ABSTRACT

Title of Dissertation: TOWARD INTEGRATING INTELLIGENCE INTO EVERYTHING AROUND US

Nakul Garg
Doctor of Philosophy, 2025

Dissertation Directed by: Professor Nirupam Roy
Department of Computer Science

The vision of ambient intelligence promises a world where computational capabilities seamlessly integrate into everyday objects and environments, creating systems that sense, learn, and adapt to human needs while remaining invisible to users. Despite significant advances in miniaturization and low-power computing, true ambient intelligence has remained elusive, hindered by a fundamental challenge: current intelligent systems require substantial energy, complex hardware, and frequent maintenance, making widespread deployment impractical. We introduce a paradigm shift in how we create intelligent systems by fundamentally reimagining sensing and computing architectures from first principles for extreme resource constraints.

This thesis centers on encoding intelligence directly into the physical domain through novel hardware-software co-design, where passive structures perform initial signal transformations without consuming power. Through novel architectures across acoustic, radio frequency, and optical domains, we demonstrate systems that achieve spatial perception, global positioning, and environmental monitoring with orders of magnitude less power than conventional approaches.

These innovations enable intelligence in previously impossible contexts: insect-scale robots that navigate complex environments, sticker-sized tags that provide GPS-like tracking for years on a single battery, and wireless sensors that monitor food quality throughout global supply chains. By bridging the gap between what intelligent systems can do and what resource-constrained platforms can support, this work establishes a foundation for truly pervasive intelligence that operates sustainably at large scale.

TOWARD INTEGRATING INTELLIGENCE INTO EVERYTHING AROUND US

by

Nakul Garg

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2025

Advisory Committee:
Professor Nirupam Roy, Chair/Advisor
Professor Ramani Duraiswami
Professor Lin Zhong
Professor Alan Liu
Professor Sennur Ulukus, Dean's Representative

© Copyright by
Nakul Garg
2025

To my family - Kavita, Sunil, and Rishabh

Acknowledgments

I would like to express my sincere gratitude to my advisor, Professor Nirupam Roy, for his exceptional guidance and mentorship throughout my doctoral studies. His support has been instrumental in shaping my research capabilities and analytical thinking. I have thoroughly enjoyed our brainstorming sessions where several creative ideas emerged and I learned how to transform research challenges into opportunities. He taught me invaluable skills in organizing complex thoughts, approaching problems with critical rigor, and communicating my ideas. Without his dedication and belief in my potential, the work presented in this dissertation would not have been possible.

I would like to express my deep appreciation to my committee members for their guidance and support. Professor Ramani Duraiswami has been incredibly supportive of my work since day one and provided invaluable feedback till the end in my academic job search. I am very grateful to have Professor Lin Zhong on my committee; his leadership and mentorship have been exemplary. He has been a role model for me and has shown unwavering support for our research directions, providing validation for the structure-assisted spatial sensing work we pursued. I extend my heartfelt thanks to Professor Alan Liu, from whom I learned how to build and scale networked systems both while taking his class on cloud networking and computing and during our discussions about the future of Edge IoT. I truly value his thoughtful feedback on this dissertation and his help in crafting my thesis story. I am fortunate to have Professor Sennur Ulukus on my

committee. As an expert in wireless, her keen insights have been particularly valuable as I explore scalable wireless systems in this work.

I am incredibly grateful to the mentors and supporters who helped me through the academic job market. Thank you to Nirupam Roy, Karthik Sundaresan, Ranveer Chandra, Lin Zhong, and Siavash Alamouti. Karthik mentored me during my internship at NEC and beyond. He is a brilliant researcher and collaborator who believed in my abilities from the start. I want to thank him for collaborating with me on the scalable UWB project and for inspiring me to aim higher. I met Ranveer during my internship at Microsoft Research. He is the best mentor anyone could ask for, and his leadership and ability to solve impactful problems has inspired me to the core. I am fortunate for his feedback and mentorship during my job search and lucky to have him as a mentor. Siavash is a rockstar researcher and has been a strong advocate for my work. Whenever I got a chance to meet him, he always gave great advice on my research directions and was especially supportive of my low power wireless work. I am lucky to know him as a friend, a mentor, and a 10x entrepreneur. I would also like to thank Akshay Gadre, Ish Jain, Nivedita Arora, Akarsh Prabhakar, Justin Chan, Vikram Iyer, Suman Banerjee, Yasaman Ghasempour, Venkat Arun, Dinesh Bharadia, Swarun Kumar, Chahatdeep Singh, Nitin Sanket, and Tara Boroushaki for their time and invaluable advice throughout my job search journey.

I am grateful to my labmates and Spacewalkers - Yang Bai, Aritrik Ghosh, Irtaza Shahid and Harsh Takawale (The iCoSMoS Lab). Thank you Yang for working with me on the spatial acoustics projects. I learned a lot about perseverance and discipline from your approach to research. Special thanks to Aritrik for collaborating on the low-power cellular work. I learned a lot about theoretical research and cherish the many deep physics conversations we had. I would like to thank Irtaza for working with me on the UWB project and spending nights to create testbeds and

calibrating our sensors. I also cherish all our fundamental discussions and recoveries from our breakdowns thinking about signal processing problems. Thanks to Harsh for working with me on the food sensing project. I have always enjoyed brainstorming with you. Thanks to the entire iCoSMoS lab for always providing me feedback on the papers, talks, and supporting me through both ups and downs in my PhD journey.

I must express my heartfelt gratitude to my amazing friends who have been my support system throughout this journey. To Satyarth, who taught me life philosophy and the positive vibe - I'm sorry I couldn't attend your wedding. To Deepak, Ravin, Smriti, Varenya, Shweta, and Abhishek - I found a family away from family especially during COVID with our Spring TP group. To Aditi for helping me with medicines in all conference travels and your perfect song suggestions. Special thanks to Meenal for her unwavering support. You are my best friend, and I am lucky to have you. I am deeply grateful to have Nitin bhaiya as a friend and a mentor - I have learned so much from our conversations, especially your advice and guidance in my early years of PhD when I came and felt lost. Chahat bhaiya, thank you for all your advice throughout my PhD journey, from my first project to my most recent one, and especially for all the job talk and interview tips. Thanks also to Anoorag, Priyal, Stella, Pooja, and Mrunal. To my dearest friends whose support and belief in me has been a driving force - Tanmay, Bhawana, Prerna, Ishani, Devina, and Sidharth, I'm super lucky to have all of you. Thank you Ayush for being a huge inspiration to pursue academia and for being my longest friend.

Most importantly, I want to thank my family. My parents (Kavita and Sunil), my brother (Rishabh), my grandparents (Ilaychi and Ramchandra), and the entire RCG family. I cannot thank my parents enough who have sacrificed so much and faced hardships just to make my life better and prioritize my education. My brother is my best friend who unknowingly taught me so much, to

dream big, never lose hope, and to smile through challenges. My grandfather is the mathematical prodigy I know and grew up with him teaching me mathematics and discipline. My grandmother, amma, taught me to be a good human being above everything else. I am lucky to have grown up learning from her. My uncle (Anil) has taught me to be selfless, to provide value and support, and to remain optimistic no matter what. Thanks to my brothers Lakshay and Gaurav for being my biggest pillars for life. My sister (Mahima) for being a role model and helping me pursue science and teaching me the fundamentals. You are the biggest inspiration for me to pursue academia. Special thanks to my superstar sisters - Himanshi, Mugdhal; my aunts - Anita, Mamta and Babita; my brother Akshay; and everyone for their love.

Table of Contents

Dedication ii

Acknowledgements iii

Table of Contents vii

List of Tables xi

List of Figures xii

List of Abbreviations xix

Chapter 1: Introduction 1
- 1.1 Challenges in Scaling Ambient Intelligence . . . 3
- 1.2 Designing Ultra-Low-Power Intelligent Systems . . . 5
- 1.3 Systems Developed . . . 8
  - 1.3.1 Structure-Assisted Spatial Intelligence . . . 8
  - 1.3.2 Scalable nextG Systems . . . 10
  - 1.3.3 Systems for Sustainability . . . 12
  - 1.3.4 Other Contributions . . . 13
- 1.4 Organization . . . 14

Chapter 2: Structure-Assisted Spatial Audio Sensing 15
- 2.1 Introduction . . . 15
- 2.2 Core Intuitions and Primers . . . 20
  - 2.2.1 Metamaterials for Passive Filtering . . . 22
- 2.3 System Design . . . 24
  - 2.3.1 Processing for DoA Estimation . . . 25
  - 2.3.2 Eliminating Source Signal Dependency . . . 25
  - 2.3.3 Eliminating Environmental Dependency . . . 28
  - 2.3.4 Synthetic Training for Deep Learning . . . 29
  - 2.3.5 Optimizing 3D Stencil Design . . . 31
- 2.4 Prototype Development . . . 36
  - 2.4.1 3D-Printing Stencil Caps . . . 36
  - 2.4.2 Calibration and Data Collection . . . 36
- 2.5 Evaluation . . . 37
  - 2.5.1 Evaluation Setup and Results Summary . . . 37

2.5.2 Impacts of external conditions . . . 39
2.5.3 Performance in different environments . . . 41
2.5.4 Impact of different sound sources . . . 42
2.5.5 Performance in known environment . . . 42
2.5.6 Localization Performance . . . 43
2.5.7 Comparison with traditional methods . . . 44
2.5.8 Comparison between learning models . . . 44
2.5.9 Energy consumption . . . 45
2.6 Discussion . . . 48
2.7 Related Work . . . 49
2.8 Conclusion . . . 51

Chapter 3: Microstructure-Assisted Vision for Ubiquitous Tiny Robots 52
3.1 Introduction . . . 52
3.2 Core Intuitions and Primers . . . 57
3.2.1 Coded Signal Projection with Structures . . . 58
3.2.2 Single-Receiver Depth Mapping . . . 60
3.3 System Design . . . 61
3.3.1 Low-Power Scene Reconstruction . . . 61
3.3.2 Directional Code Projection . . . 63
3.3.3 Optimal microstructure design . . . 68
3.3.4 Motion stacking . . . 69
3.4 Evaluation . . . 70
3.4.1 Metrics . . . 71
3.4.2 Overall Performance . . . 71
3.4.3 Impact of the Environment . . . 72
3.4.4 Impact of system parameters . . . 74
3.4.5 Impact of scene parameters . . . 76
3.4.6 Computation techniques . . . 78
3.4.7 Power consumption . . . 80
3.5 Related Work . . . 83
3.6 Discussion . . . 85
3.7 Conclusion . . . 86

Chapter 4: Ultra-Low-Power Self-Localization Using a Single Antenna 87
4.1 Introduction . . . 87
4.2 Core Intuition and Feasibility . . . 92
4.3 System Design . . . 94
4.3.1 Ultra-low Power Receiver . . . 94
4.3.2 AoA Estimation . . . 95
4.3.3 Localization with Independent Beacons . . . 97
4.3.4 Designing Programmable Directional Gain . . . 102
4.3.5 Pin-diodes as RF Switches . . . 103
4.3.6 Directional Code to AoA Mapping . . . 105
4.4 Evaluation . . . 107

4.4.1 AoA Performance . . . 109
4.4.2 Localization . . . 109
4.4.3 Impact on RF Communication . . . 110
4.4.4 Power Consumption . . . 111
4.4.5 Robustness . . . 112
4.5 Limitation and Future Work . . . 118
4.6 Related Work . . . 119
4.7 Conclusion . . . 122

Chapter 5: Scalable Asset Tracking with NextG Cellular Signals 123
5.1 Introduction . . . 123
5.2 Cellular Networks Primer . . . 128
5.3 System Design . . . 131
5.4 *LiTEfoot* Prototype Implementation . . . 144
5.5 Evaluation . . . 147
5.6 Related Work . . . 154
5.7 Discussion . . . 155
5.8 Conclusion . . . 158

Chapter 6: Large Network UWB Localization: Algorithms and Implementation 159
6.1 Introduction . . . 159
6.2 System Design . . . 163
6.2.1 Joint Range-Angle Localization . . . 164
6.2.2 Scaling to Large networks . . . 167
6.2.3 Reference Frame Transformation . . . 174
6.3 Opportunistic Anchor Integration . . . 177
6.4 Implementation . . . 179
6.5 Evaluation . . . 179
6.5.1 Localization Evaluation . . . 180
6.5.2 Latency Evaluation . . . 185
6.5.3 Micro-Benchmarks . . . 186
6.6 Discussion . . . 191
6.7 Related Work . . . 193
6.8 Conclusion . . . 194

Chapter 7: Low-Cost and Dynamic Food Quality Sensing at the Pallet-Level 195
7.1 Introduction . . . 195
7.2 Understanding Food Quality Metrics . . . 197
7.3 Physics and Core Intuition . . . 198
7.4 Design and Implementation . . . 201
7.5 Performance Evaluation . . . 203
7.6 Conclusion . . . 206

Chapter 8: Conclusion and Future Directions 207
8.1 Summary of Contributions . . . 207

8.2 Broader Impact 209
8.3 Future Directions 210
8.4 Closing Thoughts 211

Bibliography 213

List of Tables

2.1 Comparison of prototype cost, size, median error, and energy consumption of Owlet with a microphone array. 39

3.1 Breakdown of energy consumption for the hardware and software submodules. 81
3.2 Comparison of different computation optimizations showing the total energy consumed per scene reconstruction. Prototype on Raspberry Pi 4. 83

4.1 Breakdown of energy consumption. 112
4.2 Summary of related work. 119

5.1 Key characteristics of evaluation routes. 147
5.2 Comparison between latency, accuracy, and power consumption. 149
5.3 Overall power and energy per inference. 150
5.4 RF frontend power consumption. 150
5.5 Baseband power and time per inference. 151

List of Figures

1.1 Examples of intelligent, low-power, and scalable sensors developed in this thesis. 2
1.2 The intelligence-energy tradeoff: Conventional intelligent systems (top left) operate at high energy levels and offer good capabilities. Current resource-constrained systems (bottom right) provide limited intelligence. This thesis explores approaches to bridge this gap by creating systems that offer $100\times$ more intelligence at similar energy levels. 4
1.3 Our approach to sustainable ambient intelligence combines nature-inspired architectures, physics-informed AI, and scalable networks. 6
1.4 Structure-assisted sensing systems: (A) *SPiDR*'s depth imaging using 3D-printed acoustic metamaterial, (B) *SPiDR*'s acoustic stencil and its internal spatial filtering channels, (C) *Owlet*'s acoustic stencil and its internal direction-finding channels, (D) *Owlet*'s direction finding with passive acoustic structures. 8
1.5 *LiTEfoot* leverages cellular signals for nationwide tracking using an ultra-low-power miniaturized receiver. 10
1.6 *Locate3D* enables infrastructure-free tracking at scale through optimized peer-to-peer measurements. 11

2.1 The vision and technical overview of *Owlet*, a low-power and miniaturized system for extracting spatial information from sound. *Owlet* uses acoustic microstructures to embed direction-specific signatures on the recorded sound and develops a learning-based approach for signature recovery and mapping in real-time. 16
2.2 The concept of using a stencil with direction-specific hole patterns and microstructures for passive filtering of the incoming sound. The stencil embeds a directional response to the recorded signals. 20
2.3 The concept of passive directional filtering using a stencil of acoustic microstructure. The stencil embeds a directional signature to the recorded sound unique to its direction of arrival (DoA). The spectrum of complex gains represents the signature for further computation. 21
2.4 Different types of metamaterial stencils used in our experiments. 23
2.5 Angular diversity of the microphone with and without the microstructure stencil. 24
2.6 Comparison of the diversity in frequency responses (amplitude and phase) of the three types of metamaterial stencils. 24
2.7 The two-microphone model for eliminating source and environmental dependency. . . . 26

2.8 The architecture of the proposed CNN model. 29
2.9 The behavior of sound field at the outer surface of an obstacle. (a) When the object's size is much larger compared to the wavelength of the sound, the obstacle creates a shadow region. (b) When the object's size is comparable to the wavelength of the sound, the wave diffracts around the object creating high-pressure at a larger region of the surface. It also creates a high-pressure region directly opposite to the sound's directions where sound fields from the top and bottom sides meet. 32
2.10 (a) A one-hole stencil to measure surface pressure levels. (b) Sound amplitude at different angles from the sound's direction of arrival. 33
2.11 Comparison of diversity in phase and amplitude patterns for an optimal and a sub-optimal design of the stencil. 35
2.12 The *Owlet* prototype used in the evaluation experiment (left) and a 9-element uniform linear microphone array used as baseline for comparison(right). The array is $12cm$ wide, where *Owlet* is significantly smaller measuring less than $2cm$ in its largest dimension. . . 37
2.13 Various locations for system evaluations: (a) indoor laboratory, (b) indoor lobby, (c) outdoor. 38
2.14 Overall performance of the *Owlet* system compared to the traditional microphone arrays of various sizes. *Owlet* requires $100\times$ less energy than the state-of-the-art array systems while achieving better accuracy than a 9-element array. 39
2.15 Performance under external conditions: (a) The impact of varying types and loudness levels of ambient noise on the median DoA estimation error. (b) The CDF of errors when the sound source is located at varying distances from the receiver. (c) The CDF plot of estimation error for different elevation angles or the vertical positions of the sound source. (d) The CDF plots of errors that show the impact of dynamic movements in the environment. 40
2.16 The performance of sound tracking while the source is constantly moving near the sensor. The movement of the source creates a dynamic multipaths scenario. 41
2.17 The CDF of median error for (a) different environments and (b) different types of sound sources. 42
2.18 The performance for DoA estimation with known room size: (a) The confusion matrix and (b) the CDF of error in degrees of angle. 43
2.19 The localization error as (a) heatmap and (b) empirical CDF. 44
2.20 Performance comparison of *Owlet* with the implementation of beamscan, MVDR, and MUSIC algorithms: (a) The CDF of median errors, (b) The spatial spectrum for an incoming signal from 20° angle. 45
2.21 Performance comparison of *Owlet* with different deep learning models and architectures. 45
2.22 The setup for evaluating energy consumption. The setup tracks the energy requirements of *Owlet* and baseline microphone arrays under various conditions using a Keysight E6313A power supply and monitor. 46
2.23 Energy consumption of (a) MSP430FR5969 low-power ADC [1] for different sampling rates and (b) Keysight Data Acquisition System [2] for different number of microphones. 47
2.24 Overall energy consumption of array-based systems and *Owlet*. 48

3.1 *SPiDR*, an ultra-low-power acoustic spatial sensing system for mobile robots. The system uses a carefully designed 3D-printed micro-structure for projecting spatially coded signals for imaging. 53

3.2 The concept of the spatially coded channel sounding method. The received signal is the weighted linear combination of the reflections that bear the direction-specific signature. . 54
3.3 Diversity projection with the stencil with the internal channels to encode unique gains to signals to probe each pixel on the object plane with a unique signature. 58
3.4 (left) The 3D design of the stencil and (right) the internal structure showing the tubular helical paths of different lengths. 59
3.5 Ultrasound emitted from the speaker without (left) and with the stencil (right). The stencil spreads the signal energy over the region of interest. 59
3.6 The amplitude of acoustic signal at the cross section of the image plane when (left) speaker does not have any stencil, and (right) when it has a stencil. The internal micro-structure of the stencil diversifies the signal amplitude as direction codes. 61
3.7 (a) Energy consumption for different number of columns in the channel matrix H. (b) The correlation between the signals at nearby locations. The pixels within $3cm$ have correlation higher correlation higher than 0.5. 62
3.8 Row 1: Comparison of the correlation of the time domain signals project at different angles with an arbitrary stencil (left) and an optimized stencil (right). Row 2: Comparison of the location detection performance for a small object (1cm wide) placed at different angles from the sensor with an arbitrary stencil (left) and an optimized stencil (right). . . 64
3.9 Different sizes of stencils used in our experiments. 65
3.10 F values with (a) different lengths of tubes inside stencil, (b) different diameters of tubes. 67
3.11 Confusion matrix of imaging accuracy for different frequencies and after combining all frequencies together. 68
3.12 Stacking multiple frames suppresses spurious objects in the scene. Above we show the result of motion stacking of 5 frames taken as the robot moves. 70
3.13 Depth-map reconstruction using *SPiDR* for various real-world scenes. 72
3.14 Overall performance of the *SPiDR* compared to Intel Realsense lidar and ultrasound distance sensor mounted on a servo motor. *SPiDR* consumes a fraction of power compared to Lidar and motor based systems while delivering high accuracy in depth-map reconstruction. 73
3.15 Impact of varying (a) types and (b) levels of noises on depth-map reconstruction. 73
3.16 *SPiDR*'s performance in different environments. 74
3.17 Cross-sectional depth-map reconstruction performance in terms of (a) RMS error and (b) structural similarity, as a function of sparsity of the scene. 75
3.18 Horizontal localization performance in terms of (a) RMS error and (b) structural similarity, as a function of sparsity of the scene. 75
3.19 CDF plot of (a) RMS error and (b)structural similarity for depth-map reconstruction at varying resolutions. 76
3.20 Scene reconstruction results at 1cm and 0.5cm resolutions. We modify the number of columns in the channel matrix to have a higher resolution. 76
3.21 The depth-map reconstruction performance for (a) different materials and (b) different proximity between two objects. 77
3.22 Depth-map reconstruction outputs for varying proximity between objects. 77
3.23 Depth-map reconstruction outputs of different depths of the objects. 78
3.24 Performance of depth-map reconstruction with different depths of the objects. 78

3.25 Scene reconstruction of a horizontal bar with and without frequency or motion stacking. (a) Ground truth, (b) Raw output, (c) Only motion stacked, (d) Only frequency stacked, (e) Both motion and frequency stacked. 79
3.26 The scene reconstruction with and without the fractional computing method. (a) Ground truth, and the scene reconstruction (a) without and (b) with fractional computing. 79
3.27 Performance with different sizes of H_{meta}. 80
3.28 *SPiDR* prototype for power evaluation. 80
3.29 Power consumption during sampling and computation. 82
3.30 Comparison of power consumption with different sizes of H_{meta}. Prototype on MSP430FR5969. 83

4.1 *Sirius* dynamically switches the beam pattern of the antenna to embed direction specific signature to the received signal. The vector of amplitudes contain the unique signature which map to the angle-of-arrival, θ. 88
4.2 *Sirius* uses pin-diode to switch the gain pattern of an antenna. Connecting and disconnecting a conductive patch to the surface of antenna can change the shape of the gain pattern. This figure shows four different configurations of the antenna controlled using a set of two switches. 90
4.3 Feasibility study demonstrates dynamic gain pattern switching. Figure shows gain patterns for (a) our reconfigurable antenna and (b) regular antenna. 93
4.4 Passive envelope detector used by *Sirius* captures incoming signal energy without requiring power-hungry components such as oscillators and down-converters. . . 95
4.5 Gain patterns for the first 3 paths that signal takes to reach the antenna. First one is the direct path which constitutes the majority of the energy, followed by weaker reflections. 97
4.6 *Sirius* uses triangulation to localize mobile nodes. It estimates angles from at least two anchors with known locations. 98
4.7 Figure depicts envelope detector's received signal: (top) Anchor1 transmitting, (middle) Anchor2 transmitting, and (bottom) both Anchor1 and Anchor2 transmitting. 99
4.8 Anchor beacon signals designed with varying duty cycles create distinct time windows for interference-free reception. (left) Case 1: Anchor2's window as the outer neighbor of collision windows. (right) Case 2: Anchor2's window as the inner neighbor of the collision window. 101
4.9 Figure displays anchor detection algorithm outputs: (top) time domain signal, (middle) total energy from all configurations, and (bottom) signal's first derivative. 102
4.10 Using pin diode as RF switch. (a) Biasing circuit (b, c) Equivalent lumped model for on and off states. 103
4.11 Switching and sampling techniques: (a) Sequential sampling: constant-time switch, hold, and sample, (b) Uniform sampling: continuous configuration switching with constant rate, (c) Burst sampling: high-speed switching and sampling with long duty cycles for omnidirectional communication. 103
4.12 *Sirius*'s prototype for localization. 107
4.13 We evaluate *Sirius* on different antennas in ISM band 900MHz and 2.4GHz. The figure shows the fabricated reconfigurable antennas used in the prototype. 107
4.14 Setup for outdoor long-range data collection. The map shows the static anchors and node locations. 108

4.15 (a) Overall AoA estimation accuracy of *Sirius*. (b) CDF of AoA error. 109
4.16 (a) CDF of localization error (b) Localization error per location shows the impact of distance from the APs. 110
4.17 Effect of pattern switching speed on communication. 111
4.18 Impact of antenna reconfiguration on bit error rate for different antennas. 111
4.19 Energy distribution between direct path and multipath for various indoor locations. 113
4.20 AoA errors for different indoor locations. 113
4.21 AoA errors for different reflection materials. Each cluster of bars represents one type of material, and each bar within a cluster represents different indoor locations. 114
4.22 Inverse distance matrix for varying number of gain patterns (N). 115
4.23 AoA estimation errors. (left) varying levels of multipath density in the environment. (right) varying number of gain patterns. 115
4.24 (a) Correlation of gain pattern amplitudes. (b) Mean and median AoA errors showing the long-term stability of gain patterns in a dynamic environment. 116
4.25 The figure shows gain patterns collected at two different indoor locations in a dynamic environment. 117
4.26 CDF of AoA error for different environments. 117
4.27 Our experimental setup at different indoor locations in varying levels of static and dynamic multipath. 117
4.28 AoA error for varying distance from the anchor. 118
4.29 Impact of clock drift on AoA estimation 120
4.30 Impact of frequency shift on AoA estimation 120

5.1 *LiTEfoot*- an ultra-low-power wireless tracker next to a US quarter for scale. . . . 124
5.2 LTE cells and frame structure. 129
5.3 Intermodulation and spectrum folding. 132
5.4 Confusion matrix for PCI estimation using SSS^2 (i.e., the SSS signal after non-linearity). 136
5.5 LTE packets before and after stacking (a) 1.4 MHz bandwidth (b) 10 MHz bandwidth. After stacking the PSS and SSS strength increases in the 10 MHz case whereas the data fades into a DC bias. 138
5.6 (a) Blind Separation of weighted superposition (b) Linear phase change introduced due to sub-sample offset in the OFDM subcarriers. 140
5.7 The high-level circuit schematic of *LiTEfoot*. 144
5.8 *LiTEfoot* PCB tag prototype showing the low-power RF frontend. 145
5.9 *LiTEfoot*'s estimated trajectories and localization errors for in (a) urban and (b), (c) rural environments. The black lines denote the GPS trajectory and blue markers denote *LiTEfoot*'s estimated trajectory. The cell towers that are detected during the route are shown in red. (d) The empirical CDF of the localization errors. . . . 145
5.10 Diversity in downlink center frequencies used by the cells in a 200-meter radius. . 148
5.11 PCI estimation accuracy (F1 score) vs. number of stacked frames for LTE, 5G-NR, and NB-IoT. 151
5.12 (a) CDF of phase offsets in measurements before and after sub-sample offset correction. (b) Localization error for varying speed of vehicle. 152

5.13 (a) An alert is generated when the tag exits the boundary of the marked region. (b) Alert generation distance from the map boundaries. 154

6.1 *Locate3D*'s approach to include both angle and edge constraints for faster and efficient localization. 160
6.2 Comparative analysis of constraints: Incorporating angles reduces the number of edges required to attain the same level of accuracy as the "Ranges only" approach. 166
6.3 (a) Histogram of localization errors for all spanning trees. (b) Reported angles by COTS UWB sensor. 170
6.4 Different spanning trees representing rigid and non-rigid graphs. Solid lines indicate both range+angle edges, and dashed lines indicate range-only edges. (a) A connected but non-rigid graph due to missing angle information in an edge. (b) The subgraph is free to rotate. (c) Adding a range measurement makes the graph rigid. 171
6.5 The displacements corresponding to zero eigenvalues represent the translational and rotational motions that the nodes can undergo without violating any constraint. 172
6.6 (a) When Node 1 is aligned with the global frame of reference, it reports that Node 2 is positioned at angle θ. (b) When Node 1 is rotated by γ relative to the global frame of reference in a 3D space, it reports a distinctly different angle θ', for the same node. . . . 174
6.7 The relation between AoAs reported by two nodes (a) when nodes are perfectly aligned, and (b) when nodes are oriented with angle γ relative to the global frame of reference. . . 176
6.8 (a) Time to reach '0 False positives'. (b) Percentage of users registered for varying trajectory matching thresholds. 178
6.9 Room-scale evaluation: CDF of localization errors. 180
6.10 3D Localization performance of *Locate3D* compared to baseline system - Cappella [3] - which uses visual odometry along with UWB. Results show performance in different lighting conditions for (a) Moving nodes and (b) Static nodes. 181
6.11 (a) Prototype built using Raspberry Pi, UWB sensor, Intel Realsense, IR markers for ground truth and a battery pack. (b) Room-scale evaluation. (c) Building-scale evaluation (d) 3D Lidar scan of the building for reference (not used in computation) (e) Snapshot of estimated locations and MST. (f) AprilTags [4] captured by nodes for ground truth (not used in computation). 182
6.12 City-scale analysis: (a) CDF of localization errors of 30k nodes, $3800m \times 3800m$ area, and 15 anchors. (b) Errors for 30k, 60k, and 100k nodes in $1500m \times 1500m$ area and 1 anchor. 183
6.13 New York City wide-area simulation results: CDF of localization errors for 100,000 nodes using 1 (left figure) and 5 (right figure) anchors in a $22000m \times 3200m$ area. 185
6.14 Impact of submodules on (a) latency and (b) accuracy. 186
6.15 Localization errors (in meters) for infrastructure baseline [5] and *Locate3D* for varying nodes and anchors. 187
6.16 (a) Number of unreachable nodes and (b) Total number of NLOS measurements made with varying anchor area density. 188
6.17 (a) LOS and NLOS localization errors. (b) Static vs Mobile nodes localization errors . . 188
6.18 (a) City-scale results for a 1000 node topology simulation in a $200 \times 200 \times 50$ meter 3D space for a varying number of static anchors. (b) Room-scale localization error for real-world 20-node experiments with varying numbers of anchors registered. 189

6.19 Ranging and AoA errors for various range-angles. 190
6.20 (a) CDF orientation errors. (b) Percent of users registered as Virtual anchors over time. . 191

7.1 Complex permittivity of electromagnetic waves in water. 199
7.2 FreshSense: Different frequencies travel at different speeds through water resulting in an additional frequency dependent delay. 200
7.3 Estimated dispersion for varying percentage of water in the box. 201
7.4 Evaluation setup with a box of avocados. 202
7.5 Images of avocados captured across 14 days. 204
7.6 Ground truth measured for 5 individual fruits and their average trend. 204
7.7 Correlation between estimated and true DM% for the testing data. 205
7.8 Estimated and true DM% across 14 days. 205

List of Abbreviations

ADC	Analog-to-Digital Converter
AI	Artificial Intelligence
AP	Access Point
AoA	Angle of Arrival
API	Application Programming Interface
ASIC	Application-Specific Integrated Circuit
BLE	Bluetooth Low Energy
CDF	Cumulative Distribution Function
CGI	Cell Global Identity
CMOS	Complementary Metal–Oxide–Semiconductor
CNN	Convolutional Neural Network
COTS	Commercial Off-The-Shelf
DAQ	Data Acquisition System
DC	Direct Current
DM	Dry Matter
DoA	Direction of Arrival
DoF	Degree of Freedom
eDRX	Extended Discontinuous Reception
eNodeB	Evolved Node B
FCC	Federal Communications Commission
FDA	Food and Drug Administration
FDD	Frequency Division Duplex
FFT	Fast Fourier Transform
FMCW	Frequency-Modulated Continuous Wave
FoV	Field of View
GHz	Gigahertz
GPS	Global Positioning System
GS	Gerchberg-Saxton
HFSS	High Frequency Structure Simulator
IoT	Internet of Things
IoU	Intersection over Union
IQ	In-phase and Quadrature
LEA	Low Energy Accelerator
LiDAR	Light Detection and Ranging

LNA	Low-Noise Amplifier
LOS	Line of Sight
LTE	Long-Term Evolution
MCU	Microcontroller Unit
MDS	Multidimensional Scaling
MEMS	Microelectromechanical Systems
MHz	Megahertz
ML	Machine Learning
MLP	Multilayer Perceptron
MST	Minimum Spanning Tree
MUSIC	MUltiple SIgnal Classification
MUT	Micromachined Ultrasound Transducer
NB-IoT	Narrowband Internet of Things
NeRF	Neural Radiance Fields
NIR	Near-Infrared
NLOS	Non-Line of Sight
NSSS	Narrowband Secondary Synchronization Signal
OFDM	Orthogonal Frequency Division Multiplexing
PA	Power Amplifier
PCA	Principal Component Analysis
PCB	Printed Circuit Board
PCI	Physical Cell Identity
PIN	Positive-Intrinsic-Negative (diode)
PSS	Primary Synchronization Signal
QPSK	Quadrature Phase Shift Keying
R-squared	Coefficient of Determination
RF	Radio Frequency
RLS	Recursive Least Squares
RMSE	Root Mean Square Error
RSSI	Received Signal Strength Indicator
SAR	Successive Approximation Register
SDR	Software-Defined Radio
SIMD	Single Instruction, Multiple Data
SMACOF	Scaling by MAjorizing a COmplicated Function
SNR	Signal-to-Noise Ratio
SPL	Sound Pressure Level
SPI	Serial Peripheral Interface
SSIM	Structural Similarity Index
SSS	Secondary Synchronization Signal
STL	Stereolithography

SWaP	Size, Weight, and Power
SWaP-C	Size, Weight, Power, and Cost
TSS	Total Soluble Solids
ToF	Time-of-Flight
UDP	User Datagram Protocol
UE	User Equipment
UHF	Ultra High Frequency
UWB	Ultra-Wideband
VIO	Visual Inertial Odometry
VNA	Vector Network Analyzer

Chapter 1: Introduction

Imagine a world where intelligence is woven into the fabric of our physical environment. Insect-scale robots navigate disaster zones, paper-thin tags track goods across global supply chains without manual maintenance, and medical implants detect disease markers years before symptoms appear. The vision of intelligent systems seamlessly embedded in the environment has attracted sustained interest from the research community since the late 1990s, when pioneers like Mark Weiser envisioned *ubiquitous computing* [6] and Kristofer S. J. Pister, Joe Kahn, and Bernhard Boser conceptualized *smart dust* [7]. While today's smart devices offer glimpses of this future, they represent only incremental steps toward true ambient intelligence—systems that vanish into the background while autonomously sensing, learning, and adapting to human needs. The fundamental challenge lies in scaling intelligence sustainably: current approaches rely on power-hungry hardware, frequent maintenance, and expensive components, making widespread deployment impractical [8]. Achieving this vision of ambient intelligence at scale requires fundamentally new approaches that can operate within extreme resource constraints.

