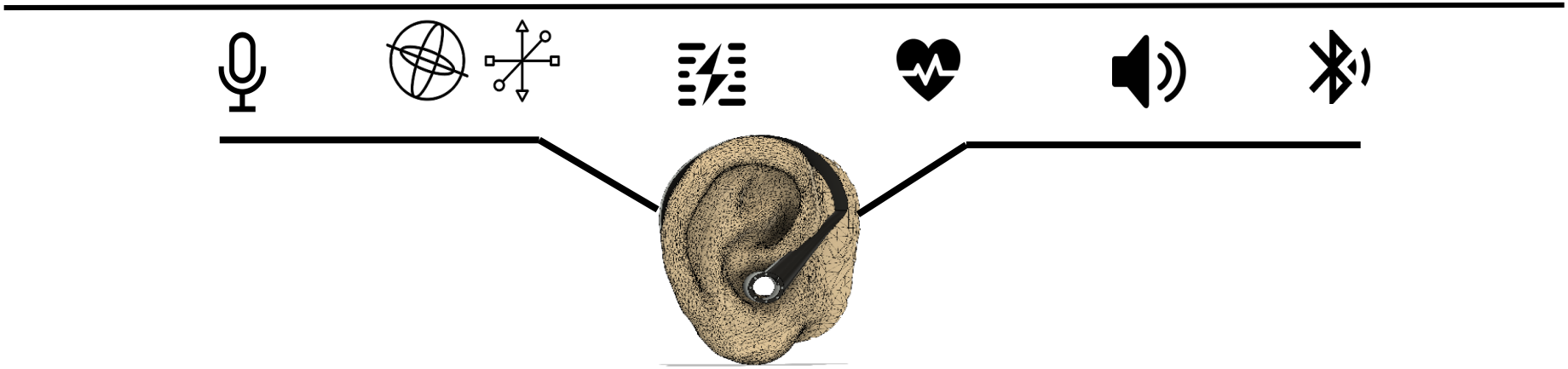
We are building a

software library

for Earable Computing

- Voice assistants
- Spatial acoustics
- Motion tracking
- Beamforming to Speech
- Noise Cancellation
- Face/Jaw Activity (eating, drinking)
- In-Ear Anatomy (health, security)

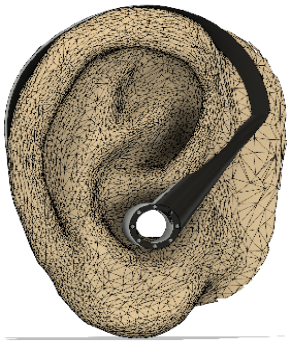
Sensor and Hardware Layer



Questions thus far …

Part II

Design a **hollow** earable that can still **cancel noise**



Existing Solutions

Existing Solutions

This Project

Better Noise Cancellation

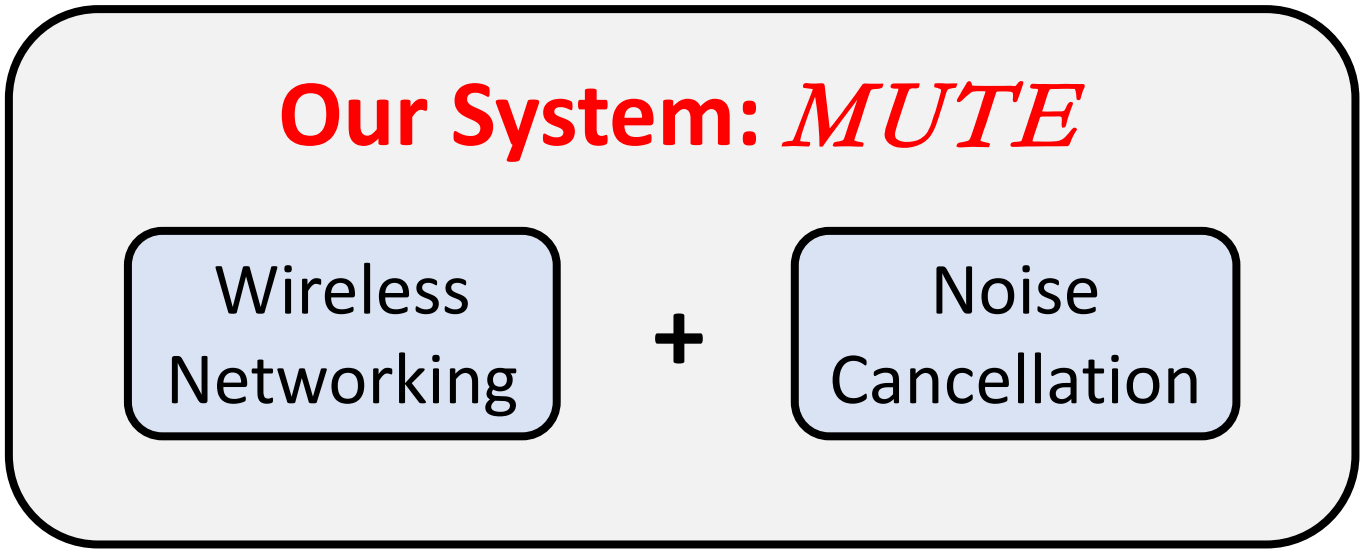
Existing Solutions

This Project

Better Noise Cancellation

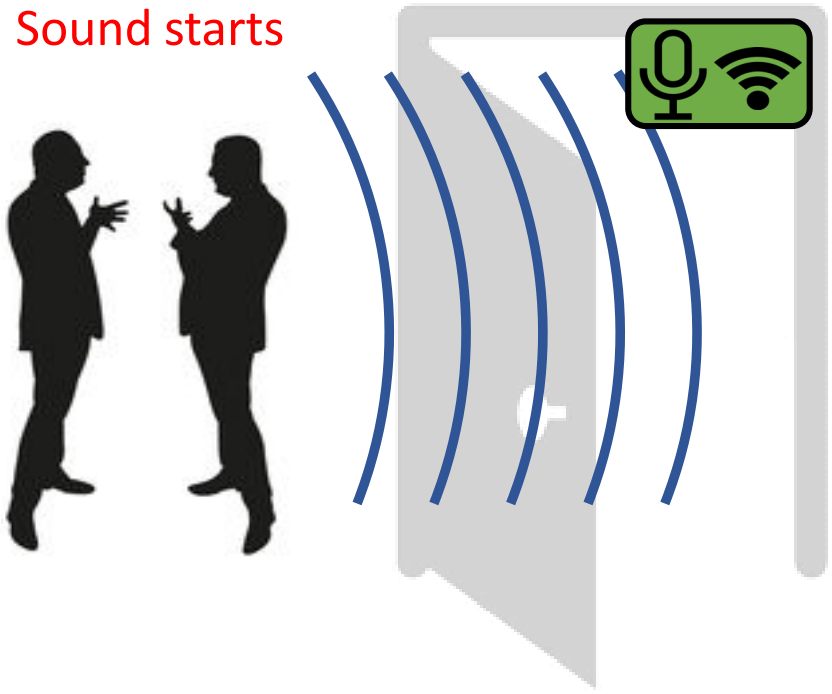
Better Form Factor (Hollow)

Comparable Performance


Our System: MUTE
Wireless Networking
+
Noise Cancellation

Alice

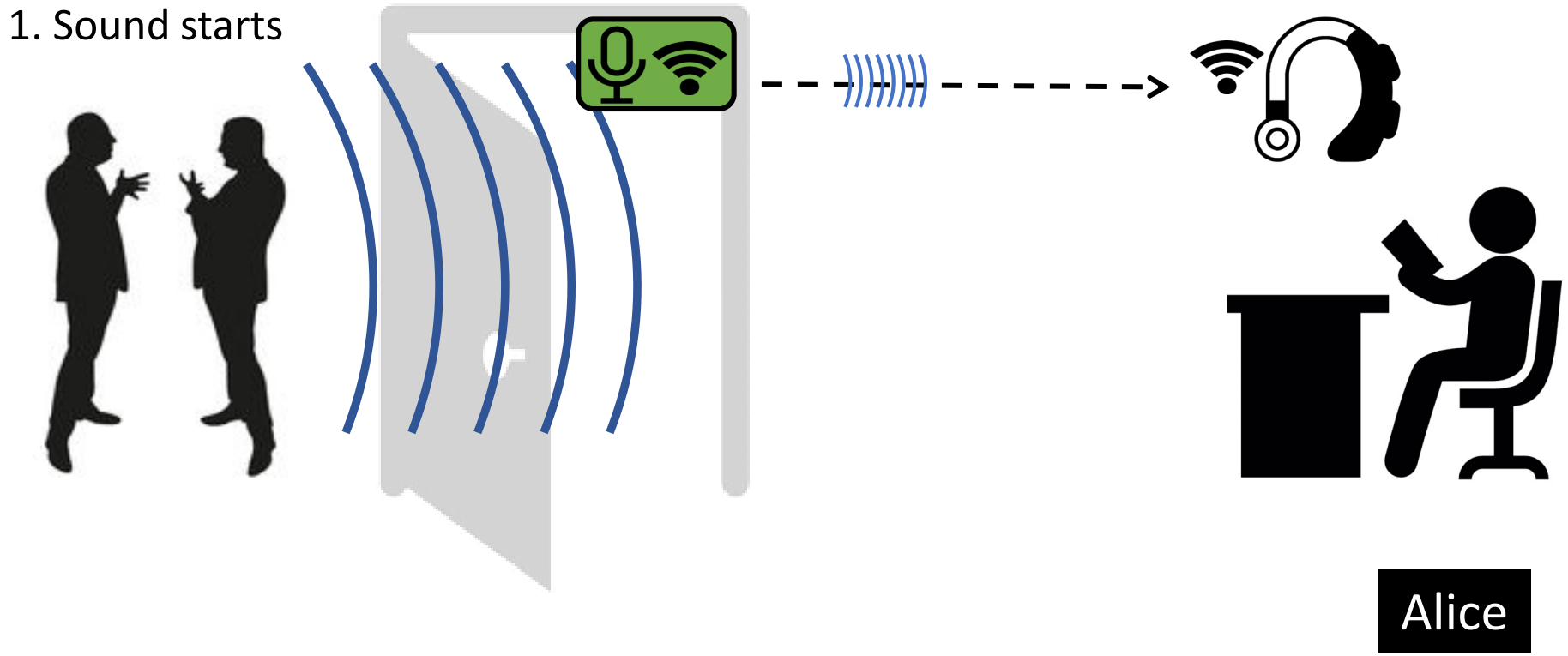
1. Sound starts



Alice

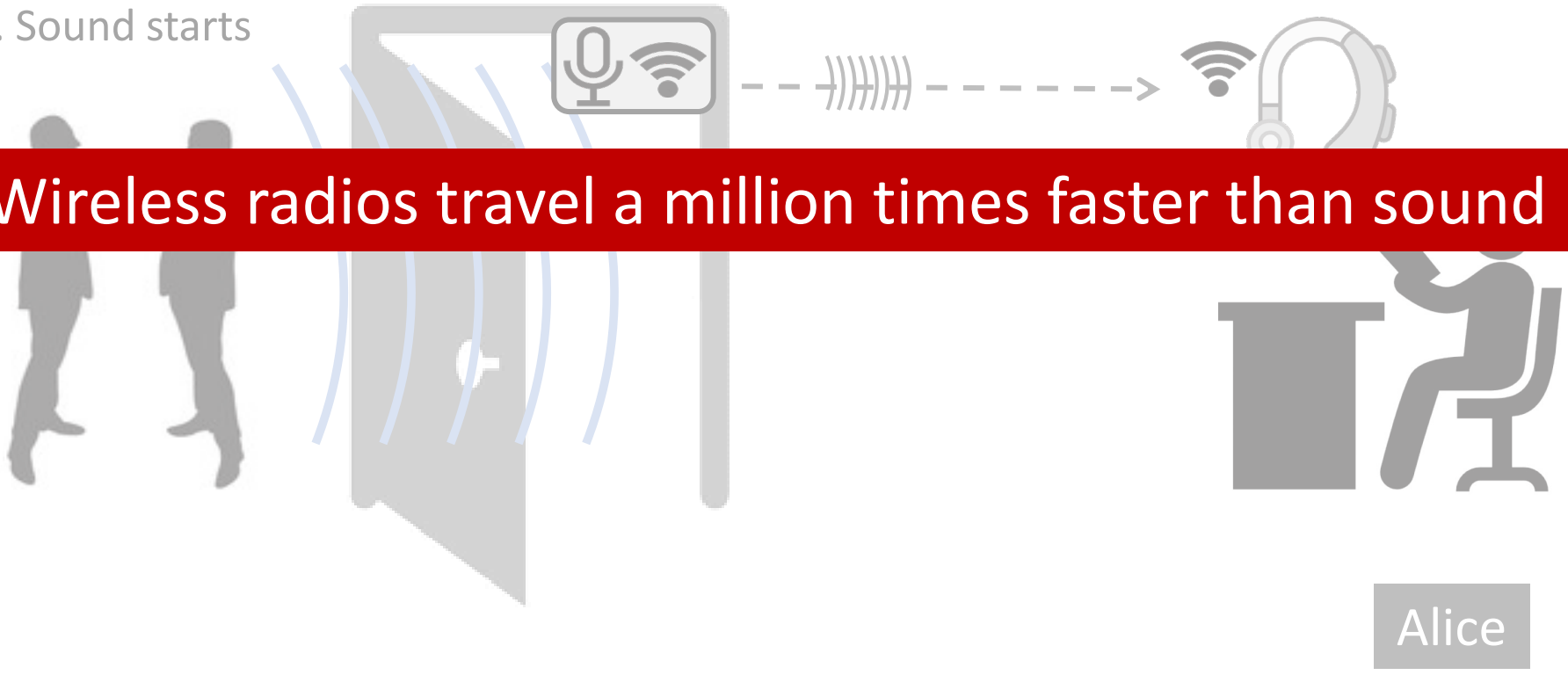
2. IoT relay forwards sound over wireless

1. Sound starts



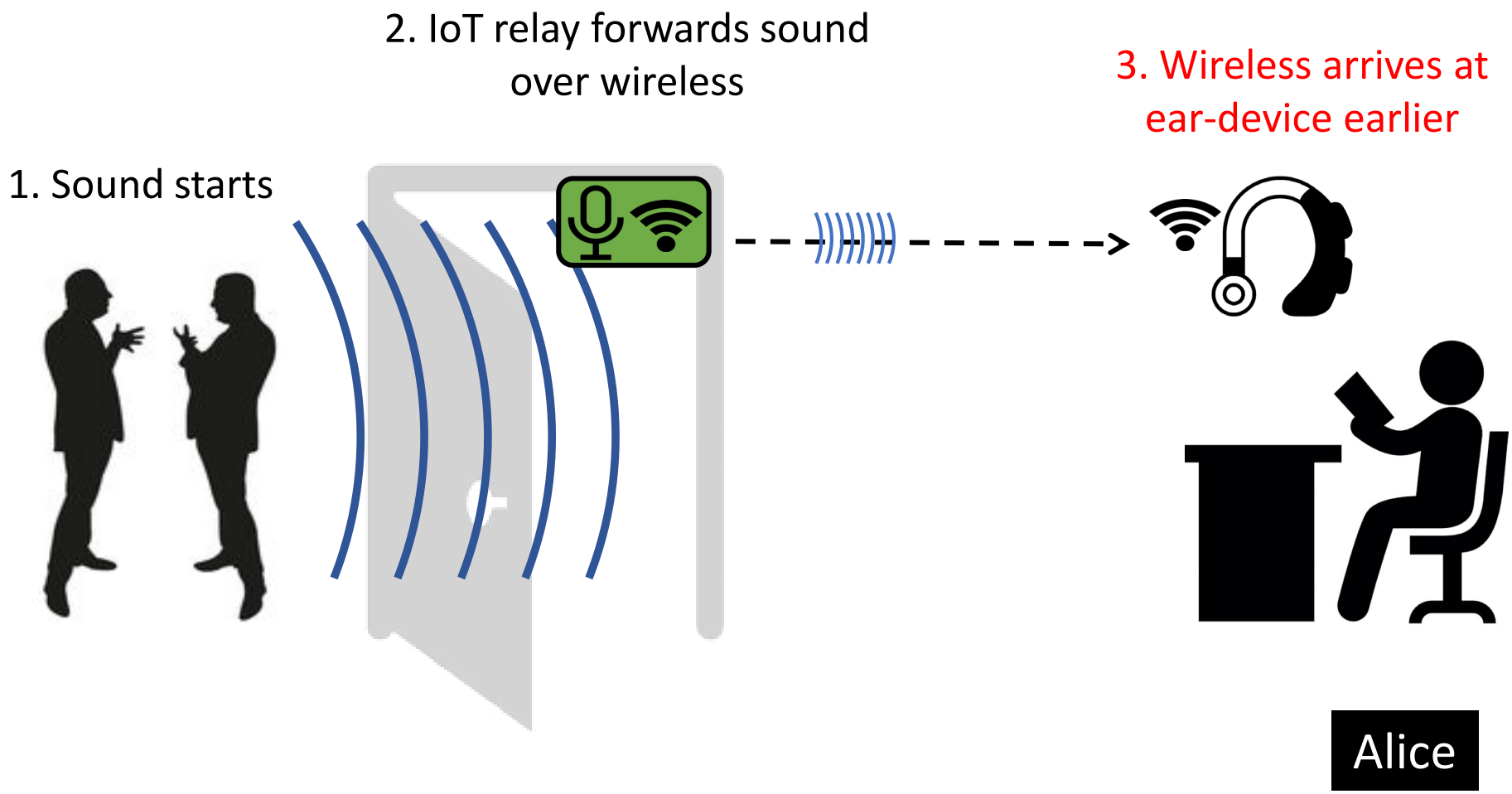
2. IoT relay forwards sound over wireless

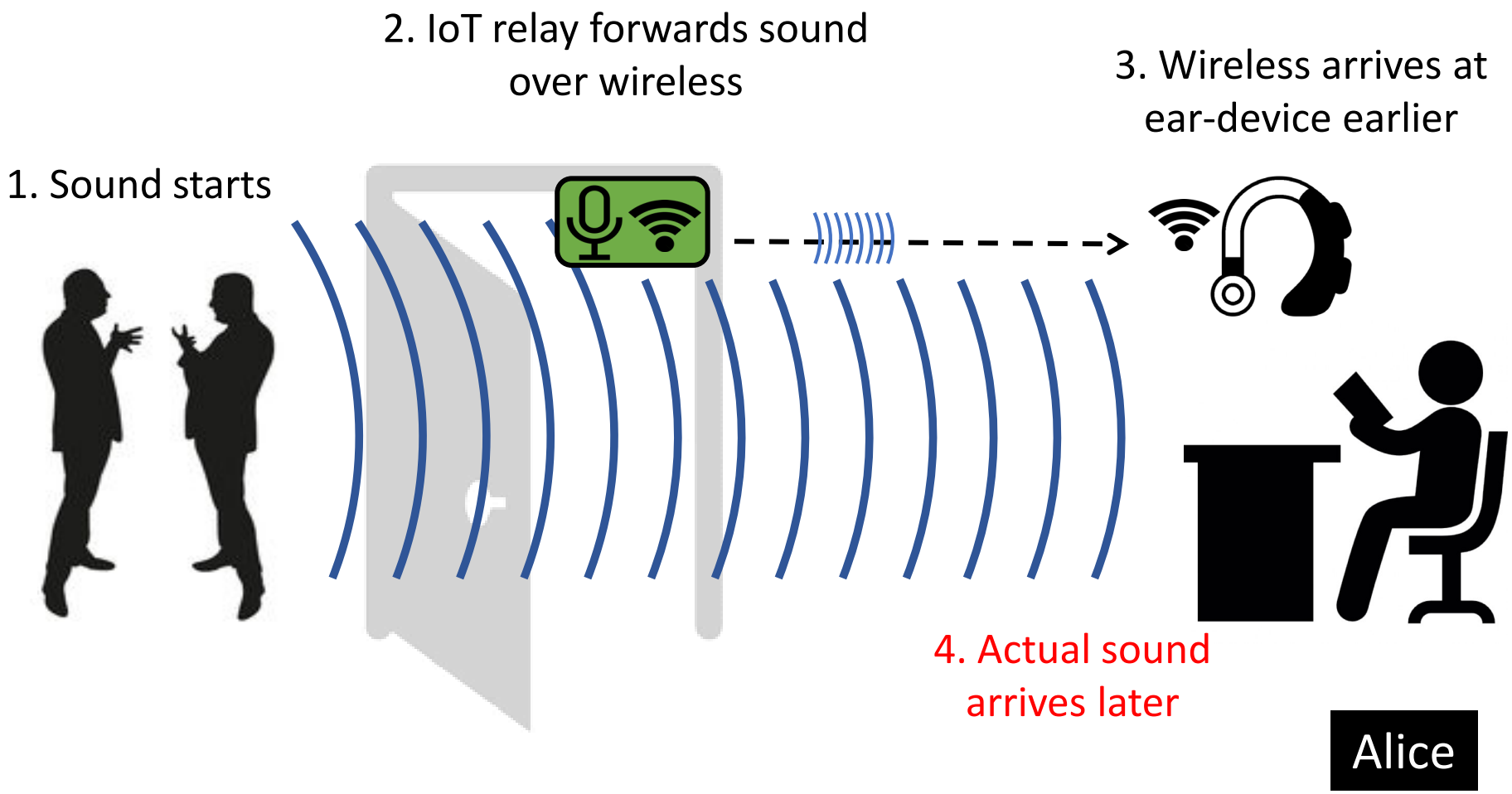
1. Sound starts

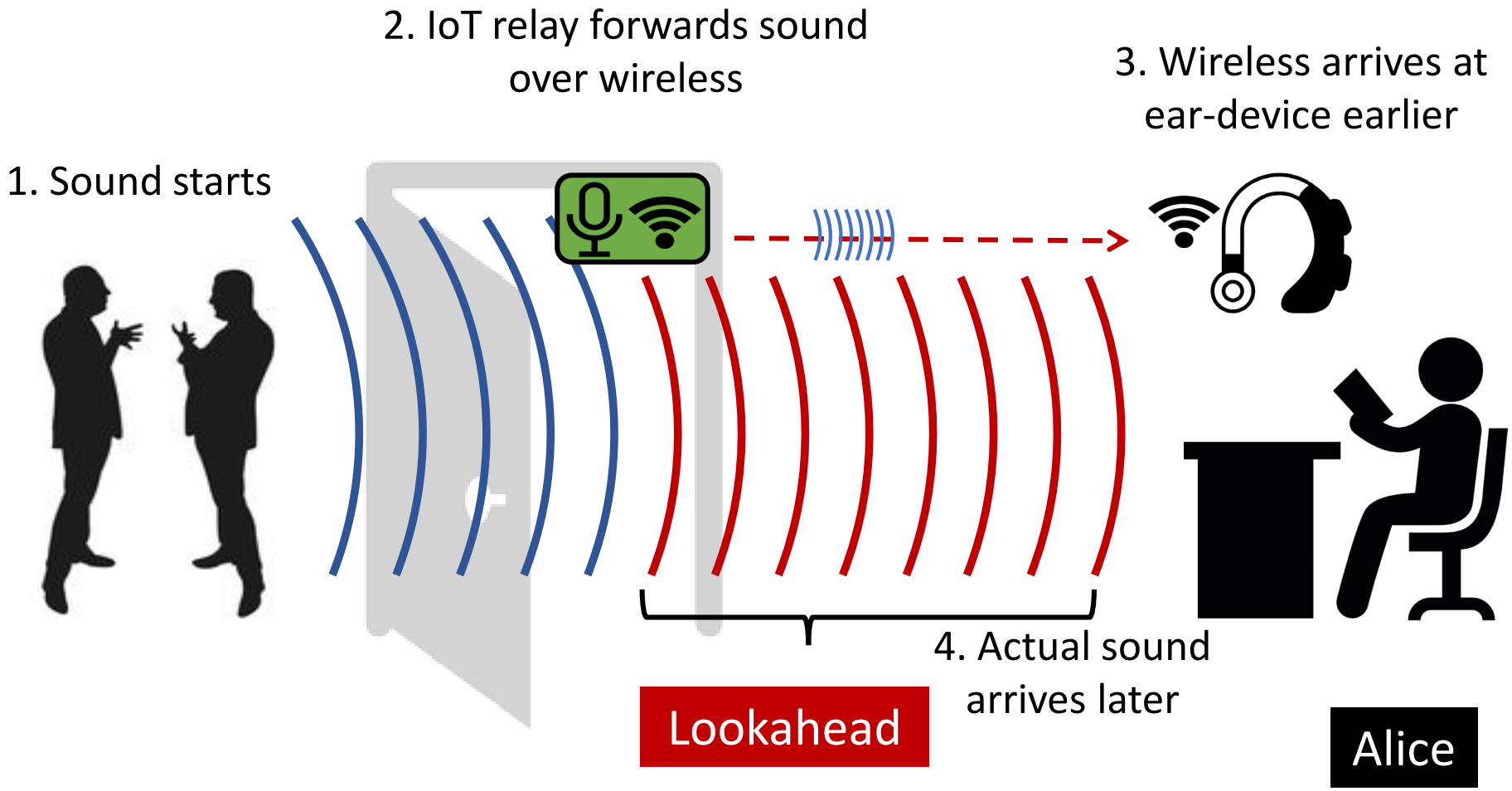


Wireless radios travel a million times faster than sound

Alice


2. IoT relay forwards sound over wireless
3. Wireless arrives at ear-device earlier
1. Sound starts
Alice


2. IoT relay forwards sound over wireless
3. Wireless arrives at ear-device earlier
1. Sound starts
4. Actual sound arrives later
Alice


2. IoT relay forwards sound over wireless
3. Wireless arrives at ear-device earlier
1. Sound starts
4. Actual sound arrives later
Lookahead
Alice

Analogy: Light travels much faster than sound

Lookahead allows us to cover ears in time

MUTE: Leverage **lookahead** for noise cancellation

Talk Outline

How can $MUTE$ leverage lookahead?

