

ROBOTIC ORIENTATION TOWARDS SPEAKER IN HUMAN-ROBOT INTERACTION

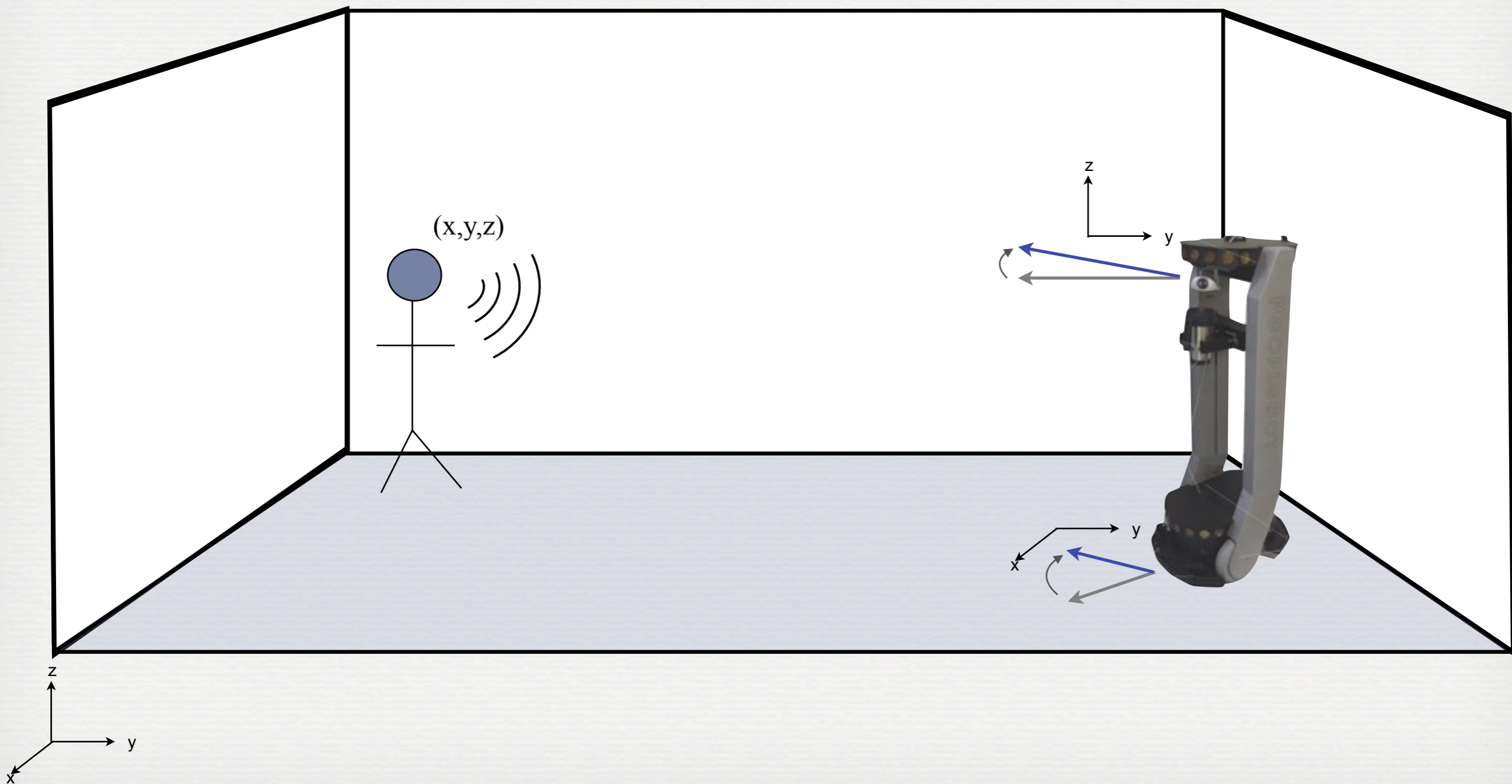
Caleb Rascón
Hector Avilés
Luis A. Pineda

IBERAMIA 2010, Session 8, Nov. 4

OUTLINE

1. Orientation in Human-Robot Interaction
2. Background and Current Issues
3. Proposed Approach and Results
4. Future Work and Conclusions

FACING THE SPEAKER IN HRI



WHY IS IT IMPORTANT?

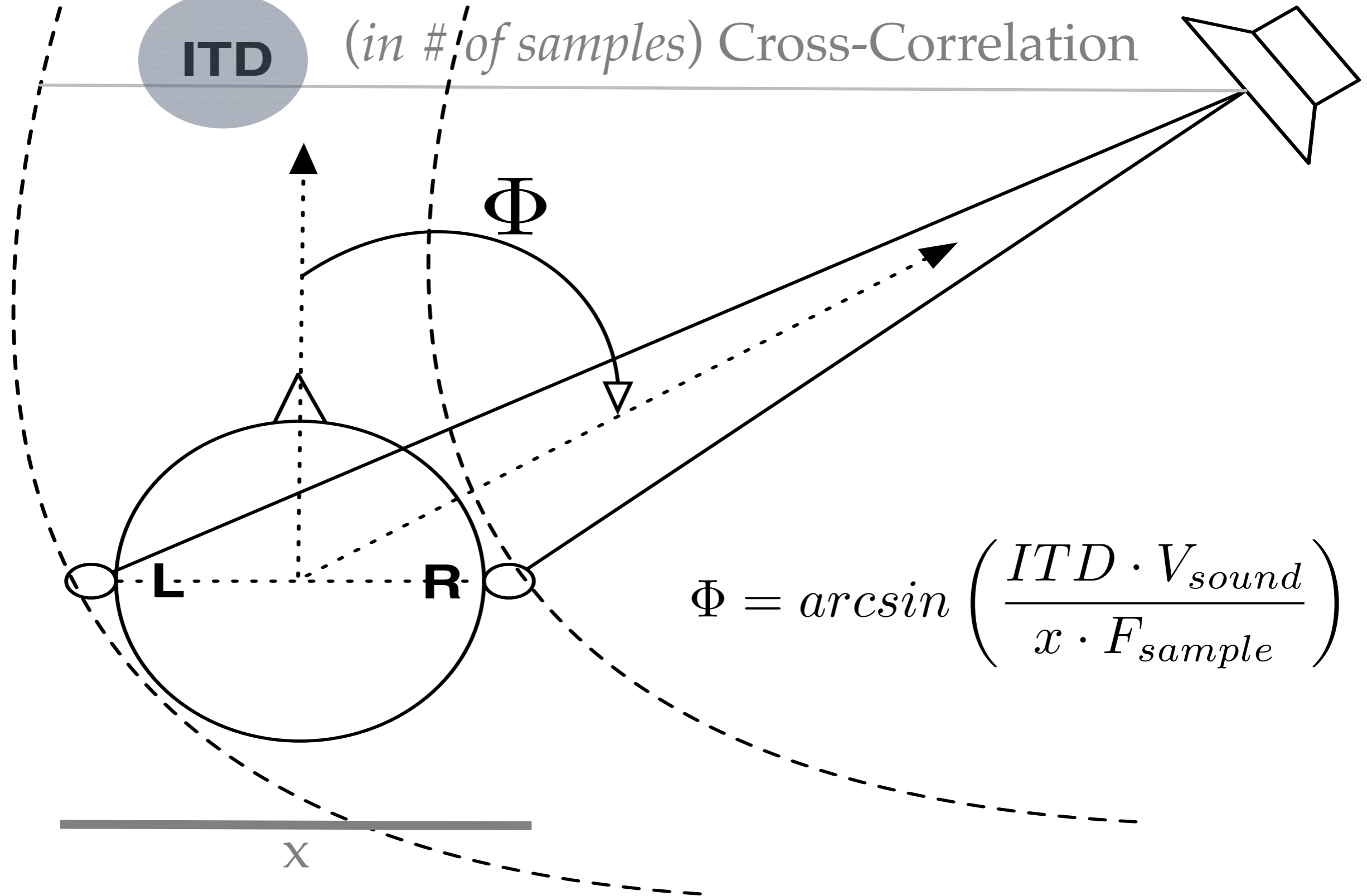
- From the user point of view:
- “Naturalness” of conversation enhanced