Over the past two decades, we have pursued miniaturization and low-power operation as primary paths toward embedding intelligence in everyday objects. This approach has yielded impressive advances in microelectronics, MEMS sensors, and efficient computing. However, simply making conventional systems smaller encounters fundamental physical barriers that force

difficult tradeoffs between intelligence capabilities and energy requirements. For instance, consider the contrast between perception systems for autonomous vehicles and insect-scale robots: self-driving cars use LiDAR sensors consuming watts of power, while a robot at millimeter scale must achieve similar functionality within sub-milliwatt constraints. Sensor arrays required for spatial perception cannot be reduced beyond wavelength-dependent limits without sacrificing resolution or accuracy. Furthermore, modern deep-learning systems demand substantial computational resources that conflict with the strict power budgets of tiny devices. Our goal, in this thesis, is not merely to reduce power consumption, but to maintain intelligent capabilities while operating within extreme resource limitations. To summarize, these challenges indicate that miniaturization alone cannot bridge the gap between current technology and true ambient intelligence. We need to fundamentally redesign our approaches to sensing, computing, and communication in resource-constrained environments.

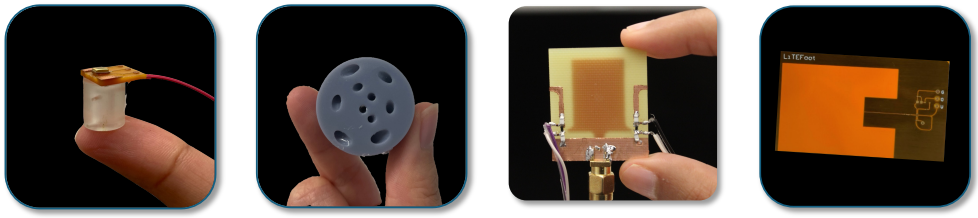


Figure 1.1: Examples of intelligent, low-power, and scalable sensors developed in this thesis.

In this thesis, we investigate how to achieve ambient intelligence by fundamentally reimagining sensing and computing architectures from first principles. We show how bio-inspired meta-structures (see Figure 1.1) enable spatial perception with single sensors that match the performance of multi-element arrays while using $1000\times$ less power; how ultra-low-power techniques can provide GPS-like positioning with sticker-sized tags operating for years on a single battery; and how novel sensing approaches can monitor food quality non-invasively throughout global supply

chains. Through innovations at the intersection of physics and computation, we create sensing systems that achieve orders of magnitude improvements in energy efficiency, size and scalability, making intelligence practical in environments previously considered impossible.

1.1 Challenges in Scaling Ambient Intelligence

The gap between our vision of integrating intelligence in everything and current reality stems from three fundamental challenges:

Nature's Efficiency vs. Artificial Systems: Nature has created intelligent systems that operate within incredible efficiency constraints. A fruit fly navigates complex 3D environments, avoids predators, and locates food with a brain of merely 100,000 neurons consuming just microwatts of power [9]. Similarly, desert ants perform precise navigation across vast distances without GPS, using minimal neural hardware [10]. In stark contrast, today's artificial intelligent systems capable of comparable perception [11], reasoning and decision-making [12] require orders of magnitude more resources—hundreds of megabytes of memory and several watts of power [13]. This fundamental mismatch prevents direct application of current AI approaches in resource-constrained devices, creating a critical barrier to embedding intelligence in everyday objects.

Physical Limitations: Traditional sensing paradigms encounter fundamental physical barriers when scaled down to micro-devices. Spatial perception systems like radar, sonar, and camera arrays rely on sampling theory principles that demand multiple sensors separated by minimal distances to avoid aliasing. The Nyquist-Shannon sampling theorem states that to accurately capture a signal with maximum frequency component f_{max}, sampling must occur at a

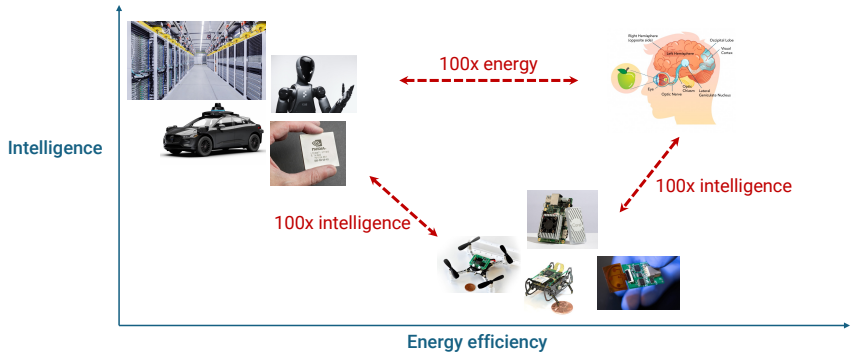


Figure 1.2: The intelligence-energy tradeoff: Conventional intelligent systems (top left) operate at high energy levels and offer good capabilities. Current resource-constrained systems (bottom right) provide limited intelligence. This thesis explores approaches to bridge this gap by creating systems that offer $100\times$ more intelligence at similar energy levels.

rate of at least $2f_{max}$ [14, 15]. This creates critical hardware requirements: higher sampling rates demand faster ADCs with correspondingly higher power consumption. Additionally, diffraction limits optical sensing according to Abbe's criterion ($d = \lambda/2NA$), while signal-to-noise ratio deteriorates as sensor size decreases. These physics-based constraints collectively create seemingly insurmountable barriers to effective sensing in tiny platforms.

Energy Sustainability at Scale: Achieving ambient intelligence requires not just creating intelligent systems but ensuring they can operate sustainably at planetary scale. As illustrated in Figure 1.2, current intelligent systems consumes order of magnitude more power, requiring frequent battery replacements or continuous power delivery. The energy gap is huge mainly because perception typically requires hundreds of milliwatts, while energy harvesting provides only microwatts in many scenarios [16]. This gap cannot be bridged through incremental improvements alone. It demands radically different approaches to energy-conscious intelligence.

These challenges collectively highlight that simply miniaturizing conventional sensing and computing approaches is insufficient. The vision of ambient intelligence requires fundamentally reimagining how we design intelligent systems from first principles, particularly for extreme

resource constraints.

1.2 Designing Ultra-Low-Power Intelligent Systems

Traditional approaches to intelligent systems follows the pipeline: collect high-resolution signals using sensors, convert analog signals to digital data, and process this data through computationally intensive large ML models. This conventional paradigm tightly integrates sensing with digital processing, requiring substantial computational resources to extract meaningful information from raw sensor data. While this approach works well for resource-rich platforms like self driving cars, drones, humanoids; it fundamentally breaks down in resource-constrained scenarios. Tiny robots, wearable devices, and battery-free sensor networks cannot support the high-dimensional data acquisition, digital conversion, and large model inferencing required by conventional intelligent systems. This mismatch between learning capabilities and hardware limitations represents a critical barrier to realizing ambient intelligence at scale.

Our approach departs from this conventional primitive by fundamentally rethinking the relationship between sensing hardware and computation. We *relax the strict hardware requirements* for sensing on resource-constrained devices and instead leverage tiny ML models that operate on *physically pre-processed signals*. The key idea is to merge the fundamental laws of physics with learning principles, creating sensing frontends that perform signal transformations in the analog domain before any power-consuming digital processing occurs. In other words, these passive structures act as physical neural encoders, encoding domain knowledge directly into hardware that requires zero power to operate.

For instance, in acoustic sensing, we implement this concept through 3D-printed metamaterials

that transform omnidirectional microphones into direction-aware sensors (Figure 1.1a, 1.1b). For wireless applications, we develop gain-pattern reconfigurable antennas that embed spatial information directly into received signal strength measurements (Figure 1.1c). By performing these transformations in the physical domain, our approach dramatically reduces both data dimensionality and energy requirements. The resulting sensor outputs contain high-level features extracted with minimal energy consumption, requiring significantly smaller neural models for downstream processing while providing information that naturally aligns with the hierarchical structure of neural networks.

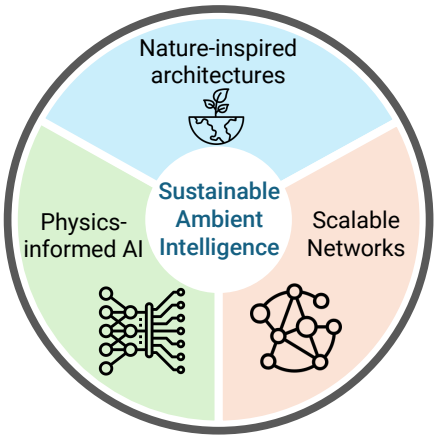


Figure 1.3: Our approach to sustainable ambient intelligence combines nature-inspired architectures, physics-informed AI, and scalable networks.

This approach combines three complementary strategies that together enable sustainable ambient intelligence, as illustrated in Figure 1.3. First, we draw inspiration from nature's efficient solutions to similar problems. Biological systems have evolved sophisticated sensing capabilities within strict resource constraints—owls can precisely locate prey in total darkness using asymmetric ear structures, insects navigate complex environments with tiny brains, and even plants respond to environmental stimuli through passive mechanisms. By studying these biological systems, we identify principles for efficient spatial sensing that can be translated into engineered

systems through biomimetic design. This bio-inspired approach leads to sensing architectures that achieve remarkable capabilities with minimal active components.

Second, we integrate physics-informed machine learning to bridge the gap between simple hardware and complex perception tasks. Our systems incorporate passive components that perform initial signal transformations in the analog domain without consuming power, effectively encoding domain knowledge directly into the hardware. These physical structures act as computational elements, extracting meaningful features before digital processing begins. The resulting signals naturally align with the hierarchical feature extraction performed by neural networks, dramatically reducing the computational burden on subsequent processing stages. This physics-informed approach enables us to extract maximum information from minimal hardware, achieving capabilities that would traditionally require complex, power-hungry systems.

Third, we develop techniques for scaling these intelligent systems across large networks and diverse environments. By reimagining how devices communicate, collaborate, and leverage existing infrastructure, we create systems that can be deployed at unprecedented scales—from city-wide sensor networks to micro-robotic swarms. This scalable approach enables widespread deployment of ambient intelligence without requiring dedicated infrastructure or frequent maintenance, making these systems practical for real-world applications.

Together, these strategies enable a new class of intelligent systems with transformative potential. They deliver extreme energy efficiency, supporting sensing functions, that typically demand complex digital systems, now with minimal power use. They improve robustness by leveraging core physical principles for better generalization. They also allow real-time adaptability through dynamically reconfigurable designs. Most importantly, they close the gap between ambient intelligence's promise and the practical limits of embedded platforms, making it feasible to scale.

1.3 Systems Developed

1.3.1 Structure-Assisted Spatial Intelligence

Spatial perception systems fundamentally rely on sensor arrays spanning multiple wavelengths, consuming hundreds of milliwatts of power and requiring substantial physical space. This dependency on arrays creates a critical barrier for resource-constrained devices like micro-robots and IoT sensors. We developed a different approach that achieves spatial sensing without arrays by combining carefully designed passive structures with minimal active components, significantly reducing both power and computational requirements.

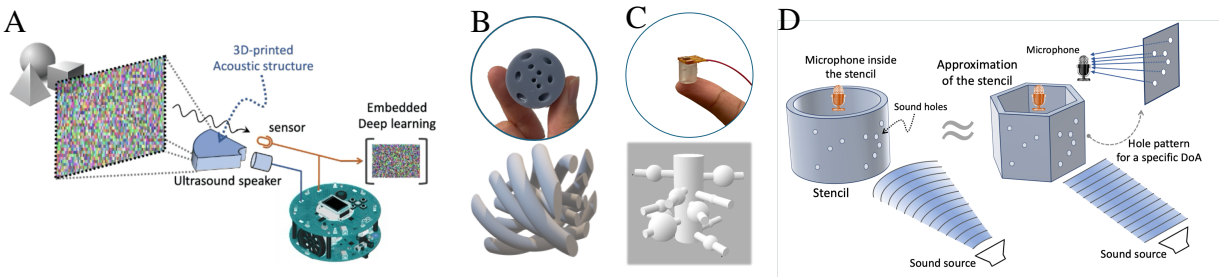


Figure 1.4: Structure-assisted sensing systems: (A) *SPiDR*'s depth imaging using 3D-printed acoustic metamaterial, (B) *SPiDR*'s acoustic stencil and its internal spatial filtering channels, (C) *Owlet*'s acoustic stencil and its internal direction-finding channels, (D) *Owlet*'s direction finding with passive acoustic structures.

Bio-inspired DoA Estimation. Conventional acoustic DoA estimation requires multiple synchronized microphones separated by half-wavelength distances. Drawing inspiration from how barn owls achieve precise sound localization through asymmetric ear structures, we developed *Owlet* [17], a system that reimagines direction estimation. The key insight lies in leveraging diffraction and Helmholtz resonance through 3D-printed metamaterials to create direction-dependent acoustic filtering (Figure 1.4 C, 1.4 D). By wrapping a single microphone with a structured pattern

of holes and resonant cavities, each sound direction creates a unique spectral signature. We solved the critical challenge of environmental robustness through a two-microphone architecture that eliminates both source and environmental dependencies. The system achieves 3.6° angular error—matching a 9-microphone array while using $100\times$ less power.

Single-sensor Depth Perception. Insect-scale robots operating in unknown environments require depth perception for navigation, but existing solutions like LiDAR and ultrasound arrays are power-hungry and bulky. We developed *SPiDR* [18], a fundamentally new approach to depth perception using a single microphone-speaker pair. The core idea is to spatially encode the acoustic channel through a metamaterial "stencil" that creates unique signatures for each point in 3D space (Figure 1.4 A, 1.4 B). These physics-optimized waveguides embed both direction and distance information in single measurements, eliminating the need for scanning or arrays. Through sparse recovery algorithms informed by wave interference patterns, *SPiDR* achieves centimeter-level depth accuracy while consuming only 0.83mJ per frame—a $400\times$ improvement over traditional solutions. This work establishes new possibilities for perception in resource-constrained robotics.

Passive Spectral Analysis. Extending our structure-assisted sensing approach to spectral processing, we drew inspiration from how the human cochlea naturally decomposes sound into frequencies. We developed *Lyra* [19], a system that uses standing wave resonators to implement FFT-like spectral analysis without power-hungry ADCs or digital processing. By leveraging wave interference patterns in carefully designed cavity structures, *Lyra* enables always-on acoustic monitoring with microwatt power consumption, making it suitable for continuous environmental sensing in battery-constrained or energy-harvesting scenarios.

RF Self-localization. Building on our success with acoustic metamaterials, we extended these principles to RF signals with *Sirius* [20]. While passive structures worked well for acoustic

sensing, the small wavelength diversity in RF necessitated a different approach. We developed a gain-pattern reconfigurable antenna that dynamically embeds direction-specific codes in received signals. Recent advances show envelope detectors enable ultra-low-power communication but cannot extract phase information needed for spatial sensing. We solved this through a neural pipeline that learns to decode spatial information directly from signal amplitudes, achieving 7° angular accuracy while consuming $1000\times$ less energy than array-based positioning systems. This enables sustainable sensor networks across agricultural fields, supply chains, and wildlife habitats where traditional GPS (25mJ per fix) would quickly deplete batteries.

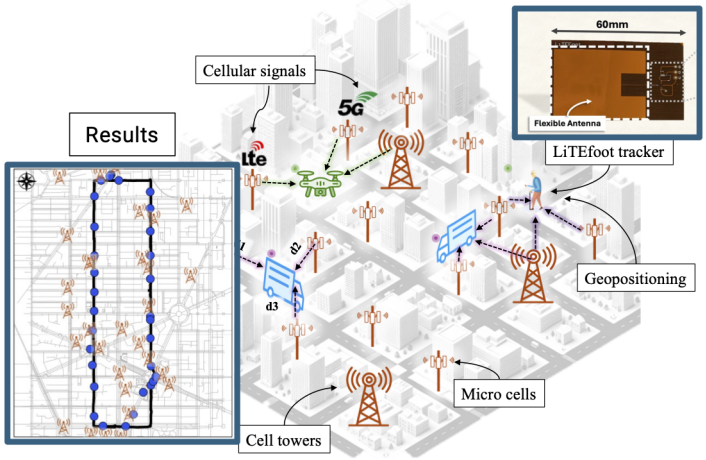


Figure 1.5: *LiTEfoot* leverages cellular signals for nationwide tracking using an ultra-low-power miniaturized receiver.

1.3.2 Scalable nextG Systems

From tracking perishable goods to monitoring elderly patients with dementia, continuous location tracking of small assets and individuals has become essential across numerous domains. However, existing solutions like GPS rely on bulky batteries or dedicated infrastructure, creating fundamental barriers to widespread deployment. We developed two complementary systems that

reimagine global positioning for resource-constrained scenarios, enabling seamless tracking from nationwide supply chains to dense urban environments.

Ultra-low-power Location Tracking. Traditional cellular localization requires frequency hopping across multi-GHz bandwidths using power-hungry oscillators and IQ demodulators that consume over 100mW. We developed *LiTEfoot* [21], a cellular-based self-localization system that uses non-linear intermodulation to simultaneously capture synchronization signals across 3GHz spectrum through a passive envelope detector (Figure 1.5). The system decodes Physical Cell Identities from the folded spectrum and performs multilateration, achieving 19m accuracy while consuming only 40μJ—a 625$\times$ reduction compared to GPS. This enables 11-year operation on a coin cell battery for nationwide asset tracking in supply chains, precision agriculture, wildlife conservation, and healthcare monitoring, all without requiring dedicated infrastructure.

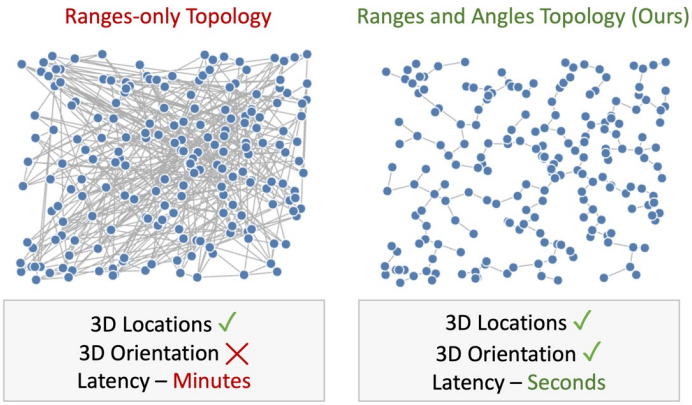


Figure 1.6: *Locate3D* enables infrastructure-free tracking at scale through optimized peer-to-peer measurements.

Infrastructure-free 6DoF Tracking at Scale. For scenarios without cellular infrastructure, we developed *Locate3D* [22], a peer-to-peer system enabling infrastructure-free 6-degree-of-freedom tracking using UWB radios for massive-scale networks. The system introduces angle measurements alongside ranging to reduce the minimum edges needed for unique topology

realization by 4× (see Figure 1.6). Using rigidity-aware spanning tree optimization and non-rigid graph decomposition, *Locate3D* achieves 0.86m accuracy in building-scale networks and 12.09m in city-scale deployments while reducing latency by 75%. The system seamlessly scales to track 100,000 devices across cities, enabling transformative applications from coordinating disaster response teams to managing autonomous delivery fleets and monitoring smart city infrastructure.

1.3.3 Systems for Sustainability

While 800 million people face hunger globally, nearly 40% of food produced is lost to waste, primarily due to inadequate monitoring during distribution. In collaboration with Microsoft Research, we developed sensing systems to enable data-driven food quality monitoring across global supply chains.

Non-invasive Food Quality Monitoring. Current food quality assessment methods require destructive testing, making continuous monitoring throughout distribution impractical. We developed FreshSense [23], a wireless sensing system that monitors dry matter content non-invasively at the pallet level. The key challenge lies in measuring subtle changes in water content through densely-packed produce where traditional RF sensing fails due to complex multi-path effects. We solved this through a dispersion-based sensing approach that exploits frequency-dependent wave propagation in water-rich environments. By analyzing electromagnetic delays with physics-informed neural networks, FreshSense achieves robust quality assessment while eliminating environmental variations.

1.3.4 Other Contributions

In my Ph.D., I have also developed solutions across security, AI verification, and audio processing, demonstrating the broader applicability of our approach.

Security and Self-defense for Drones. As drones become trusted delivery systems and law enforcement tools, they face increasing risk of mid-air attacks and vandalism. We developed DopplerDodge [24], an acoustic sensing system enabling real-time threat detection and avoidance in resource-constrained drones. Using just a single microphone and Doppler effect analysis, the system detects incoming projectiles with 100ms advance warning, enabling autonomous evasive maneuvers while consuming minimal power. This work establishes a new direction in embedded defense systems for tiny autonomous vehicles.

Side-channel Security in Embedded AI. With the proliferation of edge AI, verifying the trustworthiness of model inference is increasingly important. We developed ThermWare [25], a system that leverages thermal side-channels to detect anomalous computations in embedded AI systems. By capturing spatiotemporal heat signatures with a thermal camera, the system identifies unauthorized operations with 94% accuracy, enabling non-invasive run-time AI model monitoring without requiring system-level access or modifications.

Noise-cancellation for Wearables. To improve voice communication in noisy environments, we developed VoiceFind [26], a speech enhancement system that uses just two microphones to achieve spatial filtering of desired speech. Through a combination of harmonic-based direction finding and conditional generative adversarial networks, the system improves speech intelligibility by 16% in real-world environments. This work demonstrates how physics-informed machine learning can enable sophisticated audio processing on resource-constrained wearables.

1.4 Organization

The remainder of this thesis is organized as follows. Chapter 2 introduces a bio-inspired acoustic sensing system that achieves direction finding with a single microphone through carefully designed metamaterials, enabling spatial audio perception in wearable devices. Chapter 3 presents a novel approach to depth perception for micro-robotics that provides high-resolution imaging with minimal power requirements through physics-based spatial encoding. Chapter 4 explores self-localization for IoT devices using reconfigurable RF structures and machine learning techniques, achieving long-range positioning with significantly reduced energy consumption. Chapter 5 details a cellular-based localization framework that enables nationwide tracking with sticker-sized tags by exploiting existing cellular infrastructure for applications in supply chains and healthcare. Chapter 6 describes an infrastructure-free tracking system that supports large-scale coordination of devices in urban environments through innovative peer-to-peer algorithms. Chapter 7 presents a non-invasive approach to food quality monitoring that creates three-dimensional maps of food quality within pallets through physics-informed sensing techniques. Chapter 8 concludes with reflections on the broader implications and future directions of this thesis.

Chapter 2: Structure-Assisted Spatial Audio Sensing

2.1 Introduction

Acoustic devices are increasingly becoming pervasive in our everyday environments. Beyond just voice interfaces, a broad spectrum of applications is emerging that leverages multiple facets of context-awareness and analytics. These applications encompass indoor activity monitoring based on sound [27–29], health monitoring through acoustic signals [30, 31], speech development support and acoustic environment sensing with on-body wearables [32, 33], along with numerous outdoor use-cases utilizing distributed sensor nodes [34, 35]. With the advent of new low-power and battery-free technologies [36, 37], it has become feasible to continuously capture and process sound using independent sensing modules distributed throughout the environment. Adding spatial analysis of sound and source localization can significantly enhance the capabilities of such context-aware systems. Meanwhile, spatial sensing of sound is also critical for robotic navigation and situational awareness systems, both in aerial environments [38–40] and underwater scenarios [41, 42]. However, conventional methods for obtaining spatial sound information typically rely on capturing multiple synchronized audio streams through microphone arrays, which is a power-intensive hardware requirement that is challenging for standalone sensing modules. In this chapter, we aim to build an acoustic sensing system that enables spatial information processing on power-constrained ubiquitous devices with compact form factors.

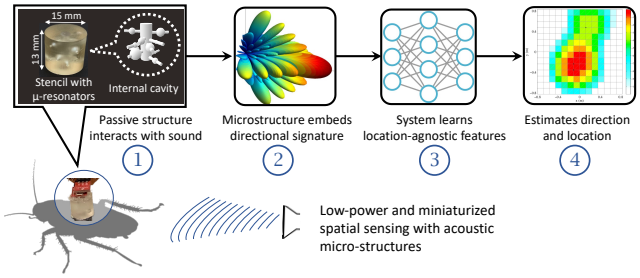


Figure 2.1: The vision and technical overview of *Owlet*, a low-power and miniaturized system for extracting spatial information from sound. *Owlet* uses acoustic microstructures to embed direction-specific signatures on the recorded sound and develops a learning-based approach for signature recovery and mapping in real-time.

Estimating spatial features of sound, such as direction-of-arrival (DoA) or source location, traditionally relies on sampling the acoustic wavefront across space using a microphone array. Since conventional DoA algorithms are fundamentally based on this spatial sampling model, both the array dimensions and the number of microphones directly impact their performance. According to the sampling theorem [14], microphones in a linear array must ideally be spaced at half the signal's wavelength ($\lambda/2$) for accurate DoA estimation. Moreover, the angular resolution, inversely proportional to the Half Power Beam-width, improves with the total length of the array aperture. Thus, achieving fine-grained DoA resolution typically demands large physical arrays. Additionally, these arrays require simultaneous sampling across all microphones, causing power consumption and hardware complexity to grow with array size. Although acoustic devices are increasingly common in ubiquitous computing, the power, hardware, and form-factor constraints limit the feasibility of traditional array-based spatial sensing. In this chapter, we explore an alternative approach to spatial signal processing that moves away from the standard spatio-temporal sampling paradigm, leveraging wave-structure interactions to enable low-power, compact, and simple designs.

Directional hearing aided by structural interactions is widespread in nature. The symmetric

placement of ears in most mammals, including humans, effectively forms a two-element array for directional processing. However, biophysical studies show that fine-grained localization relies heavily on how sound interacts with the complex three-dimensional geometry of the head [43]. Owls, for example, possess asymmetrically positioned ears along both horizontal and vertical planes [44], allowing them to precisely localize low-frequency sounds, a task difficult to achieve with symmetric structures alone. Remarkably, certain insects with body sizes much smaller than one-tenth of the sound wavelength can localize sound as accurately as mammals [45]. A grasshopper, for instance, with a body width of just $3mm$, achieves precise localization despite its small size relative to the sound wavelength. This is enabled by asymmetrical body structures that produce direction-dependent responses to incoming sounds. The sensory and neural systems of these organisms have evolved to map such responses to sound direction. Inspired by these biological mechanisms, we design a structure-assisted DoA estimation system suited for power-constrained and miniaturized sensing platforms.

In this chapter, we present the design and prototype of an acoustic localization system that introduces physical structures around a microphone to embed directional cues. As acoustic waves propagate, they interact with physical structures, resulting in transformations of the wave field. Such interactions are evident at large scales in room acoustics, where the same sound can differ based on the room's shape, size, and object placement. We demonstrate that small 3D-printed structures can similarly manipulate sound waves, imprinting unique signatures onto the passing sound. By placing a microphone within a structure only a few centimeters in size, the recorded signals inherently carry these signatures. With careful design, the structure can embed distinct signatures for sounds arriving from different angles, achieving angular resolutions of a few degrees. Our system detects these embedded signatures to infer the direction of arrival (DoA) of sound. We

name this system *Owlet*, inspired by a bird known for its exceptional auditory capabilities.

The idea of leveraging environmental variations in sound fields for localization is not new. Prior work has explored fingerprinting multipath environments and analyzing reflections for localization [46]. Closest to our approach is [47], which places objects in a 60×60 cm area with a central microphone, showing that scattered sound carries directional cues that can be used for DoA estimation. While *Owlet* builds on similar principles, it differs in two critical ways. First, we focus on creating a centimeter-scale sensing system suitable for resource-constrained robots and ubiquitous sensing applications. Our *Owlet* prototype achieves angular resolutions comparable to or better than previous work, with a compact $1.5\ \mathrm{cm} \times 1.3\ \mathrm{cm}$ sensor. Second, we address robustness to environmental changes. *Owlet* is designed to operate beyond controlled environments like anechoic chambers and does not require location-specific training data.

A primary challenge for *Owlet* is achieving sufficient multipath diversity within a small form factor. Because low-frequency acoustic signals have large wavelengths, traditional reflection-based approaches require similarly large reflectors to create diversity, directly limiting spatial resolution. We overcome this limitation by developing a diffraction-based technique for miniature acoustic structures. When sound passes through small apertures, it diffracts, effectively generating new secondary sources. We exploit this phenomenon by designing a 3D-printed cylindrical cover, called a stencil, that surrounds the microphone. These stencils incorporate optimally patterned holes that create complex but predictable multipath interference inside the structure. The resulting interference patterns carry signatures encoding the direction of arrival. To enhance angular diversity, we also integrate principles of metamaterial design into the stencil. The *Owlet* system learns these signatures through a one-time calibration and maps them to DoA estimates during operation.

Another major challenge is ensuring that the design remains robust against environmental changes that can unpredictably alter incoming sound. For practical deployment, the system must function reliably across diverse environments while requiring only a one-time calibration during manufacturing. As previously noted, room acoustics can distort the sound field and compromise the mapping between directional signatures and angles. To address this, *Owlet* introduces a reference microphone and adopts a communication-theoretic approach to suppress transient multipath effects during signature generation and mapping. This technique enhances *Owlet*'s robustness to environmental variations, making it viable for real-world applications.

This chapter explores acoustic structures as passive components for creating low-power, miniaturized solutions in ubiquitous sensing. Potential applications include wearable devices for acoustic environment sensing, such as systems for assessing speech development in infants [48,49] or personal acoustic analytics [50, 51], where sound direction is critical. Navigation in SWaP-constrained [52, 53] aerial and underwater robots can also benefit from spatial sensing capabilities enabled by *Owlet*. Moreover, *Owlet* offers a path toward directional sensing and localization in energy-harvesting systems, a task challenging for traditional microphone arrays. Figure 3.1 illustrates the broader vision and technical overview of this work. While many application opportunities emerge from this platform, this chapter focuses on developing the core capabilities and understanding the system's fundamental limits.