Timing Gain → Wideband Cancellation

Signal Processing Gain → Non-Causal Filtering

Application-Specific Gain → Sound Source Profiling

Talk Outline

How can $MUTE$ leverage lookahead?

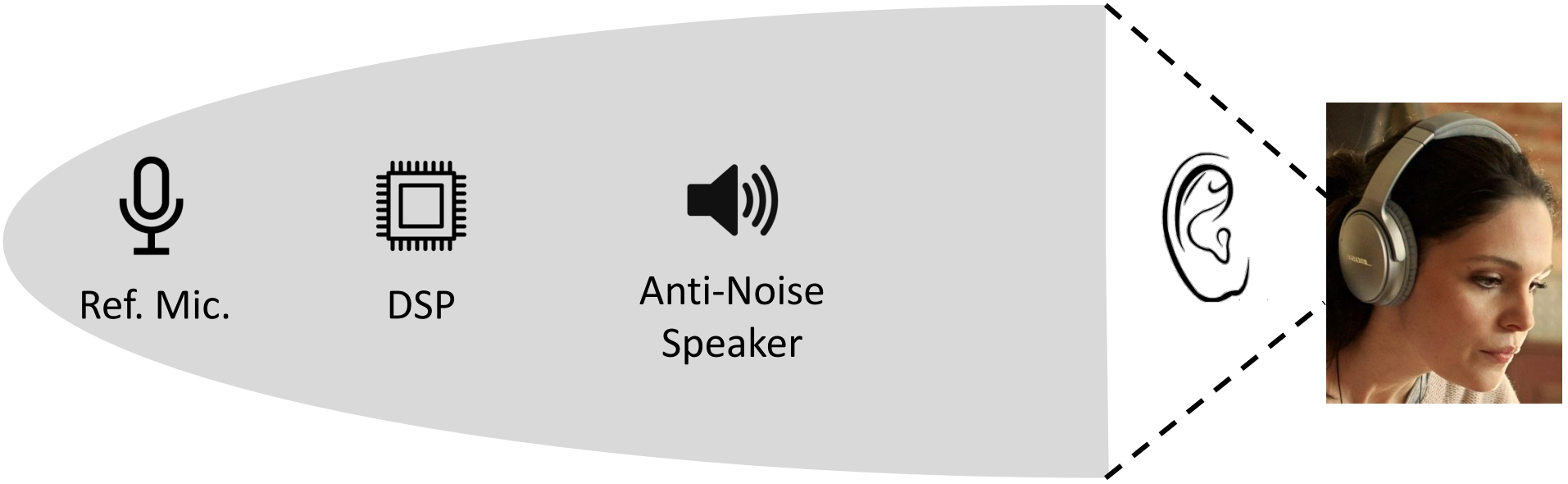
Timing Gain → Wideband Cancellation

Signal Processing Gain → Non-Causal Filtering

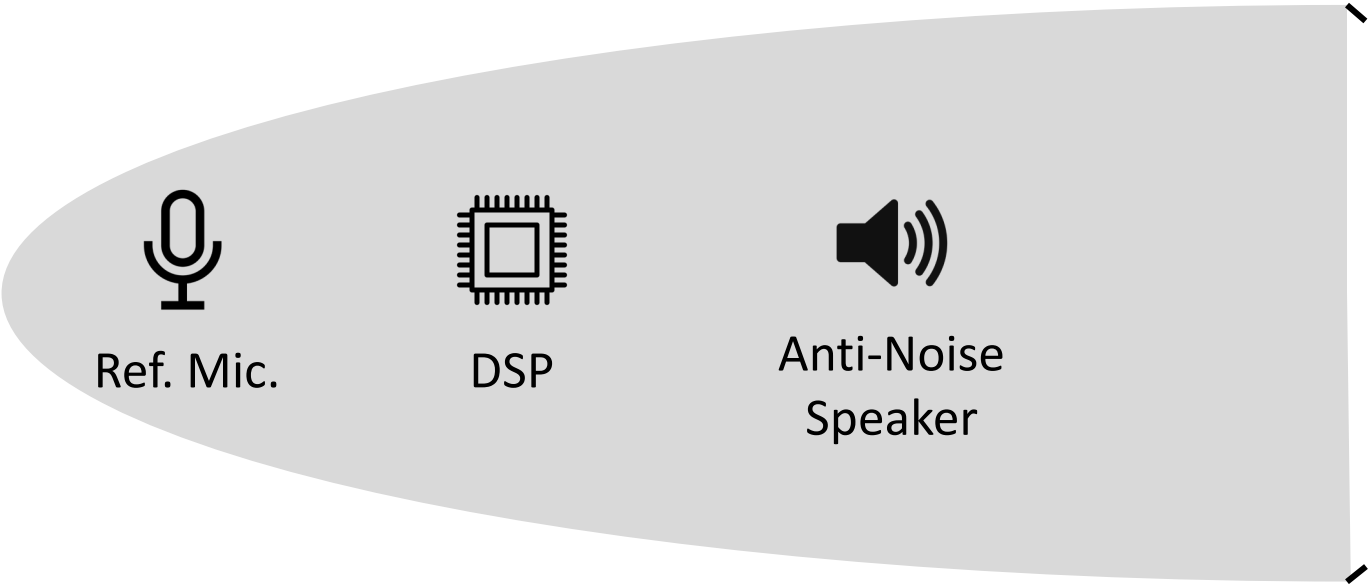
Application-Specific Gain → Sound Source Profiling

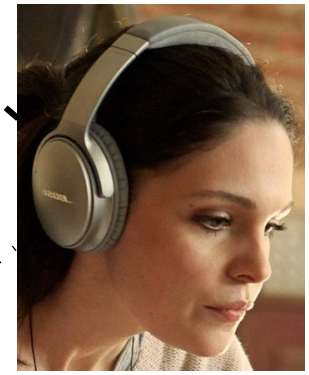
Noise Cancelling Headphones
-- What is inside?

Noise Cancellation Primer

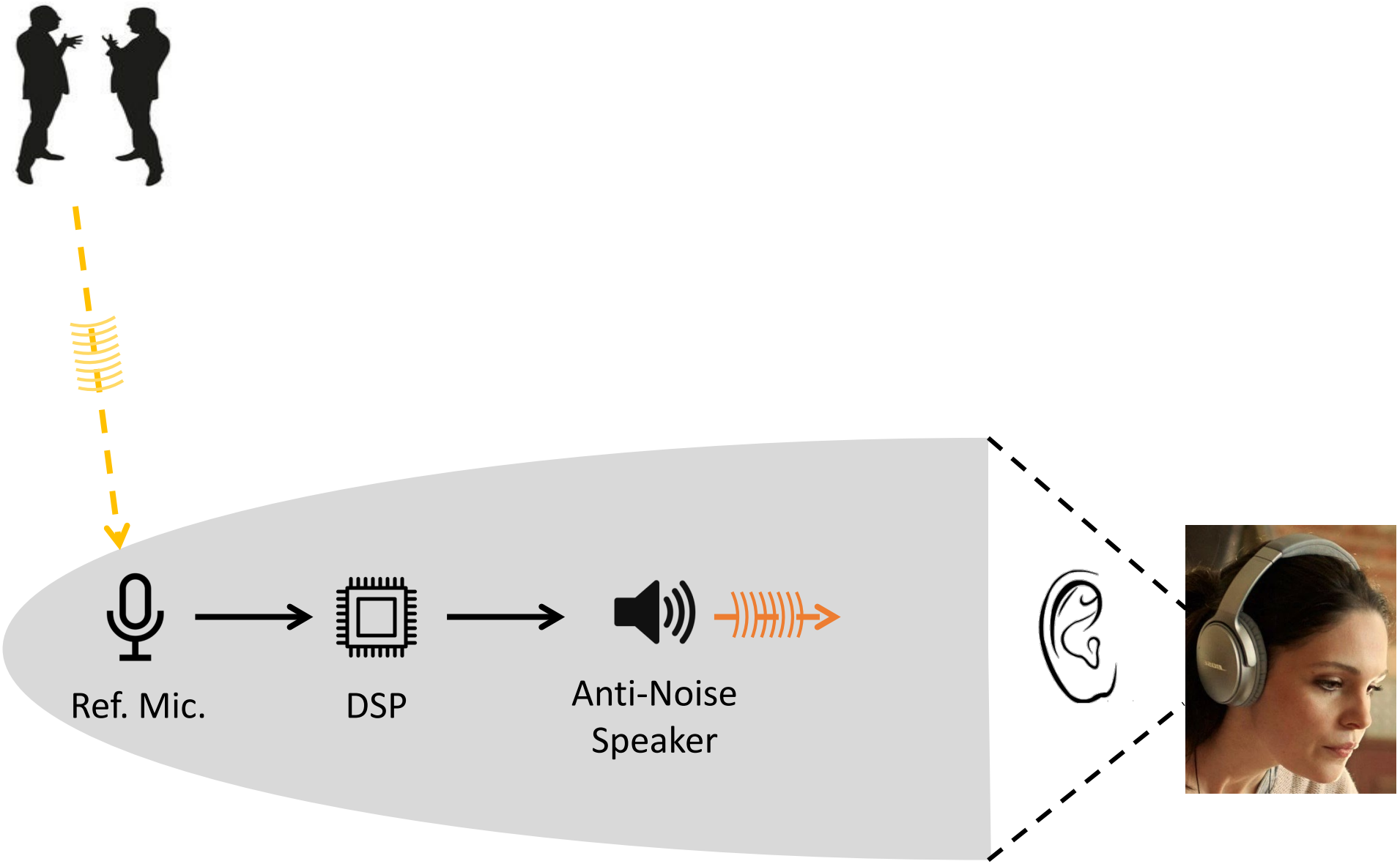


Noise Cancellation Primer

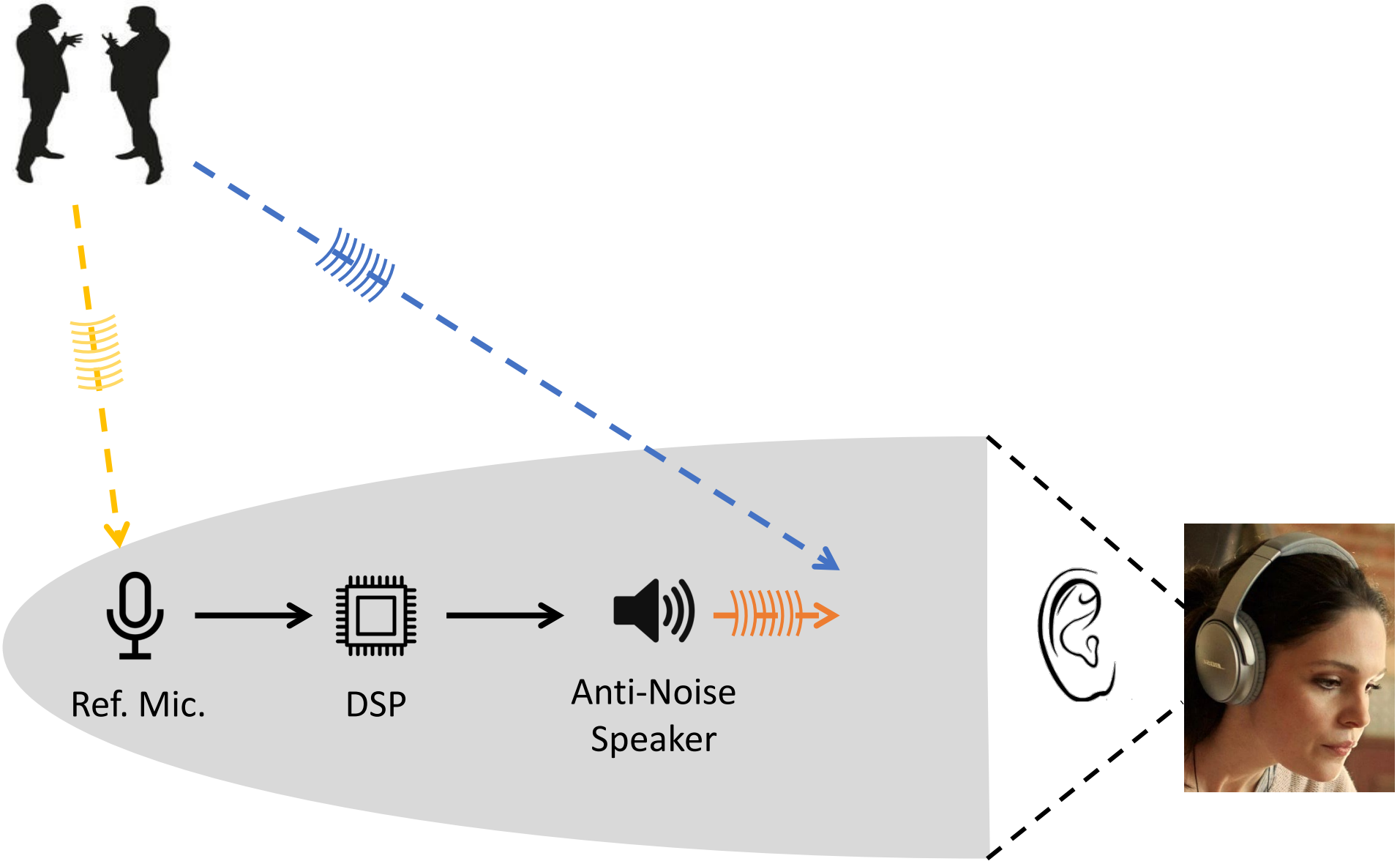




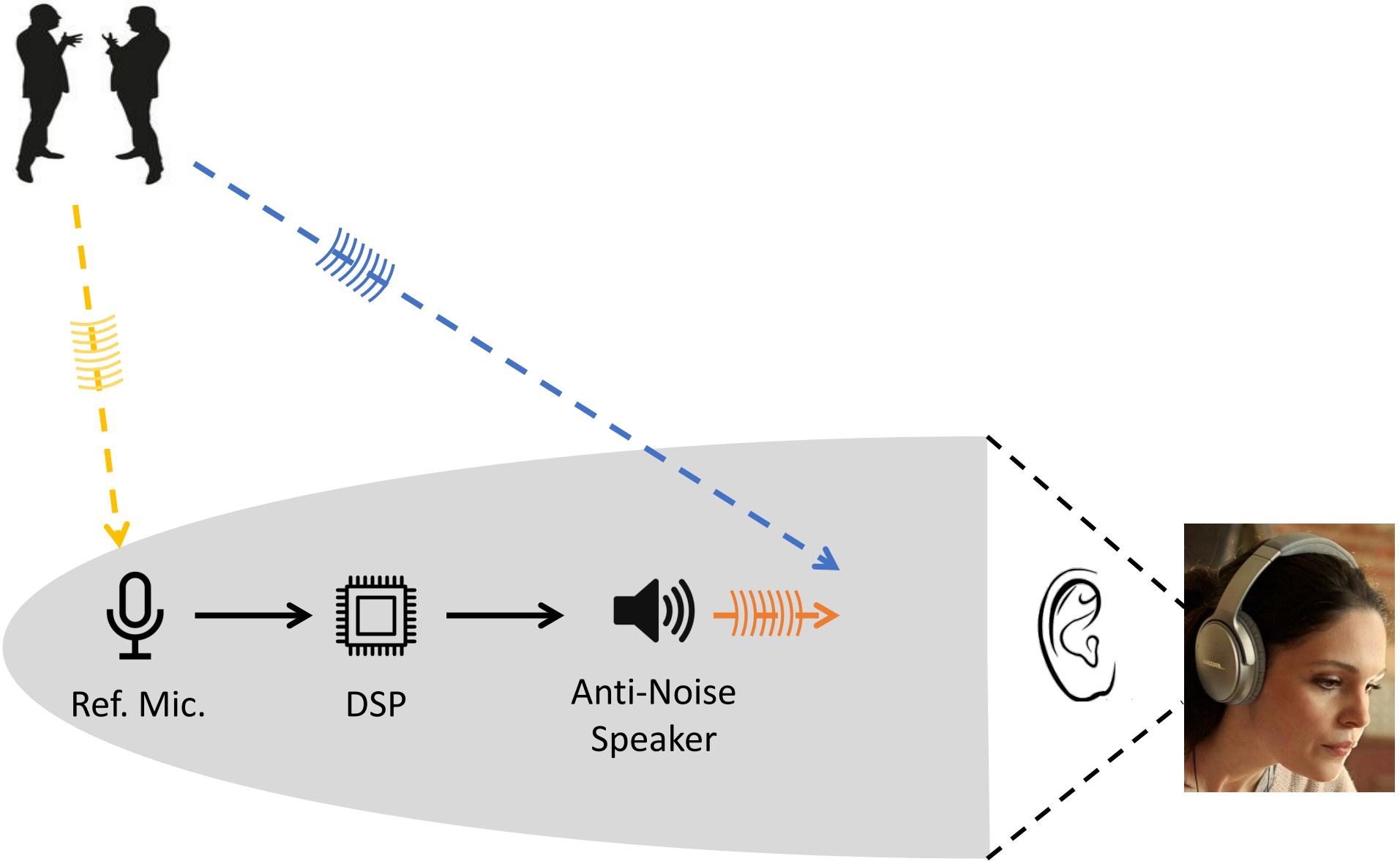
Noise Cancellation Primer



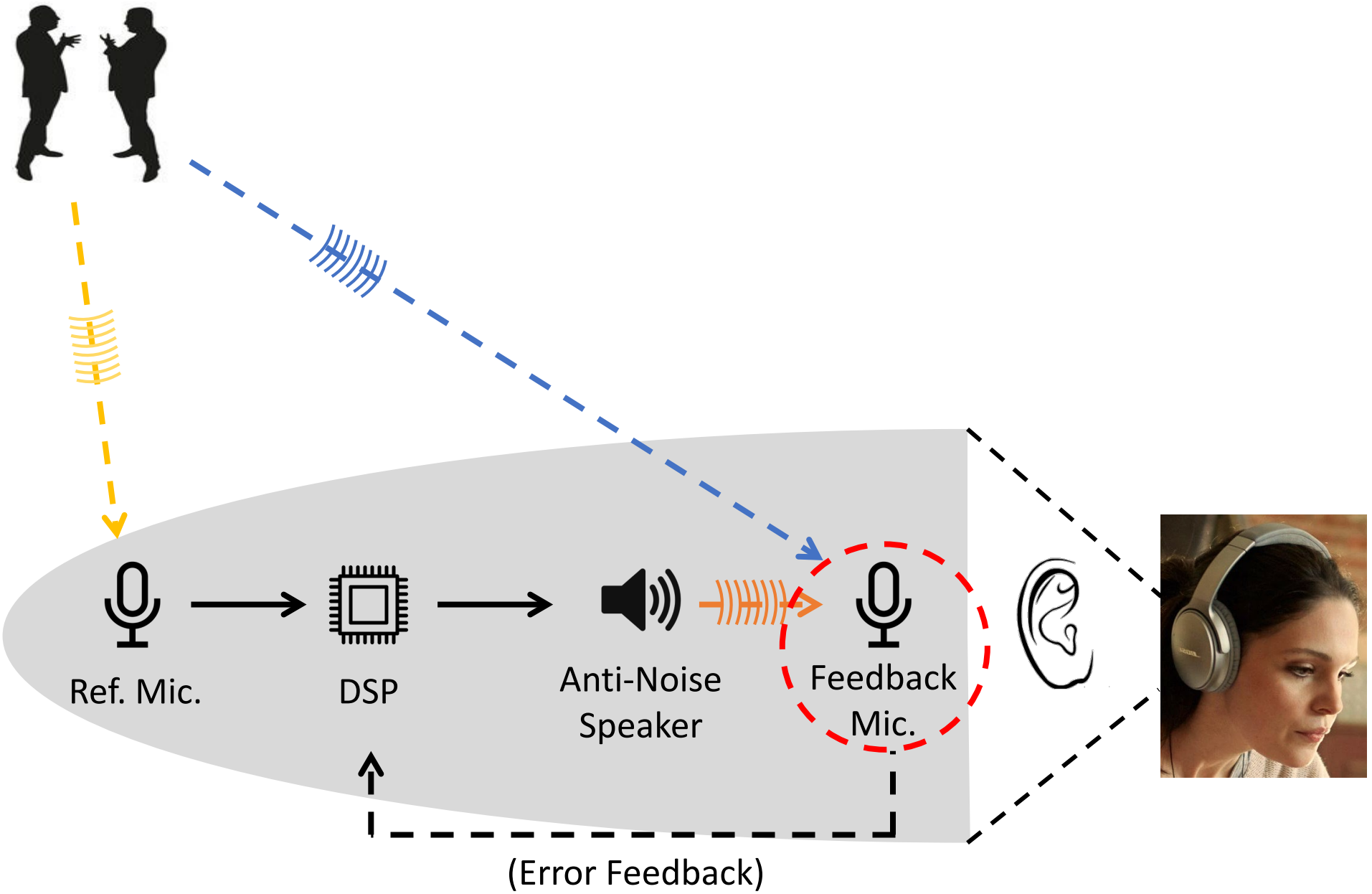
Noise Cancellation Primer



Noise Cancellation Primer



Noise Cancellation Primer



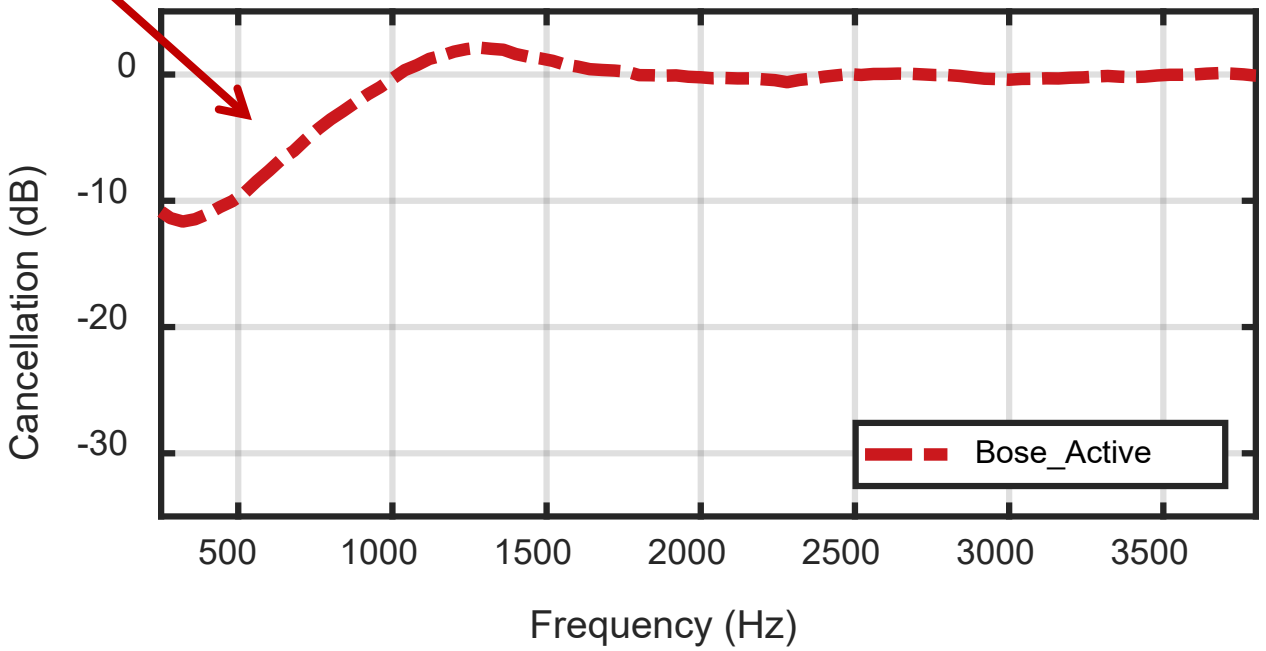
How do they perform today?