WHY IS IT IMPORTANT?

- From the robot point of view:
- Good first heuristic for user's position
 - "Robot, come here" problem
- Complements visual analysis in the case of out-of-view subjects
- Useful for direction-of-arrival filtering for speech recognition

HOW IS IT DONE?

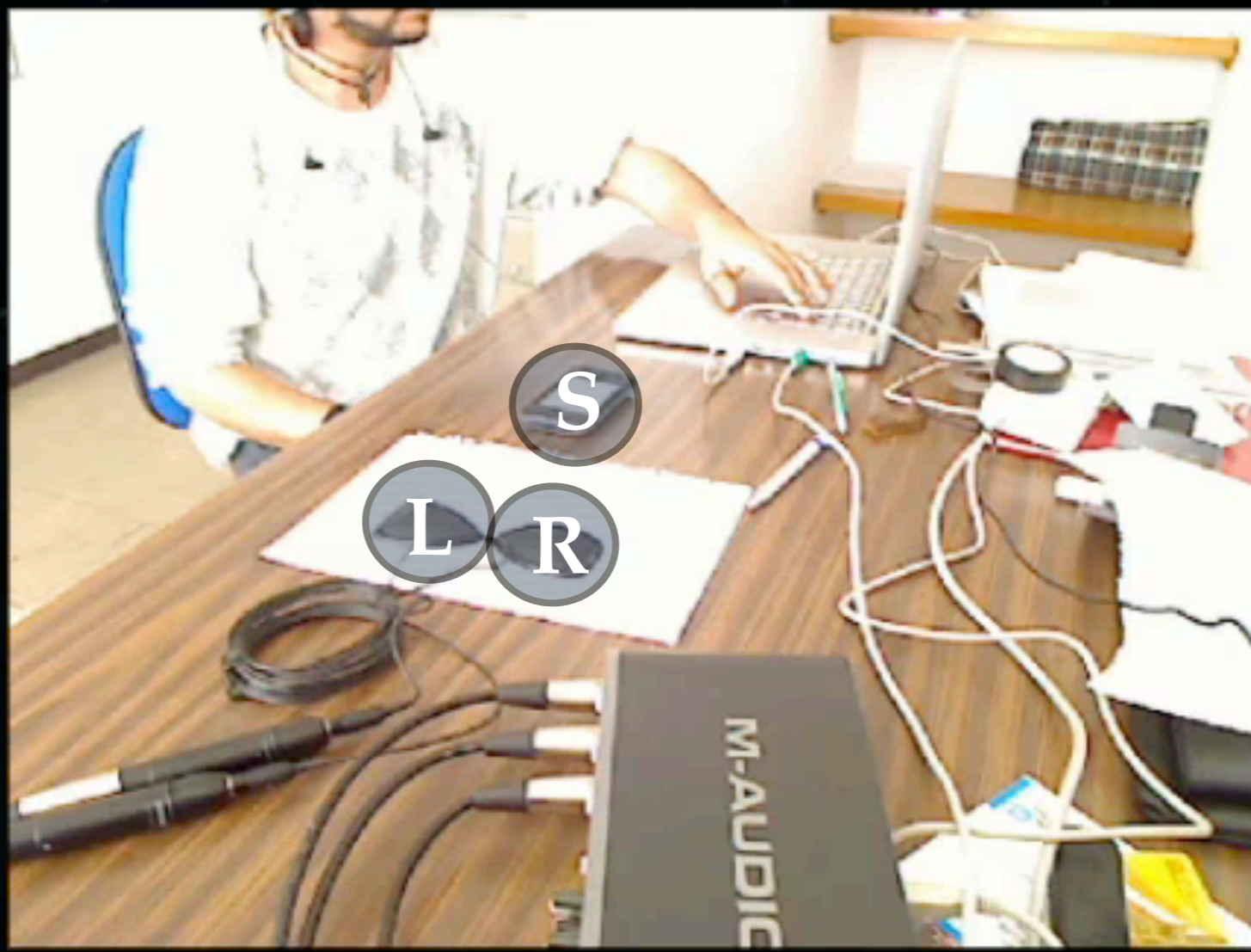
ITD (in # of samples) Cross-Correlation



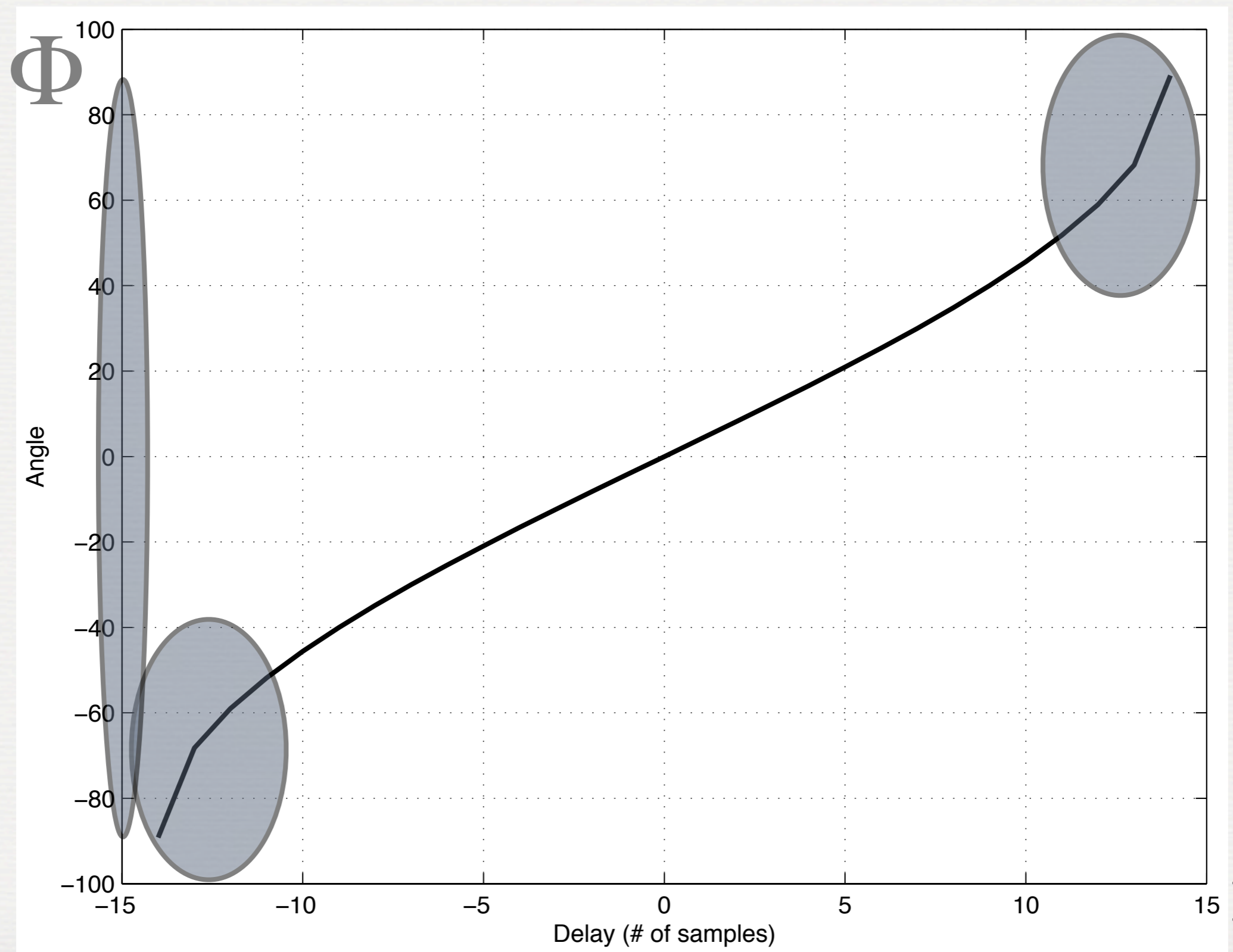