In this chapter, we make the following three contributions:

- A novel method of using passive elements for directional sensing, enabling a low-power, low-complexity, and miniaturized system for acoustic localization. The sensing and signal processing techniques ensure robust DoA estimation with a single in-lab calibration. The

system achieves a median DoA error of $3.6°$, comparable to microphone array-based solutions while significantly reducing power and size requirements.

- A replicable process for designing and 3D-printing optimal acoustic structures that encode incoming sounds with directional cues, presenting a new approach for shaping sound fields using controlled diffraction in compact metamaterial structures.

- A complete hardware and software prototype of the *Owlet* system, made available for the community to reproduce, evaluate, and extend.

In the following sections, we detail the core intuition, system design, and key findings of this work.

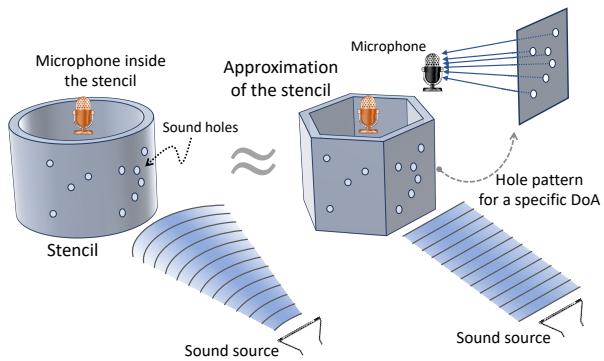


Figure 2.2: The concept of using a stencil with direction-specific hole patterns and microstructures for passive filtering of the incoming sound. The stencil embeds a directional response to the recorded signals.

2.2 Core Intuitions and Primers

The core idea of our system is to engineer a controlled environment around the microphone such that the recorded signal carries a unique, direction-specific channel impulse response. This impulse response, extracted from the microphone recording, serves as a signature of the sound's

angle of arrival. While larger objects or room acoustics naturally introduce diverse multipath effects that embed directional cues, our goal is to achieve finer-grained diversity within a compact form factor by combining principles of diffraction, interference, and structural resonance. Toward this end, we design a porous cap for the microphone, referred to as a stencil. The stencil features patterned holes on different sides, as illustrated in Figure 2.2. Incoming sound waves pass through these holes, and depending on the angle of incidence, different hole patterns influence the wave before it reaches the microphone. The holes are coupled with internal microstructures of varying parameters, imparting distinct frequency responses to the sound.

The stencil acts as a metamaterial, where the internal microstructures modulate the incoming sound and imprint a unique directional signature. Because the response of these microstructures is frequency-dependent, the directional signature is represented as a vector of complex frequency gains, G_θ. This concept is depicted in Figure 2.3.

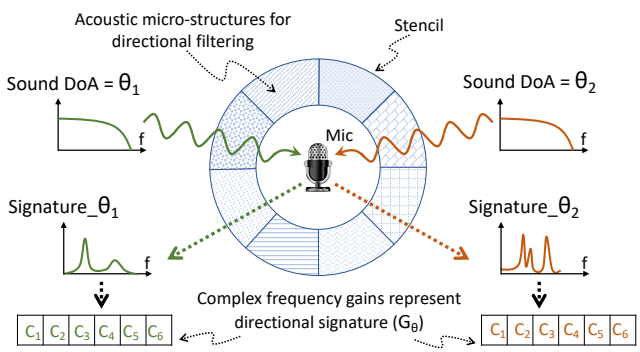


Figure 2.3: The concept of passive directional filtering using a stencil of acoustic microstructure. The stencil embeds a directional signature to the recorded sound unique to its direction of arrival (DoA). The spectrum of complex gains represents the signature for further computation.

2.2.1 Metamaterials for Passive Filtering

When sound interacts with physical structures, its frequencies are either amplified or attenuated. At larger scales, multipath reflections create such frequency variations through constructive and destructive interference. While reflections can embed directional signatures into sounds, they typically require structures comparable in size to the acoustic wavelength. Since *Owlet* targets low-frequency audible signals with large wavelengths, traditional reflection-based approaches would demand structures on the order of half a meter—prohibitively large for our goals. To achieve passive filtering within a compact design, we leverage the concept of acoustic metamaterials. Metamaterials are artificially structured materials composed of subwavelength elements that endow the material with novel properties. In constructing the metamaterial stencil, we utilize three key principles: (a) diffraction, (b) capillary effects, and (c) structural resonance.

(a) Diffraction: When waves encounter the edge of an obstacle, they bend around it—a phenomenon known as diffraction. This behavior is particularly pronounced when sound passes through an aperture smaller than its wavelength [54]. In such cases, the hole behaves as a virtual point source. When a receiver is placed behind a barrier with multiple small apertures, it observes a multipath-like environment formed by multiple virtual sources. The interaction of signals from these sources creates patterns of constructive and destructive interference, influenced by both the receiver's position and the sound frequency. We exploit this property by designing diverse hole patterns on the stencil, enabling a rich multipath environment around the microphone within a small form factor.

(b) Capillary Effect: As sound propagates through narrow capillary tubes, its acoustic impedance varies significantly [55]. The dimensions of these tubes, particularly their length

and cross-sectional area, influence the speed and phase of the transmitted sound. By integrating capillary tubes of different geometries into the stencil, we introduce controlled phase shifts between sound paths. This enhances frequency diversity, even when the physical separations between the holes are small.

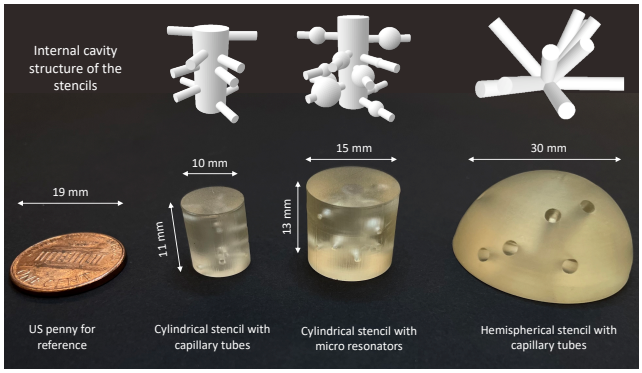


Figure 2.4: Different types of metamaterial stencils used in our experiments.

(c) Structural Resonance: Certain sound frequencies are amplified when oscillating air pressures interact with cavities along their path [56]. This phenomenon, known as Helmholtz resonance, is commonly observed in whistling bottles. We design millimeter-scale Helmholtz resonators embedded within the stencil, connected to the sound holes. By varying the shapes and dimensions of these resonators, we generate arbitrary resonance effects across different frequencies.

Figure 2.4 shows examples of 3D-printed stencils with embedded microstructures for directional filtering. Figure 2.5 illustrates the improvement in angular diversity provided by the stencil, comparing the amplitude variation of a 7 kHz tone with and without the stencil. Finally, Figure 2.6 presents the corresponding diversity in direction-specific frequency responses across different stencil designs.

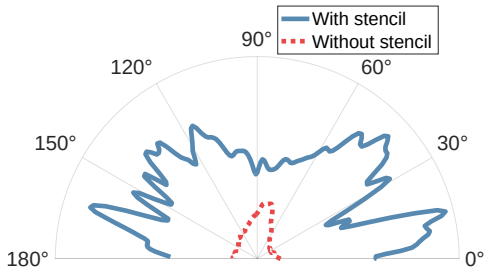


Figure 2.5: Angular diversity of the microphone with and without the microstructure stencil.

2.3 System Design

The system design focuses on two key objectives: (a) creating an optimal stencil structure that maximizes angular diversity, and (b) developing signal processing techniques to estimate the direction of arrival (DoA) from the recorded signal. Naturally, the accuracy of the system depends directly on the diversity introduced by the stencil. Our algorithms optimize the stencil design by simulating wave propagation around small structures and subsequently fabricating the optimized stencil using 3D printing. Before delving into stencil design details, we first describe the signal processing and DoA estimation techniques, providing an overview of the complete system.

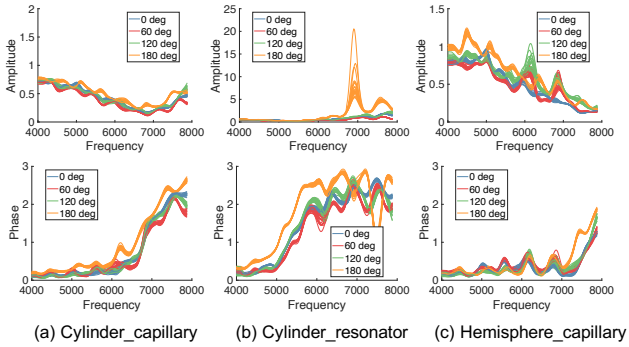


Figure 2.6: Comparison of the diversity in frequency responses (amplitude and phase) of the three types of metamaterial stencils.

2.3.1 Processing for DoA Estimation

At a high level, *Owlet*'s DoA estimation operates in two stages. First, during a one-time in-lab calibration, we generate a table of direction-specific signatures G_θ by sending known wideband signals from various directions. These signals are recorded by a microphone equipped with the stencil cap, capturing the direction-dependent modifications imposed by the structure. This calibration process is analogous to the procedures used for commercial microphone arrays.

The second stage occurs at run-time, where the system processes incoming sounds to extract the stencil-induced signature, $h_{stencil}$, and matches it against the pre-collected signature table to infer the DoA. In practice, we train a deep learning model on variations of the signature table, allowing the system to predict the DoA directly from pre-processed signals during real-world operation.

A critical aspect of this processing is accurately extracting $h_{stencil}$ from real-world signals, which must overcome two challenges: (i) separating the stencil's signature from the unknown source signal, and (ii) mitigating distortions caused by environmental multipath. We first describe our method for eliminating dependency on the source signal, followed by techniques to suppress environmental effects.

2.3.2 Eliminating Source Signal Dependency

The signal recorded by the microphone inside the stencil is the source signal modified by the stencil's directional response. Assuming no environmental effects, if $X(\omega)$ denotes the source signal and $Y_{in}(\omega)$ the recorded signal, their relationship in the frequency domain is:

$$Y_{in}(\omega) = X(\omega)H_{stencil} \tag{2.1}$$

When the source signal $X(\omega)$ is known, the stencil response $H_{stencil}$ can be directly obtained by dividing the recorded signal by the source signal, i.e., $H_{stencil} = \frac{Y_{in}(\omega)}{X(\omega)}$. In certain applications, such as navigation where a robot localizes itself using a known control signal, this assumption holds.

However, in many other applications, including ambient sound localization or user direction detection from speech, the source signal is unknown. To address this, *Owlet* introduces a secondary microphone placed outside the stencil. This reference microphone records the incoming sound without the stencil's influence, providing an independent observation of the source signal. Unlike traditional microphone arrays, the secondary microphone can be placed arbitrarily close to the primary microphone without affecting system operation. Figure 2.7 illustrates the physical setup and realistic signal model of the system.

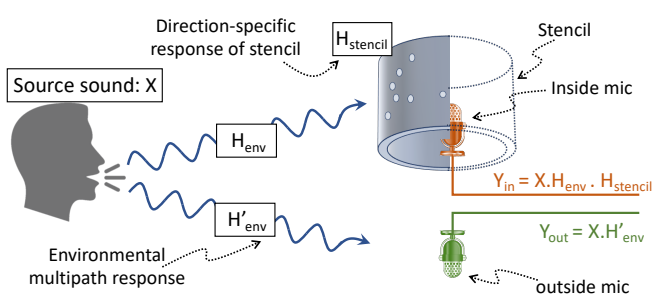


Figure 2.7: The two-microphone model for eliminating source and environmental dependency.

Consider the channel frequency responses from the source to the inside and outside microphones as H_{env} and H'_{env}, respectively. These responses capture the effects of multipath propagation, including reflections from nearby objects. The presence of the stencil around

the internal microphone introduces an additional modulation represented by the frequency response $H_{stencil}$. Assuming linearity of the system, the signal recorded by the inside microphone experiences both environmental and stencil-induced transformations, as illustrated in Figure 2.7. Thus, the signals recorded by the inside and outside microphones, $Y_{in}(\omega)$ and $Y_{out}(\omega)$, can be expressed as:

$$\begin{aligned} Y_{in}(\omega) &= X(\omega)H_{env}H_{stencil} + N(\omega) \\ Y_{out}(\omega) &= X(\omega)H'_{env} + N'(\omega) \end{aligned} \tag{2.2}$$

where $X(\omega)$ denotes the source signal, and $N(\omega)$ and $N'(\omega)$ represent independent noise at the two microphones.

Dividing $Y_{in}(\omega)$ by $Y_{out}(\omega)$ cancels the dependency on the unknown source signal but retains a residual environmental dependency:

$$\frac{Y_{in}(\omega)}{Y_{out}(\omega)} = H_{stencil}\frac{H_{env}}{H'_{env}} + N''(\omega), \tag{2.3}$$

where $N''(\omega) \ll H_{stencil}\frac{H_{env}}{H'_{env}}$.

This residual term, $\frac{H_{env}}{H'_{env}}$, implies that without further correction, the system remains sensitive to environmental variations. Thus, the stencil calibration or the training of the deep learning model would need to be performed for every new environment to ensure accurate angle prediction. Such a setup may be feasible in controlled scenarios where the positions of sound sources and sensing modules are fixed, such as object tracking on a conveyor belt. However, in most practical applications, the source location is unknown and variable, making exhaustive environment-specific training impractical. To overcome this limitation, we introduce a technique to eliminate

environmental dependency, enabling *Owlet* to operate reliably with a single in-lab calibration.

2.3.3 Eliminating Environmental Dependency

Our approach to removing environmental dependency is based on the observation that, although H_{env} and H'_{env} can vary unpredictably across environments, their ratio $H_{ratio} = \frac{H_{env}}{H'_{env}}$ remains bounded when the microphones are placed closely together.

This can be intuitively understood by considering the origin of environmental diversity. After emission from the source, sound waves reflect off various surfaces, creating multiple paths that arrive at the microphones with different delays. These differences in path lengths cause variations in the observed environmental response. When microphones are widely separated, these path differences are significant, leading to substantial variability in H_{env} and H'_{env}. However, when the microphones are placed within a few centimeters of each other, the differences in reflection paths become small and bounded. At the limit, if the microphones are exactly co-located, they observe identical environmental responses.

Therefore, the ratio $H_{ratio} = \frac{H_{env}}{H'_{env}}$ has a narrow distribution across frequencies when the microphones are closely spaced. We validate this property through both simulated ray tracing and real-world experiments, demonstrating that environmental variations can be effectively suppressed, enabling robust DoA estimation without location-specific retraining.

Once the distribution of H_{ratio} is known and $H_{stencil}$ is collected through the calibration process, we generate synthetic training data by sampling from the distribution of H_{ratio} and combining it with $H_{stencil}$. This synthetic data enables robust training of the deep learning model for angle prediction without requiring real-world sound traces. Moreover, if the dimensions of

the target environment and the positions of major reflectors are known, the synthetic training data can be tailored accordingly. Such customization accelerates model convergence and improves prediction accuracy.

At run-time, the system extracts $H_{ratio}H_{stencil}$ from the two recorded channels, $Y_{in}(\omega)$ and $Y_{out}(\omega)$. To enhance this extraction, we employ a Recursive Least Squares (RLS) adaptive filter [57] operating in system identification mode. The adaptive filter exploits the uncorrelated Gaussian noise between the two channels to estimate $H_{ratio}H_{stencil}$ by minimizing the following error term via gradient descent:

$$e(\omega) = Y_{in}(\omega) - Y_{out}(\omega)\frac{H_{env}H_{stencil}}{H'_{env}} \tag{2.4}$$

2.3.4 Synthetic Training for Deep Learning

We use the synthetic channel responses generated above to introduce environmental diversity into the training set for the neural network model. Specifically, we simulate different room environments and various source-microphone placements to create a wide range of $H_{ratio}H_{stencil}$ samples. Each sample is represented as a vector of 400 equally spaced frequency components between 0 and 8 kHz. Rather than using raw complex values, we separately extract the amplitude and phase spectra for training.

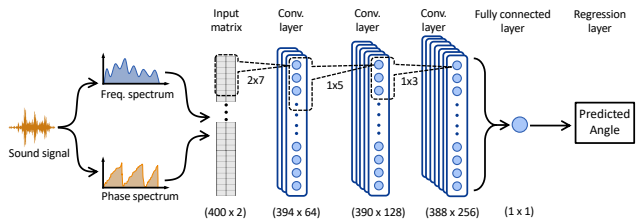


Figure 2.8: The architecture of the proposed CNN model.

We use the synthetic channel responses generated above to introduce environmental diversity into the training set for the neural network model. Specifically, we simulate different room environments and various source-microphone placements to create a wide range of $H_{ratio}H_{stencil}$ samples. Each sample is represented as a vector of 400 equally spaced frequency components between 0 and 8 kHz. Rather than using raw complex values, we separately extract the amplitude and phase spectra for training.

We adopt a Convolutional Neural Network (CNN)-based regression model for DoA estimation. CNNs are well-suited for environmental sound processing [58] and offer low-latency performance due to their compact parameter sets. Our model consists of a one-dimensional CNN with three convolutional layers followed by a fully connected layer and an output regression layer. The convolutional layers have 64, 128, and 256 filters with kernel sizes of 2×7, 1×5, and 1×3, respectively. The regression layer minimizes the half-mean-squared-error loss for angular prediction. We customize the loss function according to the target range and resolution of the directional angles. ReLU activation functions and batch normalization layers are used between convolutional layers to accelerate training, and stochastic gradient descent serves as the optimizer. The model is trained for 100 epochs with a learning rate of 1×10^{-6}. The block diagram of the CNN architecture is shown in Figure 2.8.

In addition to the regression model, we also develop a CNN-based classification model for evaluation and comparison, as described in Section 2.5.8. This model largely mirrors the regression architecture but replaces the output with a fully connected layer of size 360 (one node per degree) followed by a Softmax and classification layer.

2.3.5 Optimizing 3D Stencil Design

The performance of our system depends critically on the diversity of the frequency gain patterns across different angles. Our initial feasibility studies using randomly distributed holes on the stencil demonstrated a median DoA error of $7°$. However, the angular resolution was non-uniform, with some directions exhibiting significantly worse accuracy. This issue arises from suboptimal distributions of holes and internal microstructures, leading to similar gain patterns for different angles.

To address this, we systematically optimize the 3D stencil design to guarantee a minimum DoA detection resolution across all directions.

Ideally, the stencil should maximize the diversity of frequency gain patterns associated with each possible DoA. This design problem parallels the information-theoretic challenge of constructing maximally diverse code sequences. Each frequency gain pattern, G_θ, associated with an angle θ, can be thought of as a codeword. Our goal is to design a set of N codewords that are maximally distant from each other (i.e., maximizing the Euclidean distance between all pairs of codewords). The number of codewords, N, determines the angular resolution, $\Delta\theta = \frac{2\pi}{N}$.

Initially, we attempt to design ideal codewords of discrete frequencies and use them as guidelines for constructing desired gain patterns G_θ. These patterns are then mapped to physical arrangements of pinholes on the stencil surface at corresponding angles.

Given the number of holes N, the distances from the microphone to the holes r_n, the distance from the microphone to the stencil surface D, and the acoustic wavelength λ, the superposition of waves at the microphone can be modeled as:

$$u(\lambda) = \sum_{n=1}^{N} \frac{D}{j\lambda r_n^2} e^{\frac{j2\pi r_n}{\lambda}} \tag{2.5}$$

This equation defines the resulting acoustic field at the microphone as a sum of contributions from all the holes. For multiple wavelengths, the system becomes overdetermined, allowing approximate solutions for optimal hole placement. However, this analytical approach becomes computationally intractable even for moderate numbers of holes (greater than 10) and fails to capture the actual behavior of the wave field near small objects.

Behavior of Wave Fields Near the Stencil: In our initial model, we approximate the stencil as an N-gonal prism, assuming that only the holes directly facing the incident wavefront contribute to the recorded signal (Figure 2.2). This assumption holds when the object diameter exceeds ten times the wavelength [59]. However, for our miniaturized stencil, whose dimensions are comparable to the wavelength, this approximation fails. Waves diffract not only through the holes but also around the outer surface of the stencil, wrapping around and affecting the entire structure, as illustrated in Figure 2.9. This diffraction behavior must be carefully accounted for in the design process, motivating our transition to simulation-based optimization methods for stencil design.

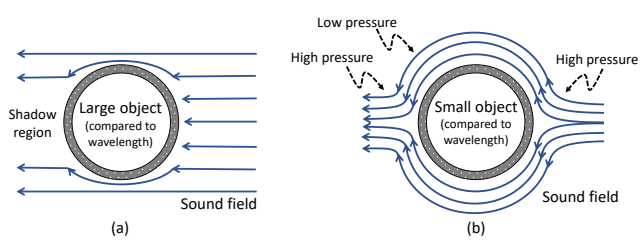


Figure 2.9: The behavior of sound field at the outer surface of an obstacle. (a) When the object's size is much larger compared to the wavelength of the sound, the obstacle creates a shadow region. (b) When the object's size is comparable to the wavelength of the sound, the wave diffracts around the object creating high-pressure at a larger region of the surface. It also creates a high-pressure region directly opposite to the sound's directions where sound fields from the top and bottom sides meet.

We verified the phenomenon of sound diffraction around small structures using a cylindrical shape with a pinhole on one side, as shown in Figure 2.10(a). We measured the sound pressure at a microphone placed at the center of the cylinder. Figure 2.10(b) shows that the sound pressure remains significantly high even when the hole is positioned more than 90° off-axis from the sound source. This confirms the bending of sound waves around the cylinder. Interestingly, we observe elevated sound pressure diametrically opposite to the incoming direction, resulting from the merging of diffracted fields from both sides of the cylinder. The angular intensity variation at the stencil surface depends on the sound frequency and strongly influences the received frequency gain pattern.

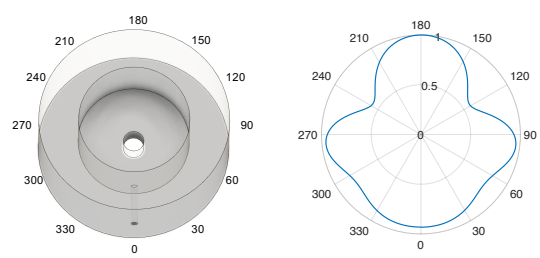


Figure 2.10: (a) A one-hole stencil to measure surface pressure levels. (b) Sound amplitude at different angles from the sound's direction of arrival.

Given these observations, we adopt a simulation-driven, forward modeling approach for optimizing the stencil design. Instead of inferring a hole pattern from a desired frequency gain profile, we generate a large number of random hole patterns and evaluate them through Monte Carlo simulations. Each stencil configuration is simulated to produce a frequency gain pattern across all possible directions (360 source locations at 1° separation). The optimal pattern is selected by maximizing the diversity of gain patterns across directions. This randomized search ensures that the final stencil design is close to globally optimal, given the cap size and other design constraints.

The optimization process proceeds as follows:

(1) Random Stencil Pattern Generation:

The diversity of directional gain patterns is directly related to the multipath diversity created by the arrangement of pinholes on the stencil. We fix the outer and inner diameters and the height of the cylindrical cap for the simulations. The pinholes are assigned a diameter of 2 mm. Random patterns of pinhole locations are generated on the side surface of the cylinder. However, simple uniform random sampling does not guarantee a minimum separation between holes, which is necessary to ensure that each hole contributes independently to the gain pattern. To enforce a minimum separation—at least half of the maximum wavelength—we modify the *Fast Poisson Disc* sampling algorithm [60]. At each iteration, the Poisson Disc method generates a two-dimensional location for a pinhole on the flattened surface of the cylinder, starting from existing seed points. The next hole is placed randomly within a circular annulus of radius 3 mm, ensuring minimum separation. The width of the annulus is varied randomly at each iteration to introduce additional diversity into the generated hole patterns.

(2) Estimating Frequency Gain Patterns:

For each randomly generated stencil, the algorithm computes the corresponding frequency-dependent gain pattern using Equation 2.5. The path differences between the pinholes and the microphone are calculated, accounting for diffraction effects around the cylinder's outer surface. We consider 400 equally spaced frequency points between 0 and 8 kHz, resulting in a 400-point complex gain pattern for each of the 360 possible source angles. Amplitude and phase corrections due to diffraction, as characterized earlier (Figure 2.10), are applied to refine the model. At the end of this step, we obtain 360 frequency gain patterns, each represented by a 400-point complex vector, which are then used for evaluating the stencil's performance.

(3) Assessing the Diversity of Gain Patterns:

Next, we assess the diversity of the gain patterns by computing the all-pair Euclidean distances, referred to as the *chord-distance*. Two highly distinguishable gain patterns yield larger chord-distances compared to similar patterns. This metric is used for the maximin decision criterion in the final stencil selection step.

(4) Stopping Criteria and Stencil Selection:

At each iteration, the algorithm records the minimum chord-distance among all pairs of gain patterns. The iterations terminate once the distribution of the minimum distances fits a Gaussian curve, indicating convergence. The stencil design corresponding to the maximum minimum chord-distance is then selected for fabrication. Figure 2.11 compares the gain pattern diversity between an optimal and a suboptimal stencil.

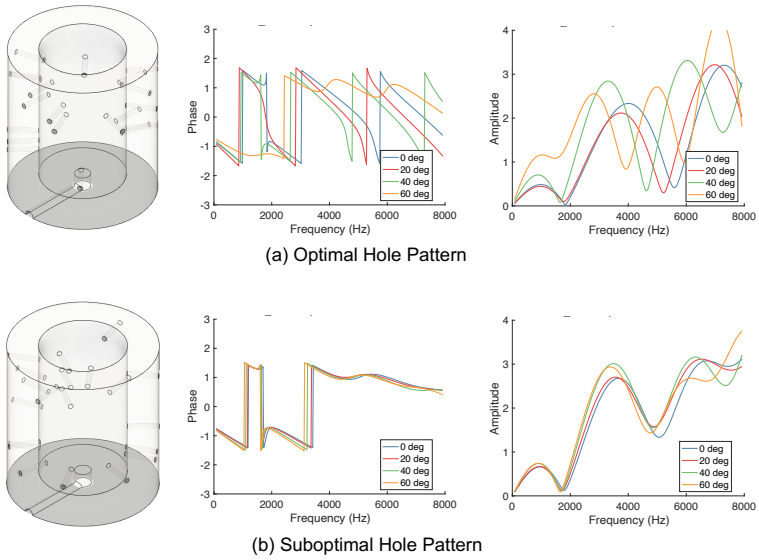


Figure 2.11: Comparison of diversity in phase and amplitude patterns for an optimal and a sub-optimal design of the stencil.

2.4 Prototype Development

2.4.1 3D-Printing Stencil Caps

We first implement the stencil design optimization algorithm in MATLAB. The resulting design parameters are fed into a Python script using the Autodesk Fusion 360 API [61, 62], which constructs a 3D model of the stencil, including the internal substructures and external pinholes. The model is exported as an STL file and sliced for 3D printing.

We use the Elegoo Mars photocuring 3D printer [63] with an ultraviolet (UV) light-curable resin of density $1.195\ \mathrm{g/cm^3}$, solidifying under 405 nm wavelength light. Compared to jetting-based printing, photocuring offers higher resolution and smoother surface finish, which is crucial for fabricating fine structural features on the stencil. Moreover, photocuring produces dense surfaces, resulting in predictable acoustic properties [64].

2.4.2 Calibration and Data Collection

We generate a wideband calibration signal in MATLAB and export it as an '.arb' file, which is loaded onto a Keysight waveform generator [65]. The signal is transmitted using two off-the-shelf speakers powered by a 40 W dual-channel amplifier. We synchronize the transmission using an external wired trigger.

The stencil is mounted on a stepper motor controlled by an Arduino [66], allowing precise rotation at 1° increments from 0° to 360°. Recordings are made using omnidirectional ADMP401 MEMS microphones [67], sampled at 16 kHz. The collected data is processed offline on a computer for further analysis.

2.5 Evaluation

We evaluate the performance of our microstructure-guided spatial sensing system by implementing a complete *Owlet* prototype and conducting experiments in various indoor and outdoor environments under different acoustic conditions. Traditional uniform linear microphone arrays (ULA) of different sizes are used as baselines to compare localization performance and energy consumption.

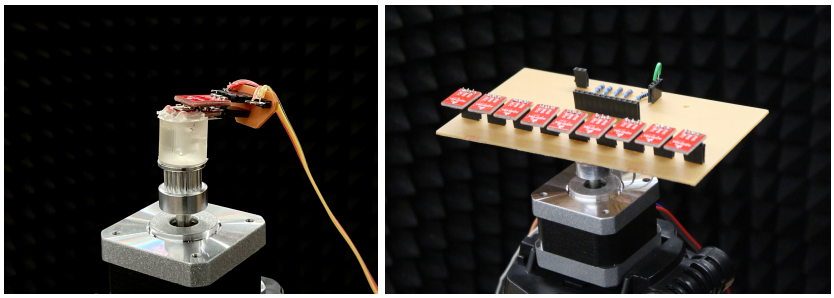


Figure 2.12: The *Owlet* prototype used in the evaluation experiment (left) and a 9-element uniform linear microphone array used as baseline for comparison(right). The array is $12cm$ wide, where *Owlet* is significantly smaller measuring less than $2cm$ in its largest dimension.

2.5.1 Evaluation Setup and Results Summary

The *Owlet* prototype consists of a 3D-printed stencil and two microphones mounted vertically, facing in opposite directions. The microphones are separated by 4 mm, with one microphone's opening covered by the stencil. For baseline comparison, we also develop a 9-element uniform linear microphone array (ULA) with 1.3 cm spacing between elements. Each microphone array element is sampled simultaneously using a multi-channel DAQ system [2]. Figure 2.12 shows the sensor front-ends for both *Owlet* and the ULA.

Both *Owlet* and the ULA use omnidirectional ADMP401 MEMS microphones [67], sampled

at 16 kHz. The recorded data is processed offline using MATLAB scripts.

The sound sources used include multi-frequency wideband signals, white noise, drone sounds, and car engine noises. Unless otherwise noted, the experiments are conducted with a multi-frequency wideband source, background noise at 40 dB SPL, a 3 ft distance between the source and the microphones, and an elevation angle of $0°$. The stencil used measures 1.5 cm $\times$ 1.3 cm and incorporates internal capillary tubes and structural resonator cavities.

Experiments are performed in several representative environments, including an indoor laboratory, a building lobby, and outdoor spaces, as illustrated in Figure 2.13. Importantly, the recorded source signal remains unaffected during structure-guided DoA estimation, as the secondary microphone is placed outside the stencil and records the sound source without modification.

Figure 2.13: Various locations for system evaluations: (a) indoor laboratory, (b) indoor lobby, (c) outdoor.

Summary: Figure 2.14 summarizes the overall performance of *Owlet* in comparison to the traditional ULA based DoA estimation technique. As elaborated later in this section, *Owlet* outperforms even a 9-element ULA running the standard multiple signal classification (MUSIC) algorithm for direction estimation, while consuming a fraction of the energy required by the array. Table 4.2 presents a comparison of estimated manufacturing costs of these prototypes in the lab, their sizes, the median DoA estimation errors, and the energy requirements.

System	Prototype cost	Size	Error	Energy
Owlet	$15	1.9cm	3.6°	16.7mJ
9-element array	$70	11.4cm	4°	2078mJ

Table 2.1: Comparison of prototype cost, size, median error, and energy consumption of Owlet with a microphone array.

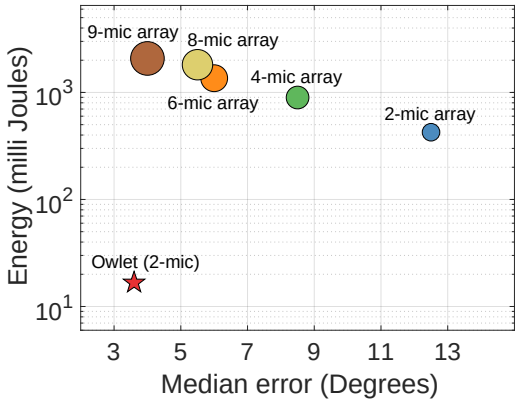


Figure 2.14: Overall performance of the *Owlet* system compared to the traditional microphone arrays of various sizes. *Owlet* requires $100\times$ less energy than the state-of-the-art array systems while achieving better accuracy than a 9-element array.