Bose QC35

Active noise cancellation
only at low frequencies

Bad

Good

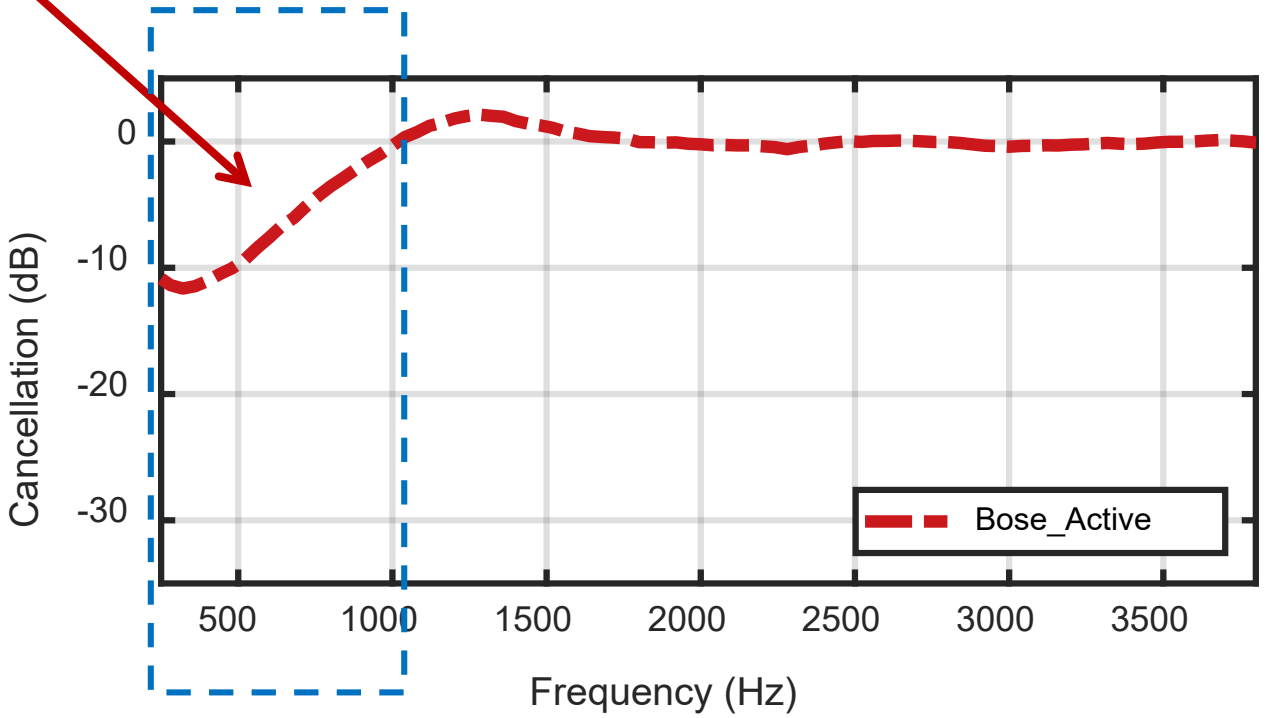


Bose QC35

Active noise cancellation
only at low frequencies

Bad

Good



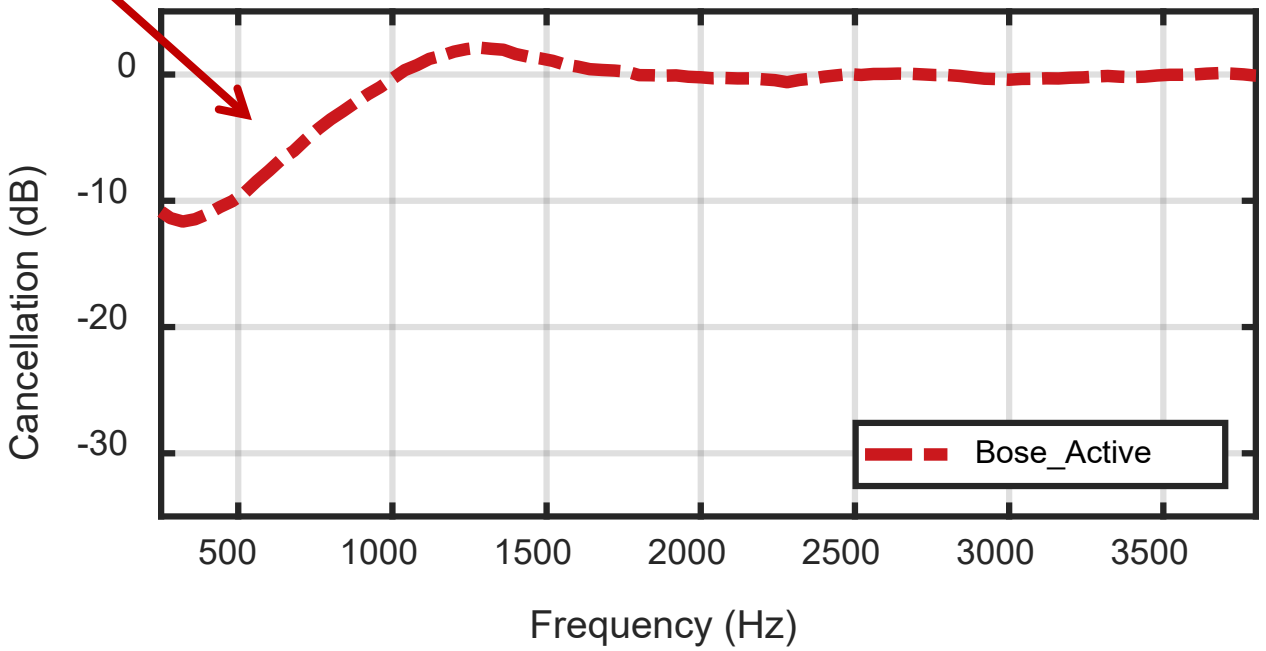
Bose QC35

Active noise cancellation
only at low frequencies

Sound absorbing material
blocks high frequencies

Bad

Good



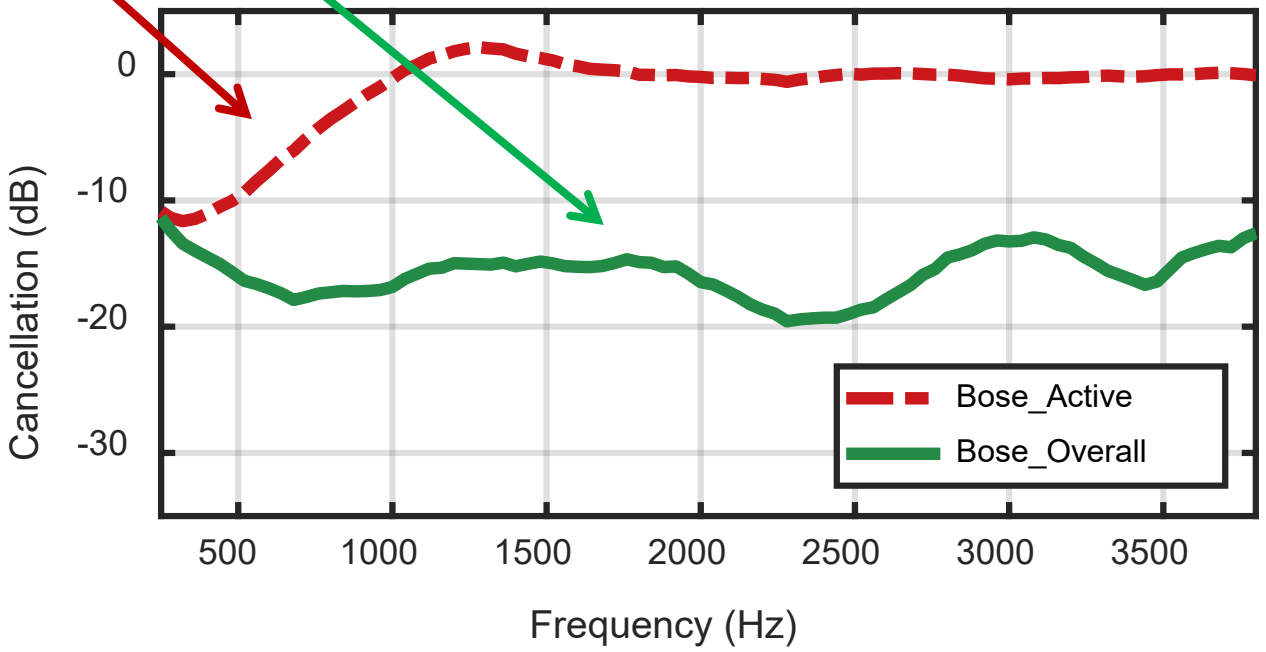
Bose QC35

Active noise cancellation
only at low frequencies

Sound absorbing material
blocks high frequencies

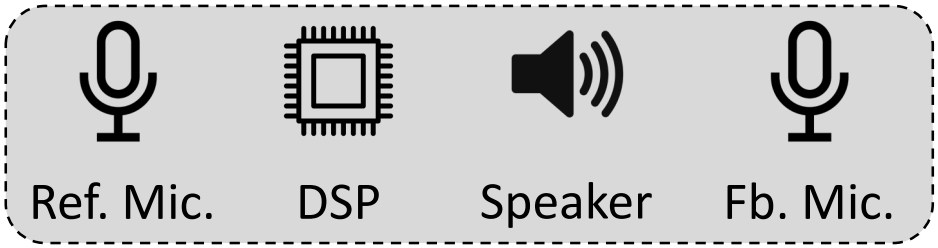
Bad

Good

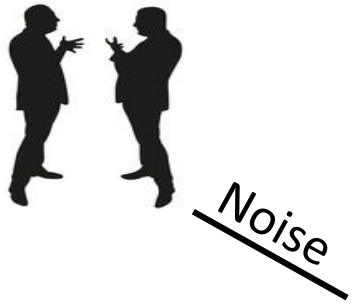


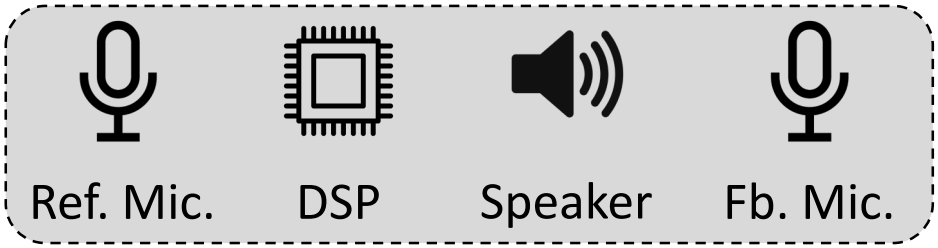
Why does noise cancellation not work at higher frequencies?

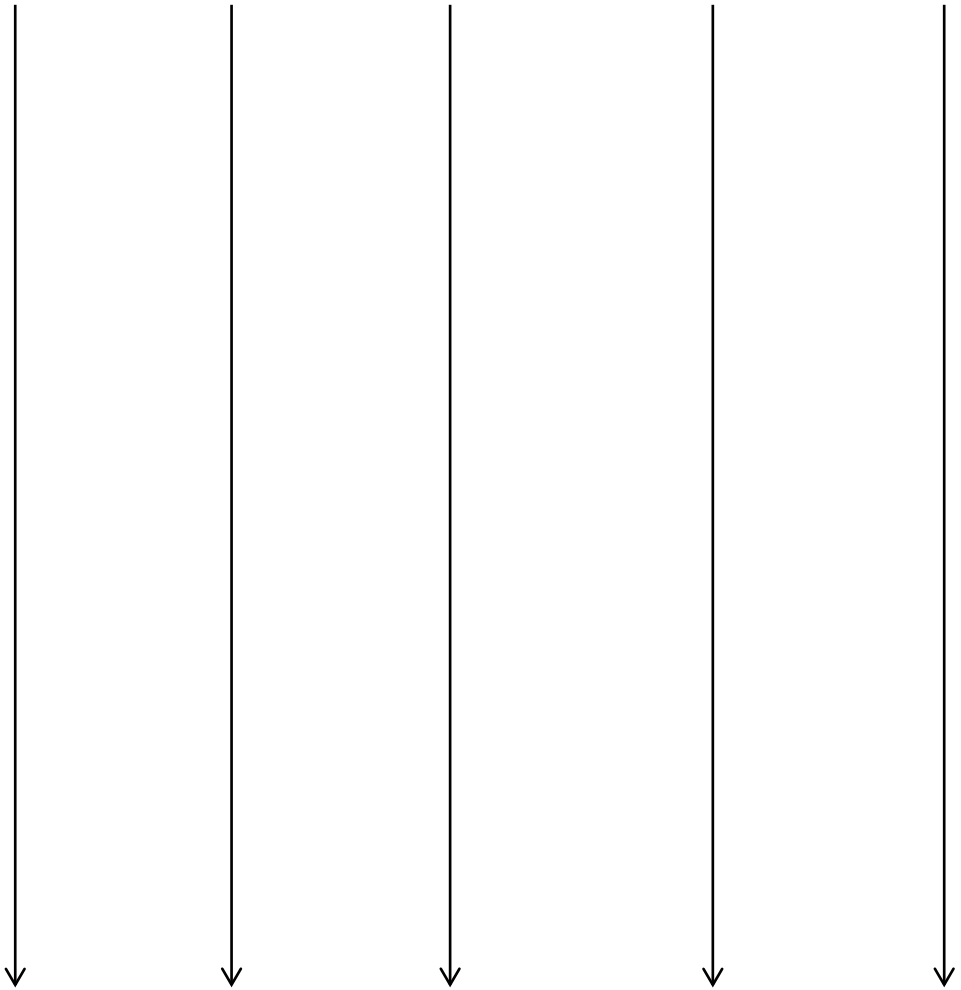
Let’s look into the headphone again.



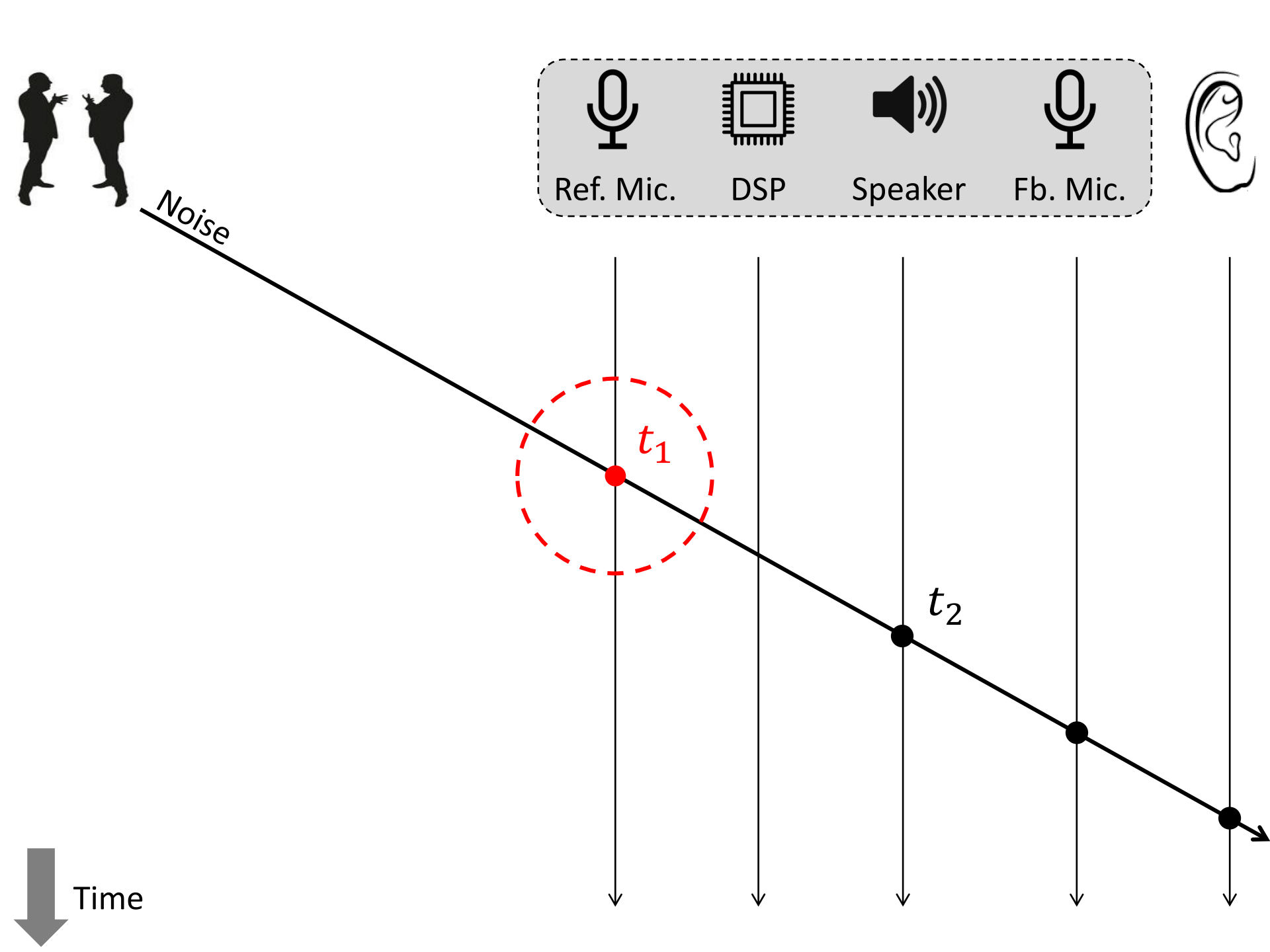
Time

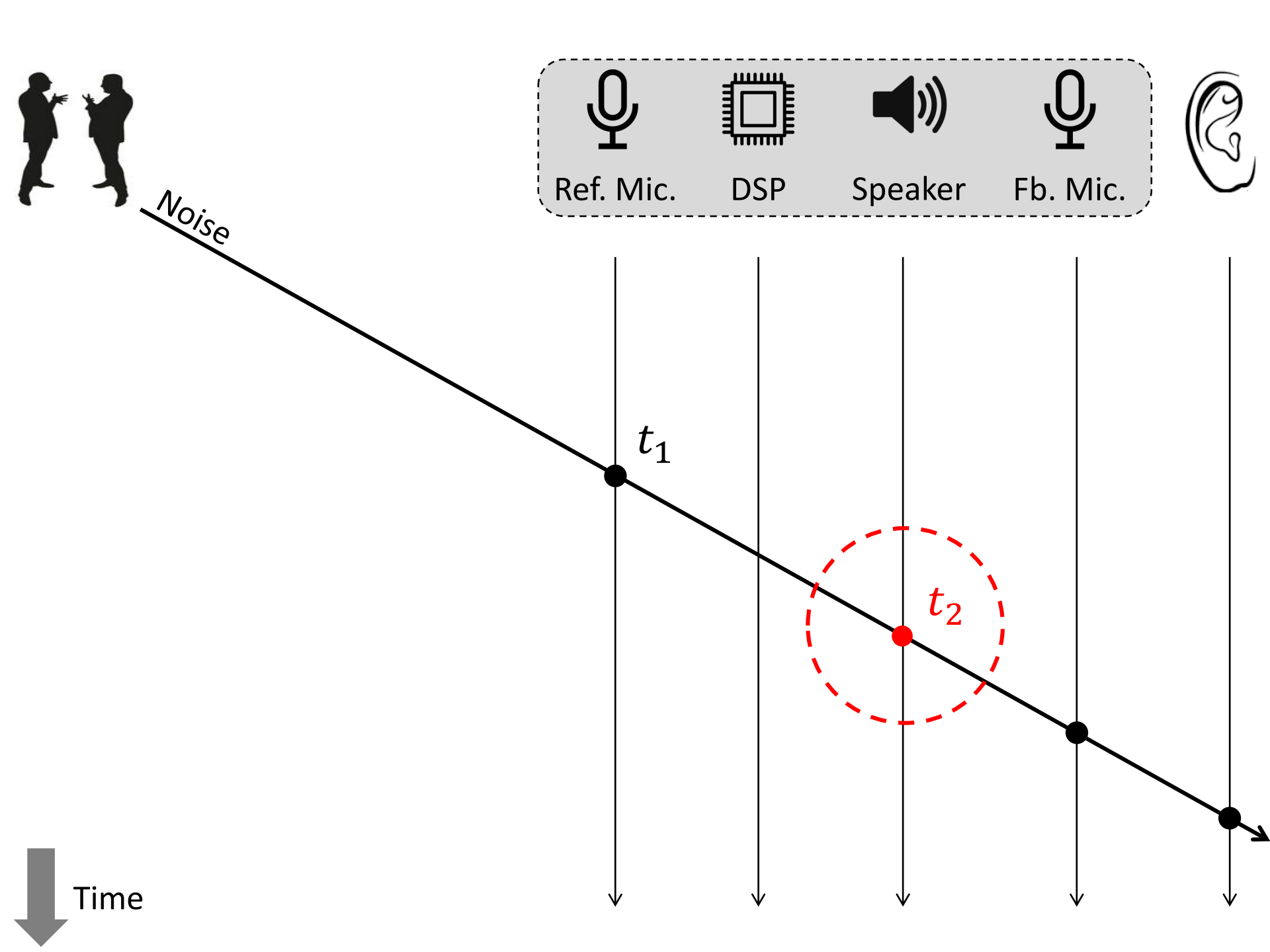

Noise

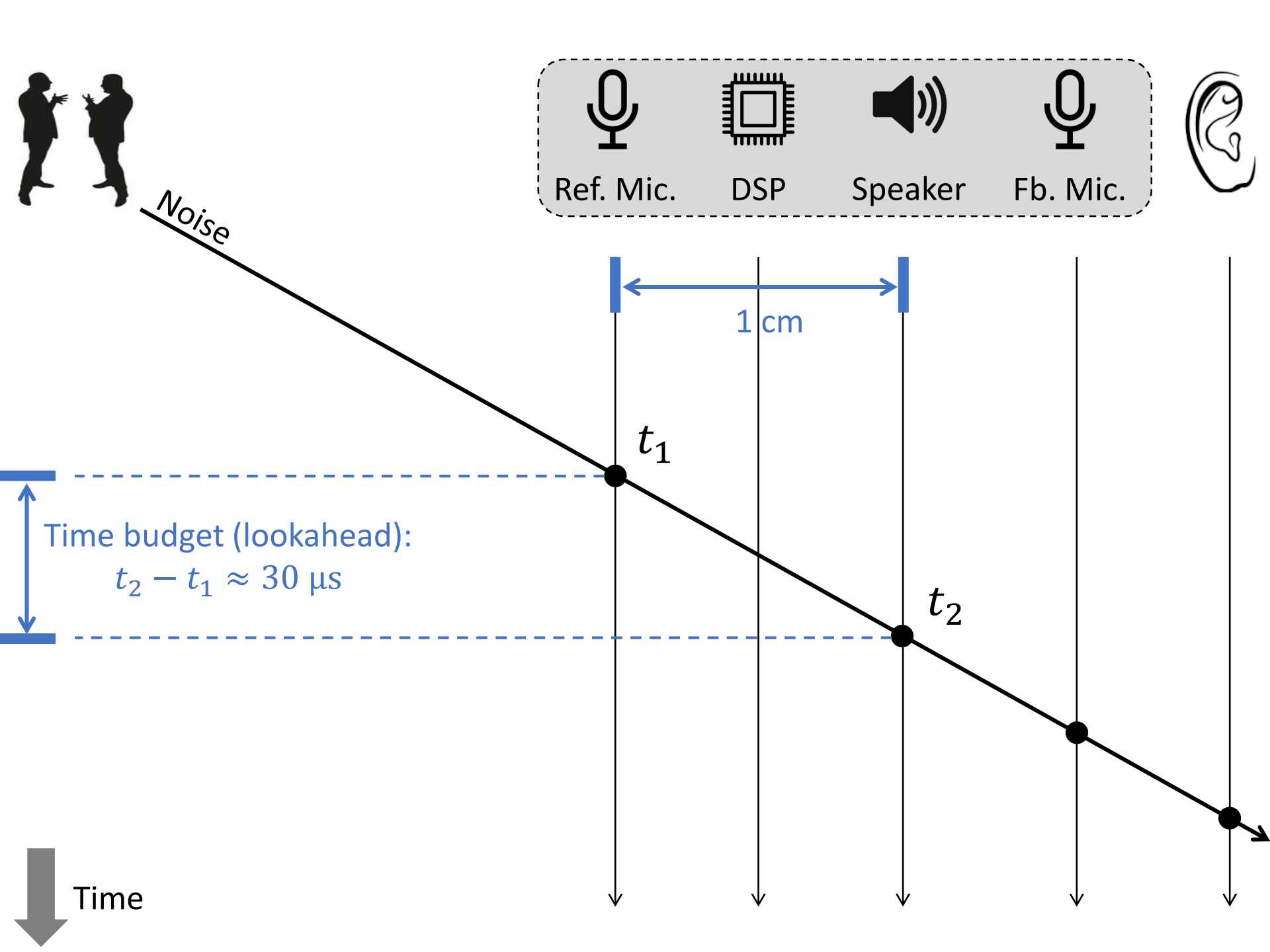

Ref. Mic.
DSP
Speaker
Fb. Mic.