bash — ttys000 — #1

```
Nihil:STK balkce$ ./soundloc90 0.1089 4410 20 11
```



ISSUES

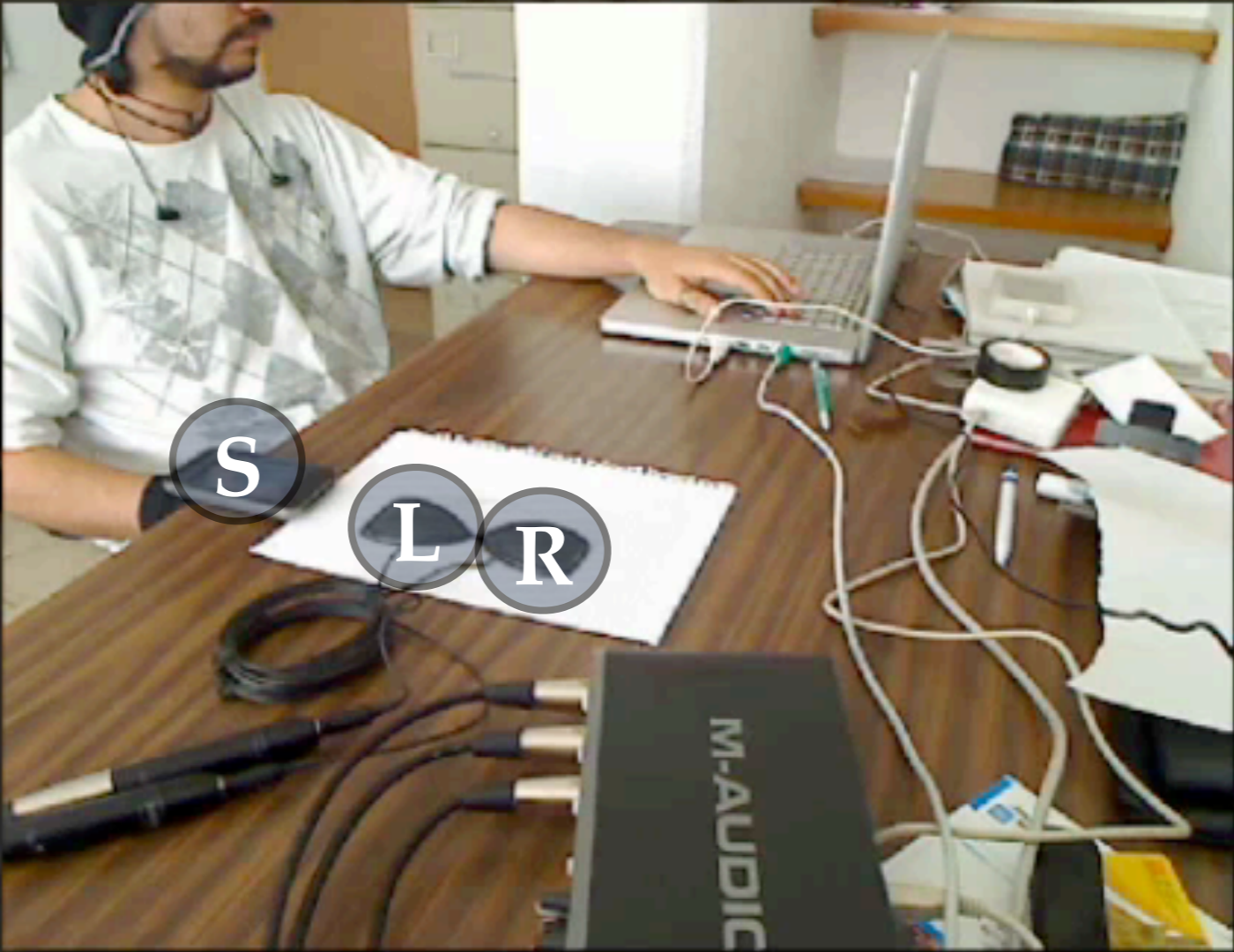


ISSUES

$F_s=44.1$ KHZ, $X=10$ CM

ITD (# samples)	Φ	ITD (# samples)	Φ
0	0°	8	34.8°
1	4.1°	9	40°
2	8.2°	10	45.6°
3	12.4°	11	51.8°
4	16.6°	12	59°
5	20.9°	13	68°
6	25.37°	14	89°
7	29.99°		

```
bash — ttys000 — #1
Nihil:STK balkce$ ./soundloc90 0.1089 4410 20 11
```