2.5.2 Impacts of external conditions

We evaluated the performance of our prototype under various adversarial conditions. We present the results below.

(a) Ambient noise: The regular noise level at the locations of experiments is around 40 dB of sound pressure level (SPL). We introduced four different kinds of noise with distinct spectral properties: (i) white noise, (ii) traffic sound, (iii) human conversation, and (iv) sound of machineries like jackhammer. We played these noise sounds from three speakers from different angles and at different levels of loudness near the receiver while performing the direction estimation. The loudness of the source sound used for DoA was 60 dB SPL, a loudness comparable to natural conversions. Figure 2.15(a) shows the median error of *Owlet*'s performance under these

experiments remains stable for a wide range of noise loudness.

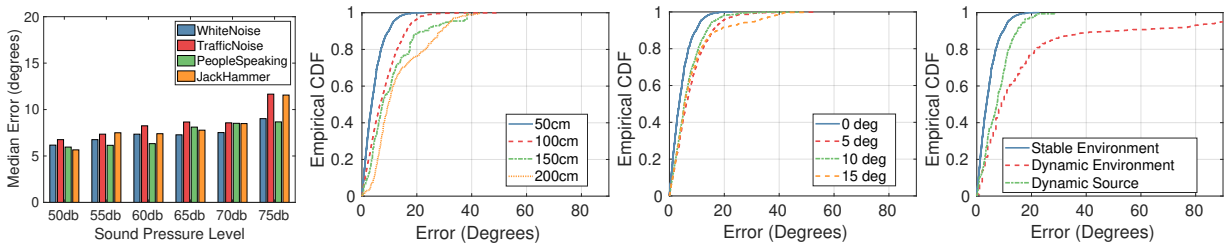


Figure 2.15: Performance under external conditions: (a) The impact of varying types and loudness levels of ambient noise on the median DoA estimation error. (b) The CDF of errors when the sound source is located at varying distances from the receiver. (c) The CDF plot of estimation error for different elevation angles or the vertical positions of the sound source. (d) The CDF plots of errors that show the impact of dynamic movements in the environment.

(b) Distance from the receiver: We measure the performance while the sound source is placed at various distances from the receiver. Figure 2.15(b) shows the median error for DoA estimation for these experiments. The error is mainly dominated by the change in signal-to-noise ratio at the receiver due to increasing distance. The intensity of the sound source was kept constant despite the location of the source. When we change the model to maintain a fixed signal loudness at the receiver, varying distance shows limited impact.

(c) Elevation angle: Current prototype of *Owlet* limits its direction estimation to the azimuth angles (directions on the horizontal plane) only. Ideally, azimuth-only DoA estimation system should not be affected by the elevation angle (i.e., vertical location) of the sound source. However, in practice microphones are not fully omni-directional and therefore regular microphone-array based DoA systems perform correctly for a certain limit on the elevation angle of the source. In addition to the microphone's limited response in the vertical plane, *Owlet* has pinhole patterns on the stencil that may be projected differently on the microphone. We evaluated the impacts of elevation angle by increasing the vertical distance of the sound source, while the horizontal distance is fixed at 150 cm. Results in Figure 2.15(c) show *Owlet*'s performance does not vary

significantly when the vertical location of the source is up to 15 cm from the center.

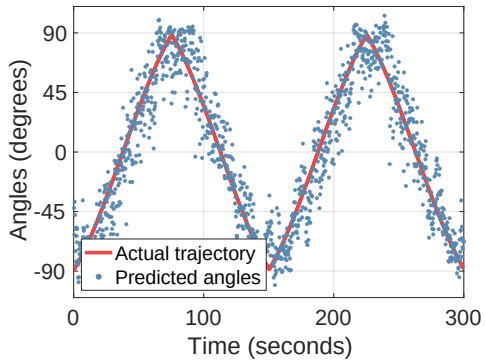


Figure 2.16: The performance of sound tracking while the source is constantly moving near the sensor. The movement of the source creates a dynamic multipaths scenario.

(d) Dynamic multipath: *Owlet* is designed to mitigate the effects of the environmental multipath. We have evaluated this feature by changing the test location in previous experiments. We further evaluate it by moving the sound source during the test to add additional disturbance to the environment. The median error remains within 7° for this test. Next, we introduce moving subjects near the setup so that the multipath environment changes during the experiment. The median error is close to 9° when three people keep walking within 3 meters from the sensor. Figure 2.15(d) shows the CDF of the error for these experiments along with the CDF for stable multipath environments. Figure 2.16 shows *Owlet*'s performance for tracking the sound source while the source is moving near the sensor.

2.5.3 Performance in different environments

We evaluated *Owlet* in several represented indoor environments, such as indoor laboratory, lobby, and open-air outdoor places as shown in Figure 2.13. To make our model robust in diverse environments, we train the deep learning model using synthetic room impulse responses

as mentioned in Section §2.3.4. Figure 2.17(a) shows the DoA estimation performance in multiple locations of these environments. The median error is within $4°$ and 90th percentile error of less than $10°$. This result shows *Owlet*'s ability to function in unknown environments with a one-time calibration during the development of the prototype.

2.5.4 Impact of different sound sources

In this experiment we evaluate the system's DoA estimation performance for parallel frequency signal and other types of signal sources. We evaluate the system's DoA estimation performance for different types of sounds. These signals are different in their active bandwidths, frequency spectrums, and loudness. Figure 2.17(b) shows comparable performance across different sounds.

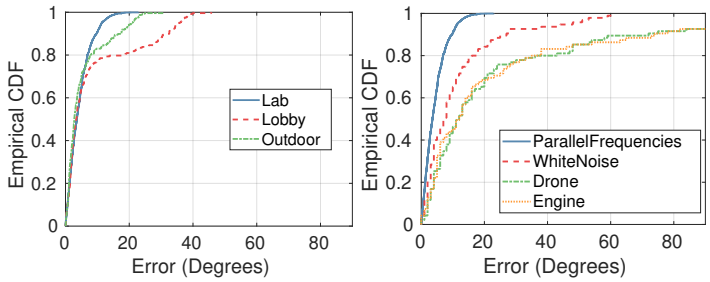


Figure 2.17: The CDF of median error for (a) different environments and (b) different types of sound sources.

2.5.5 Performance in known environment

Owlet's synthetic training data generation process can be customized if the geometry of the target location is known. We evaluate this feature by generating training data according to the test location. Figure 2.18 shows *Owlet*'s overall performance for estimating signals' direction of

arrival. In this experiment we transmitted signals from a speaker placing it at various angles with respect to the *Owlet* system. The ground truth DoAs of the signal covered the $0 - 180°$ angles in front of the receiver with $1°$ separation between locations. Note that, unlike microphone arrays, *Owlet* does not have any 'mirror location' (or front-back) ambiguity in DoA estimation by design. The confusion matrix in Figure 2.18(a) visually presents the spread of error for every ground truth angle. Figure 2.18(b) shows the empirical cumulative distribution of the error. *Owlet* exhibits a median error less than $3.3°$ and 90th percentile error of less than $10°$ in this scenario.

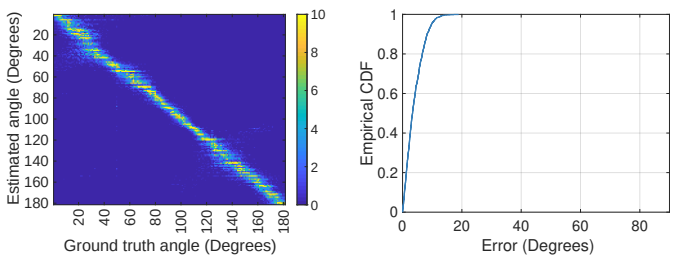


Figure 2.18: The performance for DoA estimation with known room size: (a) The confusion matrix and (b) the CDF of error in degrees of angle.

2.5.6 Localization Performance

Owlet primarily focuses on the DoA estimation of the sound. However, combining information from multiple such units can localize a sound source using triangulation. We created a setup for localization using two speakers continuously playing $50ms$ of parallel frequency pulses. We placed the *Owlet* receiver at various places within a grid in front of the speakers. *Owlet* system estimated DoA of both the speakers and estimated the location using triangulation method. Figure 2.19(a) shows the heatmap of the localization error and Figure 2.19(b) shows the corresponding CDF plot. The median localization error is 10 cm.

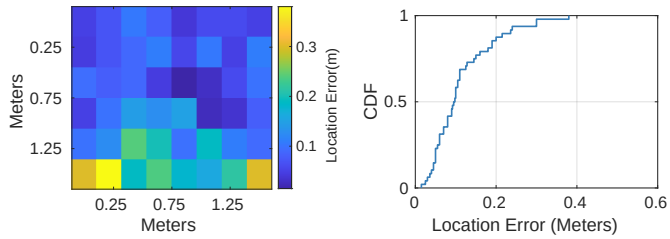


Figure 2.19: The localization error as (a) heatmap and (b) empirical CDF.

2.5.7 Comparison with traditional methods

Figure 2.20 compares performance of *Owlet* with traditional array-based DoA estimation techniques. We implement three popular and fundamentally different array-based methods – beamscan, minimum variance distortionless response (MVDR), and MUSIC algorithm. We apply these techniques on the microphone arrays having different number of elements. Results in Figure 2.20(a) show that *Owlet* significantly outperforms the other algorithms under similar conditions and signal SNR, while using only two microphones. *Owlet*'s median error is even slightly better than the MUSIC algorithm with 9-microphone array. For an estimate of the DoA resolution, we compare the spatial spectrum of each of the traditional algorithms with *Owlet*. Given *Owlet* uses a regression-based method, it does not directly produce a spatial spectrum, we rather plot the distribution of confidence score for all angles. Figure 2.20(b) shows the spatial spectrums for the signal coming at $20°$. *Owlet* exhibits narrowest beamwidth comparable to MUSIC.

2.5.8 Comparison between learning models

In Figure 2.21, we compare the performance of different deep learning models and algorithms. The regression model performs slightly better than the classification algorithm in certain scenarios.

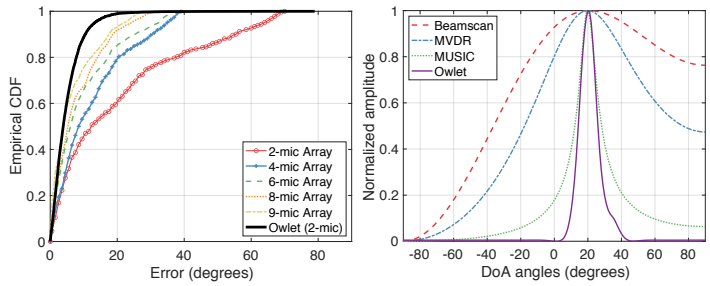


Figure 2.20: Performance comparison of *Owlet* with the implementation of beamscan, MVDR, and MUSIC algorithms: (a) The CDF of median errors, (b) The spatial spectrum for an incoming signal from 20° angle.

We also compare different architectures of the regression model. Instead of using 64, 128, and 256 filters for three convolution layers, we first reduce the filter sizes to half and got a median error of $5.6°$. We also apply only the first two convolution layers and reduced size filters, which leads to a median error of $5.8°$. When we use two convolution layers with filters 64 and 128, the median error is $7.8°$. These results show the opportunity to customize the model for resource-constrained computational environment while maintaining the target DoA performance.

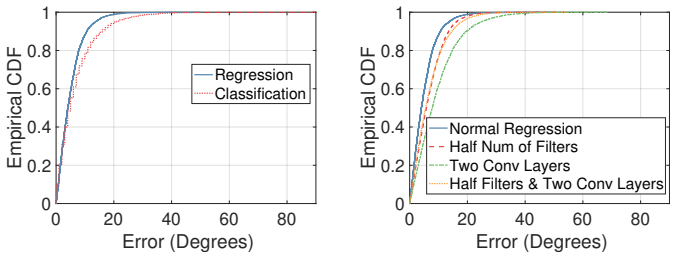


Figure 2.21: Performance comparison of *Owlet* with different deep learning models and architectures.

2.5.9 Energy consumption

In this section, we evaluate and compare the energy efficiency of *Owlet* with the traditional array-based systems. We measure the power consumption of each submodule, including hardware frontend, analog to digital conversion (ADC), and DoA computation. While we monitor the

frontend and ADC directly, the setup for the computation part requires porting the runtime codes to a Raspberry Pi 4 module and monitoring the overall power variation of the module. We used a Keysight E6313A power supply and monitoring unit [68] for precise and high-resolution tracking. The setup is shown in Figure 2.22.

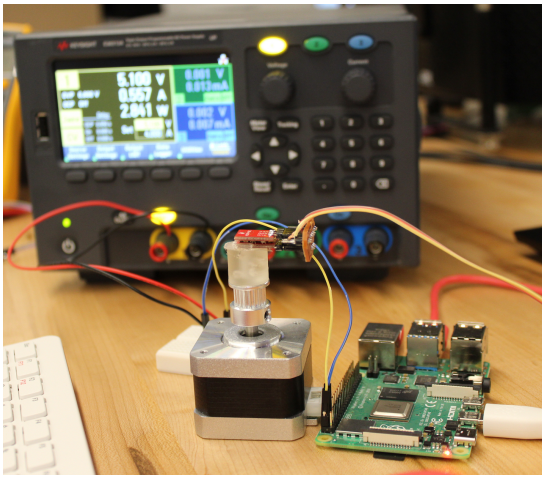


Figure 2.22: The setup for evaluating energy consumption. The setup tracks the energy requirements of *Owlet* and baseline microphone arrays under various conditions using a Keysight E6313A power supply and monitor.

Computation: We write the codes in Matlab and use Matlab Coder [69] to generate executable C files for the Raspberry Pi 4. We use Mathworks Raspbian image optimized for deep learning applications and cross compile the code for ARMv7 architecture with Neon Acceleration. This acceleration uses special registers for parallel operations which is an advantage for neural network systems over traditional techniques. We deploy the executable code and run it on offline data for 10,000 iterations. We collect the voltage and current readings from the power meter.

We also record the total time taken to complete the estimations. We learn that although the instantaneous power of *Owlet* is $1.92Watts$, which is approximately twice of traditional algorithms $1.05Watts$, but the time taken to complete the estimation is exponentially low, $8.3ms$ for *Owlet* compared to $2050ms$ for traditional algorithm. This contrast is because of the highly paralleled

operations of the neural network, which is not possible in sequential traditional algorithms.

ADC: Figure 2.23 reports the energy consumed by the ADC of an MSP430 ($3 microcontroller) [1] and a Keysight Data Acquisition System ($2500 DAQ) [2]. We use the low-power 12-bit ADC of the MSP430FR5969 and vary its sampling rate to emulate the multiplexing of multiple microphones. For the Keysight DAQ, we use the 12-bit parallel channel ADC in single-shot data acquisition mode and vary the number of channels. We record the power consumption for both the devices from the power supply and we do not connect the microphones to remove the effect of microphones and their amplifiers.

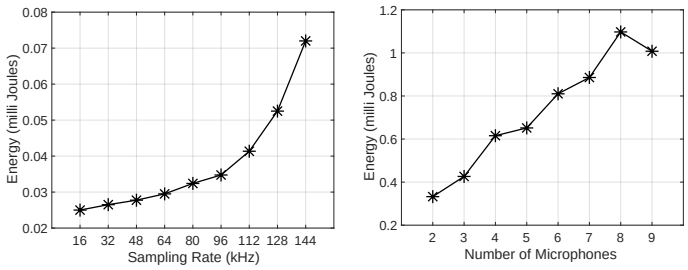


Figure 2.23: Energy consumption of (a) MSP430FR5969 low-power ADC [1] for different sampling rates and (b) Keysight Data Acquisition System [2] for different number of microphones.

Microphone frontend: We record the power consumption of the ADMP401 MEMS microphones [67]. To optimize the total energy consumption of the system, we only consider the time duration of $50ms$, a typical time duration to collect 800 samples at $16kHz$. We multiply this time with the average power consumed by the microphones to obtain the total energy consumed in Joules.

The evaluation summary in Section §2.5.1 presented a comparison of power consumption and accuracy for *Owlet* and traditional arrays. Figure 2.14 in Section §2.5.1 presents the energy consumption of *Owlet* and other state-of-the-art arrays and compares the median errors in DoA estimation. Here we show the energy consumption of each submodule, i.e., computation, ADC,

and microphone frontend, separately in Figure 2.24. *Owlet* consumes less than a 100th of the power required by the traditional arrays for similar angular resolution and accuracy.

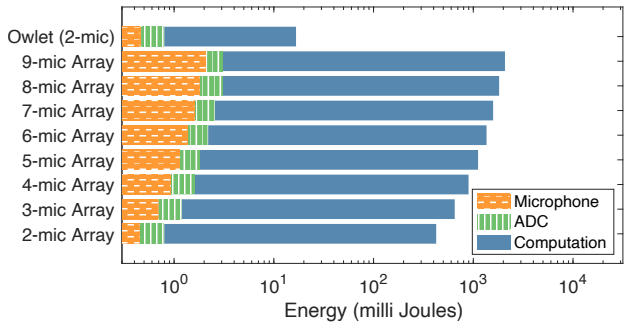


Figure 2.24: Overall energy consumption of array-based systems and *Owlet*.

2.6 Discussion

The current version of *Owlet* represents an exploratory realization of the concept, and there remain opportunities for further improvements. We discuss several aspects below.

• **Multiple Sound Sources:** We evaluated the prototype in diverse acoustic environments with various noise sources and target sounds. At present, *Owlet* assumes a single dominant sound source for DoA estimation. When multiple sources are present, the system attempts to localize the strongest signal and treats others as noise. While single-source scenarios are practical for many applications, extending *Owlet* to multiple simultaneous DoA estimations is feasible. A potential direction is the adaptation of statistical source separation techniques, followed by independent DoA detection. We leave this extension to future work.

• **Theoretical Bounds on Capacity:** Array signal processing has been extensively studied, yielding well-defined theoretical bounds on achievable spatial resolution under various constraints. Such analyses are critical for array design and performance simulation. Although *Owlet* differs

fundamentally from conventional arrays, a similar theoretical treatment is possible by analyzing the entropy of the available directional gain patterns. The stencil's shape, size, and operational frequency impose additional constraints on *Owlet*'s capacity to encode directional information. Developing a formal information-theoretic analysis would enhance understanding of the system's fundamental limits and guide future design improvements.

• **Mobility:** Doppler frequency shifts arising from fast motion may alter the frequency gain patterns used by *Owlet* for DoA estimation. Our prototype operates in the low-frequency audible range, which is relatively resilient to mobility-induced distortions. Moreover, DoA estimation based on parallel frequency signals inherently provides robustness against moderate Doppler shifts. As a result, we did not focus on mobility effects in the current study. However, extending *Owlet* to operate at higher frequencies would necessitate explicit design considerations to detect and compensate for such shifts.

• **Inaudibility of Sound Signals:** The present work focuses on audible sound frequencies for calibration and source signals. While lower frequencies offer robustness, their longer wavelengths limit achievable angular resolution due to reduced diversity in frequency gain patterns. We deliberately selected this operational range to demonstrate system performance under conservative conditions. Future iterations of *Owlet* will explore near-ultrasound (17–24 kHz) and ultrasound (> 24 kHz) frequencies to improve spatial resolution and further miniaturize the stencil design.

2.7 Related Work

There is extensive prior work on spatial analysis of sound. Classical research in direction-of-arrival estimation using microphone arrays [70–73], array signal processing for beamforming [74–

76], and subspace-based super-resolution algorithms [77, 78] has advanced this field significantly. Recent developments in ubiquitous spatial acoustic sensing [39,79–87] have opened new opportunities. Below, we briefly discuss two topics closely related to *Owlet*.

- **Acoustic Structures:**

The impact of physical structures on sound fields has a long history, from ancient architectures that amplified sound to modern designs for reverberation control and noise isolation. Relevant to *Owlet*, recent studies have explored 3D-printed acoustic metamaterials for frequency-selective absorption [88] and metasurfaces for diffraction-limited acoustic field generation [89]. The use of acoustic structures for sensing is relatively new. Li et al. [90] designed 3D-printed acoustic filters for impedance control. Other works have created acoustic barcodes using surface notches [91], and tangible user interfaces using small 3D-printed acoustic structures [92, 93]. These systems vary structure shapes to create distinguishable frequency responses, using smartphones for signal capture and classification.

- **Monaural DoA Estimation:**

Conventional DoA estimation methods primarily rely on microphone arrays [70–72, 80]. Recently, resource-efficient approaches for monaural DoA estimation have emerged. For example, [94] uses reflections in a known room environment with a single microphone for localization. Other works place small structures near a microphone to induce direction-dependent frequency responses [46, 47, 95]. While these systems achieve monaural localization, they often rely on large structures unsuitable for mobile applications. Similarly, [96] proposed learning spectral features for speaker distance estimation, but only in controlled environments.

2.8 Conclusion

This chapter presents *Owlet*, a practical system for low-power acoustic direction-of-arrival estimation and source localization using structured acoustic modulation. The core idea is to engineer the directional impulse response of a microphone by covering it with a carefully designed stencil. Sound waves diffracted by the stencil carry direction-specific signatures that encode spatial information. *Owlet* introduces robust DoA estimation techniques based on this principle and demonstrates them through a working prototype [97]. The prototype achieves angular resolution comparable to that of large aperture microphone arrays while using only two closely spaced microphones within a compact form factor. This work opens new applications in ubiquitous acoustic sensing and offers novel possibilities for sensor design through 3D-printed passive structures. This chapter represents a first step toward realizing this broader vision.

Chapter 3: Microstructure-Assisted Vision for Ubiquitous Tiny Robots

3.1 Introduction

Miniature mobile robots are rapidly advancing in capabilities and functionalities. Insect-sized robots, often only a few inches in size, are being developed for diverse applications such as search and rescue in disaster environments [98], agricultural foraging and mining [99, 100], and cooperative object transportation [101]. When equipped with autonomous navigation, these small robotic systems could enable breakthroughs in precision farming, disaster management, surveillance, and environmental monitoring. However, realizing this vision presents significant challenges, particularly in developing effective environmental perception methods suitable for the robots' stringent size, weight, power, and cost constraints (SWaP-C) [102]. In this chapter, we present *SPiDR*[1], an ultra-low-power acoustic spatial sensing system capable of generating accurate depth maps of nearby objects. An overview of the system is shown in Figure 3.1.

Traditional scene perception systems typically generate depth maps by scanning the surroundings [103]. Lidar systems [104], for example, mechanically sweep a laser beam to measure distances across the 3D scene. Radar and sonar systems [105, 106], in contrast, use phased arrays to electronically steer beams without mechanical motion, a technique known as beamforming. Mechanical scanning is power-inefficient, but beamforming demands large numbers of transducers to achieve fine

[1] SPiDR stands for Structure-assisted Perception, Detection, and Ranging

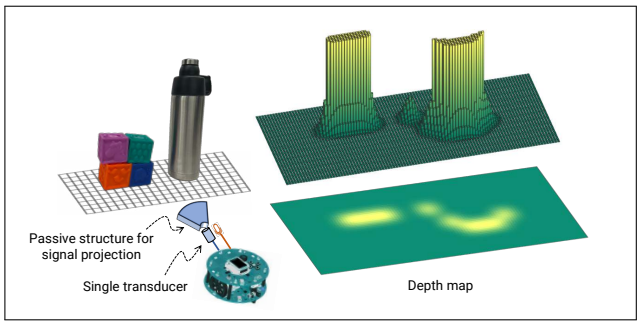


Figure 3.1: *SPiDR*, an ultra-low-power acoustic spatial sensing system for mobile robots. The system uses a carefully designed 3D-printed micro-structure for projecting spatially coded signals for imaging.

angular resolution, resulting in substantial hardware and computational complexity. Micro-robotic systems face extreme power limitations, with only 10–30 mW typically available for sensing after allocating power for actuation [107–110]. Thus, conventional scanning and beamforming methods are impractical.

In developing an ultra-low-power sensing solution, we begin by minimizing the sensing hardware requirements. We use acoustic signals at 40 kHz, corresponding to an 8.5 mm wavelength, providing fine spatial resolution while allowing simple, low-cost oscillators to generate the signal. Active sensing, which requires both signal transmission and reception, generally needs at least one transmit and one receive channel. A single-channel transmitter/receiver design rules out electronic beam steering, and even low-power mechanical steering systems consume around 33 mW [110]. To overcome this, we introduce a spatially coded signal that projects unique patterns into different directions, as shown in Figure 3.2. Reflections of this coded signal carry both direction and distance information: the direction is encoded in the received pattern, and distance is determined via time-of-flight measurement. However, using an omnidirectional transmitter, which naturally emits the same signal in all directions, presents challenges in achieving spatial coding. We address this by controlling the properties of the multipath channel using a 3D-printed microstructure.

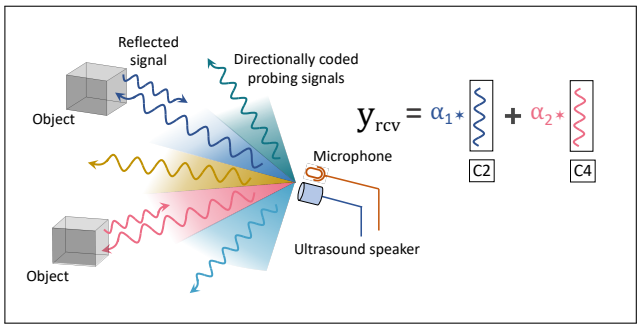


Figure 3.2: The concept of the spatially coded channel sounding method. The received signal is the weighted linear combination of the reflections that bear the direction-specific signature.

Multipath is a well-known phenomenon where signals reflect off various surfaces, traveling through different paths with distinct delays before combining at the receiver. These delays cause constructive and destructive interference, forming a unique superimposed signal at the receiver. Although environmental multipath is typically uncontrolled, we leverage the principle of multipath superposition to deliberately shape the transmitted signal.

Specifically, we design a 3D-printed cover for the speaker, termed a stencil, which splits the output into multiple paths via small internal tubes of different lengths. Each tube introduces a carefully controlled time delay before emitting the signal through distinct sound holes pointing in different directions. The resulting delayed replicas interfere predictably in space, effectively projecting direction-specific spatial codes. Conceptually, the scene is divided into uniformly spaced 3D points or voxels. If a voxel is occupied by an object, the reflected coded signal returns to the receiver carrying its unique signature. All coded reflections from the scene combine at the receiver and are processed to extract a 3D point cloud of the environment. For navigation applications, only a 2D cross-sectional view is typically needed, reducing the computational requirements.

Importantly, since the stencil is a passive structure, only the speaker and microphone in the

sensor frontend consume power, together requiring less than half a millijoule per measurement. However, low-power hardware alone does not suffice. Reconstructing the scene from the composite reflected signals typically involves a computationally expensive channel inversion. As described in Section 3.3.1, *SPiDR* adopts a fractional scene reconstruction and stitching algorithm to significantly reduce computation energy. The key idea is to generate a coarse probabilistic map of object locations, followed by iterative refinement using partial inversion of the channel matrix. This strategy exponentially reduces computational load compared to full inversion. Additionally, *SPiDR* can exploit the robot's movement to integrate multiple low-resolution maps into a high-resolution representation of the scene. With this optimization, the system can reconstruct a depth map covering a 20 cm $\times$ 10 cm area using only 0.83 mJ of energy.

Along with prior work in mobile sensing and localization, our approach builds on established theories of wireless channel estimation, multipath signal propagation, and computational sensing. The idea of performing spatial sensing through a single measurement has been explored in computational photography [111], where "coded apertures" are used to modulate spatial information onto a sensor via incoherent transformations. Coded aperture sensing is conceptually related to (and, in some sense, the inverse of) using a 3D-printed stencil to project spatially encoded probing signals. Recent advances in medical ultrasound and radio-frequency (RF) imaging have similarly utilized coded apertures to enhance imaging clarity.

The work most closely related to *SPiDR* is compressed ultrasound imaging [112], where a mechanically actuated coded aperture mask was used to image objects submerged in water. However, *SPiDR* fundamentally departs from previous approaches in four critical ways to meet the stringent practical constraints of micro-robotic platforms:

(a) No Scanning Required: Previous techniques in radar and spatial sensing with a single

sensor often rely on mechanical scanning to create a synthetic aperture [113] or incorporate moving elements [114] to introduce diversity. In contrast, *SPiDR* requires no mechanical movement, eliminating the need for power-hungry actuators.

(b) No Location-Specific Information Needed: Past approaches in echo analysis [115, 116] and location fingerprinting [117] typically involve learning signal patterns tied to specific locations. In *SPiDR*, calibration is performed once in an empty environment using an ideal point reflector, allowing the system to operate anywhere without further training. The stencil design itself is location-independent and generalizes across different environments.

(c) Power-Aware Hardware and Signaling: Although higher frequencies offer better depth resolution, *SPiDR* intentionally operates at a relatively low ultrasound frequency of 40 kHz. We use commercially available 40 kHz piezoelectric transducers, which are inexpensive and compatible with low-power oscillator circuits, supporting the system's overall energy efficiency.

(d) Power-Aware Computing: We exploit the sparsity of the target scene to minimize computational overhead. The cost of our iterative scene reconstruction algorithm scales with the fraction of the scene occupied by objects, rather than with the entire scene volume.

This work advances a broader vision of leveraging passive structures for extremely low-power sensing and perception. As the research evolves, we anticipate that the core techniques developed here will extend beyond robotic navigation to enable applications such as ubiquitous gesture recognition, long-term underwater monitoring, and soft-tissue imaging with handheld devices.

While a variety of future opportunities exist, this chapter focuses on establishing the foundational capabilities and evaluating the performance limits of the system. To this end, we make the following three specific contributions:

- A novel channel sounding method for spatial sensing, using a zero-power passive extension to a single sound source. We exploit the interaction of sound with microstructures to create a 3D-printed passive filter, termed a stencil, capable of projecting spatially coded signals with fine granularity.

- An ultra-low-power depth-map reconstruction algorithm, based on iterative refinement of a coarse initial estimate using a fractional multi-resolution reconstruction approach.

- A complete hardware and software prototype of the spatial perception system. We use this demonstrable implementation of *SPiDR* to benchmark power consumption and evaluate performance under real-world conditions.

3.2 Core Intuitions and Primers

Traditional methods for scene perception rely on scanning the environment with directional beams, either by steering the transmission, the reception, or both, one angle at a time. *SPiDR* adopts a fundamentally different approach based on *spatially coded channel sounding*. The core intuition is to transmit probing signals that carry unique signatures, or codes, associated with different spatial directions. When these signals reflect off objects, the reflected wave retains the code corresponding to the direction of reflection. Scene reconstruction thus becomes a problem of decoding these signatures from the received signal and mapping them to object positions.

While conceptually straightforward, realizing this method in a low-power, low-complexity system presents several challenges. To minimize hardware complexity and energy consumption, *SPiDR* uses only a single speaker and a single microphone for transmission and reception. We demonstrate that a carefully designed, 3D-printed passive structure can embed the required

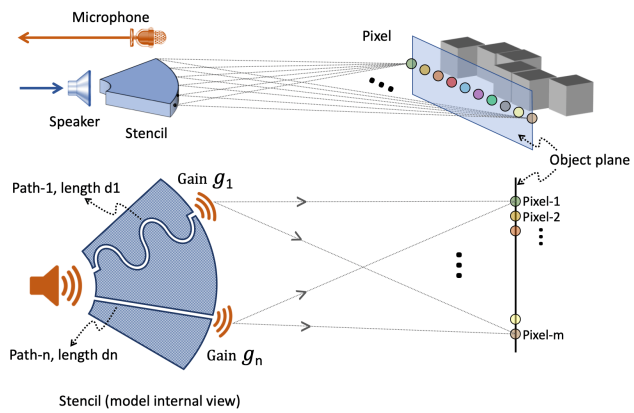


Figure 3.3: Diversity projection with the stencil with the internal channels to encode unique gains to signals to probe each pixel on the object plane with a unique signature.

directional signatures into the transmitted signal without requiring any mechanical movement or active components. If spatial coding is achieved effectively, the receiver can use a linear superposition model to recover the directional and depth information of nearby objects from the composite reflected signals. We elaborate on these two key intuitions below.