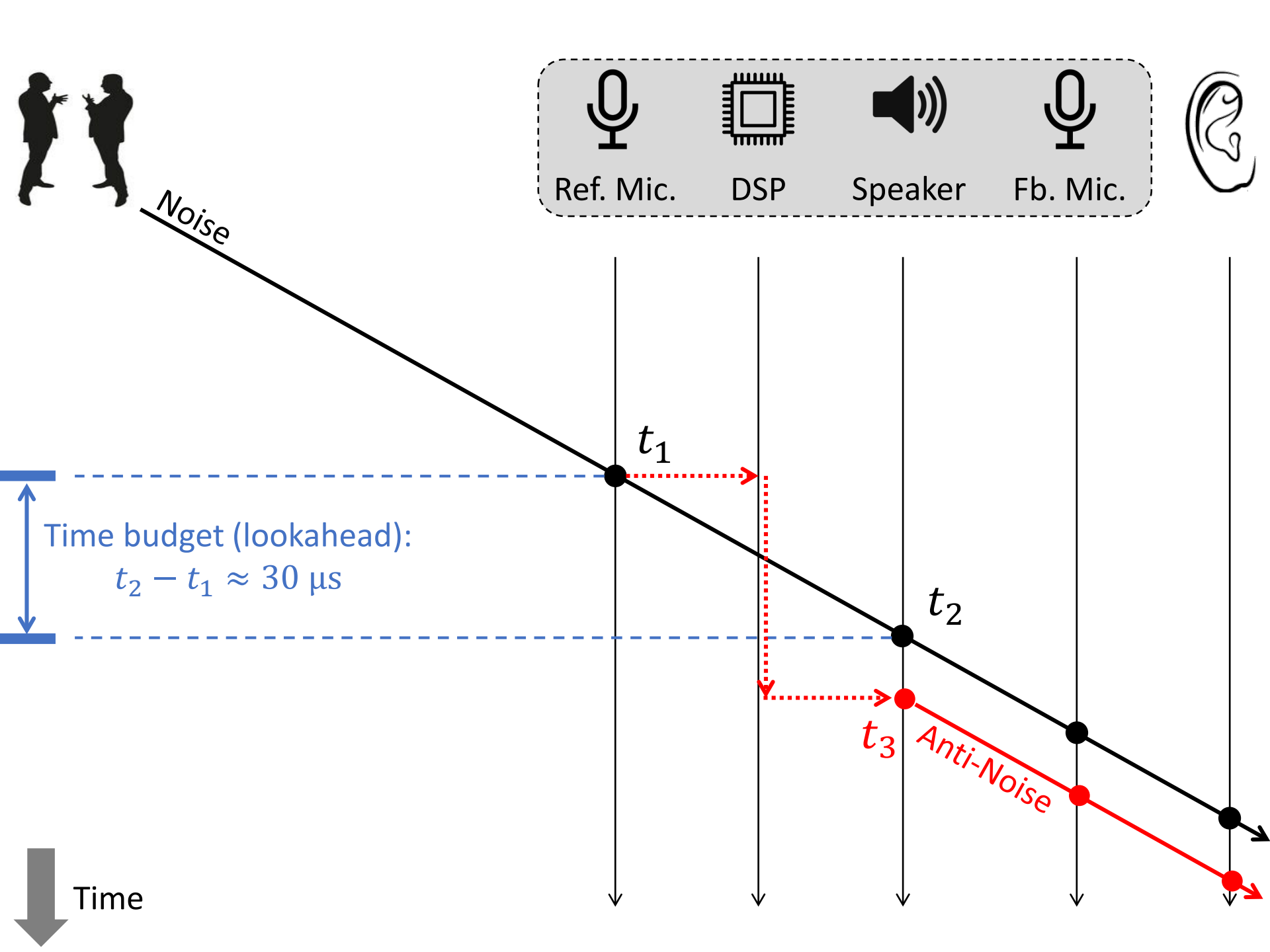


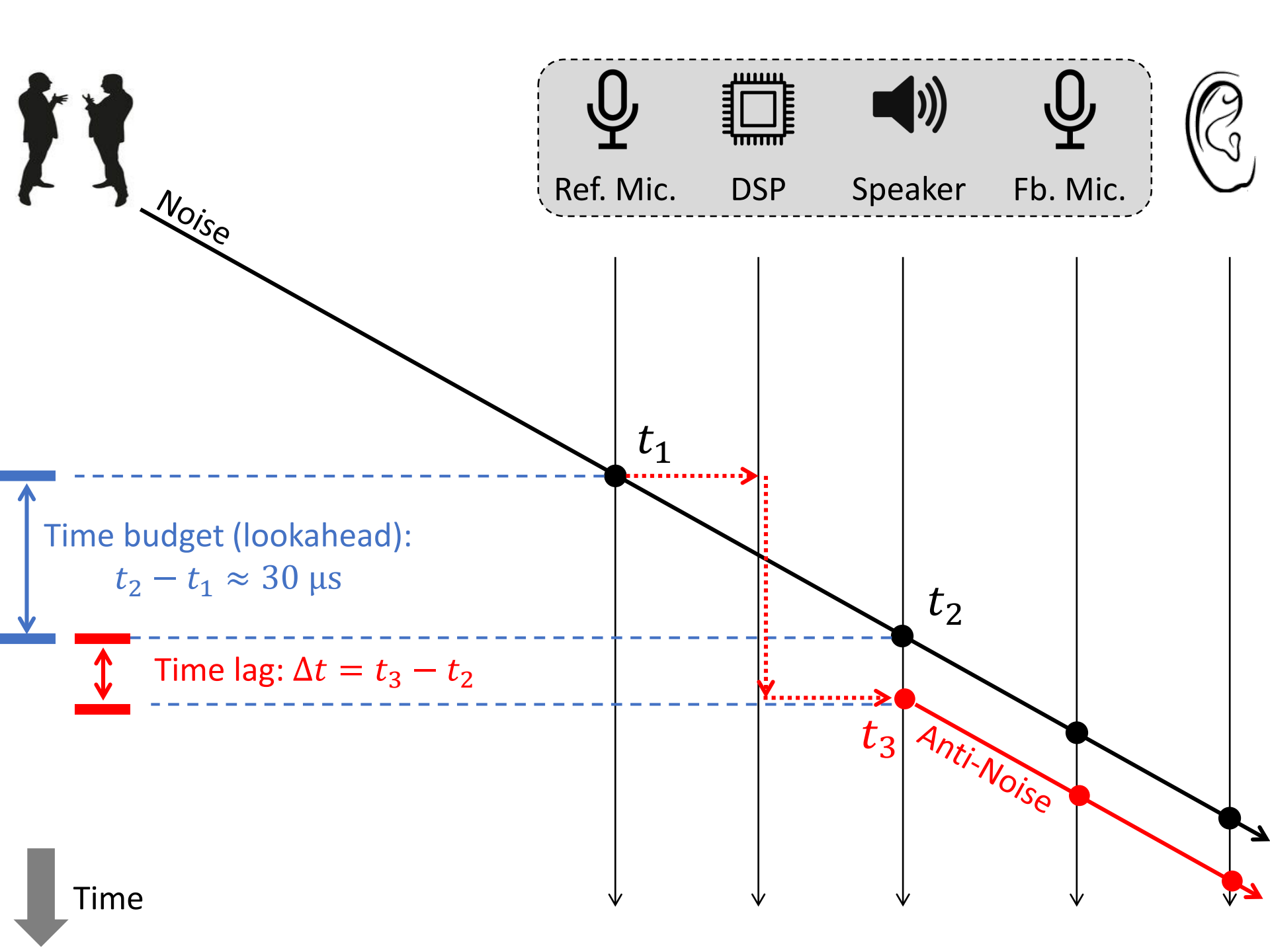

Time


Ref. Mic.
DSP
Speaker
Fb. Mic.
Noise
t_1
t_2
Time


Ref. Mic.
DSP
Speaker
Fb. Mic.
Noise
t_1
t_2
Time

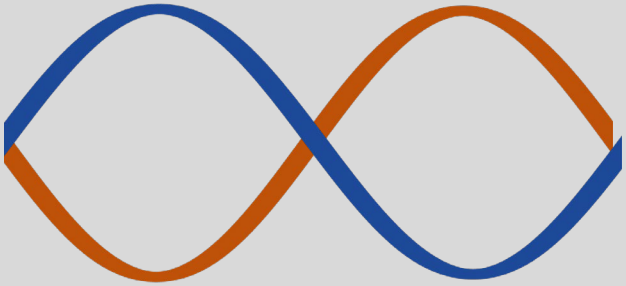

Ref. Mic.
DSP
Speaker
Fb. Mic.
Noise
1 cm
t_1
t_2
Time budget (lookahead):
$t_2 - t_1 \approx 30\ \mu s$
Time


Ref. Mic.
DSP
Speaker
Fb. Mic.
Noise
t_1
t_2
t_3
Anti-Noise
Time budget (lookahead):
$t_2 - t_1 \approx 30\ \mu s$
Time

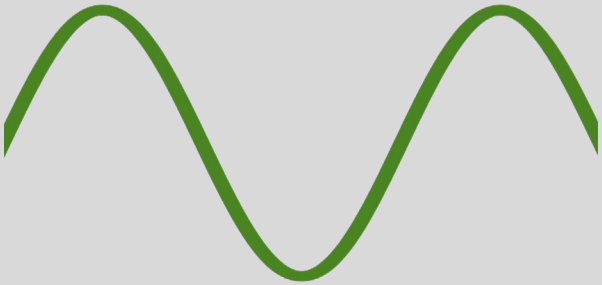

Ref. Mic.
DSP
Speaker
Fb. Mic.
Noise
t_1
t_2
t_3
Anti-Noise
Time budget (lookahead):
$t_2 - t_1 \approx 30\ \mu s$
Time lag: $\Delta t = t_3 - t_2$
Time

Low Frequency:

Noise



Anti-Noise

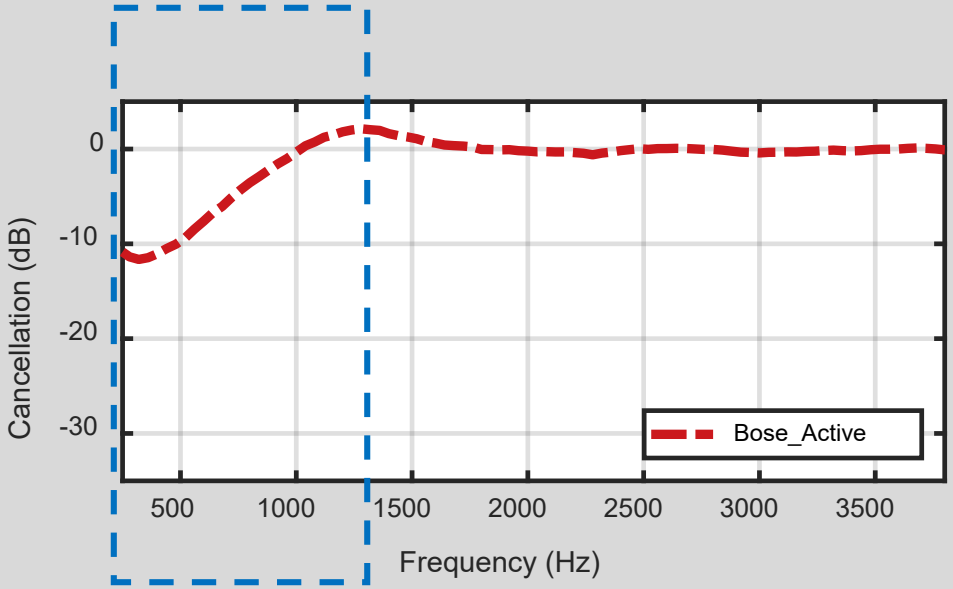


Residual Error

Time

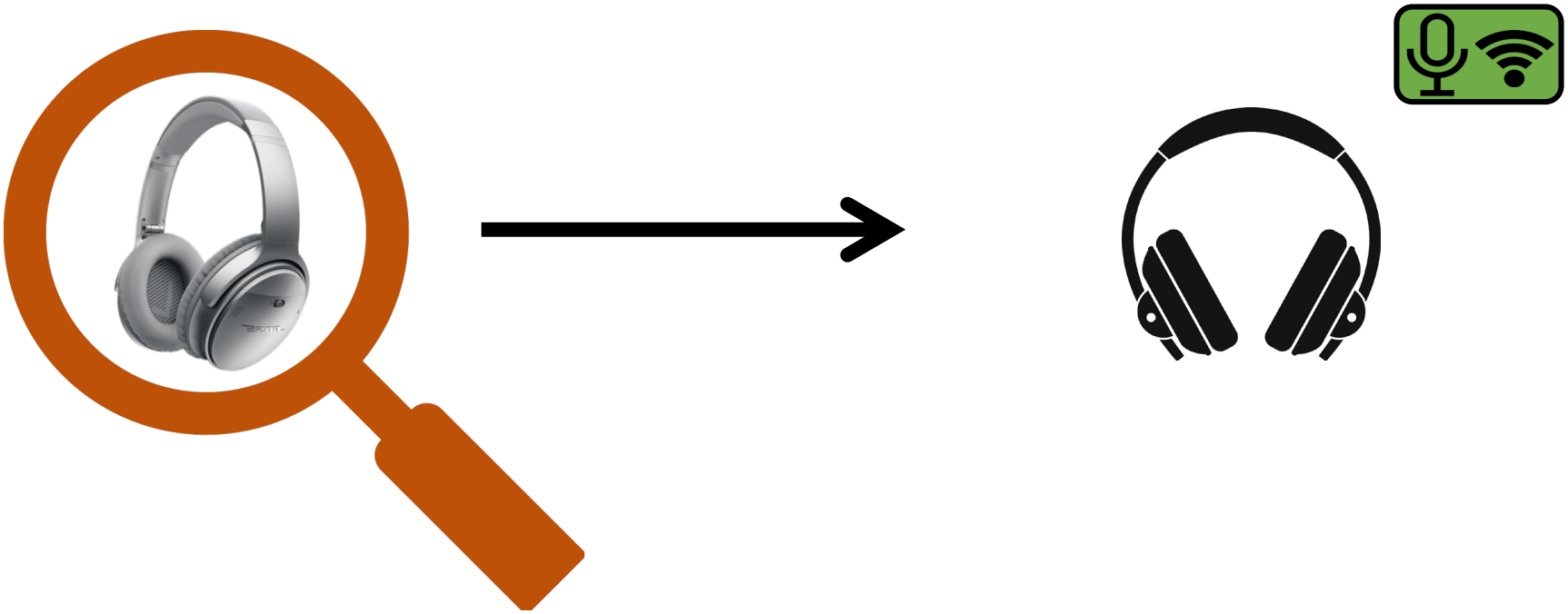
Noise

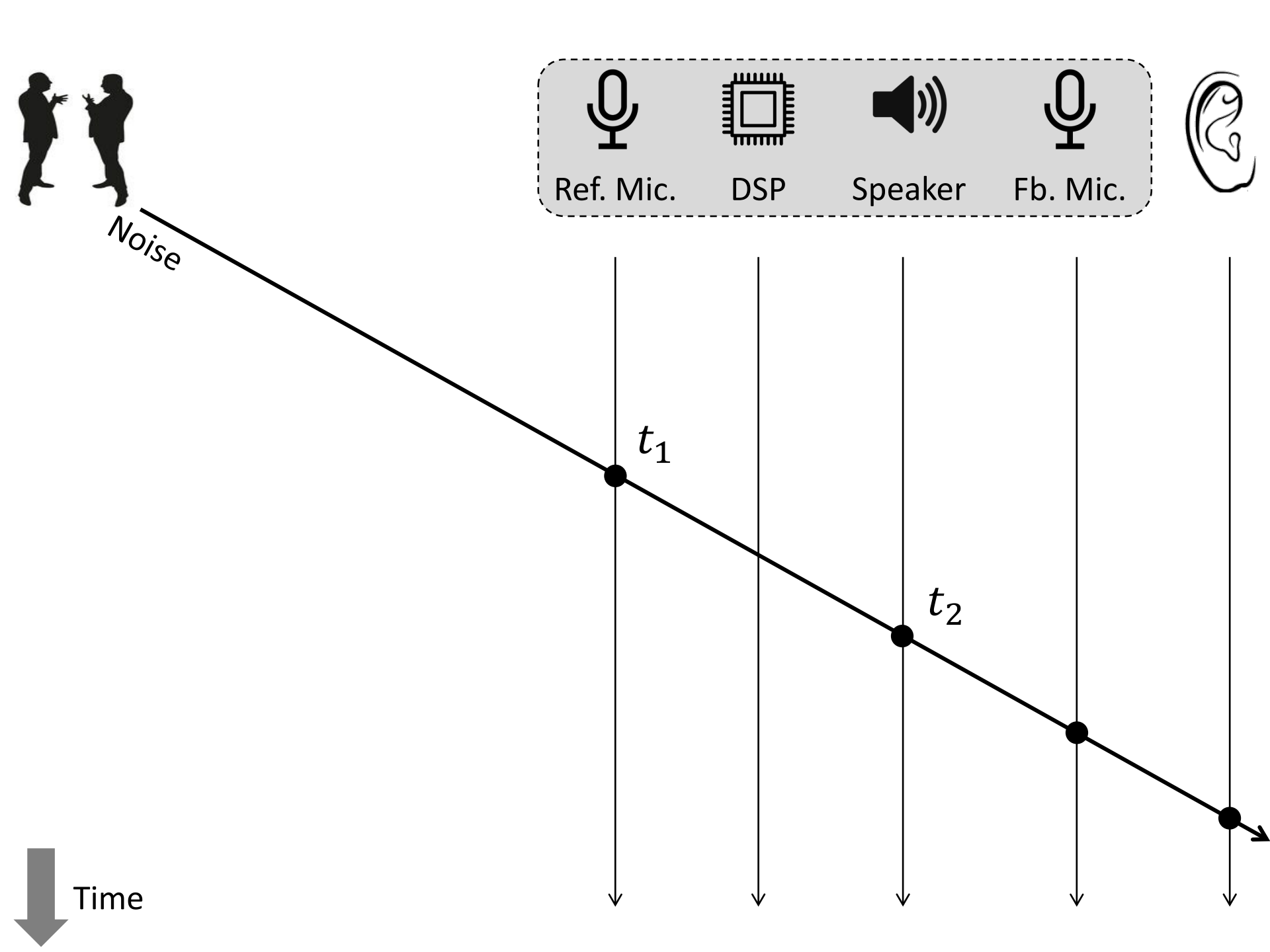
Ref. Mic. DSP Speaker Fb. Mic.

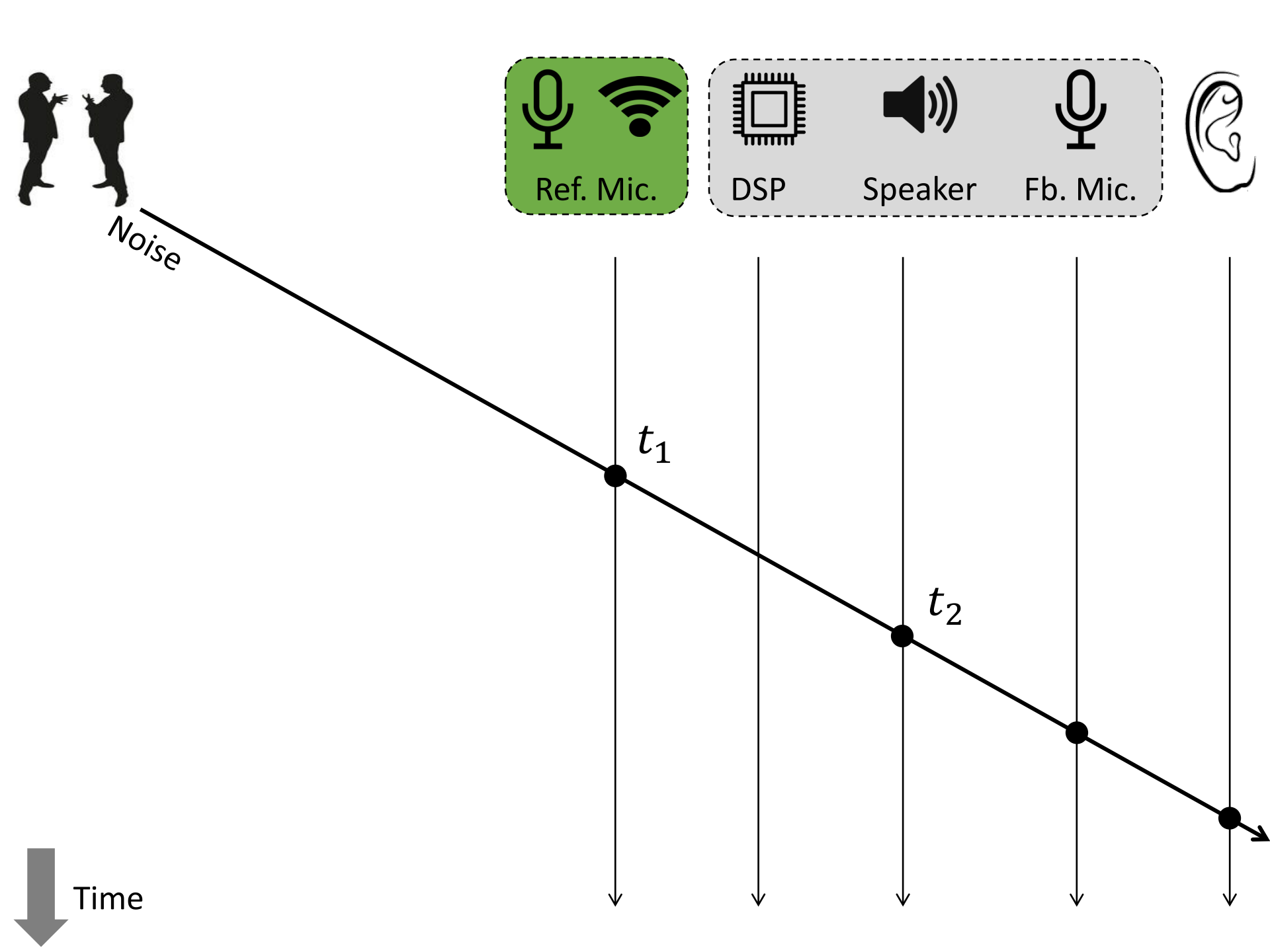


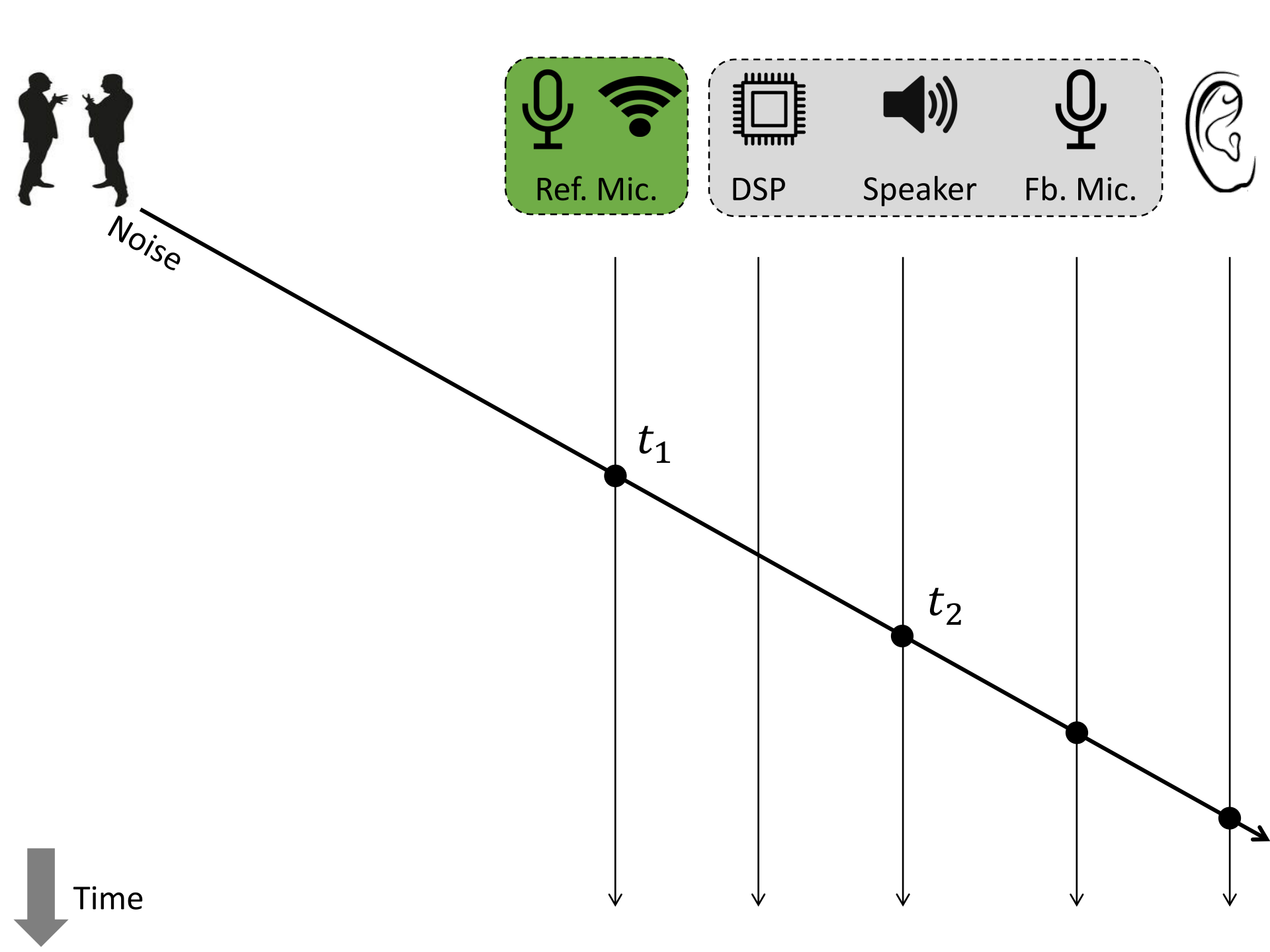
Time

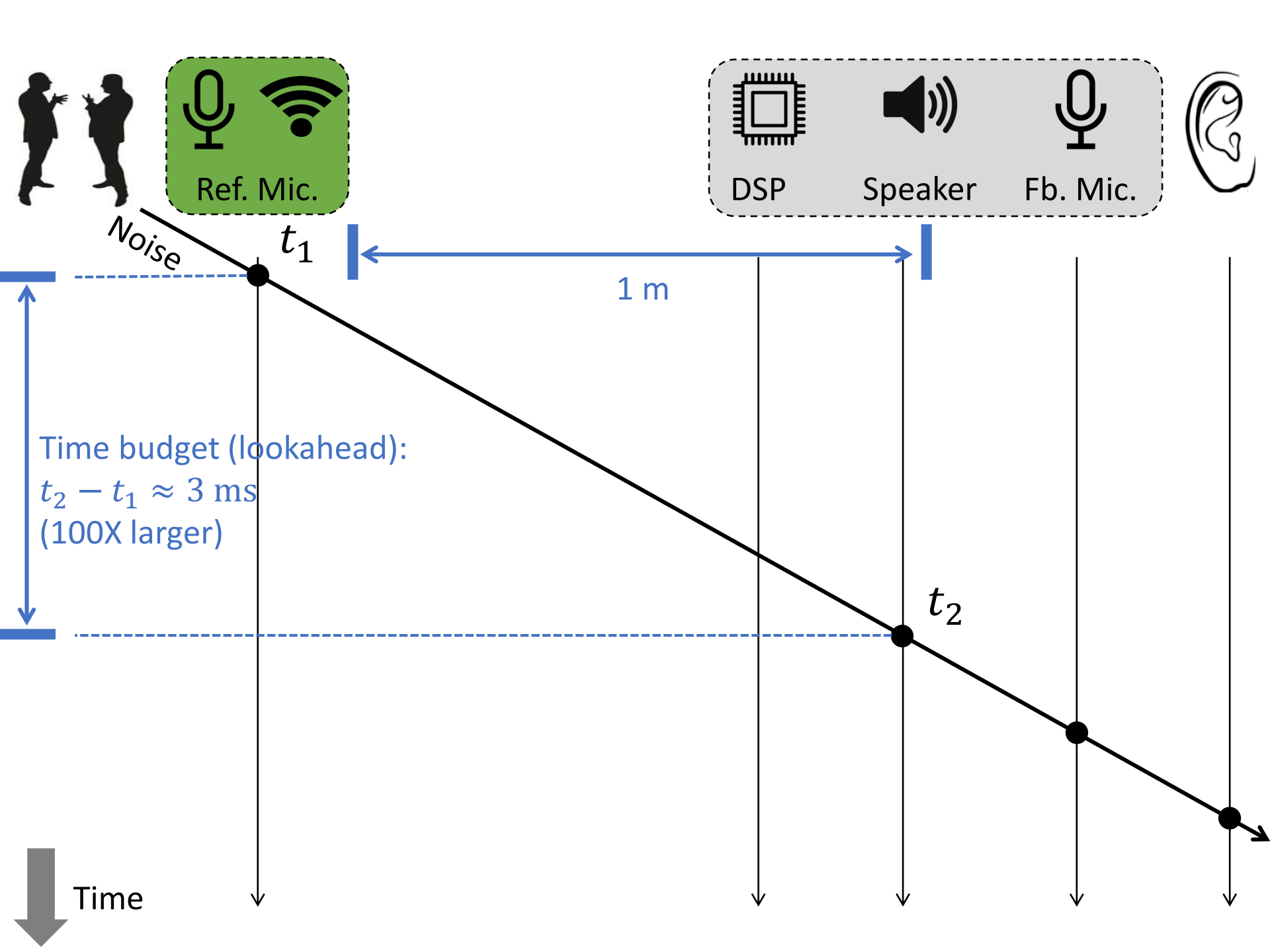
Let’s now look at noise cancellation in MUTE …