ISSUE SUMMARY

- **With a 2-mic array:**
- Only a $[-90^\circ \text{ -- } 90^\circ]$ range
- Decreased accuracy at sides ($\Phi \approx 90^\circ, -90^\circ$)
- ITD estimation frail against noise / reverb

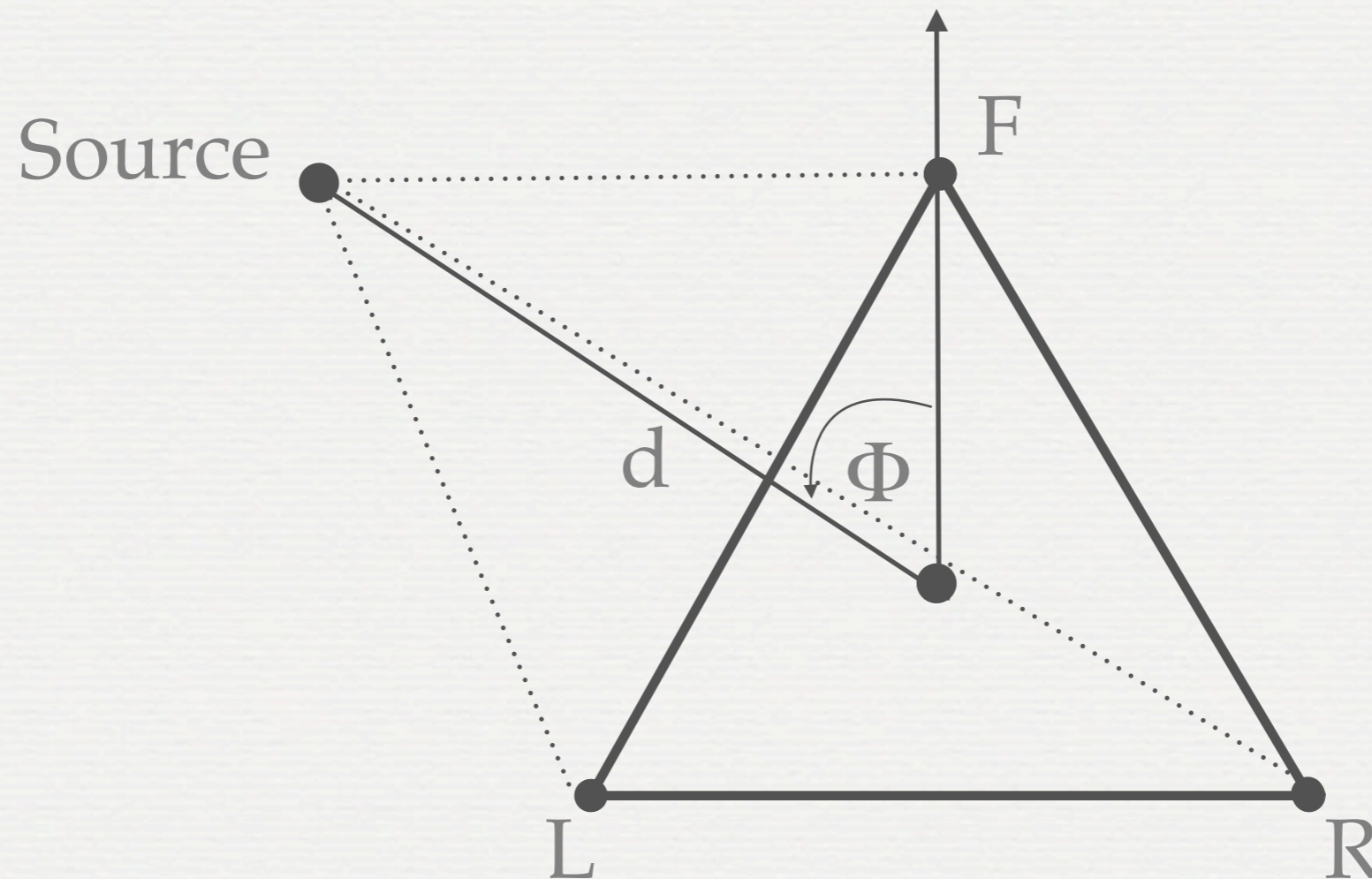
WHY ONLY 2 MICS?