3.2.1 Coded Signal Projection with Structures

The superposition of reflected signal components, commonly referred to as the *multipath effect*, alters the waveform based on the number of paths and their complex gains. While multipath is often regarded as a detrimental phenomenon in communication and acoustic applications, *SPiDR* deliberately leverages it to project spatially diverse probing signals onto the scene. However, instead of relying on uncontrolled environmental reflections, we use a carefully designed acoustic structure—termed a stencil—to create distinct propagation paths with controlled delays and attenuations.

The stencil is a 3D-printed porous cap that covers the speaker and directs the output through a set of internal tubes connected to multiple openings, each pointed toward different angular

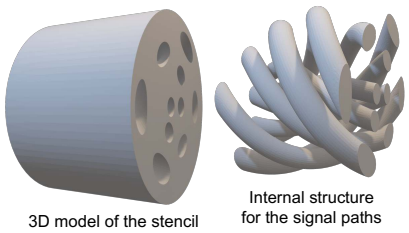


Figure 3.4: (left) The 3D design of the stencil and (right) the internal structure showing the tubular helical paths of different lengths.

directions, as illustrated in Figure 3.4. The sizes and lengths of these tubes are varied to precisely control the amplitude and phase shift of the signals emitted from each opening. Each opening effectively acts as an independent secondary source, emitting the same transmitted signal but modulated by a unique complex gain.

Figure 3.5 shows that the stencil disperses signal energy across a wide spatial region, while Figure 3.6 illustrates how it diversifies the amplitude across different angles. These directionally varied signals interfere before reaching the object plane, where they are reflected back toward the receiver. Thus, the stencil must be designed to create specific spatial patterns at the object plane, encoding unique and distinguishable codes across directions.

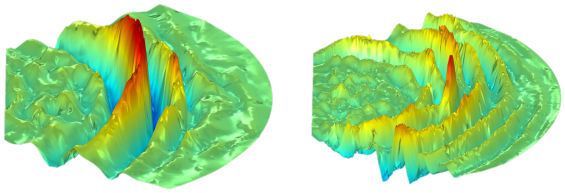


Figure 3.5: Ultrasound emitted from the speaker without (left) and with the stencil (right). The stencil spreads the signal energy over the region of interest.

3.2.2 Single-Receiver Depth Mapping

To reconstruct the scene, we discretize the environment into a collection of pixels on a vertical plane located at a fixed distance from the sensor. Multiple such planes at varying depths can cover the full three-dimensional region of interest. However, for navigation tasks, a horizontal cross-section of the 3D scene typically suffices to locate obstacles and pathways.

Accordingly, we focus on reconstructing a 2D cross-sectional view, represented by a single layer of pixels per depth plane. The probing signal reflects only from the pixels occupied by objects, and these reflections combine linearly at the single microphone.

Mathematically, the received signal y_{rcv} can be expressed as:

$$y_{rcv} = Hx \tag{3.1}$$

Here, H is a matrix whose columns correspond to the ideal reflected signals from individual pixels, and x is a vector encoding the reflectivity of each pixel. Higher values in x indicate strong reflections, corresponding to occupied pixels, while zeros indicate the absence of objects.

Thus, the vector x selects columns from H and maps the scene to a weighted sum of pixel reflections, forming the composite signal y_{rcv}. Our scene reconstruction algorithm seeks to recover x from the observed y_{rcv}, thereby inferring the structure of the environment.

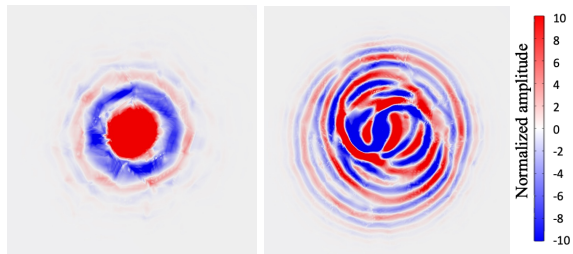


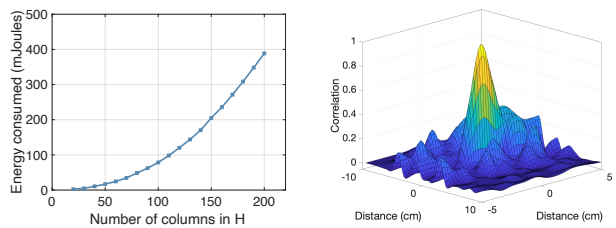
Figure 3.6: The amplitude of acoustic signal at the cross section of the image plane when (left) speaker does not have any stencil, and (right) when it has a stencil. The internal micro-structure of the stencil diversifies the signal amplitude as direction codes.

3.3 System Design

3.3.1 Low-Power Scene Reconstruction

Cross-sectional scene reconstruction, or depth-map estimation, fundamentally relies on the inversion of the channel matrix H to recover the pixel occupancy vector x, expressed as $x = H^{-1}y$, where y is the received time-domain signal of length M, and x is the vectorized representation of the 2D scene with length N. Thus, H has dimensions $M \times N$. The value of M depends on the measurement duration (t_{meas}) and the sampling frequency (f_s). For instance, a 5 ms measurement at a 200 kHz sampling rate (the Nyquist rate for low-frequency ultrasound) results in $M = 1000$ samples. N is determined by the number of pixels in the region of interest. A typical micro-robotic navigation scenario requires mapping a 20 cm $\times$ 10 cm region at 1 cm resolution, leading to $N = 200$ pixels. While manipulating a 1000×200 matrix is trivial on conventional computing platforms, it imposes significant power and latency overheads on resource-constrained single-board computers typically used in miniature robots. As shown in Figure 3.7(a), the power consumption for matrix inversion increases sharply with matrix dimensions, with a 1000×200 matrix requiring approximately 400 mJ per map. To address this challenge, *SPiDR* adopts a

multi-scale scene computation strategy, reducing power consumption by a factor of $50\times$ without compromising reconstruction quality. This strategy leverages the linearity of the signal model and the *locality of pixel values*, as described below.



(a) (b)

Figure 3.7: (a) Energy consumption for different number of columns in the channel matrix H. (b) The correlation between the signals at nearby locations. The pixels within $3cm$ have correlation higher correlation higher than 0.5.

Fractional Scene Computing: The region of interest is divided into an $M \times N$ grid, with each block corresponding to a scene pixel. Each column of the channel matrix represents the coded channel between a block and the sensor. Ideally, these channels would be uncorrelated; however, due to the diffraction limits of low-frequency signals, a degree of correlation persists between neighboring blocks, as shown in Figure 3.7(b). Rather than treating this as a limitation, we exploit the local correlations to probabilistically infer occupancy without full channel inversion. We define *meta-pixels*, groups of pixels with correlated channels. Because the projected codes are static, we can precompute the meta-pixel layout. A smaller channel matrix H_{meta} is constructed for these meta-pixels, enabling a coarse scene estimate x_{meta} at drastically reduced computational cost.

While x_{meta} lacks fine resolution, it provides probable object locations with associated confidence scores. Subsequent iterations selectively refine the high-confidence regions by zooming into smaller neighborhoods using subsets of columns from H. By iteratively stitching these

fractional maps, we reconstruct the full scene at high resolution. As evaluated in Sections 3.4.6 and 3.4.7, this method achieves more than $100\times$ power reduction compared to a full matrix inversion without sacrificing accuracy. In exceptionally dense scenes, fractional reconstruction could approach the cost of full reconstruction. However, such cases are rare in practice, and we mitigate them by employing a heuristic that monitors the total reflected signal power to detect overly dense environments.

Predictive Pre-Fetching: To further enhance energy efficiency, we precompute and store the inverses of all coarse and fine-grained measurement matrices. During real-time operation, these inverse matrices are pre-fetched based on the coarse estimated locations of objects, minimizing computational overhead. Additionally, object translation predictions based on the robot's motion are used to anticipate future pre-fetches.

Tensor Parallelization: Recent advancements in deep learning have led to hardware architectures capable of high-efficiency parallel computation. Many single-board computers now feature dedicated GPUs optimized for tensor operations [118–120]. We redesign our matrix inversion and multiplication operations to exploit this hardware via tensor parallelization. In our *SPiDR* prototype, we utilize the NEON SIMD unit [121] on the Raspberry Pi 4, converting operations into vector instructions using MATLAB Coder [69]. This significantly reduces both energy consumption and latency during scene reconstruction.

3.3.2 Directional Code Projection

The accuracy of depth-map reconstruction hinges on the ability to generate spatially diverse probing signals, ensuring that reflections from different locations encode unique signatures. While

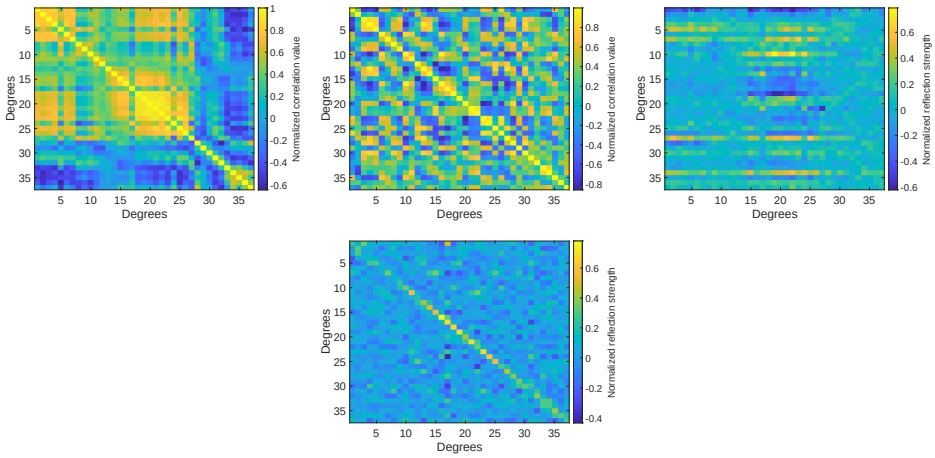


Figure 3.8: Row 1: Comparison of the correlation of the time domain signals project at different angles with an arbitrary stencil (left) and an optimized stencil (right). Row 2: Comparison of the location detection performance for a small object (1cm wide) placed at different angles from the sensor with an arbitrary stencil (left) and an optimized stencil (right).

directional speakers and beamforming arrays can achieve this, they do not satisfy the stringent latency and power constraints of micro-robotic systems. Instead, *SPiDR* uses a single 40 kHz ultrasound transducer [122], which consumes approximately 90 μJ for a 5 ms pulse, coupled with a 3D-printed passive microstructure cap, termed a stencil. This stencil modulates the outgoing acoustic wave, passively introducing spatial diversity without additional power consumption. However, arbitrary stencil designs fail to generate sufficient diversity across the scene, leading to poor reconstruction quality. As illustrated in Figure 3.8, non-optimized stencils exhibit high correlations between signals projected in different directions, degrading localization performance. To address this, *SPiDR* employs a systematic optimization process termed *multipath synthesis*, which carefully designs the internal structure of the stencil to maximize spatial signal diversity at the target scene.

Multipath Synthesis: We consider the openings on the 3D structure as secondary sound sources and assume that a specific combination of amplitude and phase distributions across these

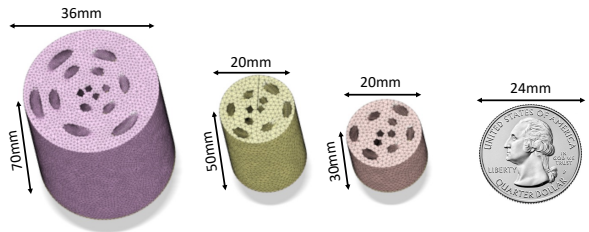


Figure 3.9: Different sizes of stencils used in our experiments.

sources can produce the desired signal pattern at the object plane. To achieve this, we first estimate the signal phases along the paths between each pair of pixels and secondary sources. Next, we compute the resulting channels by superimposing the estimated signals.

For N secondary sources, the sound pressure at the m-th pixel location, p_m, is given by:

$$p_m = i\frac{\rho c k S}{2\pi}\sum_{n=1}^{N} a_n e^{i2\pi f t}\frac{e^{-ikd_{mn}}}{d_{mn}} \tag{3.2}$$

Here:

- a_n and t are the amplitude and time delay for the n-th source,
- f is the transmitted signal frequency,
- d_{mn} is the distance between the n-th source and the m-th pixel,
- ρ is the air density,
- c is the speed of sound in air,
- $k = \frac{2\pi f}{c}$ is the acoustic wave number,
- S is the surface area of the reflecting object, assumed to be a unit constant.

This formulation follows the principles of the Weighted Gerchberg-Saxton (GS) algorithm [123] commonly used in holographic projection. Figure 3.9 shows representative examples of the metamaterial stencils used in our experiments.

The corresponding channel matrix H capturing the signal propagation between secondary sources and pixels is estimated as:

$$H(m,n) = iK\frac{e^{-ikd_{mn}}}{d_{mn}} \tag{3.3}$$

where $K = \frac{\rho ckS}{2\pi}$ is a constant.

In theory, achieving arbitrary spatial signal patterns would require a projection filter with infinitely many openings. In practice, the size and number of openings on the 3D structure are constrained, limiting the achievable diversity at the object plane and affecting pixel resolution. The signal wavelength further impacts the spatial diversity achievable for a given pixel size.

To optimize these parameters, we employ an iterative algorithm that varies structural parameters and selects the configuration that maximizes spatial diversity. We minimize the *F-measure* [124] to evaluate the quality of the channel matrix H for scene reconstruction. A smaller F-measure indicates better recovery performance, and values below 1 ensure robustness against small environmental noise.

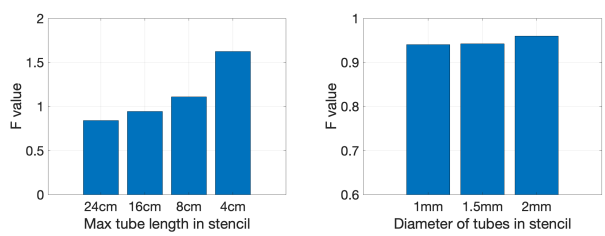
Specifically, we vary the maximum lengths of the internal tubes based on the stencil size to learn the trade-off between tube length and imaging performance. We transmit a 40 kHz, 10-cycle tone-burst signal during calibration. The stencils used in our experiments have six tubes with 1.5 mm diameters.

As shown in Figure 3.10(a), increasing the tube length decreases the F-measure, improving

reconstruction quality. For practical fabrication, a stencil of size $7 \times 3 \times 3$ cm is required to support a maximum tube length of 24 cm. Similarly, stencils supporting tube lengths of 16 cm, 8 cm, and 4 cm correspond to overall sizes of $5 \times 3 \times 3$ cm, $3 \times 3 \times 3$ cm, and $1 \times 3 \times 3$ cm, respectively.

Diffraction Effects: Diffraction at the tube openings can adversely affect signal diversity. When an acoustic wave encounters the edge of a tube with a diameter smaller than the wavelength, significant bending and spreading occur, limiting spatial resolution.

To evaluate this effect, we conduct COMSOL simulations by varying the tube diameter while keeping the maximum tube length fixed at 16 cm. As shown in Figure 3.10(b), the F-measure remains stable across moderate variations in tube diameter, indicating robustness to small manufacturing deviations.



(a) (b)

Figure 3.10: F values with (a) different lengths of tubes inside stencil, (b) different diameters of tubes.

Leveraging frequency diversity: In our measurement model, $y = Hx$, the number of columns in the matrix H indicates the number of unknowns, and the rows are the measurements. More independent measurements compared to the number of unknowns, i.e. a tall and skinny H matrix, leads to better estimation of the unknowns or better reconstruction of the map. Past works in computational photography with similar formulation attempted to increase the number of rows by taking multiple measurements while introducing controlled variations to the scene or mechanically moving the sensor for diversity. None of these approaches are feasible for

applications with small robots due to power constraints. However, we overcome this limitation of the passive structure by leveraging multipath diversity over frequency.

Note that the spatial multipath patterns per pixel created by the stencil depend on the length of its internal tubes. However, the phase introduced by these paths depends on the frequency of the signal. This means the stencil that creates a specific spatial pattern for a certain frequency, can create components of completely different phases for another frequency leading to a different spatial code. We build on this intuition to optimize the stencil design jointly for a set of discrete frequencies, such that the projected signals are maximally different at any two locations and also for any two frequencies. We leverage this frequency diversity during scene reconstruction by filtering individual frequencies from the received signal and considering them as individual measurements. We stack these measurements on top of each other to get a new measurement signal y and similarly we stack the channel matrix for different frequencies before proceeding to our scene reconstruction process. Figure 3.11 shows the reconstruction performance for different individual frequencies and for combined reconstruction.

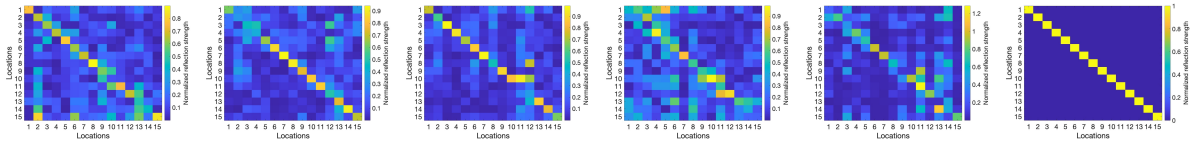


Figure 3.11: Confusion matrix of imaging accuracy for different frequencies and after combining all frequencies together.

3.3.3 Optimal microstructure design

Once the required phase delays are determined through the optimization algorithm, we make a 3D model for the stencil. To achieve mentioned phase delays, we use an automated CAD script

to design non-overlapping tubular paths inside the 3D structure. The path lengths of these tubes determine the phase difference of secondary sources, i.e., the sound signals at the output holes of the stencil. The shape of the tubes impacts the propagation of the sound waves through them due to the capillary effects and resonance of the sound column. We experimented with a number of shapes for the tubes and converge on concentric helical paths for the microstructure. Figure 3.6 shows the improvement of signal diversity distribution on a cross-sectional plane in front of a transducer with the stencil.

Next, we export the design as an STL file and slice it for 3D printing. We used the Elegoo Mars photocuring 3D printer [63] to print the stencils. We use an ultraviolet light-curable resin with $1.195g/cm^3$ density that solidifies when exposed to the light of $405nm$ wavelength. Compared with jetting-based printing, it provides a high resolution and smooth finish which is ideal for the tiny sub-structures on the stencil. More importantly, photocuring method leads to dense surfaces and makes the acoustic behavior of the stencil predictable [64].

3.3.4 Motion stacking

So far in our processing, we did not consider the movement of the robot that provides an opportunity for further improving the quality of the generated map. Perception in mobile robots requires continuous mapping of the scene during navigation. Therefore, it provides multiple snapshots where certain parts of the scene are common. It is possible to reorient the snapshots and superimpose them to generate a robust and high-resolution reconstruction of the scene. The concept of superimposition or stacking maps is a popular approach in robotic navigation as well as in photography and imaging. Some common techniques include linear and non-linear

filtering [125, 126], optical flow [127], and deep fusion [128]. However, to remain within the power budget, we borrow the technique of mean and median stacking, which runs at O(n) and O(log(n)) complexity without significantly compromising accuracy. *SPiDR* is capable of producing snapshots of the scene at a maximum speed of 170 frames per second, but depending on the speed of movement of the robot even a slow frame rate (e.g., 30 frames per second) provides an opportunity for stacking. We implemented and tested a motion-based stacking algorithm on the cross-sectional depth-maps generated by *SPiDR*. We use the knowledge about the robot's movement to compensate for the translational changes between consecutive maps before superimposition. Empirical results show that median stacking is capable of fine-grained map generated along with improvements in outlier removal. Figure 3.12 shows the performance of stacking for five frames.

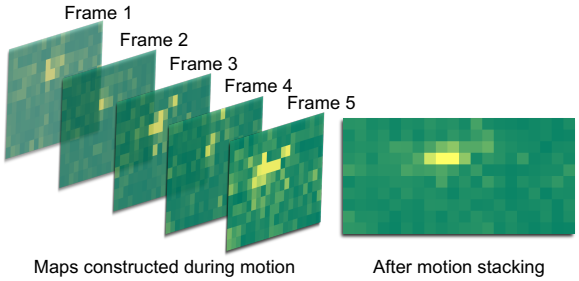


Figure 3.12: Stacking multiple frames suppresses spurious objects in the scene. Above we show the result of motion stacking of 5 frames taken as the robot moves.

3.4 Evaluation

To validate the performance of *SPiDR*, we implement a working prototype and conduct experiments across various scenes. For baseline comparison and benchmarking of energy consumption, we use a lidar system [129] and a commercial ultrasound distance sensor [130]. We first describe the experimental setup, followed by the evaluation metrics and results.

3.4.1 Metrics

We evaluate the system performance using the following standard metrics:

- **Structural Similarity Index (SSIM)** [131]:

$$\text{SSIM}(x, y) = l(x, y) \cdot c(x, y) \cdot s(x, y)$$

where $l(x, y)$, $c(x, y)$, and $s(x, y)$ represent the luminance, contrast, and structural components computed from the local means, standard deviations, and cross-covariance of the estimated image x and the ground truth y.

- **Root Mean Square Error (RMSE)**:

$$r = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (p_i - \hat{p}_i)^2}$$

where N is the total number of pixels, $\hat{p}_i$ is the binary score (0 or 1) of the estimated pixel, and p_i is the corresponding ground truth value. A value of 0 indicates the absence of an object, and 1 indicates the presence of an object.

3.4.2 Overall Performance

The *SPiDR* prototype consists of an ultrasound speaker embedded within a 3D-printed stencil and a microphone positioned above the stencil. We compare its performance and power consumption against lidar and ultrasound distance sensors. The collected signals are processed offline using MATLAB scripts. The acoustic probing signals are 10-cycle ultrasound tone bursts

at frequencies ranging from 38 to 42 kHz in 1 kHz increments, with a signal strength of 40 dB SPL. The stencil dimensions are $5 \times 3 \times 2$ cm, featuring internal tube structures. Representative scenes are shown in Figure 3.13, where the first row depicts real-world setups and the second row shows the reconstructed depth maps. We observe that *SPiDR* successfully detects scenes with up to 70% occupancy across a 20 cm wide field of view.

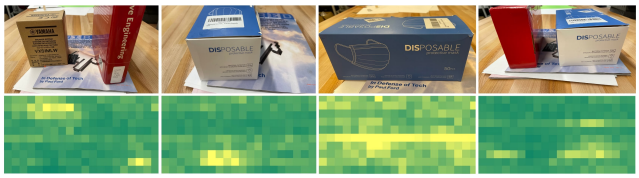


Figure 3.13: Depth-map reconstruction using *SPiDR* for various real-world scenes.

Figure 3.14 summarizes the overall SSIM accuracy and energy consumption compared with lidar and ultrasound sensors. *SPiDR* achieves comparable accuracy to lidar while consuming $400\times$ less energy. For further comparison, prior work [132] reports a radar imaging system using a 64-microphone array and a 40 kHz ultrasound source. That system consumes up to 15.7 mJ per image, which is $19\times$ higher than the energy consumed by *SPiDR*. Similarly, short-range depth cameras like Primesense Carmine [133]—designed to be $7\times$ more efficient than Kinect V2—still consume approximately 75 mJ per image, which is $90\times$ higher than *SPiDR*.

3.4.3 Impact of the Environment

Environmental Noise: We evaluate the robustness of *SPiDR* under different ambient noise conditions. Sound clips representing various indoor (household, mall, library) and outdoor (traffic) environments are played during testing. As shown in Figure 3.15(a), natural ambient noise has negligible impact on system performance, with the RMSE remaining below 0.2 and SSIM

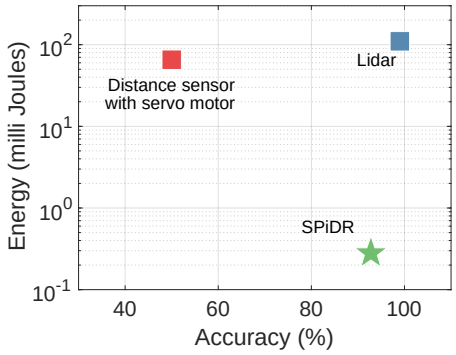
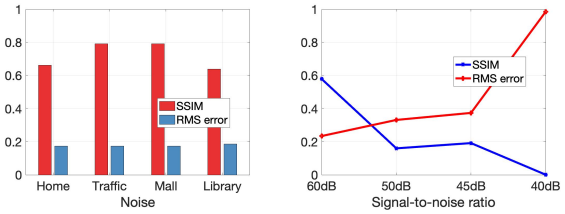


Figure 3.14: Overall performance of the *SPiDR* compared to Intel Realsense lidar and ultrasound distance sensor mounted on a servo motor. *SPiDR* consumes a fraction of power compared to Lidar and motor based systems while delivering high accuracy in depth-map reconstruction.

consistently above 0.6. This is expected, as *SPiDR* operates in the ultrasound band around 40 kHz, well above typical audible noise frequencies. We further stress-test the system by introducing increasing levels of ultrasound noise directly at 40 kHz to degrade the Signal-to-Noise Ratio (SNR). As shown in Figure 3.15(b), the system maintains robust performance up to an SNR of 60 dB. Naturally occurring ultrasound noise levels in typical environments are extremely low; even inexpensive speakers have internal noise levels corresponding to SNRs higher than 82 dB, ensuring that *SPiDR* operates reliably under realistic conditions.



(a) (b)

Figure 3.15: Impact of varying (a) types and (b) levels of noises on depth-map reconstruction.

Different locations: We evaluated *SPiDR* in three different locations – indoor laboratory, lobby, and open-air outdoor places. In these experiments, we calibrated only once in the lab

environment. Figure 3.16 shows that the performances in these environments are similar, with an RMS error below 0.2 and SSIM higher than 0.6. This result indicates that the system requires only a one-time calibration to function in arbitrary locations.

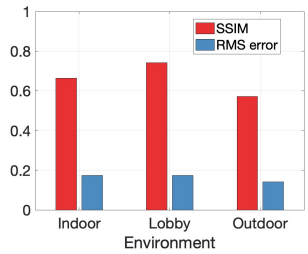


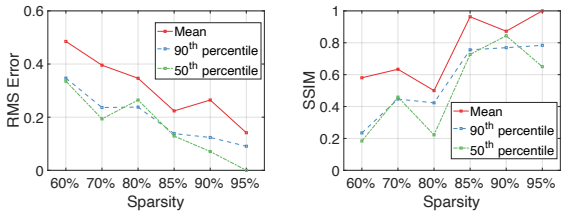
Figure 3.16: *SPiDR*'s performance in different environments.

3.4.4 Impact of system parameters

We evaluated the performance of our prototype under various conditions. We present the results below.

Cross-sectional depth-map: The width of the scene is $20cm$ and the depth is $6 - 15cm$. We place the different sizes of blocks and their combinations in different locations of the scene. 5 scenes are collected for each of the sparsity levels. As shown in Figure 3.17, the RMS error becomes lower and the IoU is higher when the sparsity is larger. The estimated scene is calculated by checking if the coefficient is larger than a threshold 0.3. The pixel with a smaller confidence score than the threshold indicates no object from the corresponding position. The coefficient equals to 1 indicates perfect matching. This means the sparser the scene, the better the performance. The performance of *SPiDR* does not vary significantly when the sparsity of the scene is within 60%. The RMS error is within 0.5cm.

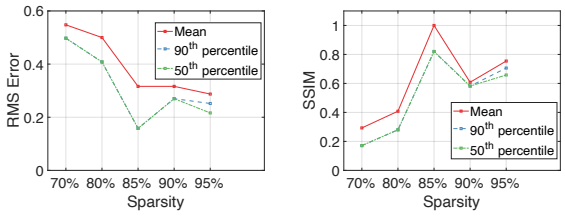
Horizontal location: Instead of imaging the 2D scene (width and depth) as 2D imaging,



(a) (b)

Figure 3.17: Cross-sectional depth-map reconstruction performance in terms of (a) RMS error and (b) structural similarity, as a function of sparsity of the scene.

we treat the depth as a constant value and only image the scene in the dimension of width. We show the result of 1D imaging when the distance between the stencil and the object layer is $10cm$. 10 scenes are collected for 1D imaging with vary number of objects and object sizes. The object sizes are within 1-6cm and there are up to three objects. As shown in Figure 3.18, all the RMS errors for sparsity $70 - 95\%$ are smaller than 0.6cm, and the SSIM is larger than 0.6 when the sparsity level is above 84%.



(a) (b)

Figure 3.18: Horizontal localization performance in terms of (a) RMS error and (b) structural similarity, as a function of sparsity of the scene.

Resolution: To test the resolution of *SPiDR*, we calibrate our system in $1cm$ and $5mm$ resolutions, respectively, and compare the scene reconstruction accuracy. We imaged 25 scenes with sparsity from $60 - 95\%$ for both resolutions, and show the CDF of the RMS error and SSIM in Figure 3.19. For both $1cm$ and $5mm$ resolutions, all the RMS errors are within 0.5cm and 50% of the scenes have SSIM larger than 0.75. As a visual example shown in Figure 3.20, the

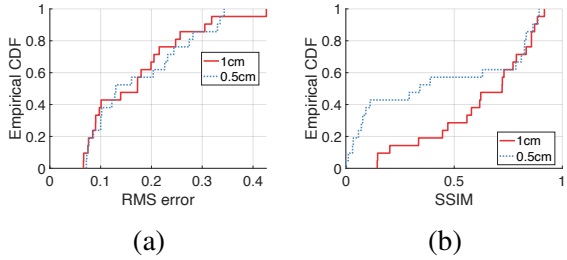


Figure 3.19: CDF plot of (a) RMS error and (b)structural similarity for depth-map reconstruction at varying resolutions.

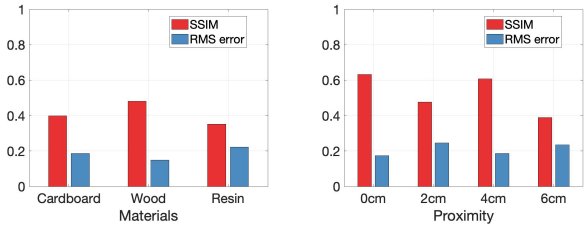
performances of these two resolutions are similar, meaning our system is able to detect 5mm-wide tiny objects.

Figure 3.20: Scene reconstruction results at 1cm and 0.5cm resolutions. We modify the number of columns in the channel matrix to have a higher resolution.

3.4.5 Impact of scene parameters

Materials/Reflectivity: The performance may vary with different types of materials since the reflection rates of sound are different. We calibrate our system with 3D-printed resin blocks and test the performance with objects made of cardboard, wood, and resin. We image 5 scenes with a sparsity $70\% - 95\%$ for each kind of material. As shown in Figure 3.21(a), the RMS errors for all the materials are below 0.3cm, which indicates *SPiDR* is robust to common types of materials.

Proximity: In this experiment, we evaluate the system's performance when the distance between two objects varies. From the experiment with an ultrasound distance sensor, we find the



(a) (b)

Figure 3.21: The depth-map reconstruction performance for (a) different materials and (b) different proximity between two objects.

distance sensor cannot detect the disconnection of two objects when the distance between them is within $3cm$. We test the system when the distances between two objects are 0,2,4,6cm, respectively. We show the quantitative scene reconstruction accuracy in Figure 3.21(b). Figure 3.22 shows the output depth-map for varying gaps. Results indicate that *SPiDR* can detect open spaces as narrow as $2cm$.

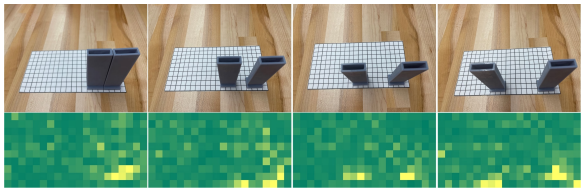


Figure 3.22: Depth-map reconstruction outputs for varying proximity between objects.

Depth: To evaluate the performance when the object is at different depths, we put a 3cm width object in depths 6,9,11,13cm, respectively. The first line of Figure 3.23 includes the ground truth of the scenes, and the second line shows the estimated images. The image becomes blurred when the distance is farther. But the location of the object is still accurate. As the system moves closer to the object, the image becomes clearer. The distance of the object from the sensor decreases the reflected signal strength which impacts the overall accuracy. The quantitative result of scene reconstruction accuracy with different depths is shown in Figure 3.24. The RMS error for

all the depths is within 0.6cm and the SSIM is larger than 0.6 when the distance is within 11cm.