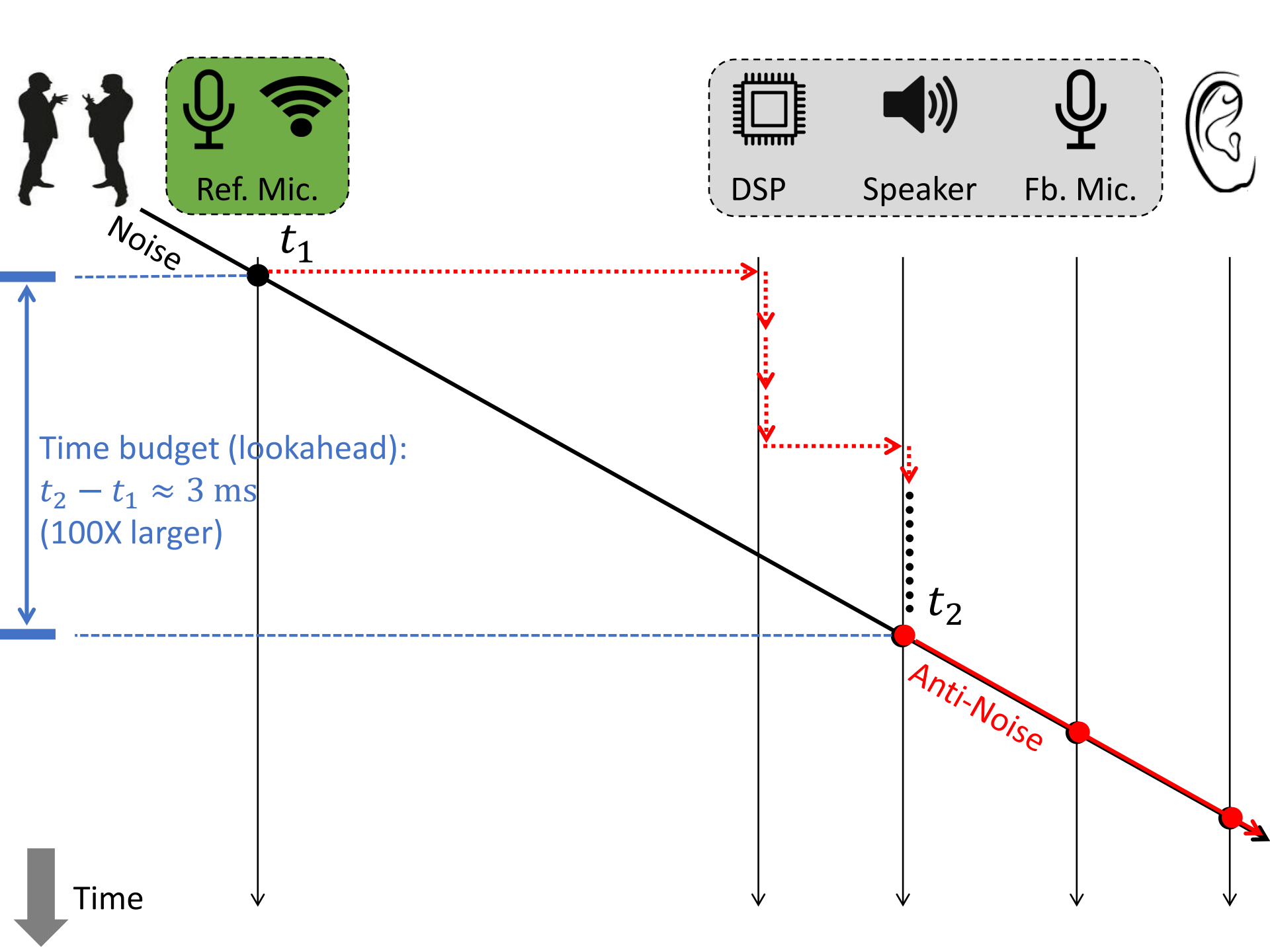


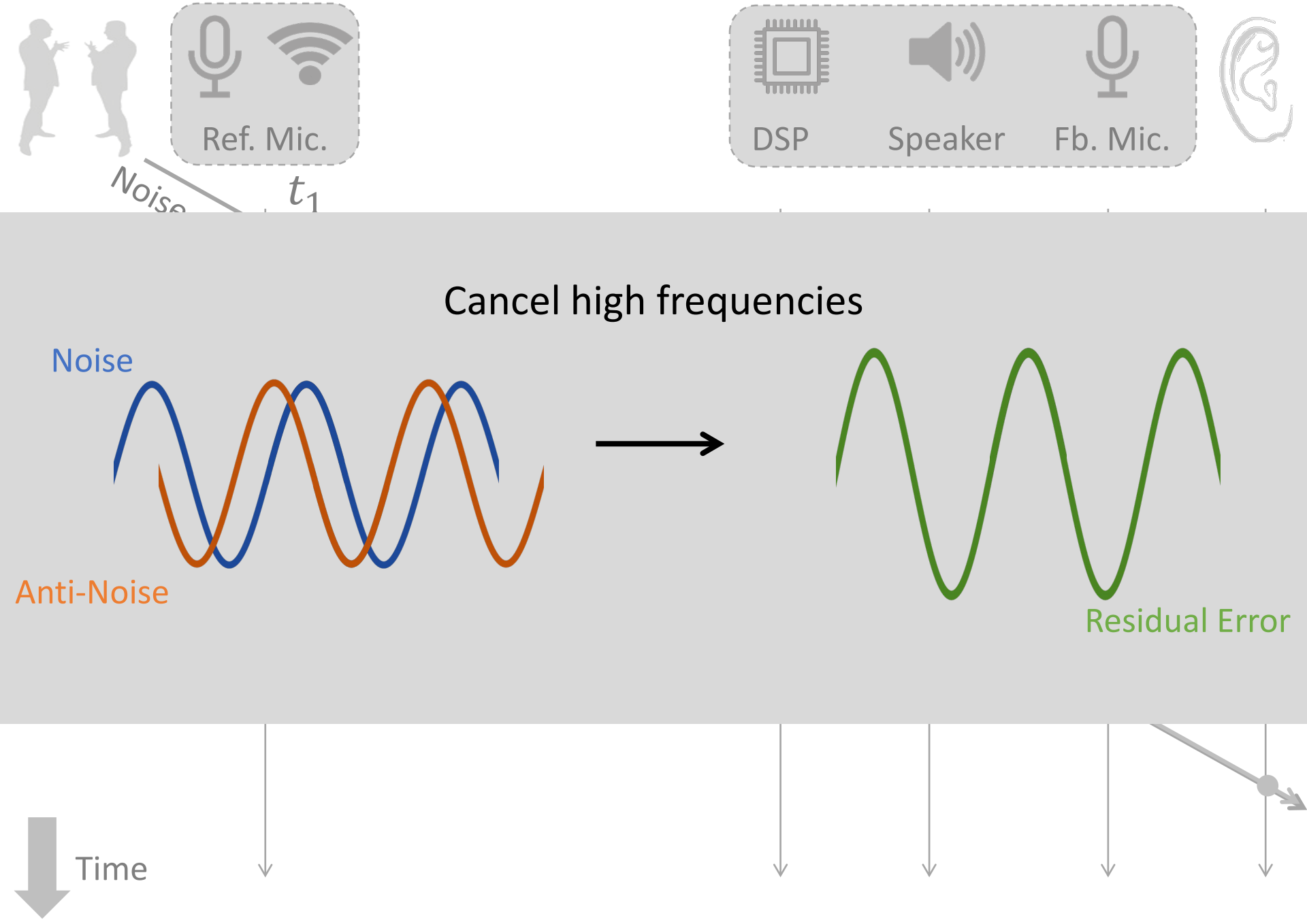

Ref. Mic.
DSP
Speaker
Fb. Mic.
Noise
t_1
t_2
Time


Ref. Mic.
DSP
Speaker
Fb. Mic.
Noise
t_1
t_2
Time


Ref. Mic.
DSP
Speaker
Fb. Mic.
Noise
t_1
t_2
Time


Ref. Mic.
DSP
Speaker
Fb. Mic.
Noise
t_1
1 m
Time budget (lookahead):
$t_2 - t_1 \approx 3$ ms
(100X larger)
t_2
Time


Ref. Mic.
DSP
Speaker
Fb. Mic.
Noise
t_1
Time budget (lookahead):
$t_2 - t_1 \approx 3$ ms
(100X larger)
t_2
Anti-Noise
Time


Ref. Mic.
DSP
Speaker
Fb. Mic.
Noise
t_1
Cancel high frequencies
Noise
Anti-Noise
Residual Error
Time

Talk Outline

How can $MUTE$ leverage lookahead?

Timing Gain → Wideband Cancellation

Signal Processing Gain → Non-Causal Filtering

Application-Specific Gain → Sound Source Profiling

Talk Outline

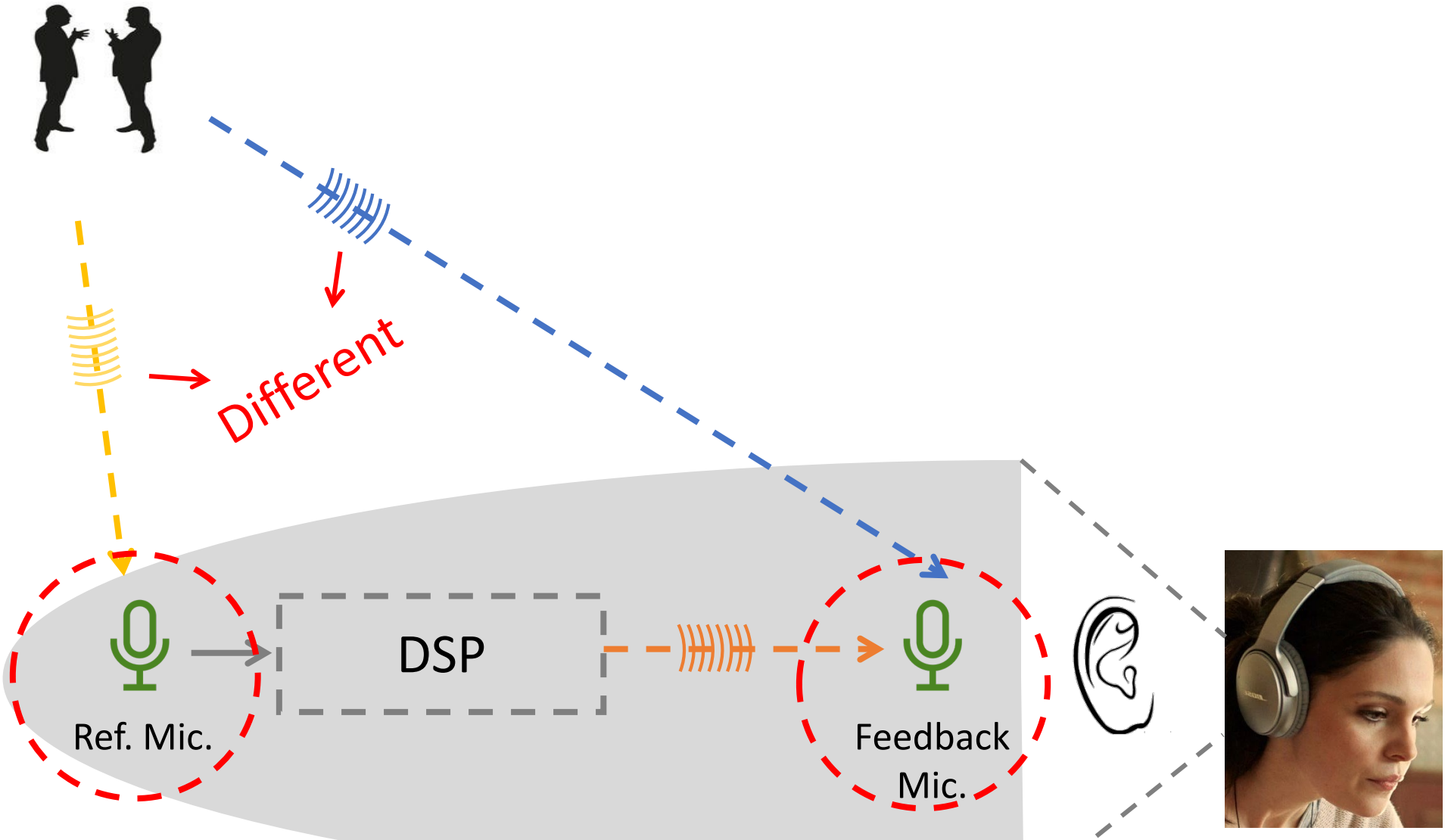
How can $MUTE$ leverage lookahead?

Timing Gain → Wideband Cancellation

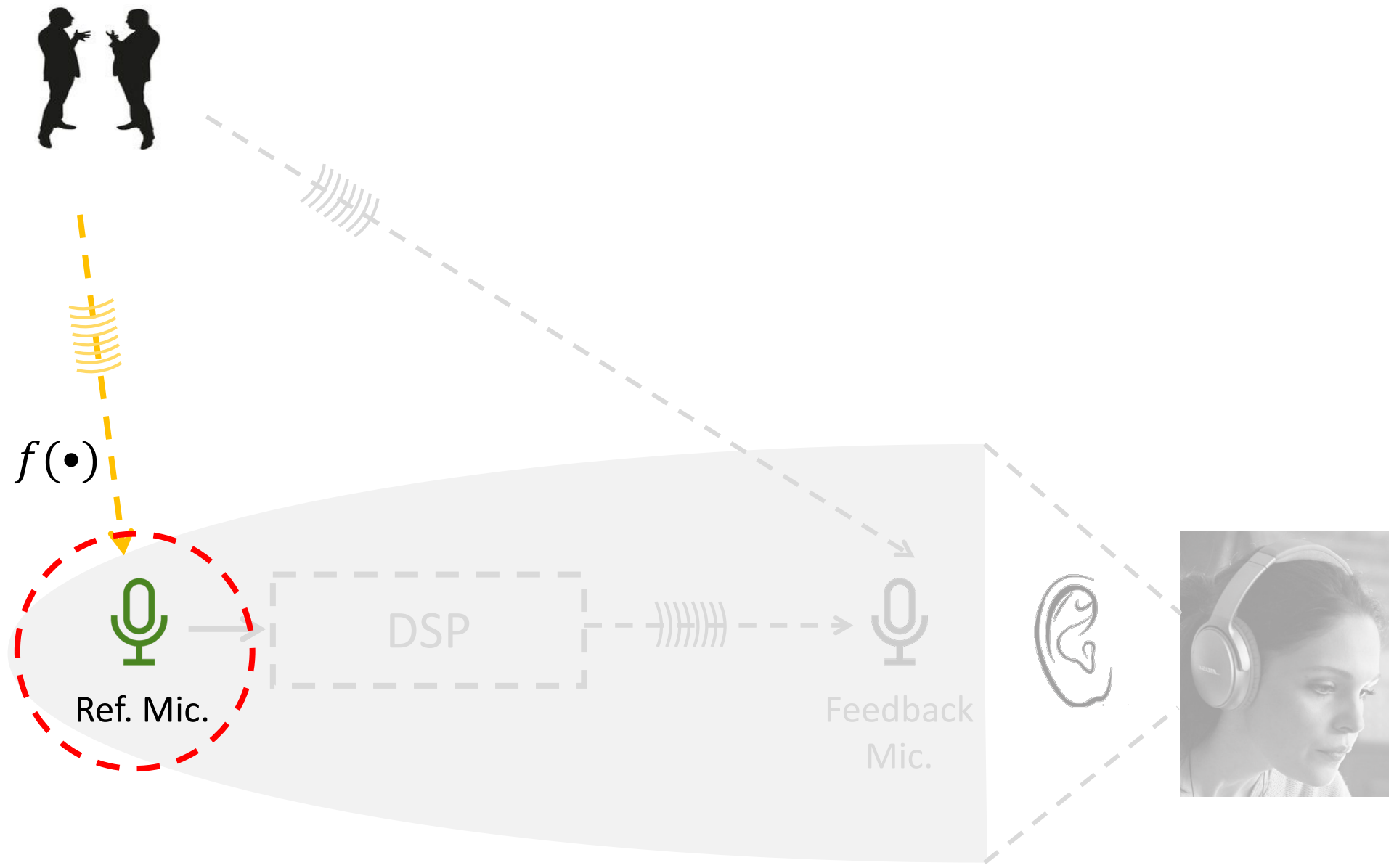
Signal Processing Gain → Non-Causal Filtering

Application-Specific Gain → Sound Source Profiling

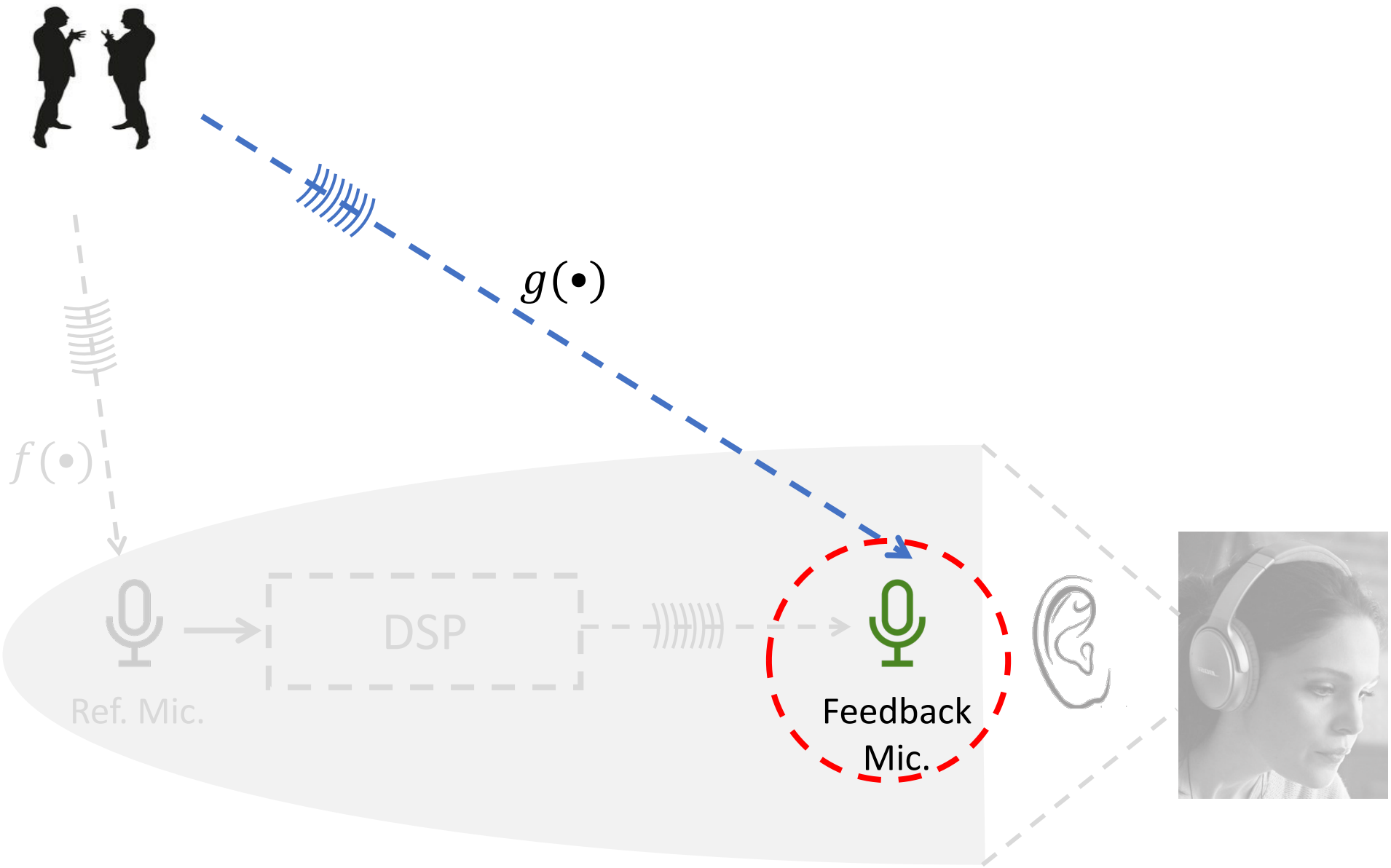
Noise Cancellation Primer



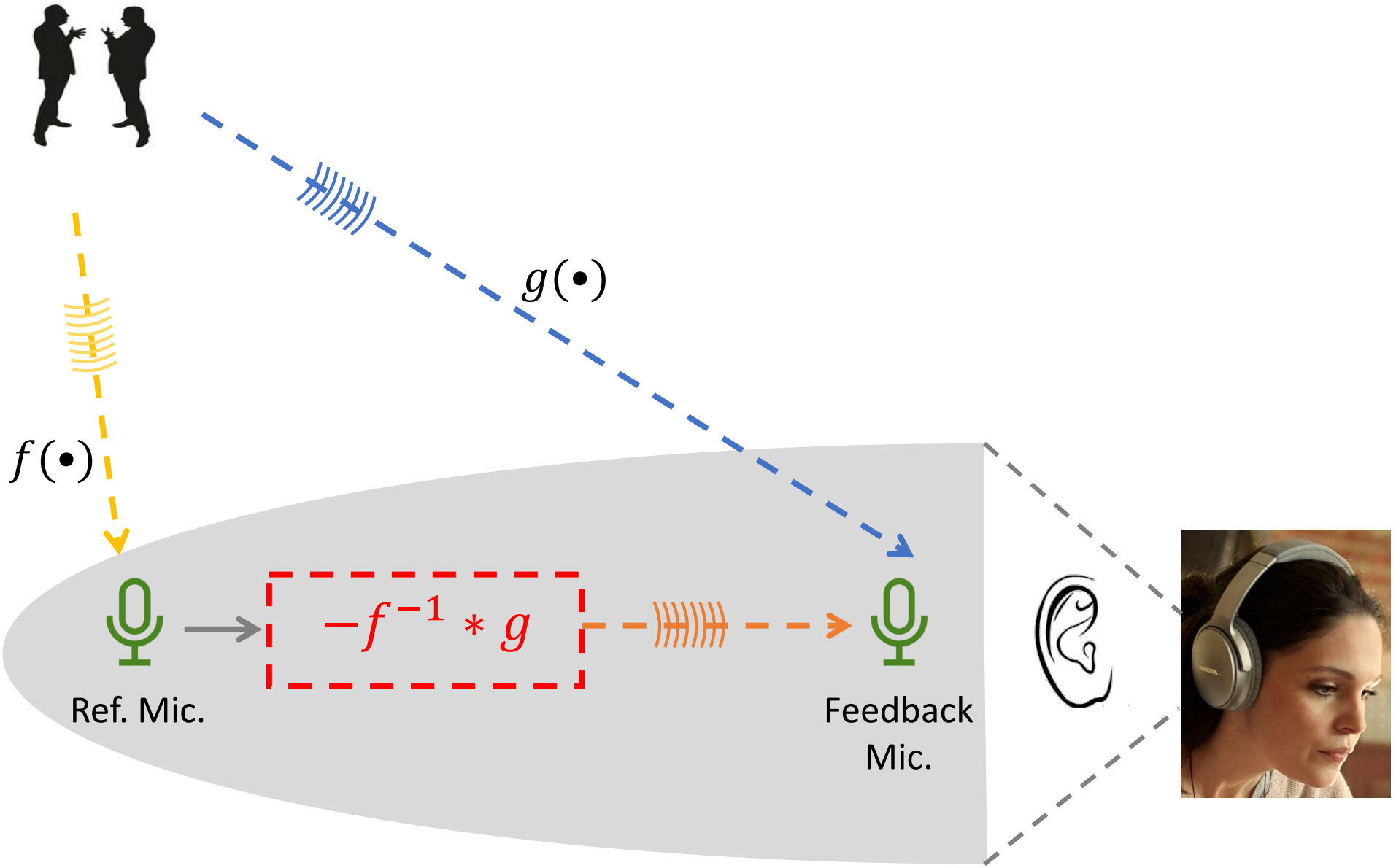
Noise Cancellation Primer



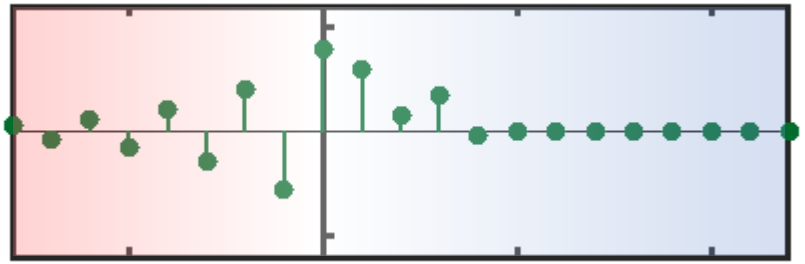
Noise Cancellation Primer



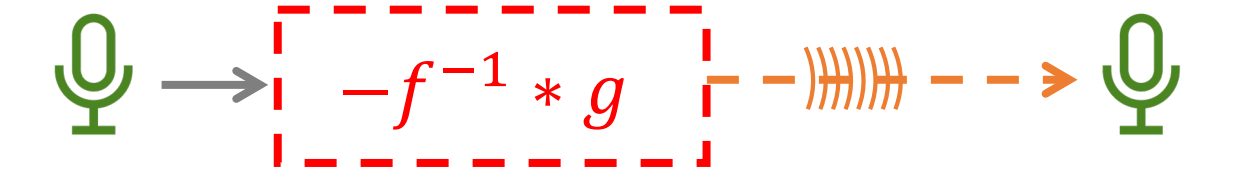
Noise Cancellation Primer



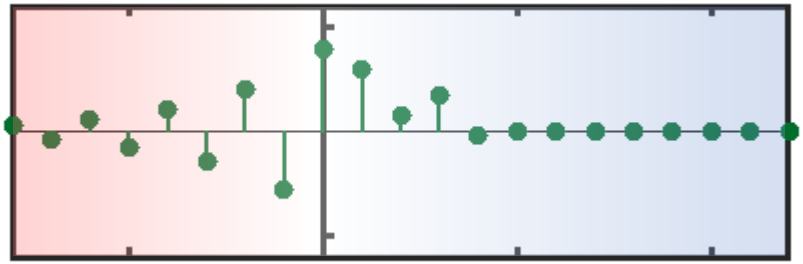
Channel Impulse Response



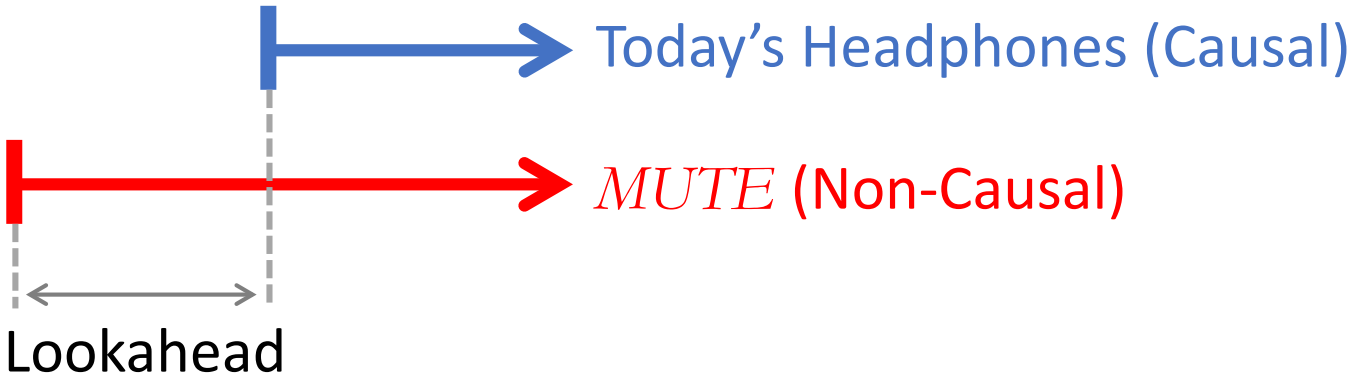
Future Current Past



Channel Impulse Response



Future Current Past



"Lookahead" → Non-Causal Filtering → Better Cancellation

Talk Outline

How can $MUTE$ leverage lookahead?