- Low cost
- Lightweight (on top of a mobile robot)
- *Biologically aesthetic*

PROPOSED ALGORITHM

a.k.a our cheat

3-MICROPHONE SYSTEM

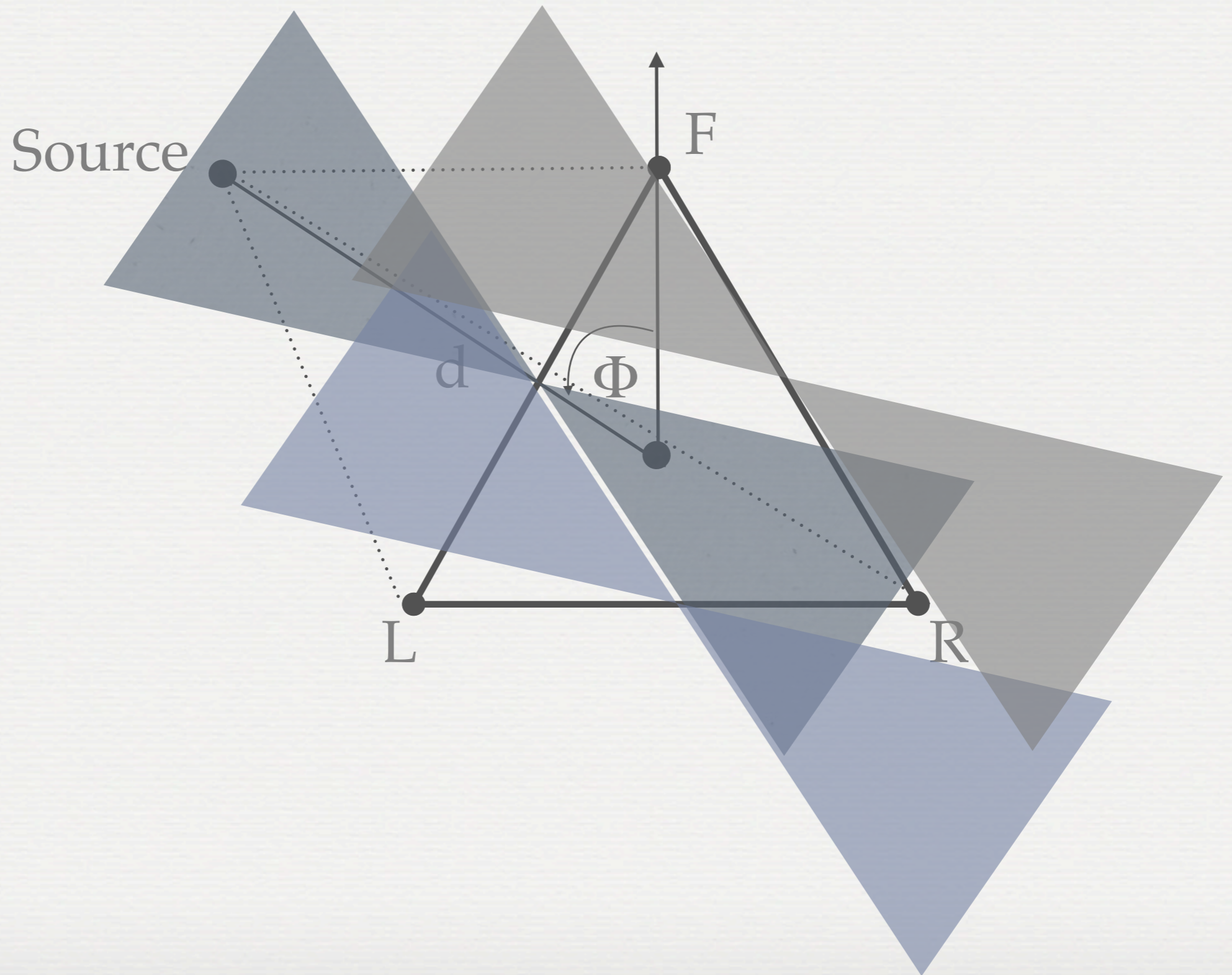


All mic. pairs are limited to estimate between $[-30^\circ -- 30^\circ]$.

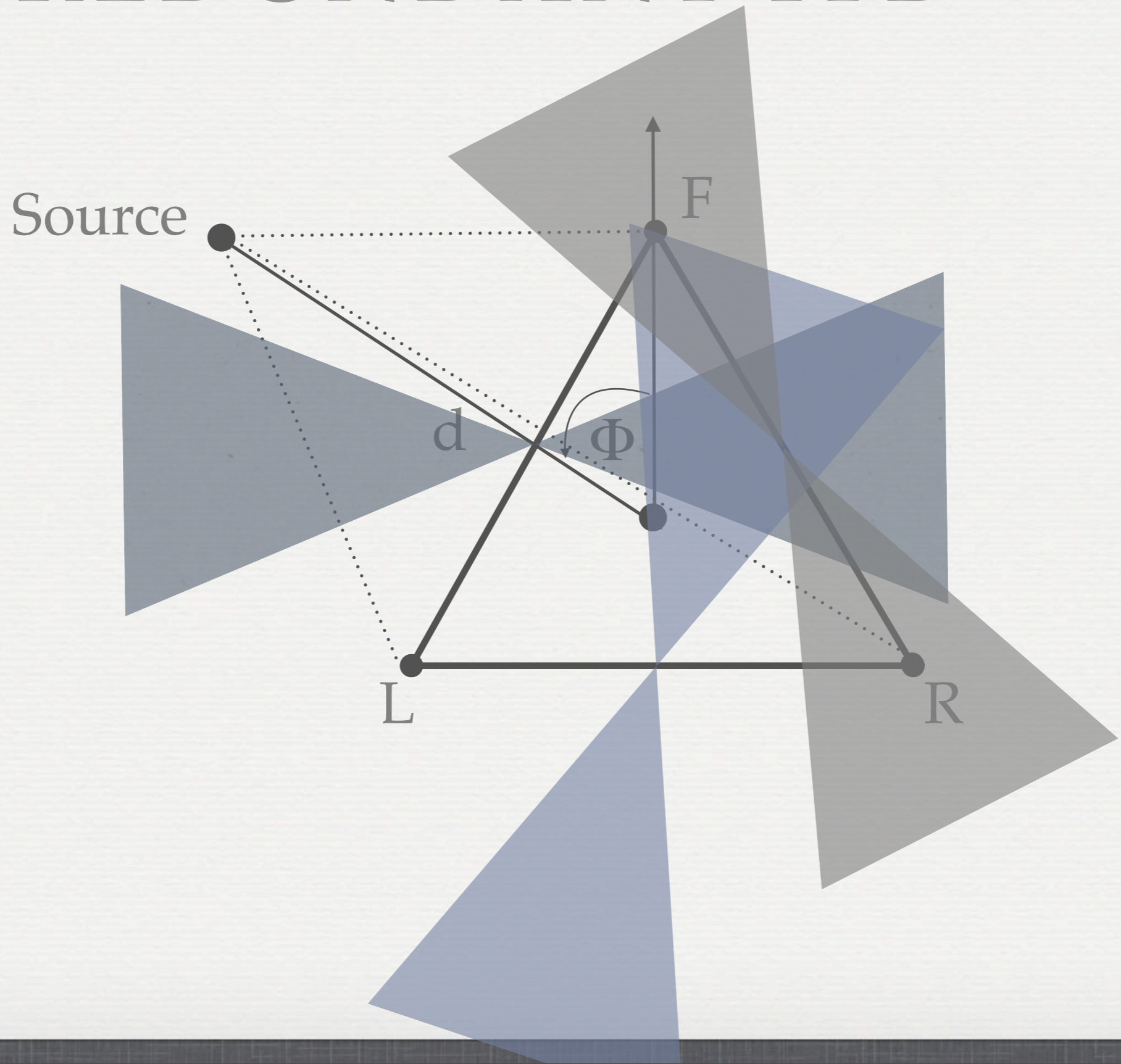
BENEFITS

- Complete angle range $[-180^\circ -- 180^\circ]$
- Semi-constant resolution throughout
- ITD estimation redundancy in every sample