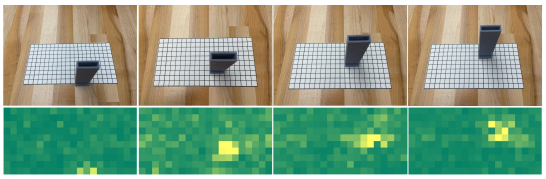


Figure 3.23: Depth-map reconstruction outputs of different depths of the objects.

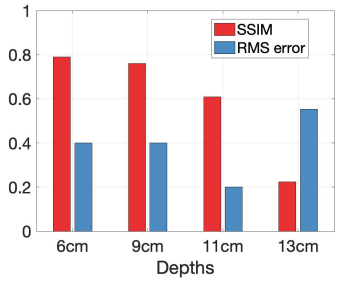


Figure 3.24: Performance of depth-map reconstruction with different depths of the objects.

3.4.6 Computation techniques

Frequency and Motion stacking: As introduced in Section 3, frequency stacking and motion stacking can increase the performance accuracy of scene reconstruction. Figure 3.25 shows an example of the stacking operations. The ground truth is a 3cm wide object located in the front middle of the scene. Before stacking, the object is recognized as two 1cm objects, and the figure has some other pixels mistakenly detected as 1cm-wide objects. After frequency stacking or motion stacking, only one object at the same location is detected, but the size is 1cm smaller or larger. If with both stacking approaches, the size and location of the object are exactly the same as the ground truth. Moreover, no points are mistakenly imaged.

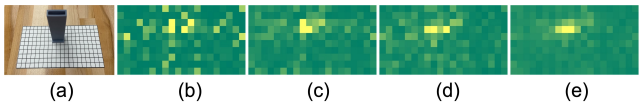


Figure 3.25: Scene reconstruction of a horizontal bar with and without frequency or motion stacking. (a) Ground truth, (b) Raw output, (c) Only motion stacked, (d) Only frequency stacked, (e) Both motion and frequency stacked.

Fractional scene observation: We propose fractional scene observation to achieve low-power scene reconstruction, as well as achieving comparable scene reconstruction accuracy. In Figure 3.26, we show two examples of scene reconstruction with one or multiple projects. We observe that the results of with and without fractional computing are comparable to each other. At the same time, we can reduce energy by 85%, as elaborated in the evaluation of the power consumption. We also evaluate the performance with different fraction rates, as shown in Figure 3.27. Naturally, the accuracy decreases with the increase of the fraction rate. In our experiments, usually a 3cm region shows a higher similarity between the signals. When the fractions become smaller, it fails to properly represent the whole scene with the H matrix after integration. This can causes error in the detection of the "region of interest".

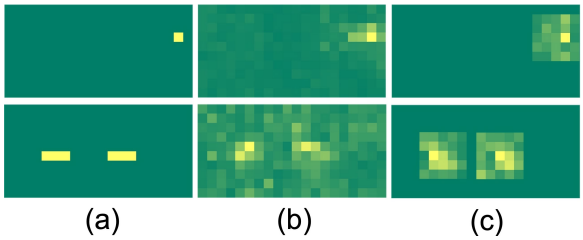


Figure 3.26: The scene reconstruction with and without the fractional computing method. (a) Ground truth, and the scene reconstruction (a) without and (b) with fractional computing.

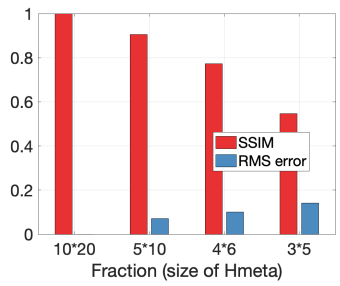


Figure 3.27: Performance with different sizes of H_{meta}.

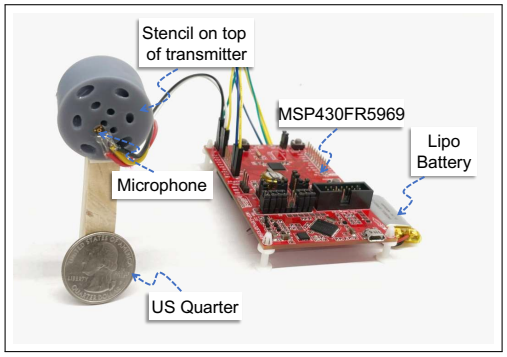


Figure 3.28: *SPiDR* prototype for power evaluation.

3.4.7 Power consumption

In this section, we develop the ultra-low-power prototype of *SPiDR* and benchmark its power consumption. Figure 3.28 shows the prototype of the setup used for evaluation. We compare its performance and energy efficiency with Intel Realsense lidar [134] and an ultrasound distance sensor [130]. We measure the power consumption of each sub-module in our system, including hardware frontend, analog-to-digital conversion, and image computation. Table 3.1 shows the breakdown of power consumption for different hardware and software sub-modules. *SPiDR* consumes a total $9.9mW$ of power and runs at a rate of 12 depth-maps per second. This results in $0.83mJ$ of energy per depth-map, which means our system can continuously run for 3 days on a common $3volts$ CR2032 coin cell battery [135].

Ultra-low-power implementation: Hardware is a bottleneck for energy consumption in most of the depth imaging techniques, like lidar and distance sensors mounted on motors. Our focus is to reduce the power consumption of the hardware front-end without losing out the accuracy. We develop *SPiDR* using TI-MSP430FR5969 MCU [1], an off-the-shelf $40KHz$ ultrasound transducer [122], and the ADMP401 microphone [136]. We use a piezo-crystal based ultrasound transducer which has a narrow resonating bandwidth at $38 - 42kHz$. We eliminate the need for a Digital-to-analog converter (DAC) by transmitting a square wave to the transducer. The impedance matching at the resonance frequency acts like a bandpass filter which converts the signal to a narrowband sinusoid. The transducer has extremely high efficiency when transmitting and receiving at the resonance frequency. We use ultra-low-power sigma-delta ADC of the MSP430 MCU to sample the received signal. The power consumption during sampling and computation is shown in Figure 3.29. We noted that the energy consumed by the ADMP401 microphone was $49mW$ so we power up the microphone only for the duration of sampling of signal, which reduces the average power per frame to $0.245mJ$. We believe that better engineered and low-power pre-amplifiers [137] can further reduce the power consumption of the system. Here we have reported the prototype power consumption from commercially available components.

Computational optimization: To measure the power consumption of the computation module with different fractional computing optimizations we implement the system on the MSP430

Submodule	**Energy consumed**
Transducer (Hardware)	$0.035mJ$
ADC (Hardware)	$0.024mJ$
Microphone (Hardware)	$0.245mJ$
Computation (Software)	$0.526mJ$
Total	$0.83mJ$

Table 3.1: Breakdown of energy consumption for the hardware and software submodules.

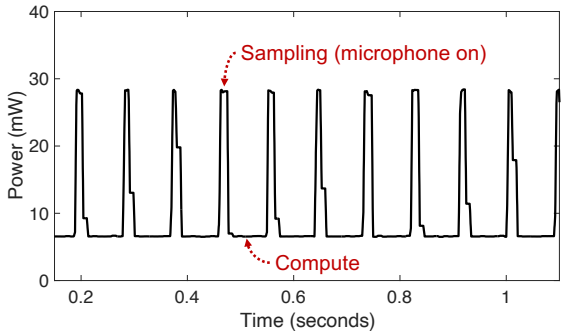


Figure 3.29: Power consumption during sampling and computation.

with varying sizes of H_{meta} and record the power consumption using the Texas Instruments EnergyTrace++ tool [138]. We run the system to compute scenes using different prefetched matrices of different sizes and record the energy consumed using the Energytrace tool. Figure 3.30 shows the average energy consumed per scene reconstruction for four different fractional computing ratios. We see that the energy drops exponentially with the decrease in size of H_{meta}. We find that 4×6 is the best size based on the trade-off between SSIM score (shown in Fig 3.27) and energy consumed. To compare the power consumption of the prefetching and fractional optimization we port the codes to a Raspberry Pi 4 [118] and record the voltage and current traces using a Keysight E6313A programmable DC power supply [68]. We write the software for *SPiDR* in Matlab and cross-compile it for a Raspberry pi 4 using Matlab Coder [69]. We use the NEON SIMD unit [121] of the Raspberry Pi 4 and export an executable C file. We run the file for 10,000 iterations and record the voltage and current traces using the power supply. We calculate the overhead of the computation module by subtracting the idle CPU power consumption. Table 3.2 shows the energy overhead with different computation optimizations.

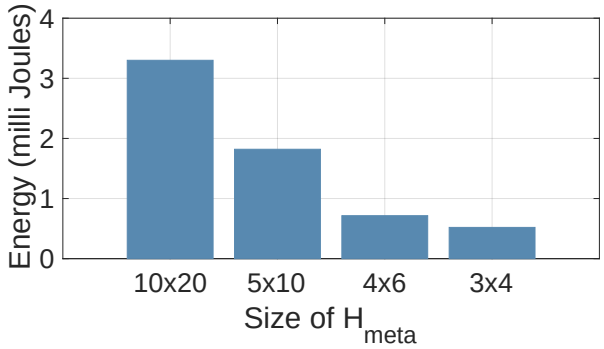


Figure 3.30: Comparison of power consumption with different sizes of H_{meta}. Prototype on MSP430FR5969.

Computation optimization	**Energy overhead**	**Gain**
None	$388.65mJ$	$1\times$
Only Fractional	$9mJ$	$40\times$
Only Prefetching	$1.42mJ$	$270\times$
Both	$0.28mJ$	$1380\times$

Table 3.2: Comparison of different computation optimizations showing the total energy consumed per scene reconstruction. Prototype on Raspberry Pi 4.

3.5 Related Work

Computational sensing has been a vibrant field for over a decade, with foundational work in compressed sensing [139] enabling a range of applications including autonomous navigation [140], contactless monitoring [141], human-machine interaction [142], and precision measurements [143]. Recent advances explore acoustic sensing within ubiquitous computing contexts [79,81,85,86,144–148]. Building on this lineage, *SPiDR* adopts a systems-oriented approach to achieve low-power spatial perception by exploiting the interaction of sound with passive microstructures. Below, we discuss research areas closely related to our work.

- **Scene Perception and Object Avoidance:** Significant progress has been made in object detection using image processing [149–151], radar [152–155], lidar [156–158], and

ultrasound [38, 73, 159–162]. However, traditional object avoidance techniques do not directly translate to microrobots due to differences in physical scaling. Recent microrobotic studies explore magnetic particle imaging for object detection [163, 164] and vision-based underwater obstacle detection [165].

• **Low-Power Sensing:** Low-power sensing is critical in applications such as body motion tracking [166, 167], chemical sensing [168–170], and health monitoring [171–174]. Recent work includes low-power signal processing techniques [175, 176], RFID-based pervasive sensing [177, 178], and RF energy harvesting transceivers [179].

• **Acoustic Structures for Sensing:** Using structures to manipulate sound dates back to ancient architecture [180]. Modern studies have explored 3D-printed acoustic microstructures for impedance control [90] and tangible user interfaces [93, 181]. Our previous work, Owlet [97, 182], demonstrated the potential of microstructures to embed directional cues into recorded signals for sound localization. While both Owlet and *SPiDR* leverage structure-assisted sound manipulation, *SPiDR* fundamentally differs by modulating the transmitted signal for spatial sensing and reconstructing depth maps.

• **Sensing with Limited Sensors:** Several works have explored spatial sensing without arrays, using structures like walls, Legos, or cubes [46, 47, 94, 95]. Other approaches use wall reflections [183], coded apertures [184–186], or holographic sound field generation [112, 187–189], mostly operating at megahertz frequencies. Directional microphone techniques [190, 191], mode beamforming [192, 193], and Micromachined Ultrasound Transducers (MUTs) [194–196] offer compact alternatives. Additional methods leverage ambient channel responses [80, 197], frequency patterns [198], array interpolation [199–201], compressive sensing [202–204], and repurposed acoustic vector sensors [205–207].

3.6 Discussion

We discuss several important aspects and avenues for future improvements.

• **Speed of the Robot:** Robot speed trades off with imaging latency. Motion stacking enhances depth-map resolution but constrains the maximum allowable speed. At 1 cm resolution, the robot must move less than 1 cm per frame for best results. Given that the current prototype produces 170 depth maps per second, the robot can achieve speeds up to 170 cm/s (approximately 6.12 km/h) without significant degradation in mapping accuracy.

• **Applicable Micro-Robots:** Most micro-robots [107, 109, 110, 208] operate within a 10–30 mW sensing and computation power budget. Available sensors in this regime are typically proximity sensors (IR or ultrasonic), which offer only limited obstacle detection capabilities. Vision-based sensors [209, 210] are hardware-efficient but computationally intensive. *SPiDR* offers a viable low-power alternative for smarter navigation. Given the stencil's compact size (2–5 cm), it can be deployed on centimeter-scale robots [211–214] and larger robots [215–219]. Moreover, as *SPiDR* operates in ultrasound, it remains functional under challenging conditions such as low illumination, smoke, and opaque obstructions.

• **Interference from Outside the Region:** We define the region of interest by considering signal properties such as frequency-dependent attenuation, which acts as a natural filter. Although unwanted reflections within 60 cm could theoretically interfere with the system (given a 3.5 ms observation window), practical measurements show negligible out-of-region signal strength. While time gating could potentially help, separating signals based solely on time is challenging due to the superimposed nature of the received signals. Addressing this interference remains an avenue for future research.

3.7 Conclusion

This chapter presents *SPiDR*, a low-power spatial sensing system designed for micro-robot navigation. Using only a single ultrasound speaker-microphone pair and a 3D-printed passive structure, *SPiDR* projects spatial diversity into the environment to reconstruct depth maps. The system consumes only 10 mW of power and generates depth maps at 12 frames per second. By leveraging structure-assisted acoustic sensing, *SPiDR* opens new possibilities for low-power perception, obstacle detection, and autonomous navigation, particularly in environments where traditional vision-based methods fail.

Chapter 4: Ultra-Low-Power Self-Localization Using a Single Antenna

4.1 Introduction

Advancements in miniaturized, ultra-low-power localization techniques have enabled new opportunities in long-term monitoring and behavior studies of wildlife and livestocks over a large open field. Sensor nodes a few centimeters in size and a few grams in weight have shown capability to self-localize and save data for several days and intermittently radio out the data to the nearby gateways. Communication being one of the core component of this model, existing solutions often compromise on the location accuracy in order to fit communication and sensing functionalities within available power budget. In this chapter, we present *Sirius*, which breaks this trade-off by using a single antenna and radio to simultaneously communicate and self-localize within millijoules of power requirement and millimeters in size.

Self-localization, where the node computes its location onboard, is particularly challenging on small and low-power nodes. Conventional sensors for location estimation are power hungry. GPS uses low-noise amplifiers requiring several milliWatts of power to receive weak satellite signals (-155 dBm) for location fixes [220]. Power requirement of Ultra-wideband (UWB) radio or Bluetooth-Low-Energy (BLE) based solutions can go up to a half of a Watt along with synchronization and other protocol overhead [221]. Fundamentally localization relies on two different spatial features extracted from the received signal – the distance from the source and the

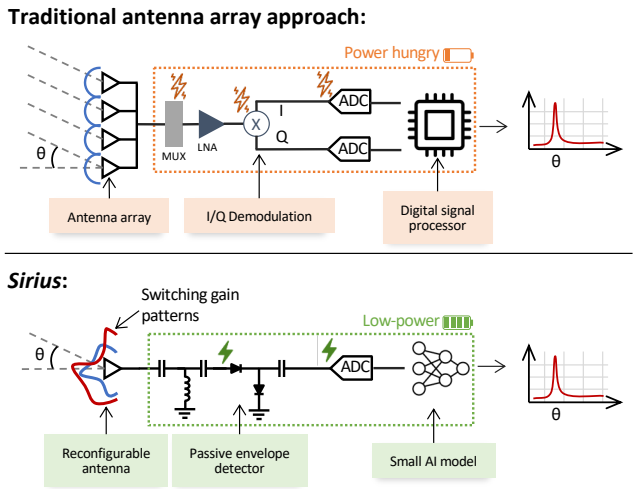


Figure 4.1: *Sirius* dynamically switches the beam pattern of the antenna to embed direction specific signature to the received signal. The vector of amplitudes contain the unique signature which map to the angle-of-arrival, θ.

angle-of-arrival (AoA). A node can use a combination of distances and/or AoAs from multiple sources to estimate its location with respect to the known locations of the sources [97, 182, 222]. Distance estimation requires strict time synchronization [222] between the node and the source so that time-of-flight (ToF) of the signal can be measured accurately. Given the signal travels at the speed of the light, a small error in synchronization can lead to several meters of error in distance and, in turn the location estimate. Onboard clock of the low-cost nodes are prone to drift and require frequent coordination with the access point, which is not a desired behavior for ultra-low-power nodes [223]. Moreover, the time resolution of the ToF estimate is inversely proportional to the radio-bandwidth. The low-bandwidth communication radios can not meet the requirement.

AoA-based localization methods, on the other hand, do not require strict coordination with the access points and are more suitable for independent self-localization on the low-power nodes [182, 197]. Triangulation using AoAs from two unsophisticated beacons can be sufficient for

the node to find its position. However, existing methods require complex hardware for AoA estimation. Spatial sampling to find the angle of arrival of the signal requires an array of antennas separated by at least half-wavelength distance. [224, 225] Antenna arrays lead to power-hungry radio frontend and a bigger space for low-frequency antenna. Spatial scanning using steerable beams – computational or mechanical – consumes several times the available power per duty cycle in these nodes. We propose an efficient method that uses angles and does not require phase information. *Sirius*'s single antenna based design reduces the form factor and power consumption of the node.

Communication and localization are often designed to work on the same radio. Passive envelope detector-based radio architectures are evolving in low-power nodes. Several recent works [226–229] focused on improvements in this communication architecture to achieve reliable data exchange at a ultra-low power budget. However, few attempts have been made on solving localization using envelope detector based architecture. This is mainly due to the reason that envelope detectors eliminate power hungry components of radio such as oscillators and high speed ADCs. But due to this, extracting the phase information of UHF waves becomes challenging.

We take a novel approach to obtain the AoA for triangulation we call 'directional code embedding'. The high level idea is to embed a direction specific code to the incoming signal during reception and later post-process the signal to recover the code unique to a specific angle-of-arrival. Imagine a radio receiver capable of fluctuating the amplitude of the signal in a subtle but consistent way that adds a signature to the signal. The fluctuation is designed to remain transparent to regular communication and signal decoding, but the signal processing can identify it as a unique code. Now if we can design this signal fluctuation to be a function of the angle of arrival, the code embedded in the signal will reveal the AoA. We leverage antenna's gain pattern for this purpose.

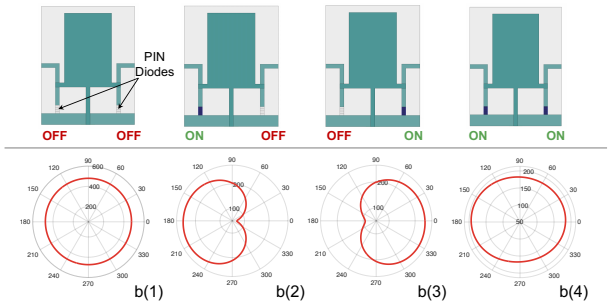


Figure 4.2: *Sirius* uses pin-diode to switch the gain pattern of an antenna. Connecting and disconnecting a conductive patch to the surface of antenna can change the shape of the gain pattern. This figure shows four different configurations of the antenna controlled using a set of two switches.

Figure 4.1 gives an overview of the system.

An antenna is basically a metallic structure that captures and transmits radio electromagnetic waves [230]. Based on the shape of its conductive structure, an antenna can have different gains of reception at different angles. In other words, a simple asymmetric shape of an antenna can amplify or attenuate signals arriving from certain angles to create a unique shape of angular gain pattern. Typically, this gain pattern is constant as the shape of the conductive structure does not change. However, we have shown that there are multiple independent conductive components on the antenna which are connected through controllable switches. These switches are turned on and off to connect and disconnect a component from the antenna leading to essentially a controllable shape of the antenna. This in turn gives a means to switch between different unique gain patterns in real-time.

Literature is sparse on self-localization methods suitable for envelope detector-based receiver. Probably closest to our work is LivingIoT [176]. It has shown an effective technique to estimate the angles of the access points with respect to the mobile node. Here two or more access points sweep the region with steerable beams in a synchronized pattern so that the mobile node estimates

the angle by observing the time of maximum beam power. It requires specialized access points with antenna arrays capable of focusing the beam and also adds a latency on location update while the beam scans the region. *Sirius*, on the other hand, is capable of estimating AoA using only a narrowband signal from the access point or even from a simple beacon. Given the system estimates the AoA from any signal using completely onboard computations, the self-localization requires only two omnidirectional beacons for the reference anchor locations. *Sirius* does not require any synchronization, not even between the anchor beacons. We have developed an ultra-low-power technique to tolerate time-frequency drifts between the transmitters and the receiver. *Sirius* can operate with any low-power envelope-detection based communication architecture while the location estimation process remains transparent to the communication and run in parallel. Table 4.2 summarizes and highlights the core differences in existing works and *Sirius*.

This chapter is a step toward a broader vision of reducing hardware complexity for sensing on the edge devices [38, 39, 231–233]. It strives to relax strict hardware requirements in spatial sensing on these resource-constrained devices by formulating it as primarily a computation and inference problem [182, 231]. This leverages the capabilities of advanced machine intelligence to bridge the limitations of unsophisticated hardware toward achievable performance. Latest microcontrollers make it possible to run these algorithms on the edge devices without frequent communication with the server to offload processing [182, 234, 235]. This approach not only enables complex sensing on battery-less platforms, but also makes them inexpensive to scale in numbers and it is easier to upgrade the sensing capabilities through software updates. *Sirius* shows a method for signal's angle-of-arrival estimation using a single antenna and an ultra-low-power radio. The core method can be applicable to a number of sensing applications beyond localization. While it opens up scopes for improvements and future research, in this chapter we focused on

assessing the possibilities and limitations of the system. We have made the following contributions.

- We present, to the best of our knowledge, the first ultra-low power self-localization system that uses a single antenna and does not require any kind of synchronization.
- We develop a novel angle-of-arrival estimation technique using a single antenna. The technique uses pin-diode based RF switches to switch between multiple beam pattern configurations to capture fine-grained AoA information.
- We develop a novel algorithm to identify anchors and classify its AoA without requiring any time synchronization.
- We implement a functional hardware-software prototype of *Sirius*. We extensively evaluate the prototype on its real world performance and power consumption.
- We open source our hardware and software for the community to replicate and build. Code available at https://doi.org/10.5281/zenodo.7922283.

4.2 Core Intuition and Feasibility

Traditional methods of AoA estimation have relied on spatial sampling or scanning each angle using a directional receiver. A scanning beam is either generated by a highly directional antenna which is mounted on a motor or generated using a phased antenna array. However, both of these approaches are power expensive. Motors consume huge amounts of energy to rotate and are generally limited in motion and speed. On the other hand, antenna arrays fundamentally require spatial sampling of the signal using multiple antennas to extract phase and perform beamforming [224, 225, 236]. This requires power-hungry components such as oscillators, down

converters, multiplexers and ADCs which are now becoming obsolete for the future ultra-low-power IOT nodes [229, 237, 238].

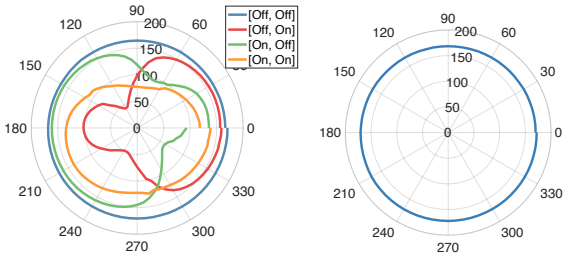


Figure 4.3: Feasibility study demonstrates dynamic gain pattern switching. Figure shows gain patterns for (a) our reconfigurable antenna and (b) regular antenna.

Our goal with *Sirius* is to enable localization on ultra low-power nodes comprising of passive envelope detectors. The key intuition is to embed a direction specific code using subtle amplitude fluctuations in the received signal, so that even the envelope detectors can sense these amplitude changes without requiring to sample the phase of the signal.

To embed direction-specific fluctuations, we dynamically change the gain value of the antenna. We design a reconfigurable antenna that can switch between diverse gain patterns on the fly. Figure 4.1 shows the reconfigurable antenna receiving the signal with multiple gain patterns to create a vector of gain values which uniquely maps to the AoA of the signal.

To reconfigure the gain pattern of an antenna, we change the shape of the antenna. We achieve this by carefully connecting and disconnecting patches of conductor to the antenna using RF switches. To check the feasibility of our idea, we first run a simulation experiment. Figure 4.2 shows the gain pattern of the simulated antenna where we use pin diodes to connect patches of conductor to either side of the antenna.

We also make a prototype of the designed antenna and perform a simple experiment to verify the switching in the real world. We place a stationary antenna transmitting a pure tone of

900MHz and our reconfigurable antenna a few meters apart mounted on a turntable. We collect the gain pattern of the antenna multiple times while switching between four different antenna configurations. Figure 4.3 shows the polar plot of the gain pattern of our antenna along with the gain pattern of a regular off-the-shelf omnidirectional antenna. We clearly see the diversity in our antenna which we will use to design our AoA estimation technique.

While the concept looks promising, there are several challenges in designing a practical localization system. Next, we discuss the system challenges in detail.

4.3 System Design

4.3.1 Ultra-low Power Receiver

In order to minimize the power consumption of *Sirius*, we eliminate the power-consuming parts from the front-end radios like high-frequency oscillators and down converters. We do not sample the amplitude and phase of the signal. Instead, we only sample the energy of the incoming signal using a passive envelope detector. As we discussed in the previous section, we could capture the direction-specific signature by just extracting the energy of the signal. To this end, we use a passive envelope detector to receive the signal which measures the absolute energy of the signal. The output of the envelope detector is rectified voltage which is proportional to the energy of the signal. We capture the voltage using an ultra-low power sigma-delta ADC. The sampling frequency of the ADC is extremely low, in the kHz range, which is orders of magnitude less than the MHz ADCs found in conventional radio receivers. This provides a significant reduction in power as oscillators and ADCs are the most power-hungry components of the radio front-end. Figure 4.4 shows the circuit diagram of our receiver. *Sirius* does not require any amplifier to

increase the SNR of the signal, rather it uses an impedance-matching network to decrease the reflection of the signal from the antenna. We use SMS7630 schottky diodes [239] and microwave grade capacitors and inductors to develop a sensitive and passive front-end for *Sirius*.

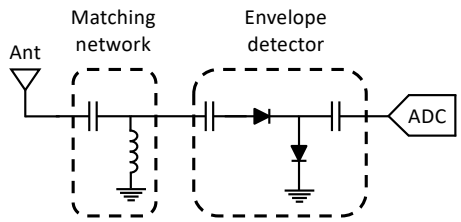


Figure 4.4: Passive envelope detector used by *Sirius* captures incoming signal energy without requiring power-hungry components such as oscillators and down-converters.

4.3.2 AoA Estimation

When a signal, $x(t)$, arrives at the antenna surface, the amplitude of the sampled signal, $y(\phi, t)$, is proportional to the gain of the antenna, $G(\phi)$, at a particular angle-of-arrival, ϕ, of the signal.

$$y(\phi, t) = G(\phi)x(t) \tag{4.1}$$

While most antennas have a fixed gain pattern, we leverage the reconfigurable antenna that can dynamically change its gain pattern using RF switches. We use this antenna to sample the signal with multiple gain patterns, each with its own unique gain value. This gives an array of sampled signals, $[y_1(\phi, t)y_2(\phi, t) \ldots y_n(\phi, t)]$, all with different gain patterns but same angle-of-arrival. Here n is the number of unique gain patterns.

$$\begin{bmatrix} y_1(\phi,t) \\ y_2(\phi,t) \\ \vdots \\ y_n(\phi,t) \end{bmatrix} = \begin{bmatrix} G_1(\phi)x(t) \\ G_2(\phi)x(t) \\ \vdots \\ G_n(\phi)x(t) \end{bmatrix} \tag{4.2}$$

The amplitudes of this vector, $[|y_1(\phi,t)|, |y_2(\phi,t)| \ldots |y_n(\phi,t)|]$, contains a unique signature about ϕ. However, it is dependent on $|x(t)|$ which can change with transmitting signal strength or the distance between the antennas. We cannot compute AoA, unless we precisely know $|x(t)|$. To solve this, we divide each element of the array with the first element, making the signature independent of the transmitted signal.

$$C = \begin{bmatrix} 1 \\ |y_2(\phi,t)|/|y_1(\phi,t)| \\ \vdots \\ |y_n(\phi,t)|/|y_1(\phi,t)| \end{bmatrix} = \begin{bmatrix} 1 \\ |G_2(\phi)|/|G_1(\phi)| \\ \vdots \\ |G_n(\phi)|/|G_1(\phi)| \end{bmatrix} \tag{4.3}$$

We extract the signature code induced by the antenna, C, and look it up in a table of pre-learned signatures that corresponds to the AoA, ϕ.

Effects of multipath: Our formulation for AoA estimation assumes that the gain pattern of the antenna can be used for decoding. However, the presence of reflecting surfaces in the environment can affect the gain pattern amplitudes due to multipath. We conducted simulation and empirical experiments, which reveal that when an unobstructed line-of-sight path exists, the energy variations of the gain pattern due to multipath are minimal. This scenario is common in outdoor spaces and certain large indoor spaces (e.g., warehouses). To validate our hypothesis,

we conducted a simulation study and observed that the energy contrast between the direct path and reflected paths is extremely high, owing to the losses incurred by the inverse-square of the distance factor and the absorption of the signal by reflecting surfaces. The gain pattern of the received signal, as demonstrated in Figure 4.5, reveals that the direct path accounts for 92% of the energy, while all reflections combined constitute only 8% of the total energy. We evaluate *Sirius*'s performance in different levels of multipath and non-line-of-sight (NLOS) scenarios in the evaluation section, where we discuss the effectiveness of AoA estimation under such conditions. We conclude that if the direct path accounts for 80% of the signal energy, then gain codes can be effectively employed to extract the AoA.

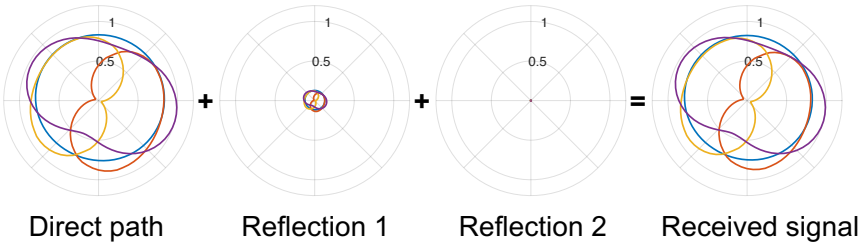


Figure 4.5: Gain patterns for the first 3 paths that signal takes to reach the antenna. First one is the direct path which constitutes the majority of the energy, followed by weaker reflections.

4.3.3 Localization with Independent Beacons

Sirius uses the triangulation method for localization, which requires angle estimations from at least two known anchor locations as shown in Figure 4.6. We aim to make the localization process independent of the communication to the access points (AP). This allows our system to operate in parallel to any unmodified communication protocol using the same receiver radio and *Sirius* antenna. We rather use two or more unsophisticated narrowband beacons as the anchors

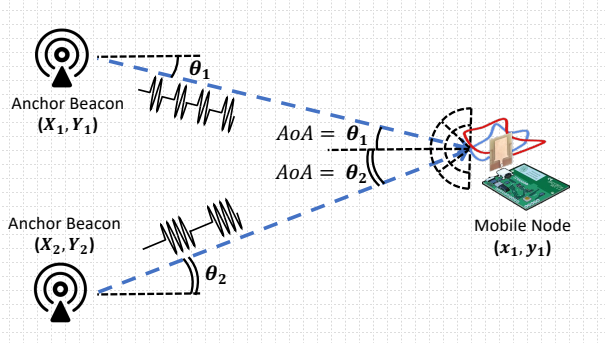


Figure 4.6: *Sirius* uses triangulation to localize mobile nodes. It estimates angles from at least two anchors with known locations.

without any synchronization. These anchor-beacons require a one-time parameter selection setup and their hardware consists of simple oscillator-based frequency generators, which further reduces the cost and power consumption for long-term maintenance-free deployment in scale. However, this simple localization infrastructure requires addressing two challenges on the ultra-low-power receiver node: (a) identification of signals from individual anchor-beacons and (b) detecting and avoiding collisions of unsynchronized beacons signals.