Timing Gain → Wideband Cancellation

Signal Processing Gain → Non-Causal Filtering

Application-Specific Gain → Sound Source Profiling

Talk Outline

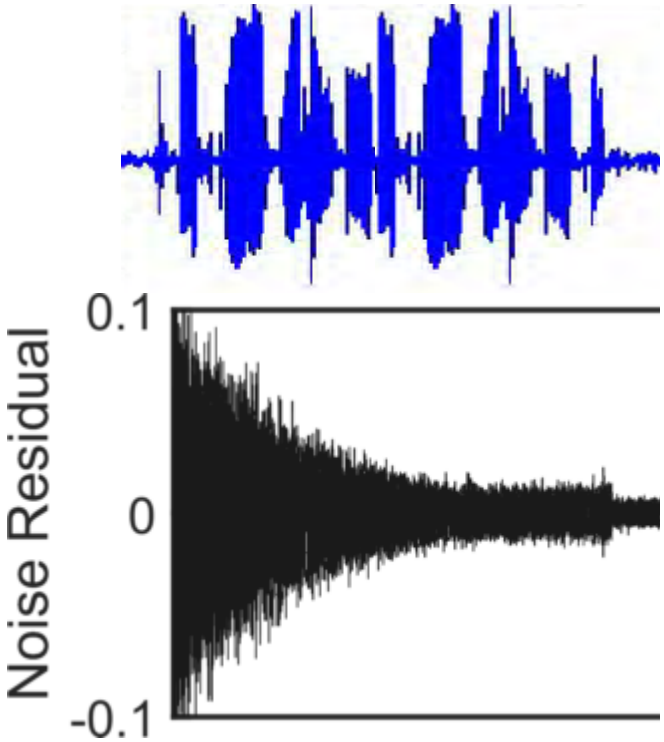
How can $MUTE$ leverage lookahead?

Timing Gain → Wideband Cancellation

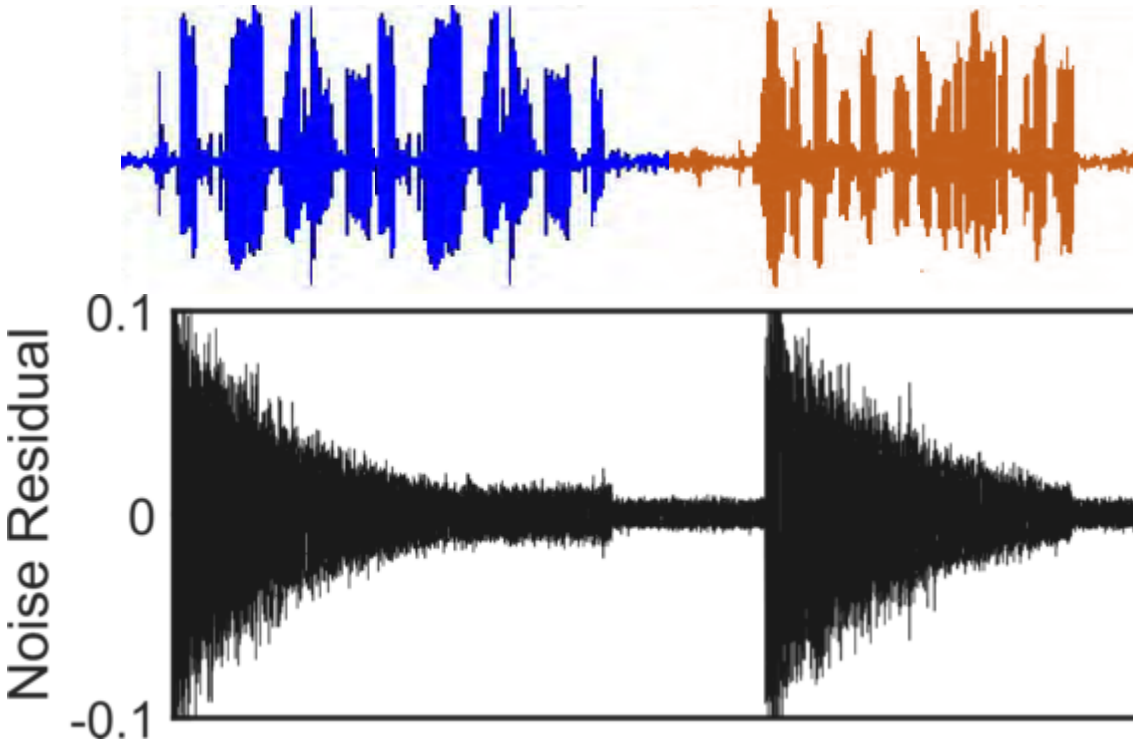
Signal Processing Gain → Non-Causal Filtering

Application-Specific Gain → Sound Source Profiling

Alice

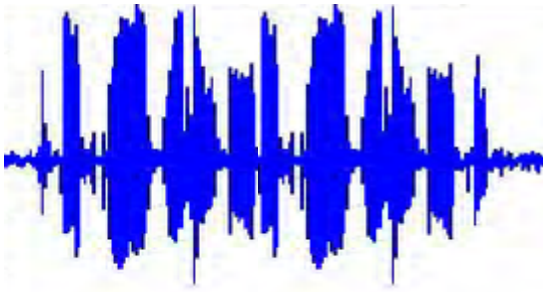


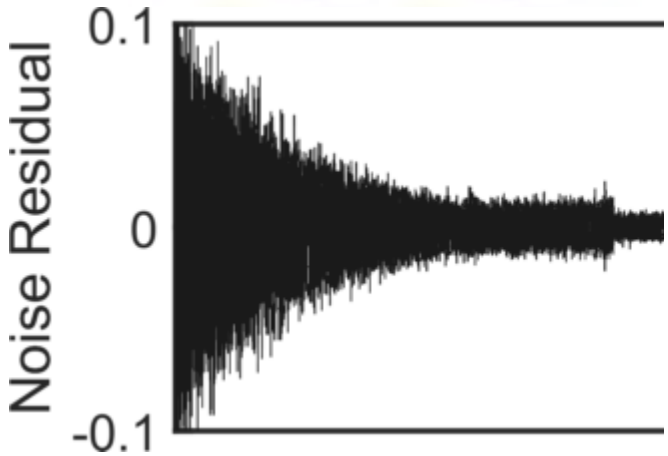
Alice Bob



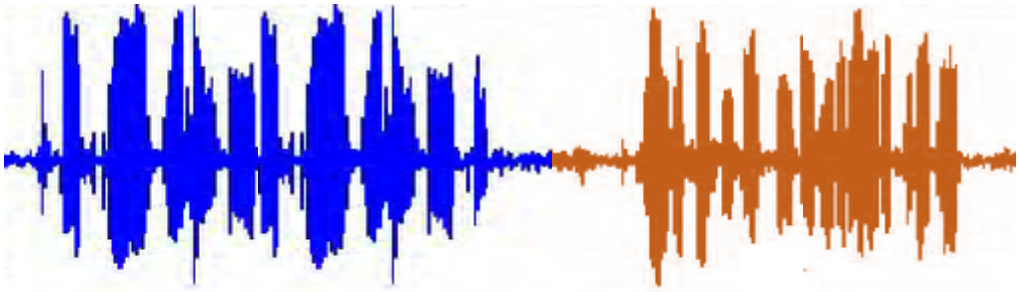
With lookahead …

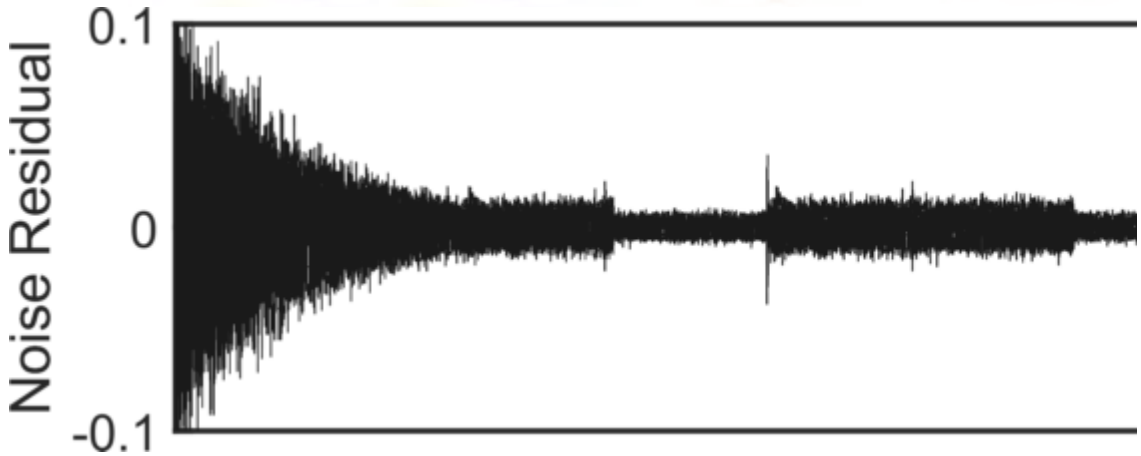
Alice





Alice Bob





Talk Outline

How can $MUTE$ leverage lookahead?

Timing Gain → Wideband Cancellation

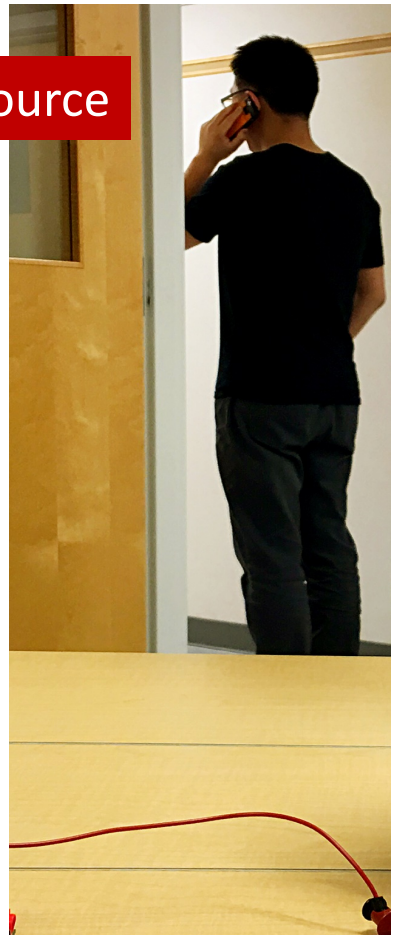
Signal Processing Gain → Non-Causal Filtering