REDUNDANT ITD



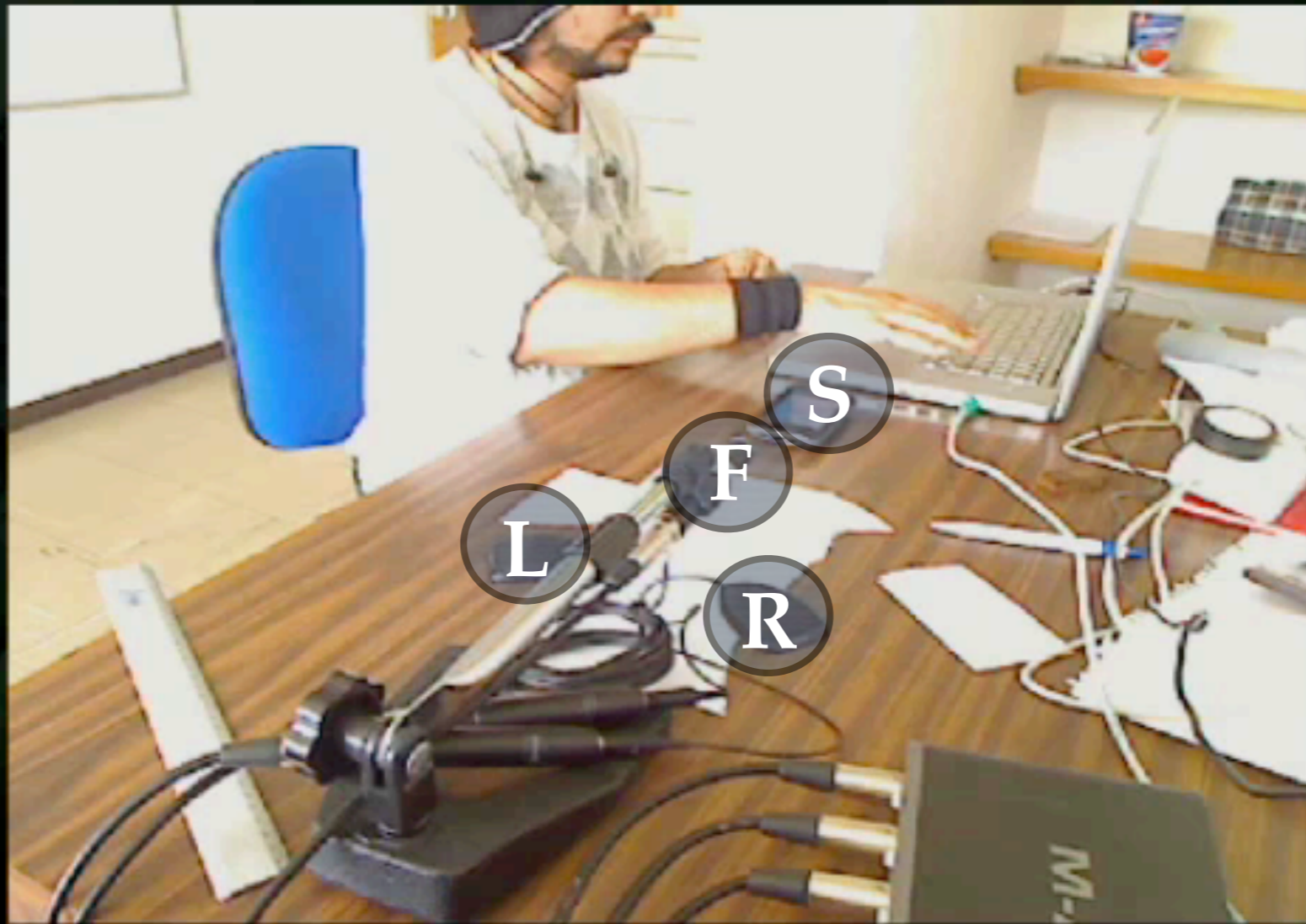
REDUNDANT ITD



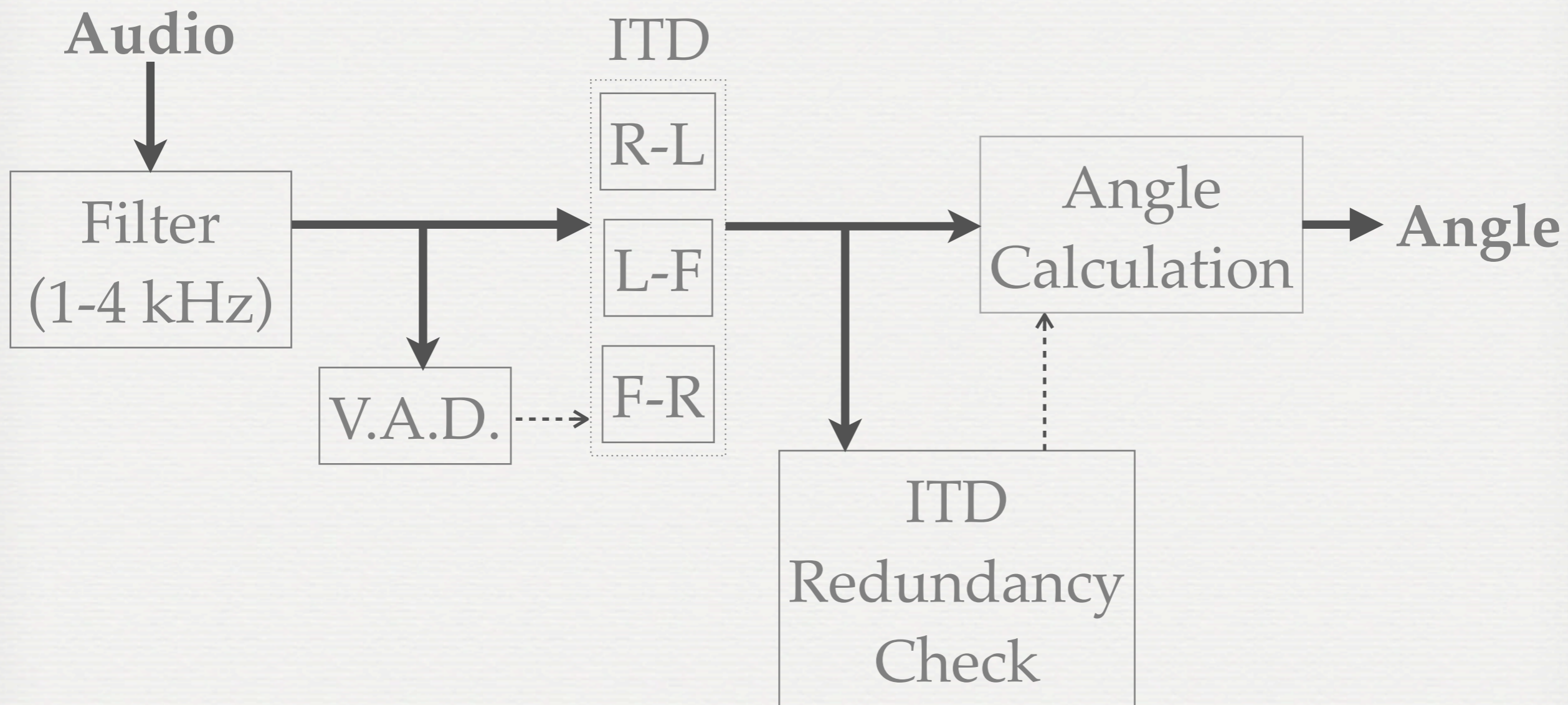