To make sure that the node receives the signal from each anchor beacon without interference, we design the transmitting signal using On-Off keying. Our key observation is that if the duty cycle of the anchor signals is different from each other, then no matter what time offsets they have, there will be a repeating time window in the received signal where only Anchor1 and Anchor2 signals are received. Additionally, if we set the duty cycle of Anchor1 equal to twice that of Anchor2, then it results in equal time windows for both anchors. Figure 4.8 shows the received signals in different cases of time offsets between the anchors.

Anchor identification: Let us assume that we have an algorithm to detect collision windows, which we will explain later. Our core idea here is that if the duty cycle of the anchors are set, then no matter what the time offset is, there can only be two possible cases of collisions. Case 1,

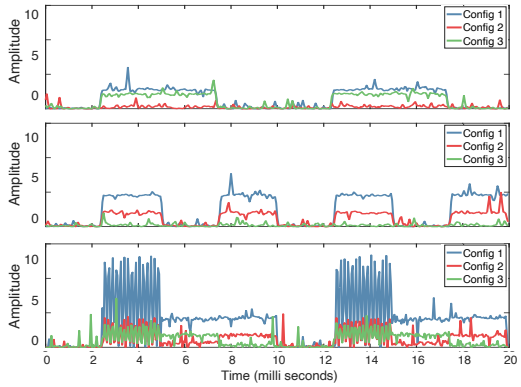


Figure 4.7: Figure depicts envelope detector's received signal: (top) Anchor1 transmitting, (middle) Anchor2 transmitting, and (bottom) both Anchor1 and Anchor2 transmitting.

where the length of collision is equal to anchor1's time period, and Case 2, where it is smaller than anchor1's time period. Figure 4.8 shows both cases where collisions are represented by a sinusoid wave.

Our algorithm first aims to identify the case of collision (Case 1 or Case 2) by looking at the length of the collision windows. We empirically found that a tolerance threshold of 95% works best to identify the collision case. After the identification of the type of case, we assign the exterior neighbors of collision windows as anchor2 and interior neighbors as anchor1 if Case 1 is detected. Whereas, if Case 2 is detected, we assign the interior neighbors of collision windows as anchor2 and exterior neighbors as anchor1.

Identifying collisions: If we successfully find the collision window, then the Anchor detection algorithm can label time windows to their corresponding anchors. However, detecting collisions is not straightforward in *Sirius* because we do not sample the incoming signal. As discussed in the previous section, we use an envelope detector which only gives the amplitude of the incoming signal. To solve this challenge, we leverage the non-linear properties of the envelope detector. We transmit a single tone of frequency, f_1 from anchor1 and, a slightly different

frequency, $f_2 = f_1 + \Delta f$ from anchor2. Due to the non-linear property of the rectifier, the output contains a beat frequency component equal to Δf. Let us take a closer look at this.

Let $x_1(t) = A_1 e^{j2\pi f_1 t}$ and $x_2(t) = A_2 e^{j2\pi f_2 t}$ be two sinusoidal signals with frequency f_1 and f_2 be transmitted from the two anchor beacons respectively. When received by *Sirius*'s receiver antenna, the signal will become a superimposition of $x_1(t)$ and $x_2(t)$ which can be written as

$$y(t) = A_1 e^{j2\pi f_1 t} + A_2 e^{j2\pi f_2 t} \tag{4.4}$$

$$y(t) = [A_1 e^{j2\pi(\frac{f_1 - f_2}{2})t} + A_2 e^{-j2\pi(\frac{f_1 - f_2}{2})t}] e^{j2\pi(\frac{f_1 + f_2}{2})t} \tag{4.5}$$

In other words, $y(t)$ is an amplitude modulated signal with modulation frequency $\frac{f_1 - f_2}{2}$ and carrier $\frac{f_1 + f_2}{2}$.

After the envelope detector, the output signal, $y_{env}(t)$, becomes the absolute of $y(t)$

$$y_{env}(t) = |y(t)| = |A_1 e^{j2\pi f_1 t} + A_2 e^{j2\pi f_2 t}| \tag{4.6}$$

$$y_{env}(t) = A_1^2 + A_2^2 + 2A_1 A_2 cos(2\pi(f_1 - f_2)t) \tag{4.7}$$

We clearly see that in the case of superimposition, or collision, a beat frequency, $\Delta f = |f_1 - f_2|$ appears in $y_{env}(t)$. Figure 4.7 shows time domain signals collected in the real world where the beat frequency is shown due to the collision of anchor signals.

Note that Δf is hardly affected by the clock drifts of unsynchronised beacons. A typical 10Mhz oscillator drifts at a rate of $2 - 6ppm$ [223, 240]. For $\Delta f = 10kHz$ centered around

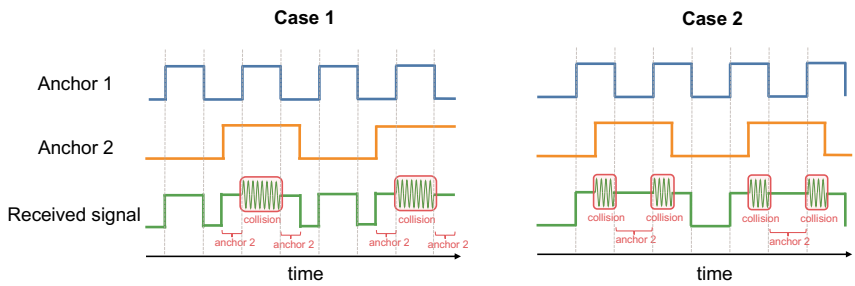


Figure 4.8: Anchor beacon signals designed with varying duty cycles create distinct time windows for interference-free reception. (left) Case 1: Anchor2's window as the outer neighbor of collision windows. (right) Case 2: Anchor2's window as the inner neighbor of the collision window.

2.4GHz, the frequency can be deviated by only $500 - 600Hz$ which is a negligible difference (¡6%) in the beat frequency. We have also shown the impact of clock drifts and frequency offsets on *Sirius* in the evaluation section.

Beat frequency detection: We use a sliding window to classify each window as "No signal", "Signal detected" or "Collision detected". We first apply a threshold on the signal energy to identify if the received signal has at least one anchor's signal and label it as "Signal detected". Then we apply a beat frequency detection algorithm to identify collision regions and label them as "Collision detected".

To make a robust beat frequency detector, we leverage synchronization in antenna configuration switching. We sample all the configurations in a burst, which makes the signals from the antenna configuration synchronous. To identify a beat frequency, we find the first derivative of the signal by taking the difference between adjacent samples. In the case of DC, the result is zero, but an oscillating signal is preserved with only DC bias removed. We find the energy of this oscillating signal around zero by taking an absolute sum. Figure 4.9 shows the steps of finding beat frequency. Due to some nulls and low gain points in the gain pattern, the beat frequency might be weak and hard to detect. To solve this, we also take the product of the energy obtained from all the

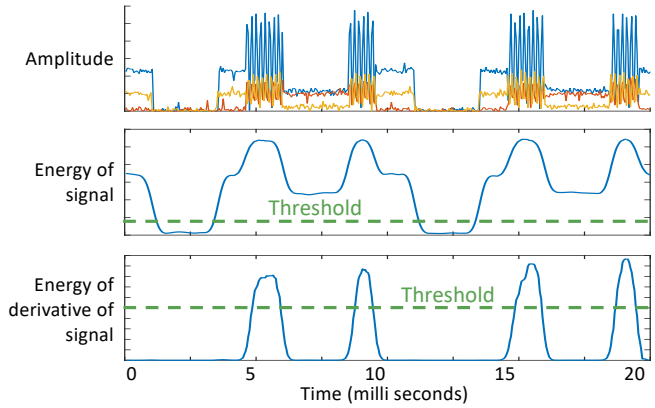


Figure 4.9: Figure displays anchor detection algorithm outputs: (top) time domain signal, (middle) total energy from all configurations, and (bottom) signal's first derivative.

configuration signals. This suppresses false positives and amplifies the beat frequency confidence value. Then we simply apply a threshold to identify if the time domain window is a collision or not.

4.3.4 Designing Programmable Directional Gain

The main focus of designing the antenna is to have a reconfigurable and programmable gain pattern. The gain pattern is determined by various characteristics like the patch shape, feed locations, and substrate parameters of the antenna. The most efficient way to make a single antenna reconfigurable is to change its shape. We use switches to connect or disconnect patches of conducting elements to the antenna. Depending on the location of these patches, it alters the flow of current inside the antenna. This changes the gain pattern as shown in Figure 4.2. However, adding a random patch to the antenna can lead to impedance mismatch and decrease the sensitivity of the antenna. We use HFSS simulation to carefully design the antenna and tune the location and shape of each additional patch. We design reconfigurable antennas in two ISM bands (900MHz and 2.4GHz) to show that the antenna can be tuned to any desired frequency while achieving

reconfigurable gain patterns. In our system, we use pin-diode-based RF switches which only require two different voltage levels (e.g., 0V or 3V) to toggle between open and closed states. The switches are located on the antenna surface between the patches. We control the state of the switch using dc voltage which is applied from an external microcontroller.

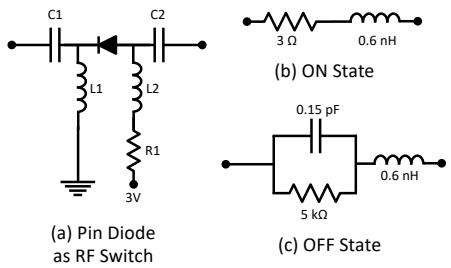


Figure 4.10: Using pin diode as RF switch. (a) Biasing circuit (b, c) Equivalent lumped model for on and off states.

4.3.5 Pin-diodes as RF Switches

Pin diodes, unlike conventional diodes, act like current-controlled resistors for microwave frequencies. The value of resistance is set by the DC current flowing through the diode. The PIN diode behaves like a closed circuit in forward bias, when current is flowing through it, and has a very high resistance to microwave frequencies otherwise. We can control the state of the switch by

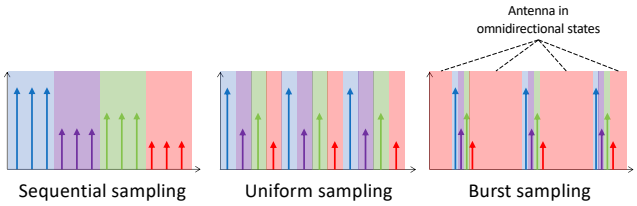


Figure 4.11: Switching and sampling techniques: (a) Sequential sampling: constant-time switch, hold, and sample, (b) Uniform sampling: continuous configuration switching with constant rate, (c) Burst sampling: high-speed switching and sampling with long duty cycles for omnidirectional communication.

applying a constant voltage across the two terminals of the diode as shown in Figure 4.10(a). We limit the amount of current flowing through the diode using a resistor, $R1$. Figure 4.10(b) shows the equivalent lumped circuit model of the pin diode in On and Off states. In forward bias, which is the on state, the PIN diode is modeled as an inductor and resistor in series. In off-state, it is modeled as an inductor in series with a paralleled capacitor and resistor. As shown in the figure, we use 3Ω, $0.6nH$, $0.15pF$ and $5k\Omega$ as our values to model the lumped components. These values are obtained from the datasheet [241] provided by the manufacturer.

Pin diodes are two-terminal devices and using them as a switch requires merging the control terminals with the RF terminals. This merging leads to a problem of RF signal leaking out of the antenna. The switching circuit and wires become a part of the antenna and contribute to signal reception and transmission, which affects the overall gain pattern of the antenna. To isolate the biasing circuit from the RF terminals, we use inductors and DC-blocking capacitors. The capacitors prevent the DC bias current from entering the antenna and the inductors act as RF chokes which prevents the RF signals from passing through the switching circuit while providing a clear path for the DC bias current. Figure 4.10(a) shows the biasing circuit with the isolating components.

- **Can *Sirius* antenna be used for communication?**

A lot of works on low-power IoT use an envelope detector-based radio receiver to reduce power consumption. *Sirius* aims to complement these IoT systems with spatial sensing and localization. We aim to design the sensing pipeline in such a way that any existing communication pipeline is unaffected by the amplitude fluctuations introduced due to gain pattern switching.

One of the core challenges in designing such a pipeline is to switch the antenna configuration fast enough. Low-speed switching hinders the communication abilities of the antenna since

changing the configuration of the antenna can also affect the wireless channel. As we discussed in the previous section, *Sirius* uses pin diodes as switches instead of conventional RF switches. The fast response time of pin diodes enables us to switch the antenna configuration at extremely high speeds. In our prototype, we have used the Infineon BAR50-02v [242] pin diode that can change its state in less than 1.1 microseconds, achieving a maximum switching rate of 1MHz.

We leverage this high-speed switching of pin diodes to sample the signal in bursts. As shown in Figure 4.11 traditional sampling techniques either use sequential sampling i.e., sampling one configuration at a time, or uniform sampling i.e., continuously switching the configuration and sampling. Both techniques require repeated changing of the configuration which constantly disturbs the wireless communication channel. In burst sampling, we keep the antenna in a default state where both the switches are off and the antenna has an omnidirectional gain pattern. This consumes no energy since the switches are off and the antenna behaves like a regular antenna. When the system needs to find the AoA it switches the antenna configuration in a fast burst fashion and collects one sample at each configuration. This burst switching repeats at a time period defined by the user. Due to the extremely high-speed switching rate of pin diodes, the antenna can stay in the default omnidirectional state for the majority of the time, which can lead to uninterrupted communication. We evaluate the communication quality later in the evaluation section.

4.3.6 Directional Code to AoA Mapping

Sirius embeds directional codes by dynamically switching the gain pattern of the antenna. These codes are unique for each angle and thus map to a specific angle-of-arrival. In our design process, we first design the antenna in HFSS simulation and optimize the diversity of all gain

pattern configurations. In an ideal system, we should use the simulated gain patterns to learn and store the mapping function from the extracted codes to angles. However, due to imperfections in real-world fabrication process and assumptions in antenna modeling, the simulated gain patterns might not match with the measured gain patterns, which leads to errors in AoA estimation.

Therefore, in *Sirius*, we do a *one-time* calibration and record the real gain patterns of the antenna. We perform the calibration only once, which can be used for all the other antennas fabricated with the same parameters and materials. In other words, for *Sirius* to become a product, the manufacturer only needs to calibrate once and produce multiple antennas using the same calibration data.

There are multiple ways to store and learn the mapping function from extracted codes to angles. For example, one can use look-up tables, euclidean distance-based estimation, SVM or neural networks to find the angles. In our experiments, we compared the euclidean distance-based method and the neural network-based method. We found that neural networks are robust to environmental changes even though they consume more power for computation. We also show the performance results of euclidean distance-based system in section 4.4.5.

In *Sirius*'s prototype, we train a small multilayer perceptron with 4 fully connected layers of size 16x8x16x1. Each layer is followed by a relu activation function to capture the non-linear gain patterns and the last layer is a softmax regression layer to estimate the angle in degrees. To reduce the power consumption of computation, we optimize the inference operations for a low-power microcontroller. We use the Low-Energy Accelerator (LEA) library provided by MSP430 [243]. The library supports dot multiplication of matrices, which accelerates forward propagation. We also developed a tool in Matlab that can take any trained model as an input and produce a header file in C containing the architecture of the model, the learned weights and biases, and the functions

required for the angle inference. We will also open-source this tool for the community along with the other source codes of *Sirius*.

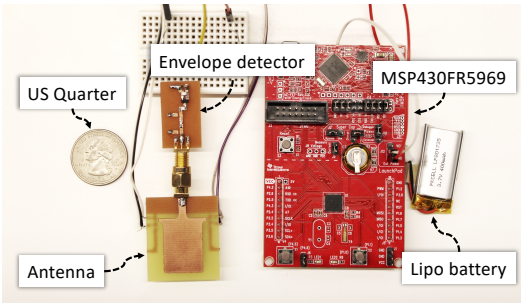


Figure 4.12: *Sirius*'s prototype for localization.

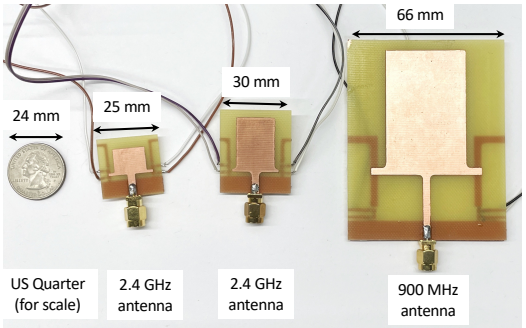


Figure 4.13: We evaluate *Sirius* on different antennas in ISM band 900MHz and 2.4GHz. The figure shows the fabricated reconfigurable antennas used in the prototype.

4.4 Evaluation

To this end, we implement a prototype of *Sirius* and evaluate the key components.

Node setup: Figure 4.12 shows the implemented prototype of our mobile node. It consists of a reconfigurable antenna, an envelope detector, and a low-power microcontroller. The microcontroller takes care of switching the gain pattern of the antenna and sampling the output voltage of the envelope detector. We have used MSP430FR5969 [244] as our microcontroller due to its low power footprint and optimized LEA library [243] for matrix multiplication. Figure

Figure 4.14: Setup for outdoor long-range data collection. The map shows the static anchors and node locations.

4.13 shows the lineup of fabricated antennas centered at 900MHz and 2.4GHz. The antennas are designed and optimized using HFSS simulation software [245] and etched on FR4 PCB. Each antenna consists of two pin diodes, one on each side of the antenna, acting as RF switches. The circuit diagram of the RF switch is shown in Fig 4.10, where R1 is a 100ohm resistor, L1 and L2 are 200nH chokes, and C1 and C2 are 100pF DC blocking capacitors. We used the Infineon Bar50-02v [242] as the pin diodes since they have a very fast switching speed and low insertion loss. The antenna is directly connected to the envelope detector which is a single-stage dickson rectifier consisting of two Skyworks SMS7630 schottky diodes [239]. To increase the efficiency of the envelope detector, an impedance matching network is used to match the input impedance of the circuit to 50Ω impedance of the antenna.

Anchor beacon setup: In the performed experiments we have used the following configuration. We use the SDR, USRP N210, as our transmitting anchor beacons. The output of the SDRs are connected 3dBi omnidirectional antennas Ettus VERT900 [246] and VERT2450 [247] for 900MHz and 2.4GHz experiments respectively. We use a power amplifier RF5110G [248] and tune the output power to 30dbm which is within the allowed power limits by FCC.

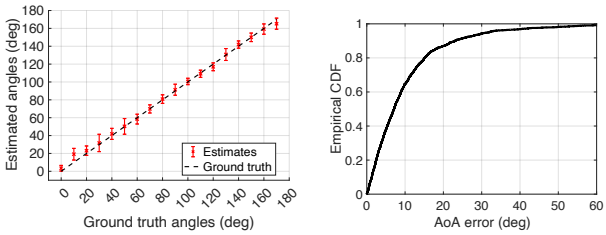


Figure 4.15: (a) Overall AoA estimation accuracy of *Sirius*. (b) CDF of AoA error.

4.4.1 AoA Performance

To perform controlled AoA experiments, the antenna was mounted on top of a programmable turntable controlled by an arduino and a stepper motor. We used MATLAB to communicate with the turntable and control the precise rotation of the antenna. Finally, the output from MSP430 was communicated to MATLAB using a USB-to-serial interface and stored for evaluation. We performed the experiment outdoors for varying locations and environments throughout the measurements. Fig. 4.15(a) shows *Sirius*'s overall performance in estimating AoA. The figure presents the average estimated angle at each measurement along with error bars indicating standard deviation. 4.15(b) shows the cdf of the overall error. The median AoA error is less than 7 degree.

4.4.2 Localization

We evaluate the localization performance of *Sirius* in a long-range outdoor environment which is closer to real-world deployment. The experiments are conducted in an open field surrounded by multiple trees on the boundary, and a parking lot with multiple cars. The setup used in this evaluation is similar to the previous experiment. We place our mobile node on a portable cart to emulate a moving vehicle or a drone. Fig. 4.16(a) shows the setup in the open field. We place two AP in the field and move the cart to the marked locations shown in Fig. 4.16(b). The

maximum distance between the AP and the node was 26 meters. We collect the estimated locations for all the possible orientations of the receiving antenna at all the marked locations. Figure 4.16(c) shows the CDF of the estimated AoA error. The median error in location is less than 2.5 meters while the error in 1D localization is less than 1.3 meters.

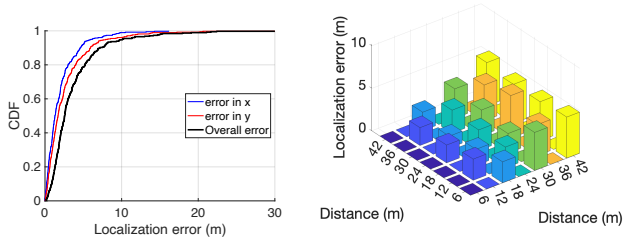


Figure 4.16: (a) CDF of localization error (b) Localization error per location shows the impact of distance from the APs.

4.4.3 Impact on RF Communication

We evaluate the ability of *Sirius* to support wireless communication along with angle-of-arrival estimation. *Sirius* uses a reconfigurable antenna that repeatedly changes the gain patterns of the antenna. This can lead to disruption in communication due to changes in the effective channel. To evaluate the impact of switching, we set up an experiment using our antenna as a receiver and a COTS 900MHZ antenna, VERT900 [246] as the transmitter antenna. Both antennas are plugged into the SDR USRP-N210 [249] interfaced with Matlab. The transmitter USRP device continuously transmits messages at a bit rate of 2Mbps and the receiver USRP calculated the total number of bit errors while addressing problems such as carrier frequency offset, timing offset, and frame synchronization [250]. The modulation used in this experiment was QPSK of order 4. While receiving the messages, we run *Sirius* system with varying rates of switching using an MSP430. Figure 4.17 shows the effect of beam pattern switching on BER. We observe that even

at switching speeds of 480 configurations per second, the BER is < 0.001. Figure 4.18 compares the BER of our antenna with the COTS antenna. We observe that our antenna running at 100Hz of AoA has a similar BER with the COTS antenna.

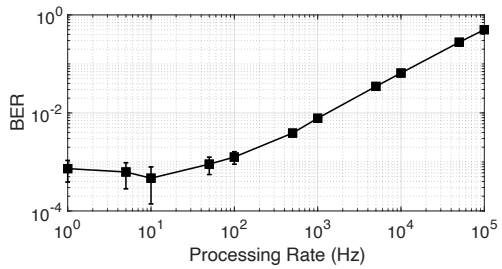


Figure 4.17: Effect of pattern switching speed on communication.

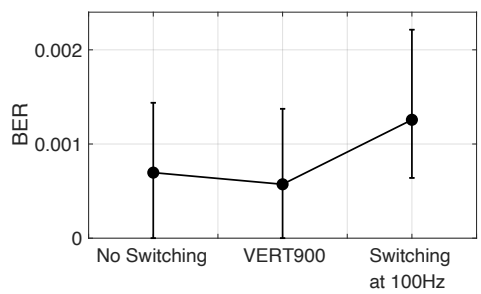


Figure 4.18: Impact of antenna reconfiguration on bit error rate for different antennas.

4.4.4 Power Consumption

In this section, we benchmark the power consumption of *Sirius*'s prototype. We measure the energy consumption of each sub-module of the system. Table 4.1 shows the breakdown of the energy consumption for one location estimation. The switching and sampling sub-module consists of the energy consumed by the pin diode-based RF switch and ADC and of the microcontroller. The beacon detection algorithm consumes the least energy and it consists of rearranging the samples of different antenna configurations and classifying the signal to corresponding anchor

beacons. The AoA estimation sub-module is the forward propagation of a floating point precision neural network. We run the AoA sub-module once for each beacon to get the x and y coordinates. The total energy consumed per location estimate is $36mJ$.

Submodule	Energy consumed
Switching and sampling	$0.32mJ$
Beacon detection	$0.09mJ$
AoA estimation	$19.50mJ$
Total energy per AoA	$19.81mJ$
Total energy per Location	$36.71mJ$

Table 4.1: Breakdown of energy consumption.

4.4.5 Robustness

In this section, we evaluate the robustness of *Sirius* in varying practical conditions like varying level of multipath, NLOS cases, number of gain patterns, reflections material, clock drift, frequency deviations, and distance from the anchor beacons.

Impact of multipath: *Sirius* assumes that the line-of-sight path (direct path) exists between the transmitting beacon and the receiving antenna. However, due to the presence of reflecting surfaces in the environments, the radio signals also reflect multiple times before reaching the antenna through other paths. Due to the loss of energy during propagation, reflections, and absorption, radio waves undergo exponential attenuation. Since the magnitude of attenuation is proportional to the inverse square of the distance traveled and the number of reflections, the energy contrast between the direct path of the signal versus the reflected paths is extremely high. We perform an extensive simulation study, using the RF communication toolbox in Matlab [251], in various multipath scenarios to analyze the energy distribution of the signal between the direct path rays and multipath rays. Figure 4.19 shows the distribution of the energy between these two types

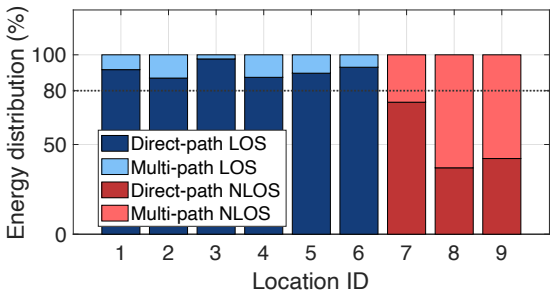


Figure 4.19: Energy distribution between direct path and multipath for various indoor locations.

of paths as percentages. We clearly see that in LOS cases when the direct path is unobstructed, more than 80% energy is concentrated in the direct path. However, in NLOS scenarios, the direct path becomes weaker and the contrast of energy decreases to less than 50%. This impacts the overall AoA estimation errors.

Impact of NLOS: To compare *Sirius*'s performance in LOS and NLOS scenarios, we simulate both scenarios. The LOS cases have a clear direct path between the transmitting beacon and the receiver, whereas NLOS cases have an obstacle like a wall or a furniture item between the direct path. Figure 4.20 shows the median errors in the AoA estimation for both scenarios. The results indicate that *Sirius* can estimate angles within 10 degrees of median error in all 6 LOS locations. However, in NLOS scenarios, where the reflections overpower the direct path's energy, the estimation errors increase up to 50 degrees.

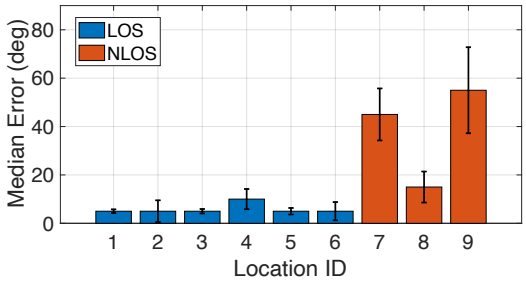


Figure 4.20: AoA errors for different indoor locations.

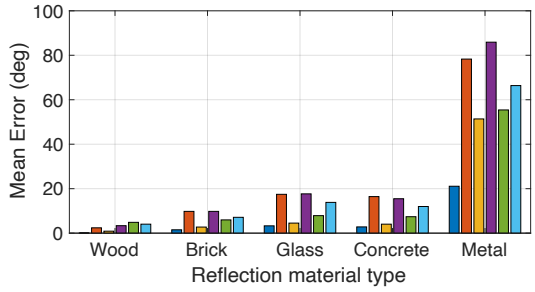


Figure 4.21: AoA errors for different reflection materials. Each cluster of bars represents one type of material, and each bar within a cluster represents different indoor locations.

Impact of reflection surface material: A significant portion of energy loss occurs during the reflection of radio waves from an object. The relative permittivity and conductivity of the material of these objects determine the amount of radio signal which gets absorbed and reflected. To study the impact of this material type, we simulate the five most common types of material found in an indoor environment and run simulations at various indoor locations. Figure 4.21 shows the mean AoA estimation errors for wood, brick, glass, concrete, and metal.

Impact of dynamic environment: Since the performance of *Sirius* depends on the movement of reflecting surfaces and multipath signals, we perform a real-world evaluation in three different levels of dynamic multipath, as shown in Figure 4.27. We conducted low dynamic level experiments in an outdoor environment, adjacent to a glass building with 2-3 people moving around. Medium dynamic-level experiments were carried out in an indoor hall with 5-10 people around, while high dynamic-level experiments were also conducted in the same indoor hall but with a large crowd (¿20 people) moving around naturally. Figure 4.23 shows the mean AoA error in increasing level of dynamic levels. We clearly see that when the environment is less dynamic, the AoA error is low with limited standard deviation. However, in more dynamic conditions, where the reflection paths fluctuate in close proximity, the mean AoA errors increase up to $20\ degrees$.

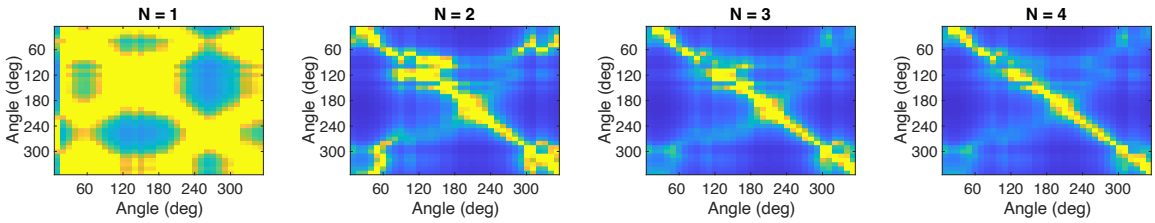


Figure 4.22: Inverse distance matrix for varying number of gain patterns (N).

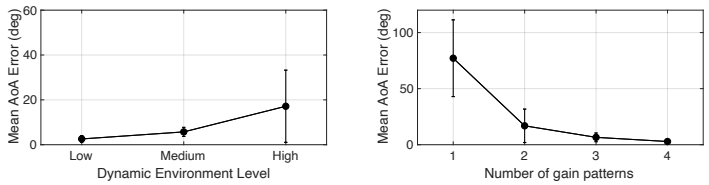
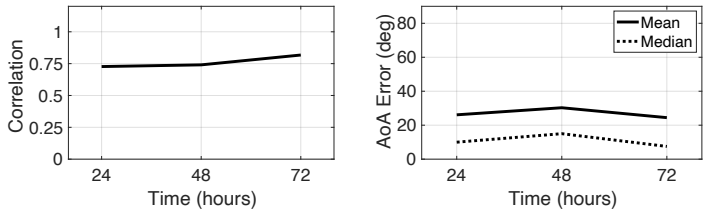


Figure 4.23: AoA estimation errors. (left) varying levels of multipath density in the environment. (right) varying number of gain patterns.

Impact of number of patterns: *Sirius* aims to achieve ideal diversity by optimizing for a diagonal distance matrix of gain-pattern features. In this experiment, we investigate the impact of the number of switches and gain patterns on overall diversity and AoA estimation accuracy. Specifically, we examine the correlation between the number of gain patterns, diversity, and AoA estimation errors. As depicted in Figure 4.22, we observe that increasing the number of reconfigurable gain patterns leads to a larger distance between the features and nearby angles. Notably, the confusion matrix displays a faint diagonal, indicating ambiguity in 180-degree apart angles, such as 120 and 300 degrees (i.e., 120+180). This confusion is more frequent in angles where the on-off and off-on are similar. However, as we increase the number of switches, the ambiguity decreases, resulting in a more accurate estimation. Figure 4.23 shows the mean AoA estimation accuracy for varying number of gain patterns. The error sharply drops as we increase the number of patterns from 1 to 3 and then it slowly saturates as the change in diversity is not as much.