Application-Specific Gain → Sound Source Profiling

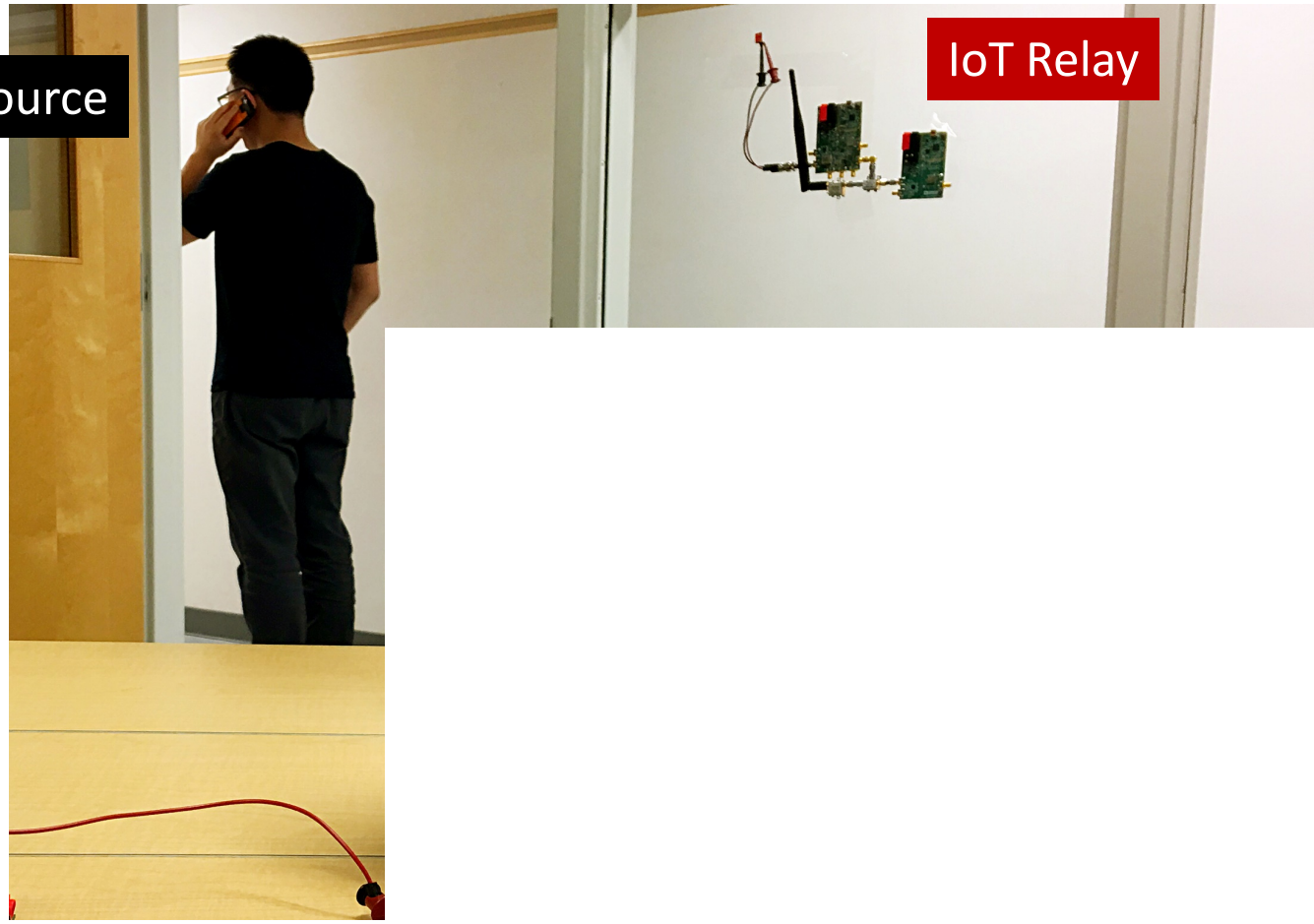
Implementation & Evaluation

MUTE Platform

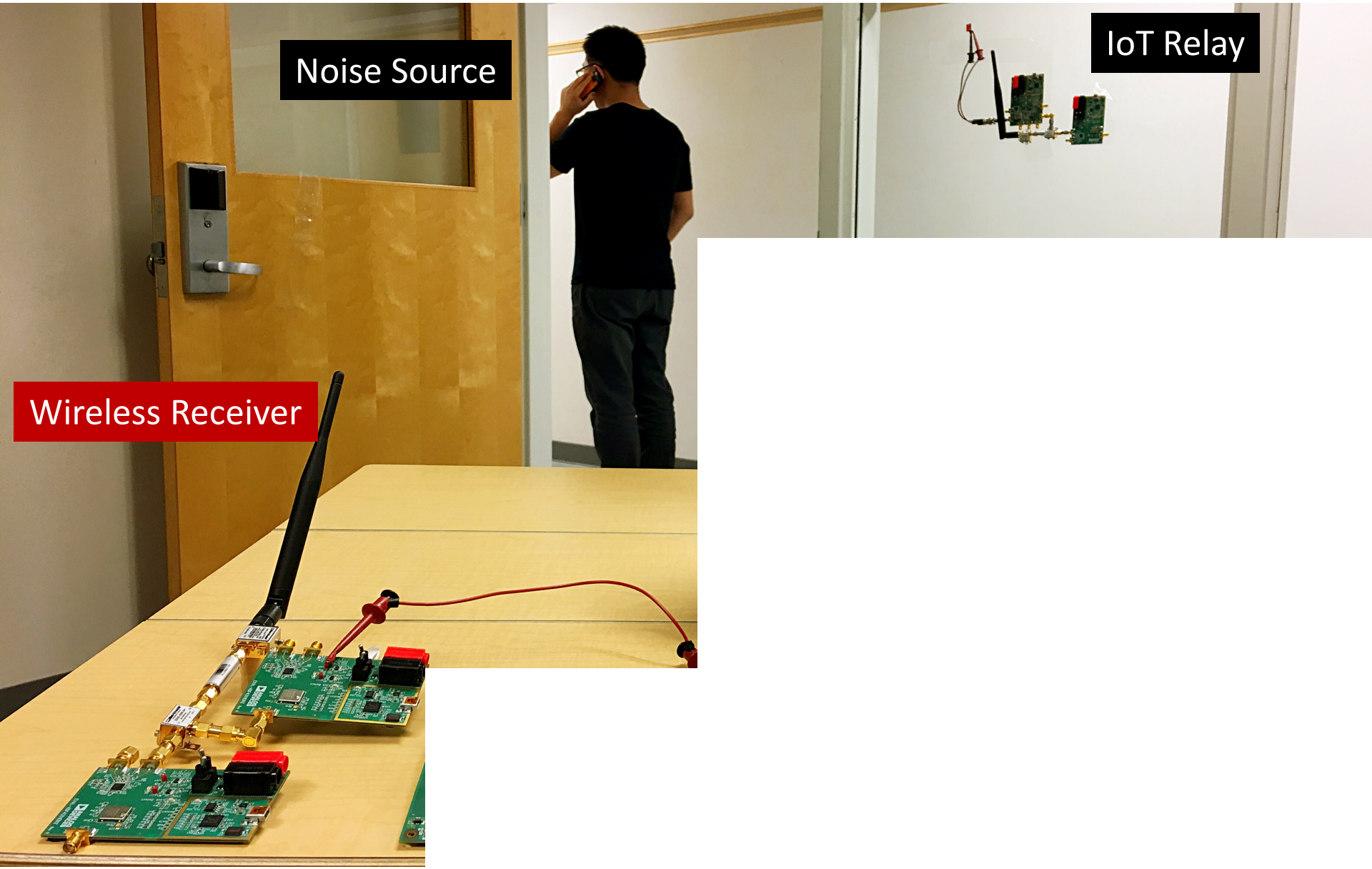
Noise Source



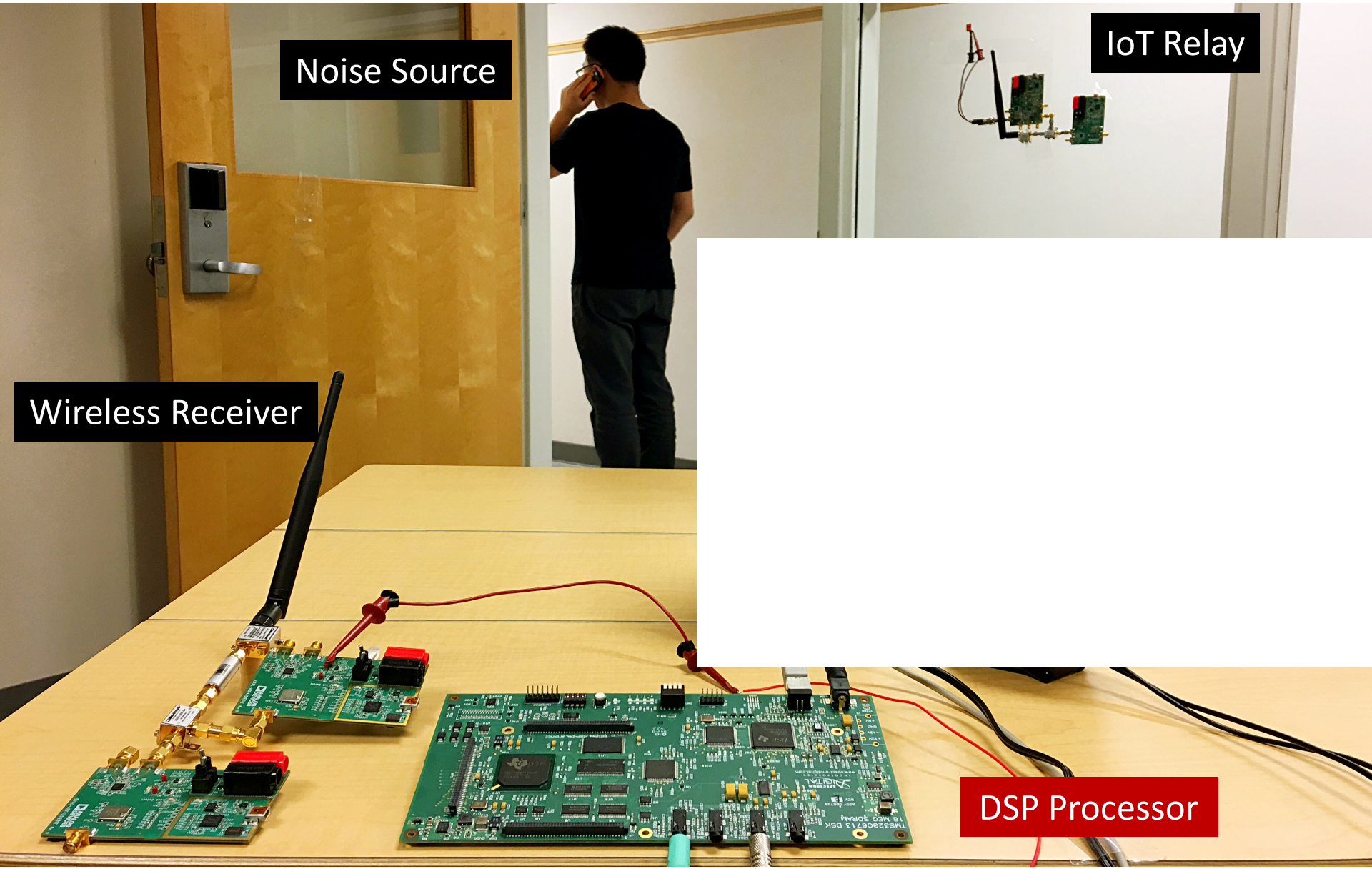
MUTE Platform



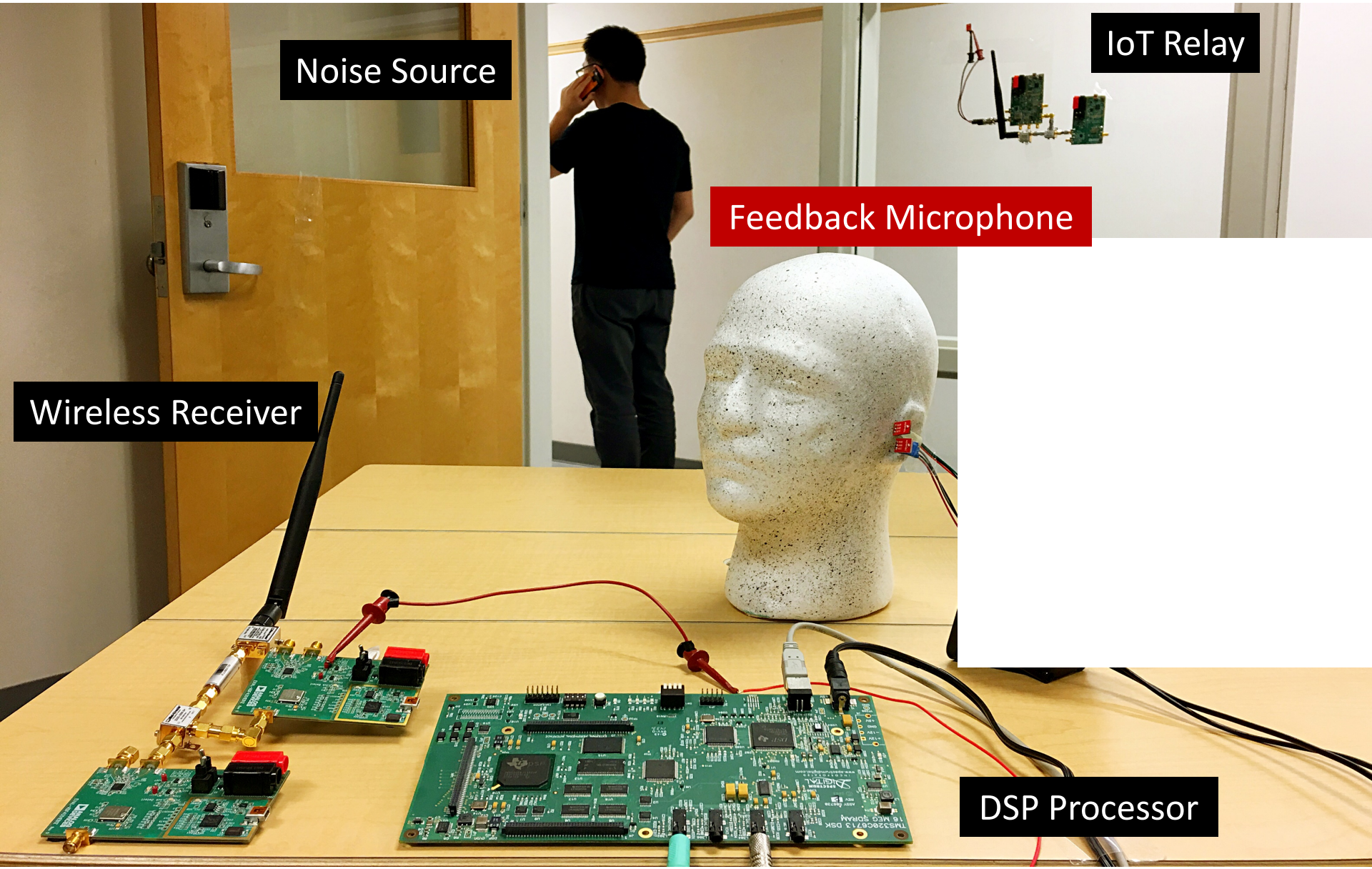
MUTE Platform



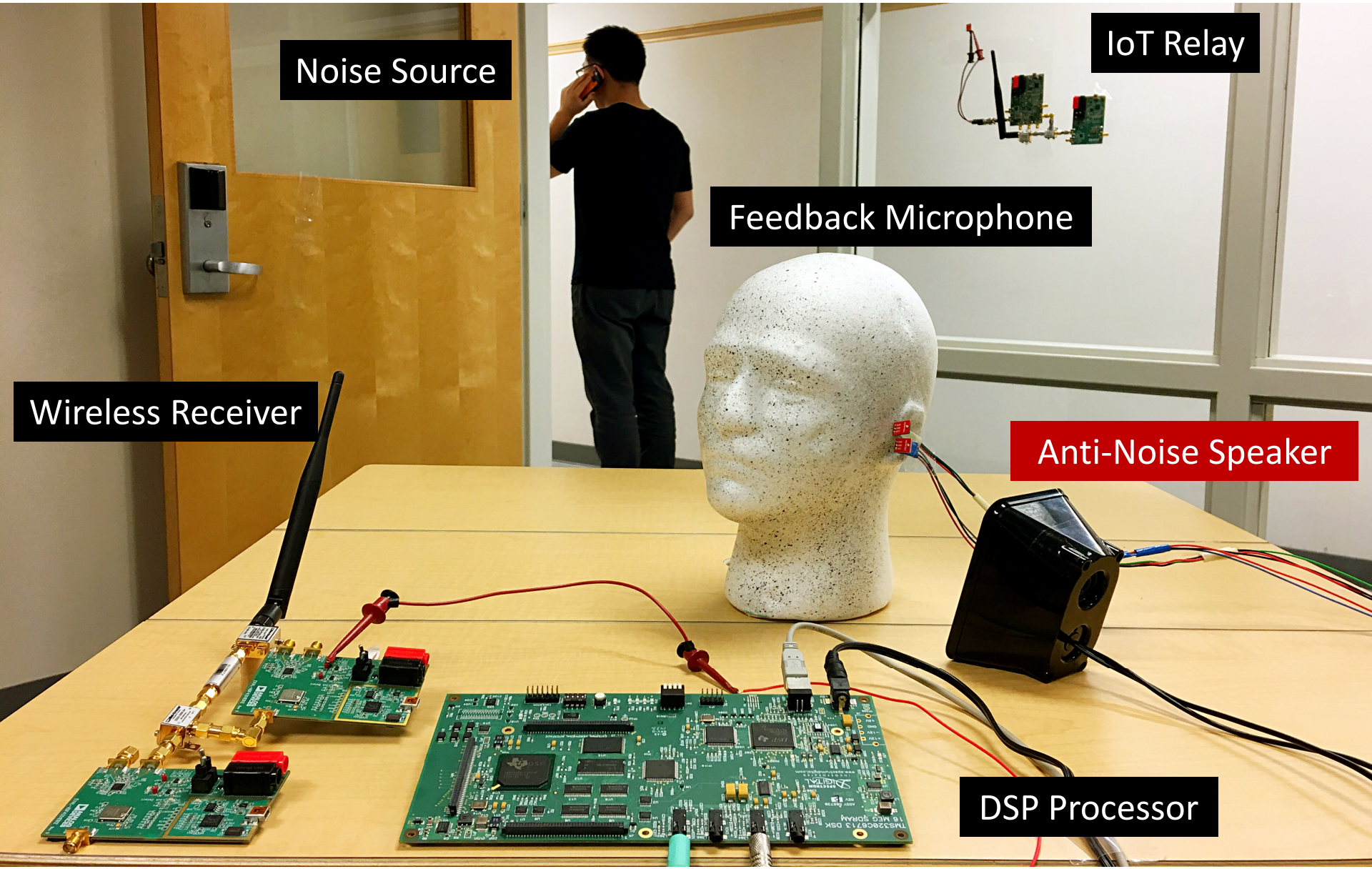
MUTE Platform



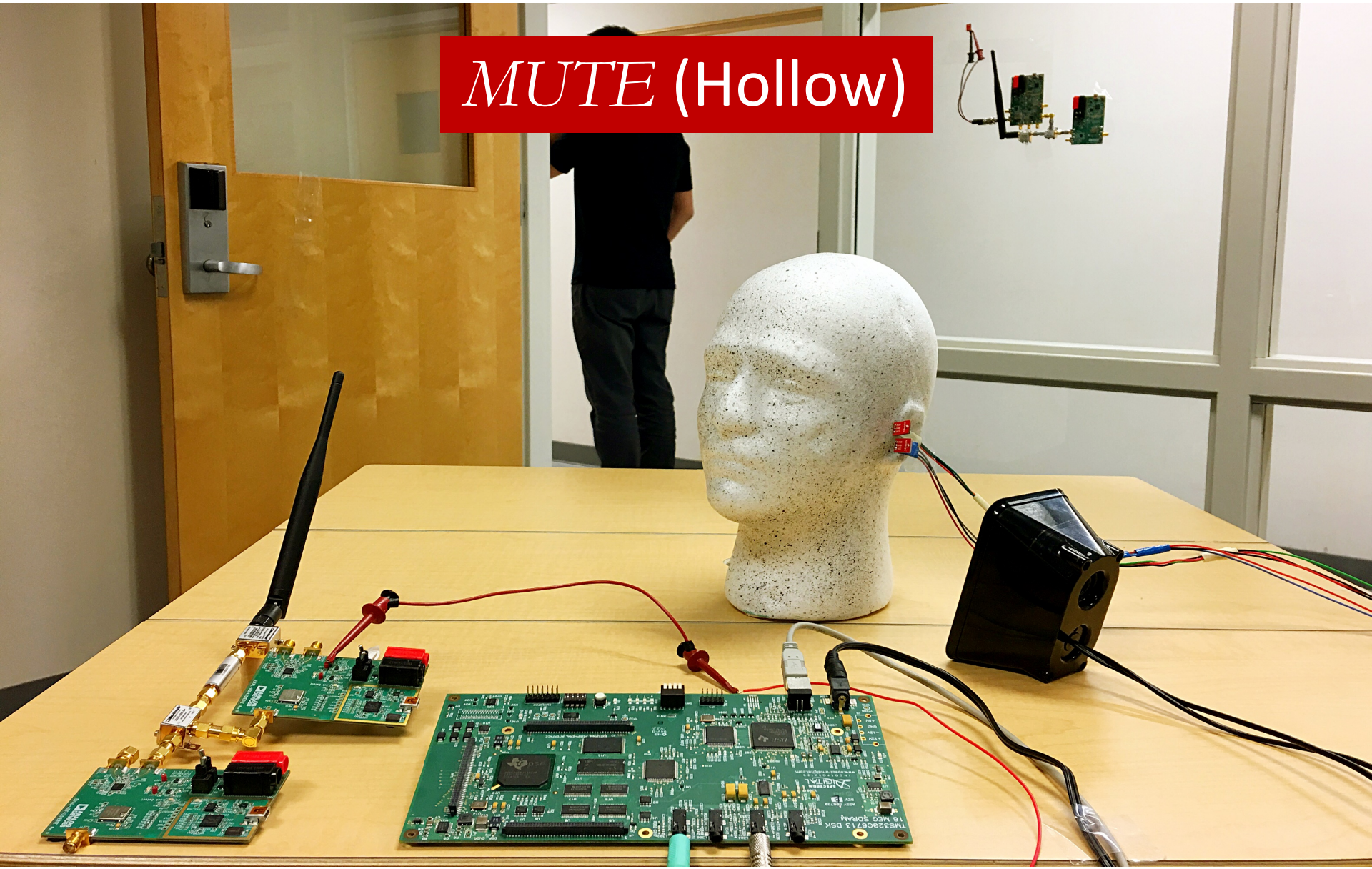
MUTE Platform



MUTE Platform

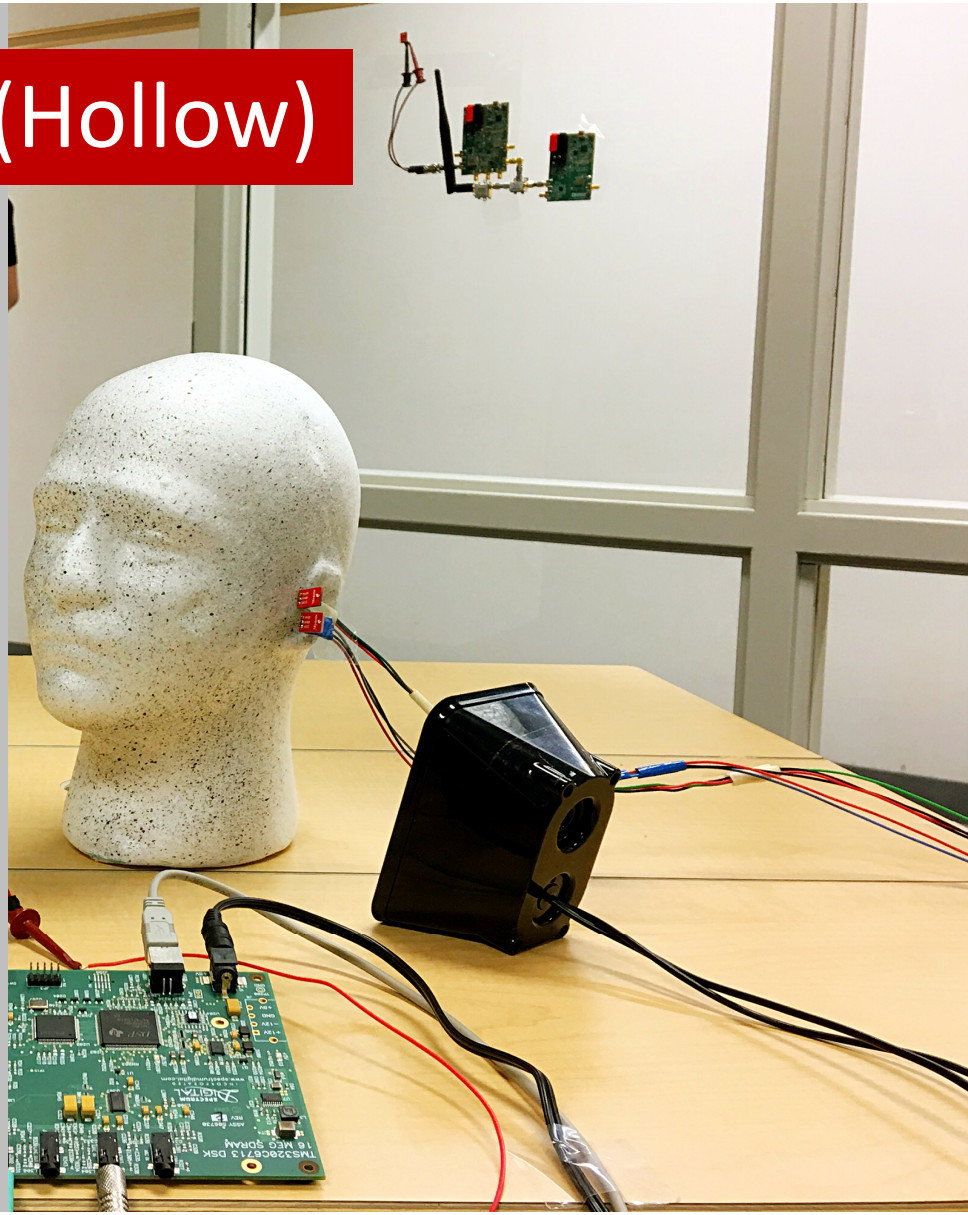


MUTE Platform



MUTE Platform

MUTE (Hollow)



Evaluation

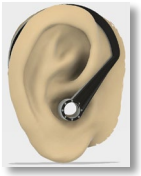
Goal: Comparable Performance

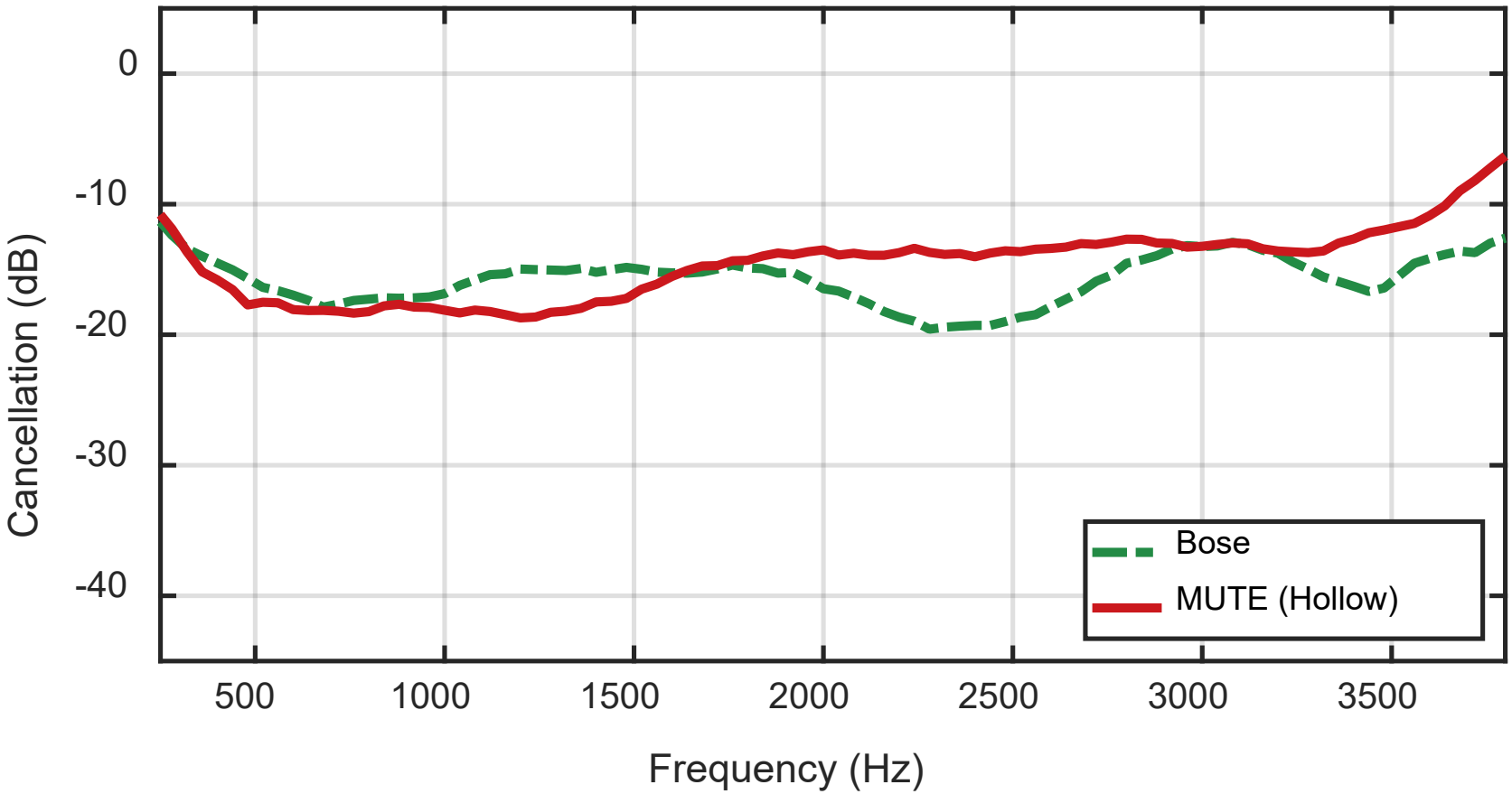
MUTE (Hollow)

Ear Blocking Headphone

1. Bose headphone

2. *MUTE* hollow design (comfort)

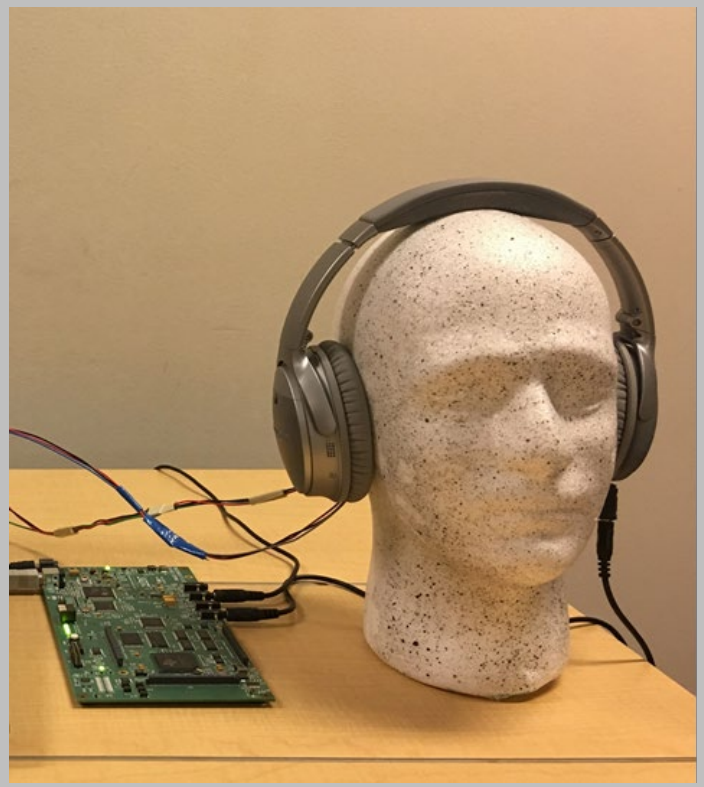




Evaluation

Goal: Better Performance

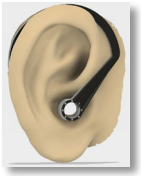
MUTE (Non-Hollow)



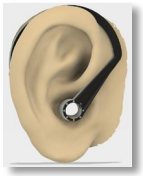
Ear Blocking Headphone

1. Bose headphone

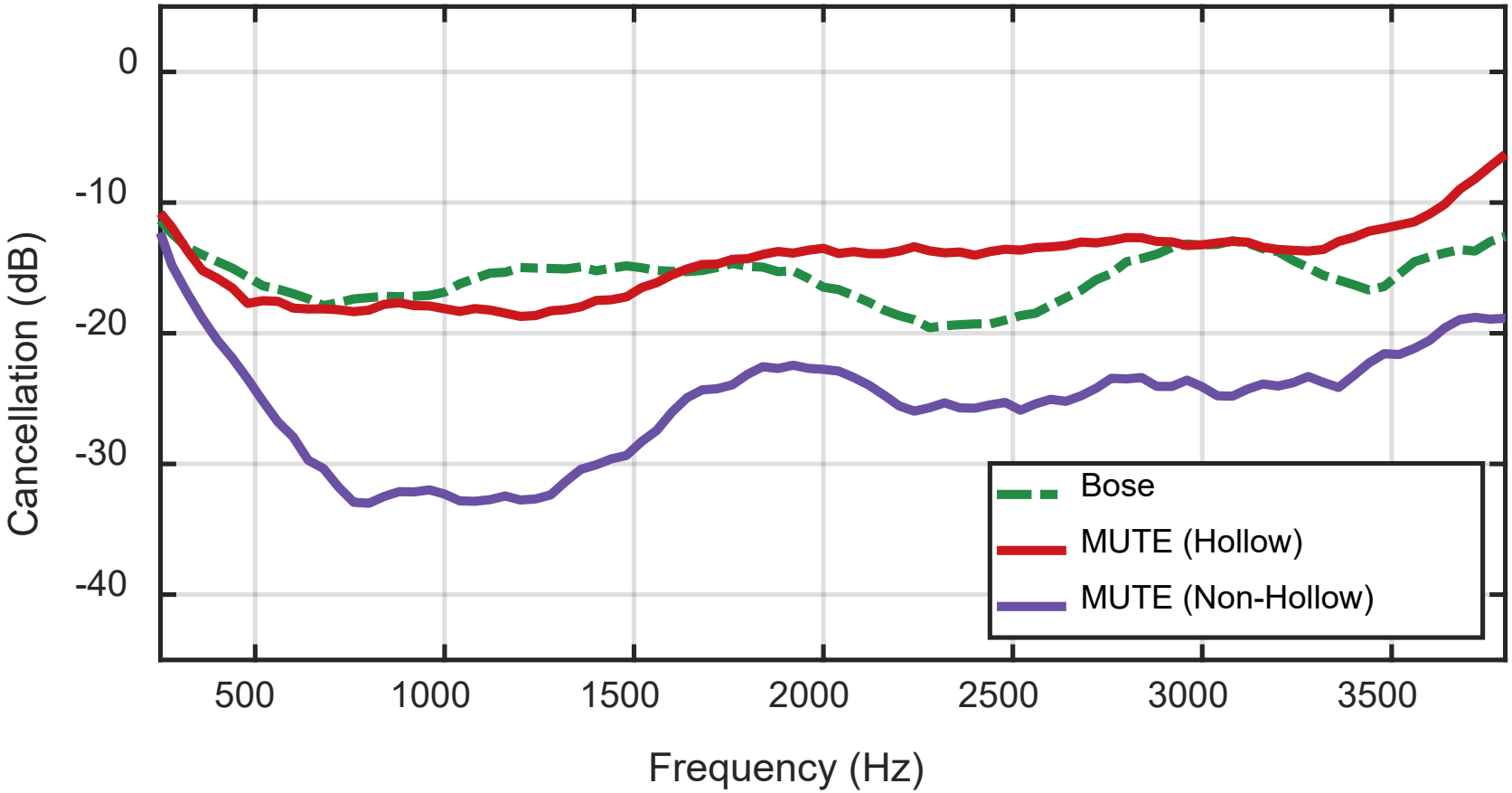
2. *MUTE* hollow design (comfort)