bash — ttys000 — 31

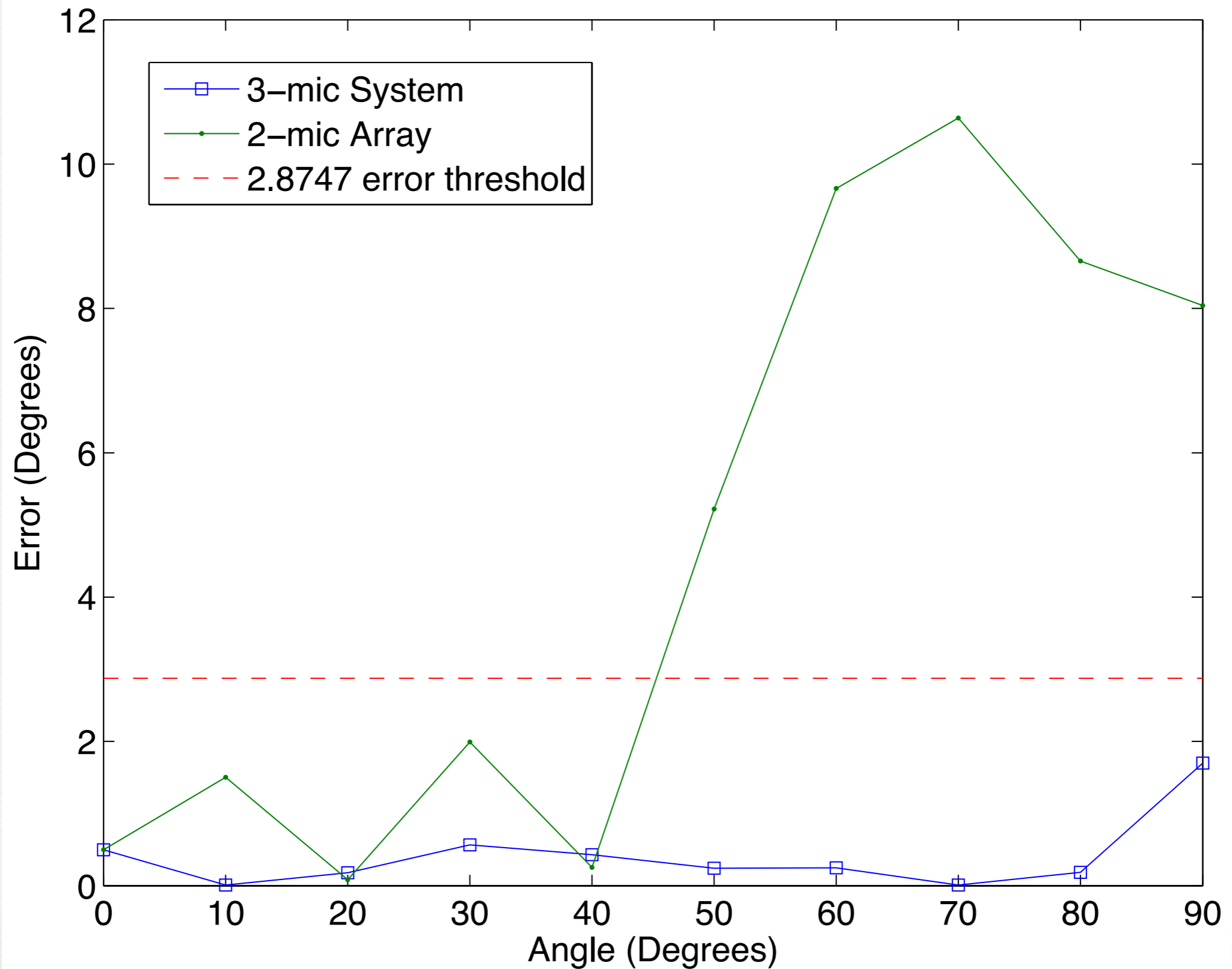
```
Nihil:STK balkce$ ./soundloc 0.179 4410 20 11
```