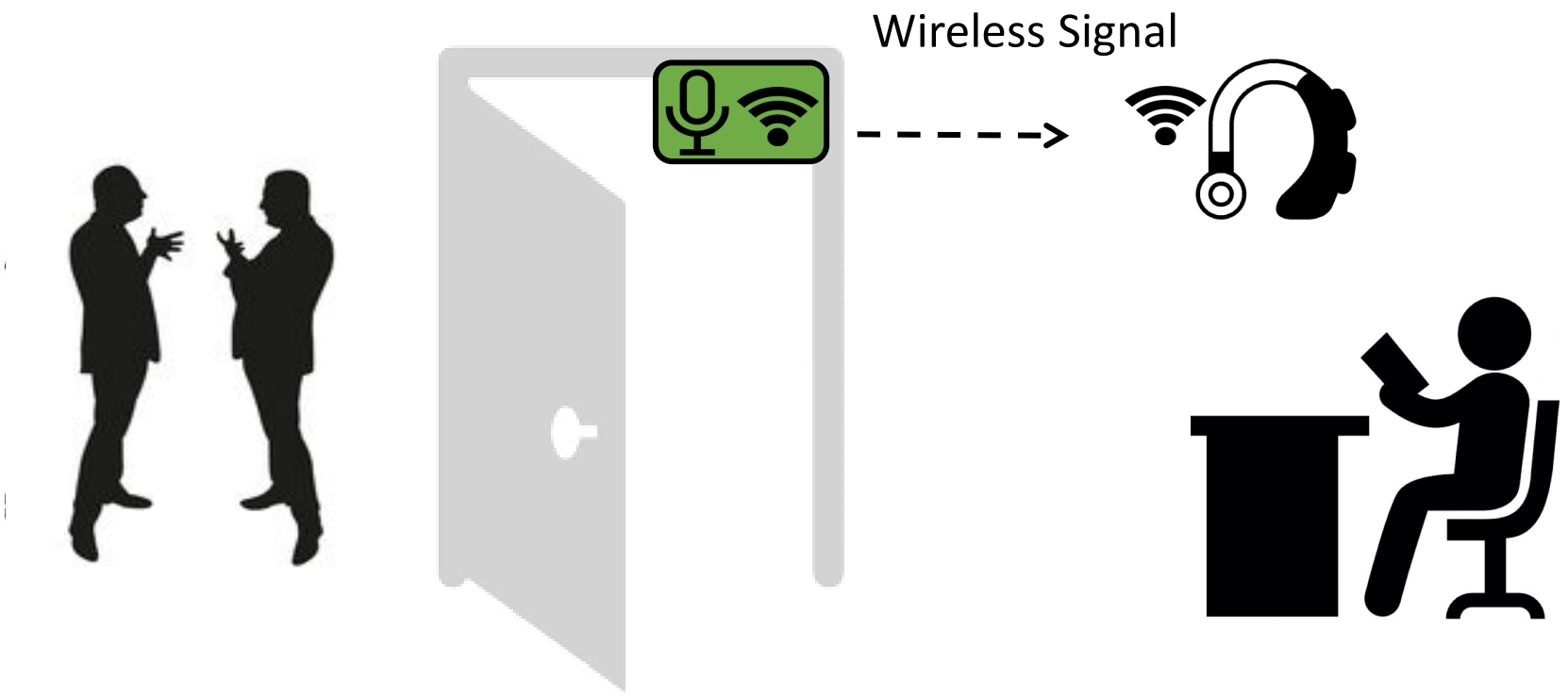
3. *MUTE* non-hollow design (performance)



+



Zooming Out

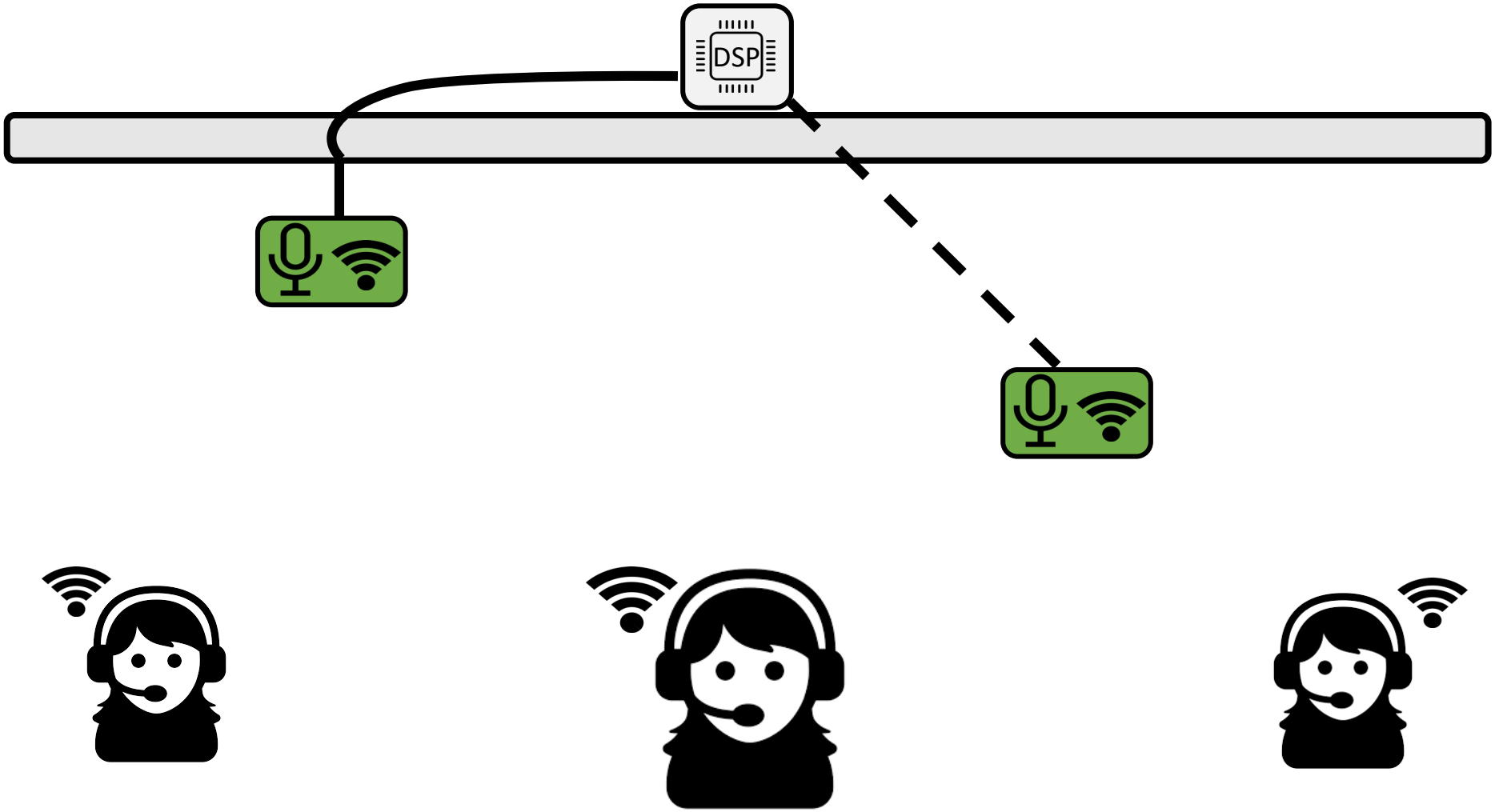

Wireless Signal

MUTE Tabletop Relay

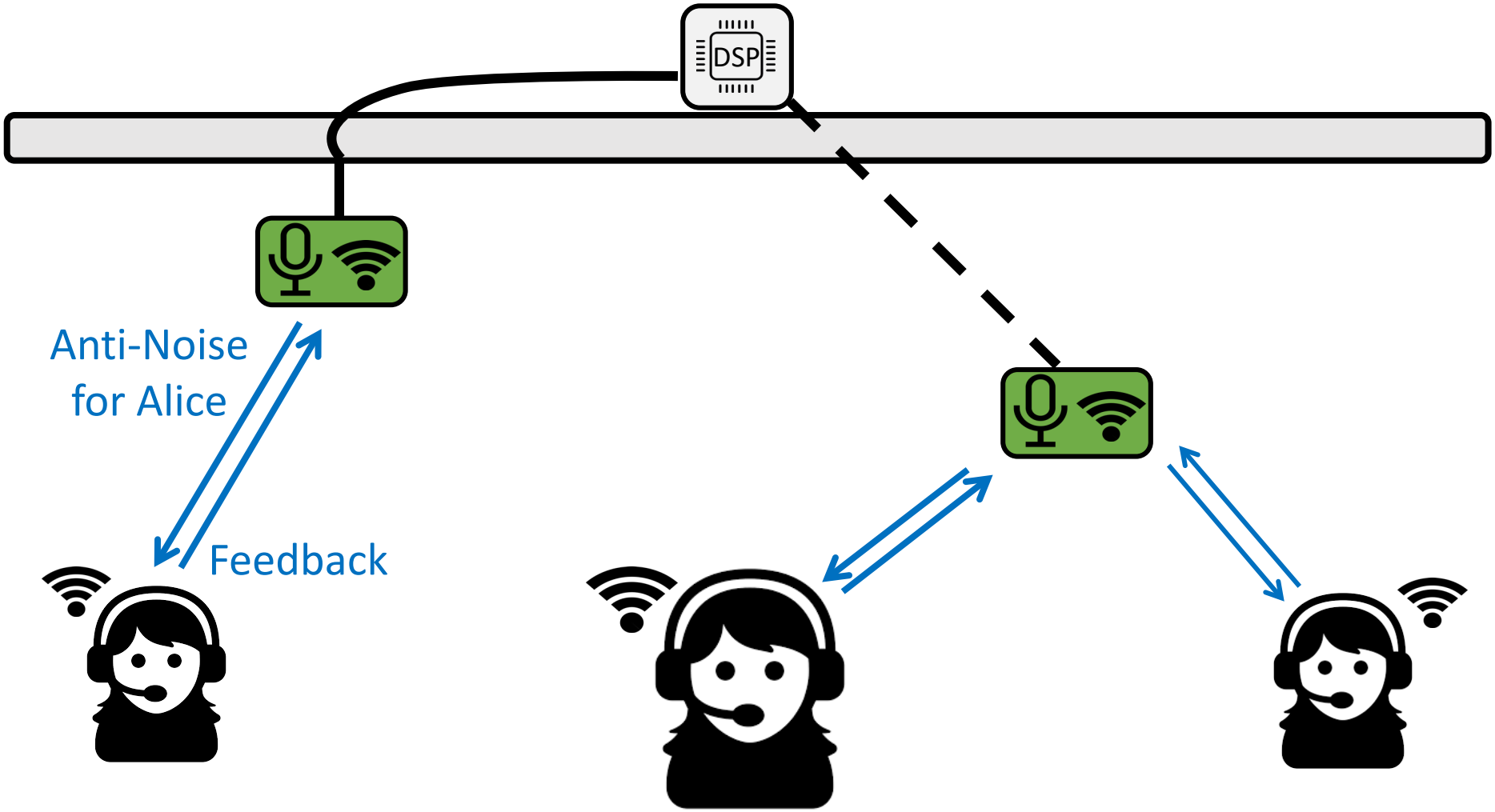
MUTE Tabletop Relay

MUTE Tabletop Relay

Noise Cancellation as an Edge Service

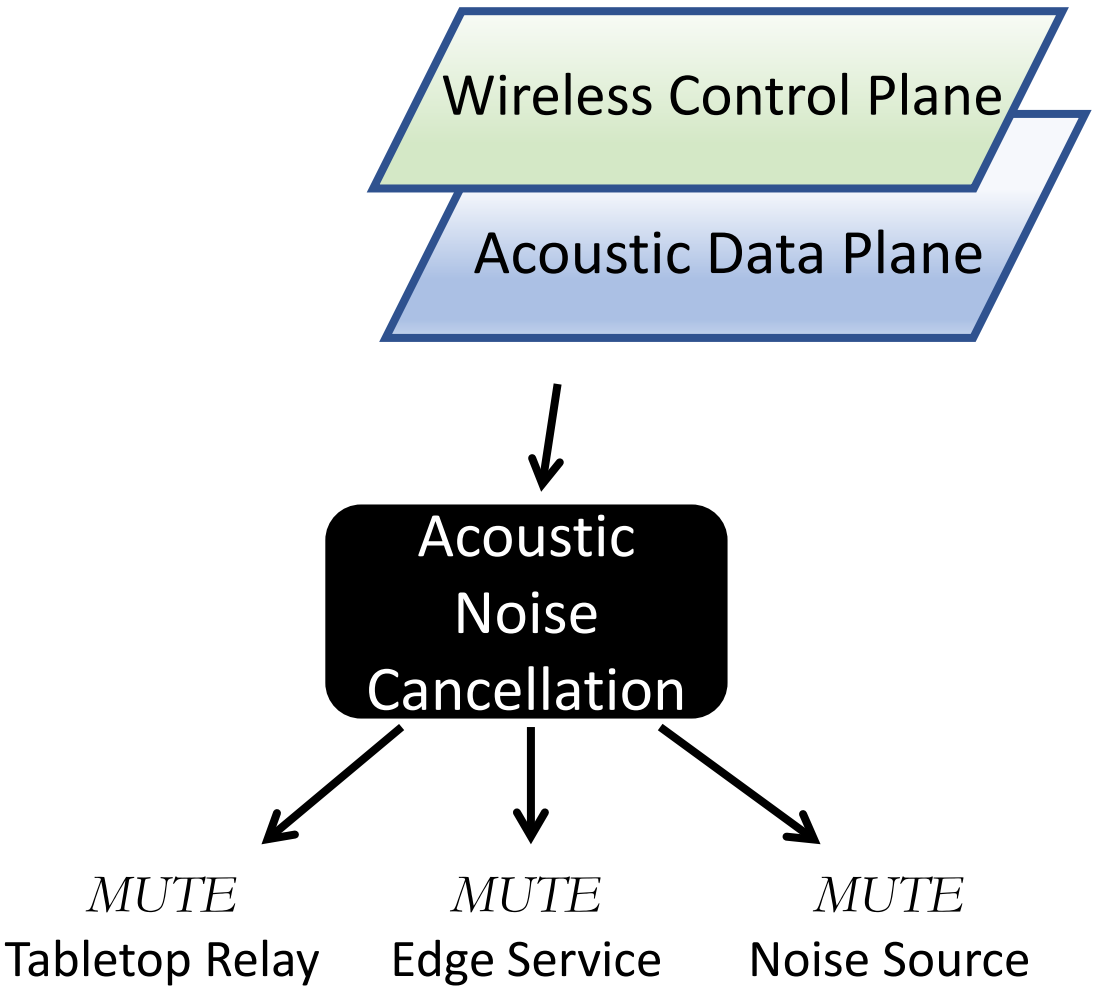


Noise Cancellation as an Edge Service

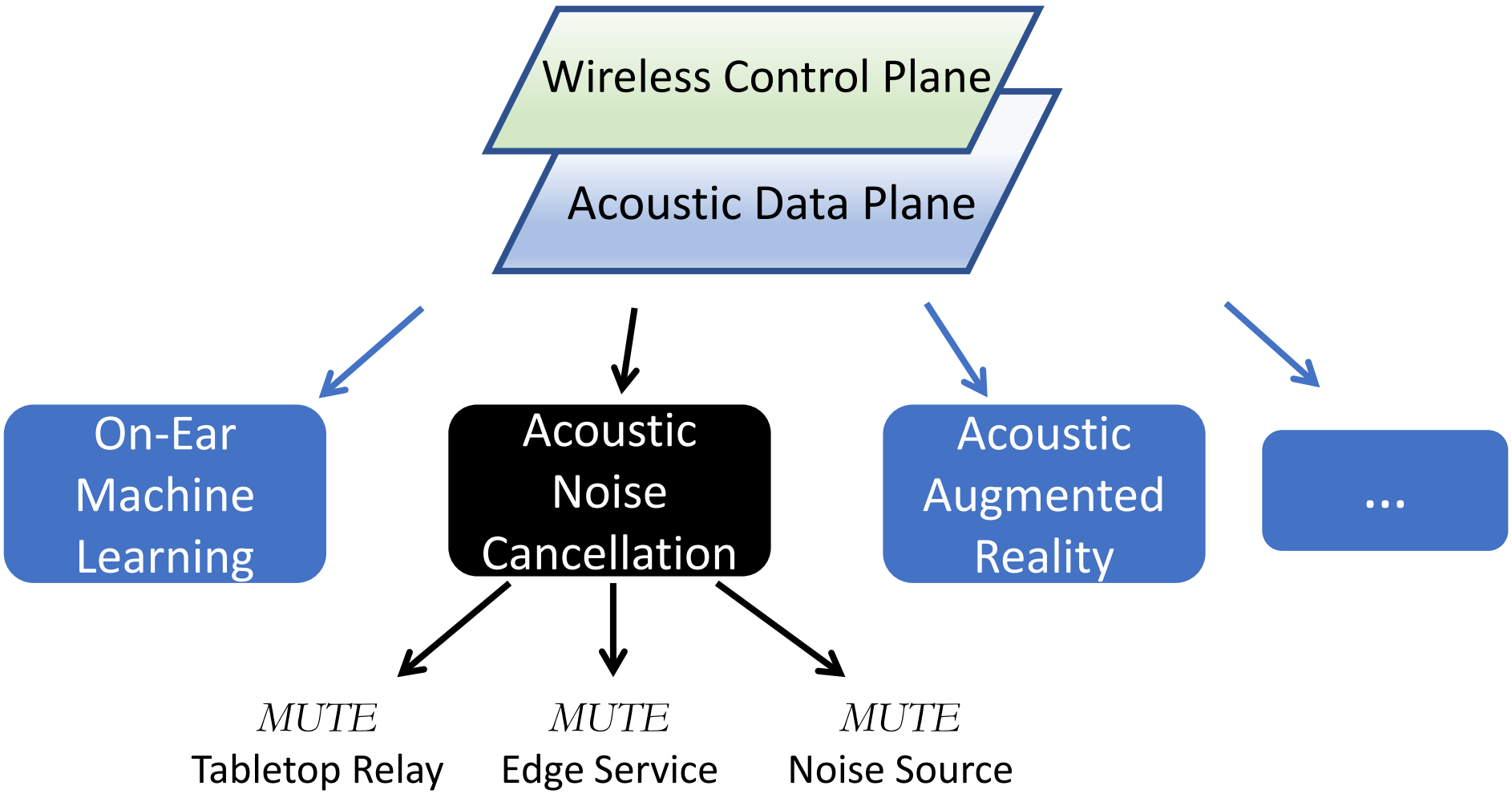


MUTE Enabled Noise Sources

Conclusion

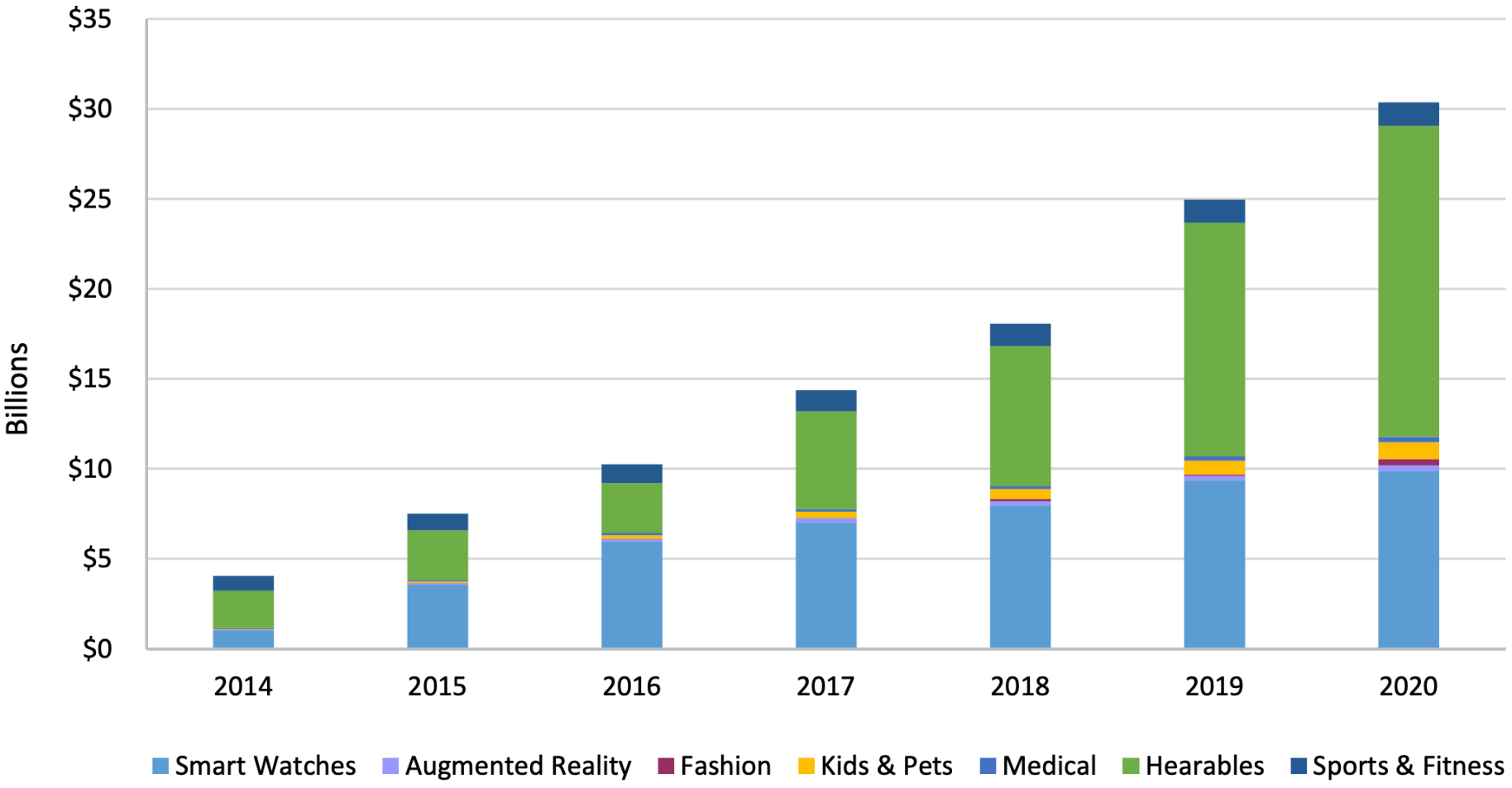

Wireless Control Plane
Acoustic Data Plane
Acoustic Noise Cancellation
MUTE
Tabletop Relay
MUTE
Edge Service
MUTE
Noise Source

Conclusion


Wireless Control Plane
Acoustic Data Plane
On-Ear Machine Learning
Acoustic Noise Cancellation
Acoustic Augmented Reality
...
MUTE Tabletop Relay
MUTE Edge Service
MUTE Noise Source

Closing Thoughts …

The Wearable Market Projections



WiFore: https://www.nickhunn.com/wp-content/uploads/downloads/2014/08/The-Market-for-Smart-Wearables.pdf
https://www.nickhunn.com/wp-content/uploads/downloads/2014/08/The-Market-for-Smart-Wearables.pdf

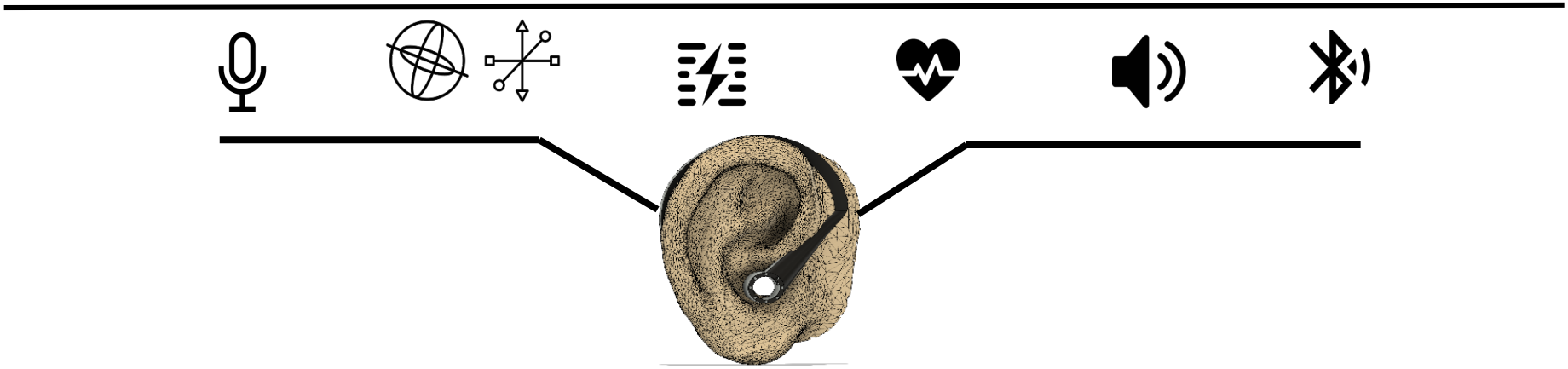
We are building a

software library

for Earable Computing

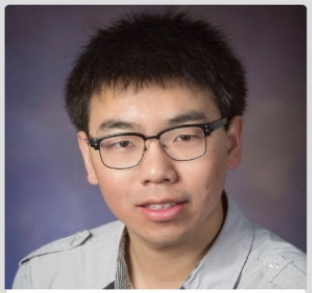
- Voice assistants
- Spatial acoustics
- Motion tracking
- Beamforming to Speech
- Noise Cancellation
- Face/Jaw Activity (eating, drinking)
- In-Ear Anatomy (health, security)

Sensor and Hardware Layer



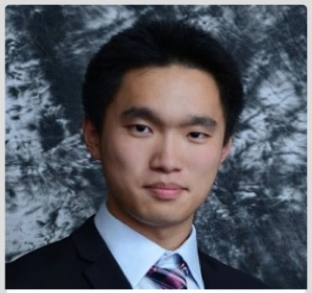
Haitham Hassanieh
Assistant Professor
Dept. of ECE & CS

Romit Roy Choudhury
Professor
Dept. of ECE & CS



Sheng Shen
PhD Student
Dept. of ECE

Suraj Jog
PhD Student
Dept. of ECE



Junfeng (Jayden) Guan
PhD Student
Dept. of ECE

Jiaming Wang
PhD Student
Dept. of ECE

Yu-Lin (Wally) Wei
PhD Student
Dept. of ECE

Zhijian Yang
PhD Student
Dept. of CS

Sohrab Madani
PhD Student
Dept. of ECE

Jay Prakash
Visiting Student

Thanks to my excellent students and collaborators …

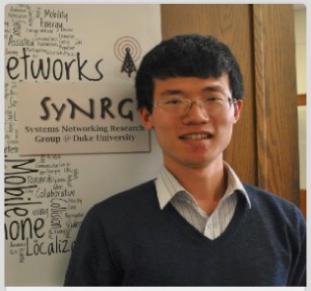
Thanks to all of you

Alumni in Academia

Ashutosh Dhekne
PhD 2019
Assistant Prof., CS
Georgia Tech

Nirupam Roy
PhD 2018
Assistant Prof., CS
U Maryland, College Park

Mahanth Gowda
PhD 2017
Assistant Prof., EECS
Penn. State Univ.



He Wang
PhD 2016
Assistant Prof., CS
Purdue Univ.

Jie Xiong
M.S. 2009, PhD at UCL
Assistant Prof., CS
UMass Amherst