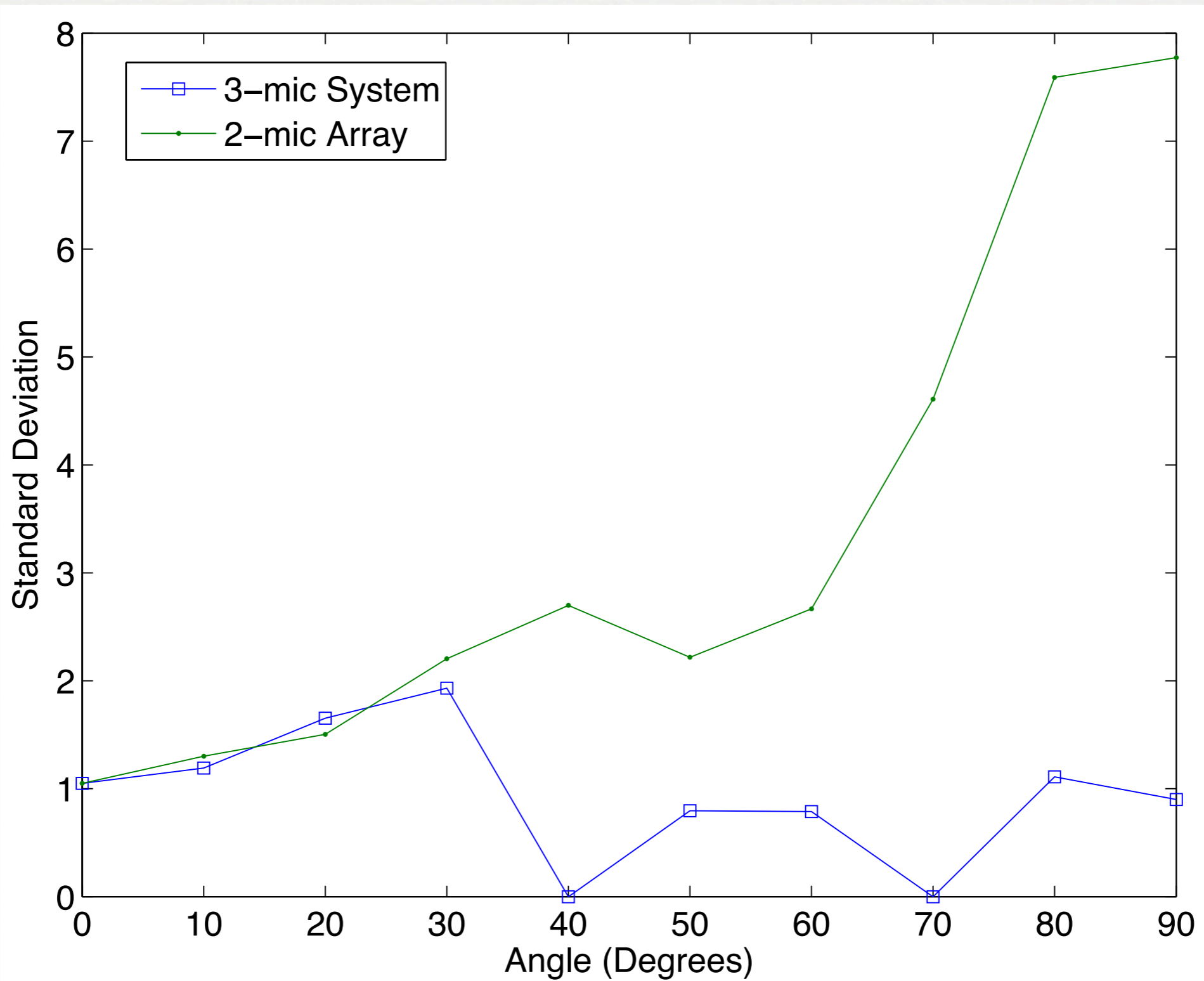
SYSTEM SUMMARY



ERROR COMPARISON

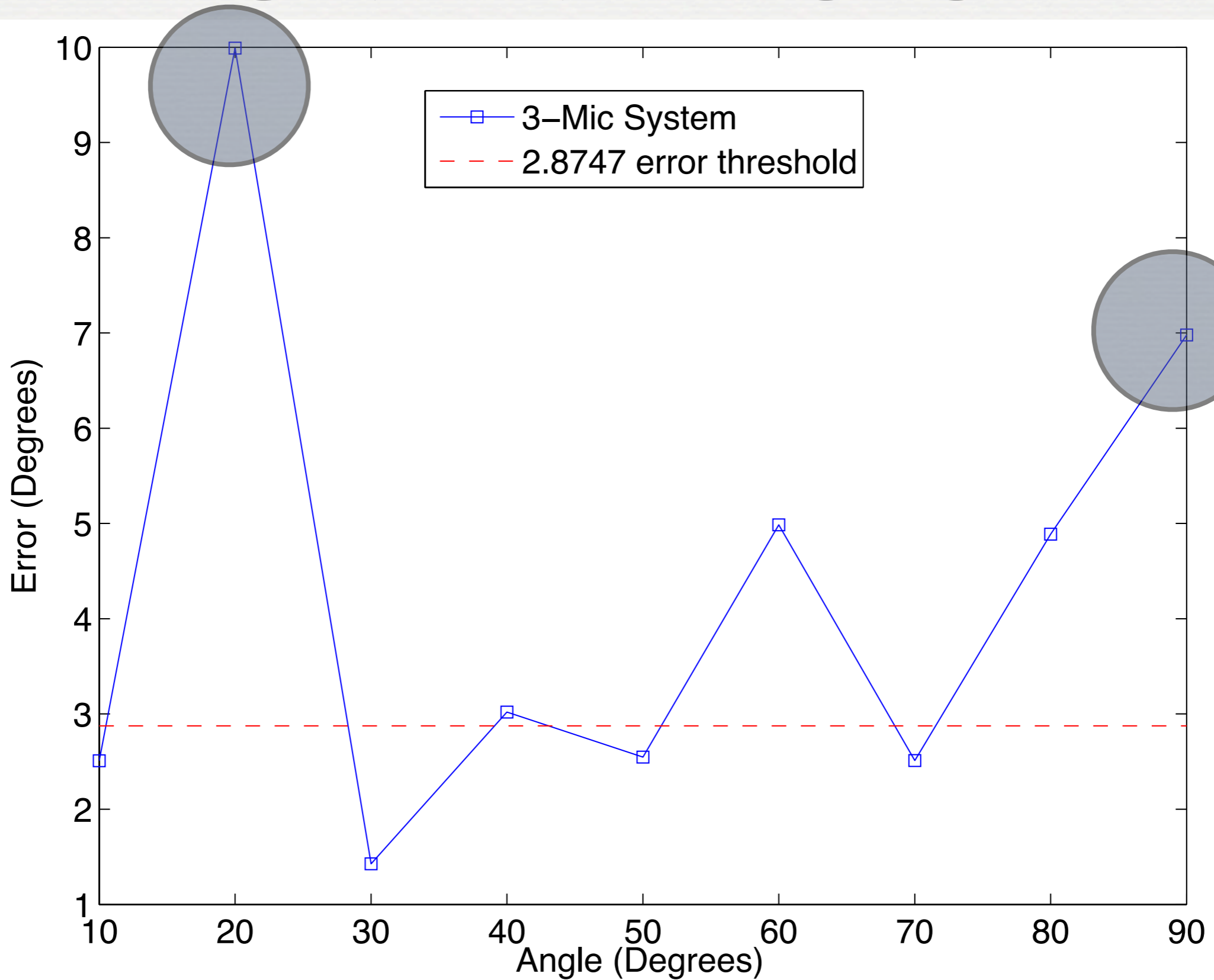


VARIABILITY COMPARISON





ONLINE TESTS



CONCLUSIONS

- Robot orientation important in HRI
- 2-mic arrays may be biological aesthetic, but have many issues
- 3-mic system outperformed 2-mic array while still being lightweight
- **Future work:** ASR is still sensitive towards reverb

THANK YOU

¿Questions?
Tips are welcome...