Long term stability: Figure 4.24 shows the temporal stability of the antenna gain patterns over 72 hours. Our experiment tested the system every 24 hours for 360 degrees of AoA estimation at 10-degree steps in an indoor environment with the natural movement of people and furniture. Figure 4.24 (a) presents the correlation metrics for gain patterns at varying time intervals, while Figure 4.24 (b) shows mean and median AoA estimation results. We conclude that the gain patterns are predictable and stable for long time intervals.



(a) Gain pattern correlation (b) AoA estimation results

Figure 4.24: (a) Correlation of gain pattern amplitudes. (b) Mean and median AoA errors showing the long-term stability of gain patterns in a dynamic environment.

Indoor vs outdoor: We also compare the performance of AoA estimation in indoor vs outdoor setups. We record the gain pattern of the antenna at multiple indoor locations to see the effect of multipath on the gain pattern. Figure 4.25 shows an example of the gain patterns in two different indoor locations. We show the setup for this experiment in Figure 4.27. We also evaluate the AoA performance in indoor vs outdoor setups. Figure 4.26 shows the CDF of AoA error. We clearly see that *Sirius* performs equally well with a median AoA error of 7-8degrees in all environments.

Impact of distance: Increasing the distance between the anchor beacon and the node changes the incoming signal strength and the absolute gain pattern. We evaluate its impact on the AoA performance. Figure 4.28 shows the median error in AoA for varying distances. Even though the absolute gain pattern changes, the relative gain pattern still remains the same. The median

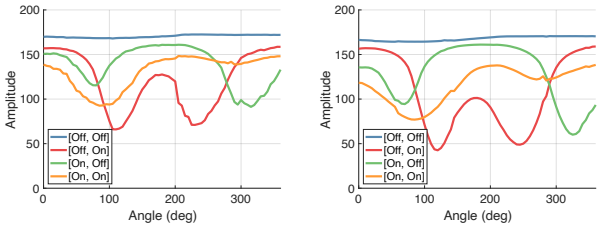


Figure 4.25: The figure shows gain patterns collected at two different indoor locations in a dynamic environment.

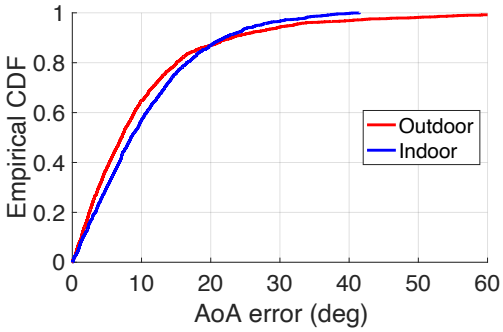


Figure 4.26: CDF of AoA error for different environments.

error is 7 degrees with a standard deviation of 3 degrees.

Impact of clock drift: We evaluate the effect of unsynchronized clocks on AoA error. To see the impact of clock drift we collect data over a period of five hours without any external reference clock. We find the average drift as 5-7ppm. Figure 4.29 shows the confusion matrix of euclidean distance-based estimation between the two sets of gain patterns collected in five hours. We also show the spectrum of an example test signal where we clearly see the peak.

Figure 4.27: Our experimental setup at different indoor locations in varying levels of static and dynamic multipath.

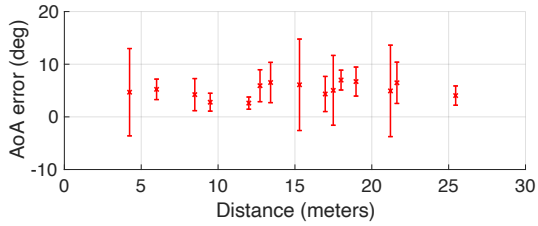


Figure 4.28: AoA error for varying distance from the anchor.

Impact of frequency shift: Perceived frequency of a node can be different from the transmitted frequency due to various reasons like doppler effect, carrier frequency offset, or clock drift. In this evaluation, we test the impact of any frequency shift on the gain pattern. We collect a set of gain patterns for 2.400GHz and 2.401GHz (1MHz difference). Figure 4.30 shows the confusion matrix between the two frequencies. We see a strong correlation between the two gain patterns. Even a euclidean distance-based estimation gives us a 10deg median error.

4.5 Limitation and Future Work

NLOS scenarios: Despite its promising performance, *Sirius* has a limitation of assuming the existence of a line-of-sight (LOS) path between the transmitting beacon and the receiving antenna. This assumption may not always hold true due to the presence of obstacles in the environment. To overcome this limitation, one possible solution is to extract the channel impulse response and pick the first peak corresponding to the direct path signal. However, implementing this solution in low-power hardware would require significant innovation, which we leave to future work.

Dense multipath: Through our experiments in dense indoor environments, we have found that *Sirius* can successfully estimate the angle of arrival if the direct path energy accounts for at least 80% of the total energy. However, this condition may not hold true in scenarios with

extremely closely placed reflectors or near-perfect reflecting surfaces, where the proportion of direct path energy may decrease, leading to higher AoA estimation errors.

Accuracy and power consumption tradeoff: While increasing the number of gain-pattern configurations improves diversity and accuracy, it also requires more pin diode switches, leading to increased power consumption. Alternatively, one could employ low-power methods to configure antennas such as switching feed points or polarization [252, 253]. Our evaluation results demonstrate that, with up to 480 reconfigurations per second, the bit error rate (BER) of a communication link remains below 0.001.

4.6 Related Work

We have summarized the related works in Table 4.2.

Past work	Self-localization	Single-antenna node	Single-antenna anchor	Low-power	Sync-free	Accuracy
[222, 254–256]	×	✓	✓	×	×	50 cm
[224, 225, 236, 257]	×	✓	×	×	✓	6 degrees
[258]	✓	×	✓	×	✓	7.5 cm
[176]	✓	✓	×	✓	×	4.6 degrees
Sirius	✓	✓	✓	✓	✓	7 degrees

Table 4.2: Summary of related work.

RF Localization: The topic of RF Localization has received a lot of attention from the academic community over the past years. Most of the works either rely on either ranging [222, 254–256] or angle-of-arrival [224, 225, 236, 257] to localize a node. However, achieving good accuracy has required complex systems architectures like large antenna arrays or tight time

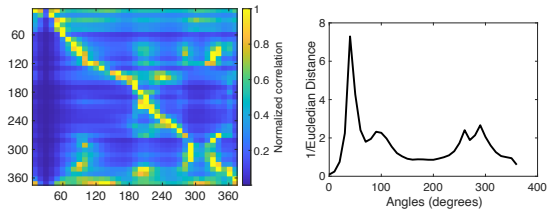


Figure 4.29: Impact of clock drift on AoA estimation

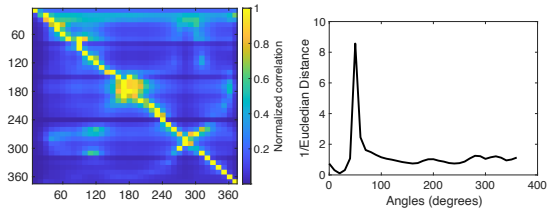


Figure 4.30: Impact of frequency shift on AoA estimation

synchronizations for extracting phase information. The AoA resolution is limited by the number of antennas in the arrays and their separation distance. Some papers [254, 259, 260] have emulated an array of virtual antennas by leveraging the motion of the node. *Sirius* proposes a miniaturized *single-antenna* system that can estimate AoA with multiple APs without requiring any kind of synchronization. *Sirius* uses a passive envelope detector and low-power ADCs to sample the energy of the signal. μLocate [175] and TagFi [261] backscatter RF signals from a low-power tag and receive from one or more APs to estimate the location. This provides an energy-efficient way to localize multiple low-power nodes but the location is estimated on the AP side which limits the scalability. Moreover, the SNR of backscatter signal is weaker compared to active radios which limit its application to small-range indoor environments. RFID-based works like RF-IDraw [262] use antenna arrays to find the angle-of-arrival of the tag and RFind [263] shows that RFID tags can backscatter wide bandwidths, up to 220MHz, to achieve fine-grained resolution ranging. While RFID tags consume less power they do not have the resources to do self-localization and their ranges are limited to 3-6 meters.

RF Self-localization: Living IoT [176] designs a self-localization system that can be attached to bumblebees. Similar to *Sirius*, LivingIOT uses multiple APs and angle-of-arrival to find the location of the node. However, the technique uses AP beamforming at the AP which requires multiple antennas. Moreover, all the APs must be time synchronized with each other and with the nodes to avoid interference. *Sirius* uses a single antenna and does not require any synchronization. This allows the system to be easily deployed in remote locations. *Sirius* also differs from [176] in localization latency and fast mobility. LivingIOT system takes $100ms$ to sweep the AP beams and one location. This limits the node's mobility relative to the APs. *Sirius* tackles this problem by using high-speed pin diode-based switches which can go up to 10^3 AoAs/second without trading off the communication link. ISLA [258] proposes a self-localization system that leverages the high bandwidth of 5G signals to perform better ranging. It leverages MEMS acoustic resonators instead of high-speed ADCs to improve the ranging resolution. ISLA differs from *Sirius* as it is a time-of-flight-based technique whereas *Sirius* is an angle-of-arrival-based technique. ISLA nodes have two receiving antennas on the node and consume high power because of the bulky RF front end. In *Sirius* the nodes have a single antenna and an envelope detector and use no power-consuming front end like clocks and down converters, which dramatically reduces energy consumption.

Reconfigurable antennas: The applications of reconfigurable antennas is multi-fold. An antenna can be reconfigured in frequency [264, 265], gain pattern [266–268], polarization [269–271] or a combination of all [272–275]. In prior works, researchers have used multiple ways to reconfigure antennas. [276–279] use actuators and motors to physically change the structure or shape of the antenna. [265, 268, 280] uses electrical switching to connect or disconnect patches to the radiating surface of the antenna. RF-MEMS switches [281–284] use mechanical movement

which gives good isolation and insertion loss, however, they have high power consumption. Pin-diodes based switching [264, 265, 268, 285, 286] offer low power footprint and are very fast, in order of nanoseconds. The idea to use reconfigurable antennas for angle-of-arrival estimation has been explored in [280, 287–289]. However, existing works require sampling of IQ samples of the signal. Whereas, *Sirius* uses an envelope detector to only sample the energy of the signal. Moreover, most of the system implementations are limited to simulation models and anechoic chamber experiments. [285] is the closest work with *Sirius* in terms of antenna design since both the works use monopole antenna design and pin diodes to switch the configuration. However, [285] is a composite frequency and pattern reconfigurable antenna tuned to operate at 3.1 GHz and 6.8 GHz. In our prototype, we design the antenna at 900MHz and 2.4GHz ISM bands.

4.7 Conclusion

This chapter presents *Sirius*, a low-power self-localization system. The core idea of the chapter is to embed directional codes to any received signal by dynamically switching the gain pattern of an antenna. The received signal carries the code in amplitude fluctuations which uniquely maps to the angle-of-arrival. *Sirius* develops a robust AoA estimation technique using envelope detector-based receiver and creates a prototype for demonstration and performance evaluation. Our evaluations show a median AoA error of less than 7degrees of localization error of 2.5meters while consuming only 19mJ and 36mJ of energy for the respective modules. The chapter breaks away from traditional array-based AoA methods and opens up new possibilities for low-power IoT nodes.

Chapter 5: Scalable Asset Tracking with NextG Cellular Signals

5.1 Introduction

Cargo theft is one of the major perils to supply chains accounting up to $30 billion in cost per year to transport and retail businesses [290, 291] adding to inflation and economic crisis. Existing asset tracking solutions, that focus on large containers or transport vehicles for cost-effectiveness [292, 293], remains mostly ineffective as the most common method is 'pilfering' or stealing a part of the cargo than the entire shipment. Organized crime proliferates in this space leveraging the gap in technology for long-distance asset tracking at smaller granularity and a lower cost. On the other hand, a diverse set of other applications await a maintenance-free solution for wide-area location tracking, including livestock monitoring, warehouse logistics, and locating children, pets, and people with mental illness [294, 295]. A location tracking tag of size of a coin cell that remains operational for years and capable of monitoring self-location across all roads in the country with little infrastructure cost seems to be a fitting technological answer. *But is it even possible to meet all these requirements in a practical asset tracking solution?*

Asset tracking, geofencing, and location-based services are a major part of the growing industry that targets smart commercial and personal application spaces [296]. The solutions generally consist of a tag that contains a localization module to estimate its current location and a communication module to periodically update the location status to the monitoring server [297].

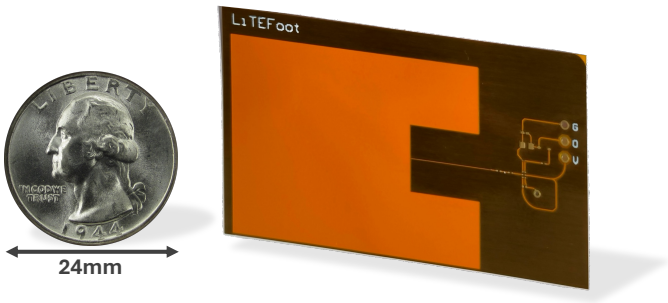


Figure 5.1: *LiTEfoot*- an ultra-low-power wireless tracker next to a US quarter for scale.

A significant majority of the solutions relies on Global Positioning System (GPS) for location estimation for its infrastructure-free self-localization [298, 299]. However, GPS tracking over a long period of time is challenging due to its high power consumption. [300] A GPS-enabled tag, therefore, often comes with sizable battery pack making its overall form factor to be similar to a smartphone and weights over 200 gram [301], which makes it challenging to incorporate in a low-power device. Moreover, traditional GPS-based localization methods, while effective in open-sky environments, falter in urban canyons and dense cities with skyscrapers [302].

The need for small, low-cost, and low-power alternatives for asset tracking has led to the next dominant techniques for commercial tags which implements proximity-based localization [303, 304]. In this design the tag can not self-localize, rather it uses a short range communication with nearby infrastructure which in turn estimates the location of the tag [305]. For instance, location of an RFID-enabled tag can be predicted when it comes in range of RFID readers deployed at specific locations in a building [306]. Naturally, the infrastructure costs of such solutions are suitable for small controlled environments. Apple AirTag expands the scope of proximity-based location using the 'Find My' network of Apple as the infrastructure [307]. Here the tag basically functions as a Bluetooth Low Energy (BLE) beacon to interact with mobile device in range and the phone updates the tag's approximate location on the server using its own GPS. Unfortunately,

its location fixes are unpredictable across long distances as it is conditioned on the presence of a GPS-enabled Apple device with in a few meters of range [305]. While the proximity-based location techniques can lead to a very low-power and small tags for the simplicity of functionality, it offloads the responsibility to infrastructure. The lack of standalone location estimate makes it infeasible for asset tracking over large geographic regions.

Global coverage of cellular infrastructure has opened an interesting possibility of wide-area localization without investing on dedicated infrastructure. Popularity of 4G-LTE and enthusiasm around 5G networks have led to a widespread cellular coverage around the world. In the United States, mobile networks cover more than 98% of the population and over 90% of road networks, as reported by the FCC [308]. Several research works [309–311] have explored this opportunity for nation-wide localization techniques [300] and multiple commercial products are available cellular localization on smartphones and with tags [312]. However, these systems encounter following two critical challenges in adapting cellular localization on a low-power platform.

(a) First challenge is the latency and computational overhead of scanning wide bandwidth. Active cellular bands are spread across several GHz of bandwidth and naturally, nearby cells use different bands. A self-localizing tag aiming to identify signals from all available cells will require to scan the entire bandwidth. Low-cost narrow-band radios operating in tune-and-capture mode can take tens of minutes to complete one round of scanning and not applicable to dynamic scenarios where the tag can move to different cells within that time [313]. A real-time localization of a tag moving on a vehicle in an urban area will require the tag to scan the spectrum within a few milliseconds. (b) The second challenge is the downconversion of the passband signal for synchronization and cell identification. Traditional radios require power-hungry local oscillators at high-frequencies and sophisticated mixing hardware to convert the signal to baseband before

applying correlation on synchronization signals [226, 314]. This operation can itself takes tens of milliWatts of power, which makes the tag last for only a few hours on a battery source like a coin cell.

In this chapter, we proposed an architecture for passive cellular localization within the power limit of 1 millijoule and with location update the latency of around 10 milliseconds. This radio architecture exploits the core opportunity from the non-linear transform of the cellular downlink broadcast signals. We explore the behavior of the subcarriers in the cellular synchronization signals under this transformation and show methods to recover cell identity from the systematic superposition of intermodulated frequencies, which we call spectrum folding. Our low-power radio frontend uses a simple envelope detector to realize spectrum folding and onboard algorithms can identify any number of cells operating at any part of the 3GHz spectrum. We build on radio to realize a localization system leveraging the ambient 4G-LTE signals. We call this system *LiTEfoot*. The prototype achieves a median localization error of 22 meters in urban areas and 50 meters in rural areas. It senses a 3GHz wideband LTE spectrum in 10ms using non-linear intermodulation, consuming 40 J of energy per location inference in CMOS simulation, and 0.8 mJ in our PCB prototype. This results in a 625$\times$ reduction in energy consumption for the CMOS design and a 27$\times$ reduction for the PCB prototype, compared to GPS, which consumes 25 mJ per location inference [315].

LiTEfoot presents a new class of low-power receivers tailored to cellular synchronization signals (see Fig 5.1). We aim to offer a solution that ensures broad coverage and operates on minimal energy, opening up possibilities for applications requiring extensive geographic reach and energy efficiency. The core ideas of spectrum folding and the low-power radio can be adapted for applications in various low-power location-based services and also for channel synchronization

battery-less communication platforms.

The use of non-linearity for implicit signal mixing or passive downconversion is a proven technique. It has been explored extensively in various low-power applications to replace power-hungry oscillators in radio receivers [316–320]. Notably, *MIXIQ* [226] proposed a low-power 802.11ax receiver that modifies OFDM packets to send a helper signal, preserving IQ data for such implicit baseband downconversion. *Saiyan* [314] demonstrated a low-power, long-range receiver for frequency-modulated signals. These lines of work have shown remarkable power efficiency when custom base stations are available to broadcasting specially designed signals. Our work *LiTEfoot* is unique in leveraging non-linear transform with unmodified cellular signals. The core insight lies in a theoretical analysis of certain auto-/cross-correlation properties that survive non-linear intermodulation and baseband superimposition of a signal. *LiTEfoot*'s system design leverages this insight for identifying standard cellular synchronization signals (PSS and SSS) despite their noisy baseband mixture after the non-linear transformation. This novel architecture enables a single-antenna ultra-low-power radio, made from simple off-the-shelf components, to simultaneously detect multiple cell towers' identity without requiring to decode the cellular packets. By allowing non-linearity to intermodulate and fold the spectrum to baseband, *LiTEfoot* sacrifices IQ data but gains the ability to sense the entire 3 GHz LTE spectrum in just 10 ms, a capability crucial for low-power wide-area localization and wide-band spectrum monitoring as well.

While several opportunities exist in research and application, this chapter focuses on developing the core capabilities and assessing the limits of the systems. To this end, we have made the following three specific contributions:

- A novel method for signal searching over a wide frequency spectrum. We presented the theoretical foundation and the practical implementation for the non-linear transformation of spectrum and synchronization signal detection on the transformed narrow-band signal.
- Development of a signal strength multilateration-based localization with the low-power radio.
- Prototype development and evaluation of the system with real-world existing 4G-LTE infrastructure.

5.2 Cellular Networks Primer

A. Structure of a Cell

In cellular networks, "cells" are the fundamental geographic segments served by a stationary transceiver, called a base station or an eNodeB. In an LTE cellular network, base stations typically use 3-sector cells served by three antennas each covering 120° angular sectors of the cell for improved communication and capacity gain (see Fig. 5.2). While omnicells with single antennas covering 360° field are also common. LTE defines four types of base stations with different power and coverage areas – macro, micro, pico, and femtocells with up to 20 km, 2 km, 200 m, and 30 m of coverage respectively.

Cells are organization-level concepts and therefore different service providers can have overlapping regions of their cells. In fact, a physical cell site or cellular tower can be used by multiple service providers operating independent base stations each having one or more antennas serving different sectors of the corresponding cells. These base stations obviously operate on

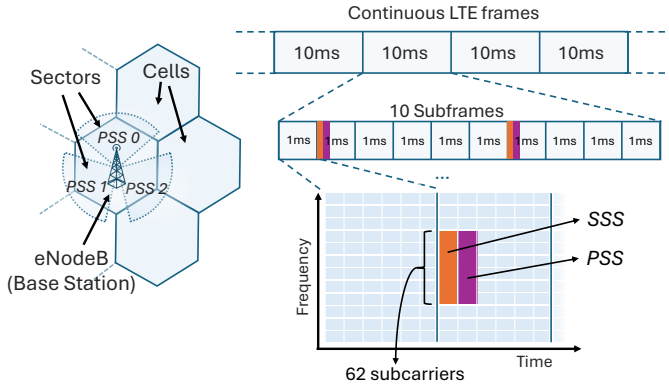


Figure 5.2: LTE cells and frame structure.

independent frequency bands and are non-interfering to each other while communicating with mobile clients called User Equipments (UE). However, signals from multiple base stations of the same or different service providers are often available at a physical location and can be observed by scanning across frequency bands.

Operational frequencies: LTE operational frequencies are organized in over 70 bands allocated with different bandwidths over a large spectrum ranging from 450 MHz to 3.8 GHz. Some service providers also use parts of their underutilized 3G bands for 4G LTE transmissions. Although, typical frequency bands used in the USA lie between the range of 700 MHz to 3.5 GHz. Within a frequency band, the bi-directional communication between the base station and the mobile device is defined by two different duplex modes – frequency and time division duplex (FDD and TDD). However, almost all service providers in the USA use frequency division duplex (FDD) modes where the uplink (i.e., transmission from the mobile device to the base station) and downlink (i.e., transmission from the base station to the mobile device) use separate frequencies predefined per band.

B. Cell Identity

Each individual cell in LTE networks can identified using the Cell Global Identity (CGI) which is a globally unique identifier [321]. While the CGI is useful at the network-level operations and management, at the physical layer mobile devices rely on the Physical Cell Identity (PCI) for cell identification. The PCI is a 9-bit number ranging from 0 to 504 that locally identifies each sector antenna. Naturally, this limited number of PCI values are reused over different cells, but the numbers are carefully assigned to distant cells during LTE site planning to avoid confusion between nearby cells [322]. For practical purposes, the PCI numbers can be successfully used for cell selection and handoffs.

Synchronization Signals: The PCI numbers are created by combining two numbers between 0-2 and 0-167 which are in turn used as seeds to two pseudorandom sequences, called Primary and Secondary Synchronization Signals (PSS and SSS), used for synchronization and PCI number identification. The Primary Synchronization Signal (PSS) utilizes the frequency-domain Zadoff-Chu sequence and is numbered from 0 to 2, while the Secondary Synchronization Signal (SSS) is formed using maximum length sequences (m-sequences) and numbered from 0 to 167. Mobile devices utilize these synchronization signals, which are emitted by base station antennas, to determine their PCI values. The PCI is calculated using the formula $PCI = 3 * SSS + PSS$.

C. Base Station Signal Format

The LTE physical layer maintains the continuous connection model, especially on the FDD downlink, which operates as a continuous stream. Apart from carrying user data, the downlink transmissions contain system information, network parameters, and control signals that

are essential for connection establishment and handoffs, and maintaining synchronization. The communication stream is organized as a time-frequency grid of resources, called a radio frame. Along the time axis, a frame has a duration of 10 milliseconds as shown in Figure 5.2. Each frame is subdivided into 1-millisecond subframes, which are further divided into two slots, each lasting 0.5 milliseconds. One slot contains 7 or 6 Orthogonal Frequency Division Multiplexing (OFDM) symbols, depending on whether the normal or extended cyclic prefix is used. Along the frequency axis, the frame length is defined by the number of OFDM subcarriers placed at 15 kHz spacing. The total number of subcarriers depends on the bandwidth of the channel, ranging from 72 in 1.4 MHz to 1200 in 20 MHz channel bandwidth. For the FDD downlinks, the PSS signal is broadcasted at the central 62 subcarriers of the last symbol of time slots 0 and 10, and the SSS signal is broadcasted one symbol before PSS.

5.3 System Design

In this section, we will discuss in detail the design of *LiTEfoot* and the technical contributions.

A. Intermodulated Spectrum Folding

Downconversion, a critical process in RF architecture, traditionally relies on a mixer to lower high-frequency signals to the baseband by multiplying them with a carrier wave generated by a local oscillator. However, an alternate approach involves utilizing a squaring operation-multiplying it by itself to achieve downconversion without the need for an external carrier. This technique facilitates simultaneous wideband mixing and eliminates the need for a traditional oscillator, which simplifies the overall RF design. By removing the oscillator, the system becomes less complex,

making it an alternative to modern RF architectures.

We introduce a technique called ”intermodulated spectrum folding” that is built upon the distinct properties of LTE synchronization signals (PSS and SSS), which consistently exhibit a narrow bandwidth of 1.4 MHz (1.08 MHz without guard bands) across all LTE frame bandwidths, ranging from 1.4 MHz to 20 MHz. By implementing a squaring operation, this method isolates and folds LTE signals from various bands into the baseband. This process results in a superimposition of signals through both in-band and cross-band mixing, as illustrated in Figure 5.3. These superimposed signals can then be selectively extracted using a narrowband filter.

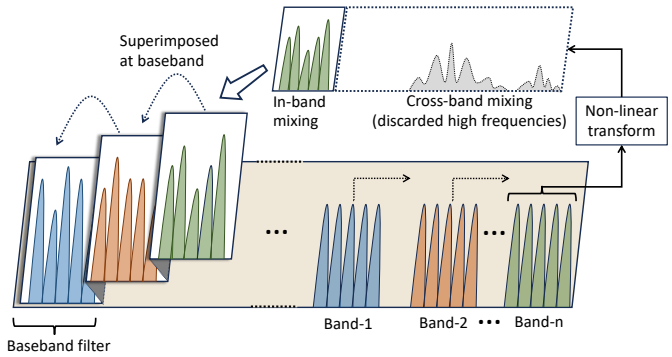


Figure 5.3: Intermodulation and spectrum folding.

Before reaping the benefits of intermodulated spectrum folding, it is pivotal to address its inherent challenges, notably the side effects of intermodulation and signal intermixing. For the effective implementation of spectrum folding in *LiTEfoot*, three critical properties must be ensured: (1) *Correlation Preservation Post-Intermodulation:* The technique must maintain the correlation characteristics of PSS and SSS after intermodulation, essential for accurate PCI estimation. (2) *Interference Mitigation:* The system must be designed to effectively withstand interference from adjacent carriers and bands that are folded into the baseband, ensuring that it can clearly distinguish the synchronization signals. (3) *Relative Signal Strength Extraction:* Despite signal squaring

and superimposition, extracting relative signal strengths from overlapping base stations is crucial, enabling applications such as multilateration for precise location estimation.

Solving these challenges is fundamental to the success of intermodulated spectrum folding as a strategy for LTE. The subsequent system design section will explore how *LiTEfoot* builds upon spectrum folding, addressing these challenges.

Envelop Detector as Non-linear Channel: The envelope detector is an efficient way to introduce non-linearity in RF circuits using a simple passive diode. Consider two sinusoidal signals, $sin(2\pi f_1 t)$ and $sin(2\pi f_2 t)$, where f_1 and f_2 denote their respective frequencies. Let $y(t) = sin(2\pi f_1 t) + sin(2\pi f_2 t)$ be the incident signal at the envelope detector. The envelope detector subjects $y(t)$ to a non-linear squaring operation producing

$$\begin{aligned} y^2(t) = \frac{1}{2}[&\cos(2\pi(f_1 - f_2)t) - \cos(2\pi(f_1 + f_2)t) \\ &+ 2 - \cos(2\pi 2 f_1 t) - \cos(2\pi 2 f_2 t)] \end{aligned} \tag{5.1}$$

The output signal $y^2(t)$ produces a baseband component $f_1 - f_2$, alongside the harmonics $2f_1, 2f_2$ and $f_1 + f_2$.

Non-linearity in OFDM signals: Extending $y(t)$ to an OFDM signal consisting of N subcarriers, each centered around a carrier frequency f_c, we can write it as:

$$y(t) = \sum_{n=1}^{N} sin(2\pi(f_c + \Delta f_n)t), \tag{5.2}$$

where Δf_n is the frequency offset of the n-th subcarrier from the carrier frequency. The squaring operation results in a sum of products of all pairs of subcarriers, which, when expanded, produces

terms like:

$$y^2(t) = \frac{N}{2} + \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N}[\cos(2\pi(\Delta f_n - \Delta f_m)t) \\ - \cos(2\pi(2f_c + \Delta f_n + \Delta f_m)t)], \tag{5.3}$$

The output includes baseband components, $(\Delta f_n - \Delta f_m)$, independent of the carrier, and components at twice the carrier frequency, $2f_c$. To isolate the baseband components, we simply use a low-pass filter, eliminating all the higher frequency terms. Finally, the filtered output, $y^2_{lowpass}(t)$, retains only the baseband frequencies:

$$y^2_{lowpass}(t) = \frac{N}{2} - \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N}\cos(2\pi(\Delta f_n - \Delta f_m)t), \tag{5.4}$$

The generation of baseband frequency components results from the non-linear intermodulation of the subcarriers. Essentially, it contains the frequency differences of the incoming signals, effectively downconverting or 'folding' the higher frequency elements into the baseband region.

B. Signal Detection on Folded Spectrum

Mobile devices rely on correlation with the cell-specific synchronization signals, PSS and SSS, for initial cell identification and frame timing alignment. The pseudorandom sequences used in these signals show excellent auto-correlation for noise resilience and controlled cross-correlation property for robustness to the efficient separation of signals from different cells and support the unambiguous identification of the PCI. Spectrum folding intermodulates the subcarriers of the PSS and SSS signals and superimposes signals from different cells at the baseband. In this section,

we show that these signals sufficiently maintain these properties for successful cell identification in practice.

Correlation properties of PSS^2: After the non-linear operation, the number of unique PSS in LTE reduces from three to two. Specifically, the Zadoff–Chu sequences for PSS_1 and PSS_2, originally distinct due to their respective roots $q = 29$ and $q = 34$, selected for their good auto and cross-correlation properties under various timing and frequency offsets [323], become indistinguishable after undergoing the squaring operation. This is because these sequences, when squared, produce identical signals, effectively merging PSS_1 and PSS_2 into a single, non-distinguishable entity. This reduction, however, does not compromise *LiTEfoot*'s performance in PCI estimation. PSS signals are primarily utilized for synchronization offset estimation, while the SSS serves as the key determinant for actual PCI estimation. Therefore, despite the non-linear transformation's effect on PSS orthogonality, the integrity of PCI estimation remains unaffected, ensuring robust signal identification and time synchronization.

Correlation properties of SSS^2: An intriguing aspect of SSS signals is their resilience in maintaining excellent cross-correlation properties even after undergoing non-linear transformations. The SSS in LTE systems employs M sequences, renowned for their wide and flat spectrum due to their pseudorandom nature [324]. The non-linear transformation of SSS preserves the flatness and wideband characteristics which helps in maintaining good correlation properties even after the spectrum is intermodulated. This persistence of correlation is the key in *LiTEfoot* ensuring the robustness of our PCI estimation.

PCI Estimation after Spectrum Folding: To estimate PCI from intermodulated signals, we employ a time-domain correlation approach similar to a matching filter. This method extracts all the PCIs by correlating the received signal with squared versions of the predefined PSS and

SSS:

$$PCI = \{i \mid \text{corr}\left(x, PSS_i^2 \oplus SSS_i^2\right) > thresh, i = 1, ..., N\} \tag{5.5}$$

Here, $\oplus$ is the concatenation operation and x is the received signal. The PSS and SSS have a constant bandwidth of 1.08 MHz which is crucial for maintaining signal integrity across LTE's variable bandwidth scenarios.

We experimentally validate the correlation performance of the intermodulated signal (SSS^2) for all 504 PCIs in LTE across varying bandwidths, as shown in the confusion matrix in Figure 5.4. Our analysis confirms that the non-linear squaring operation preserves the orthogonal nature of the PSS and SSS signals, maintaining their unique cross-correlation and auto-correlation properties. Notably, we observe a *mod 4* ambiguity in the correlation spectrum, seen as secondary peaks at the ± 4 indices relative to the original index. These secondary peaks, while present, exhibit significantly lower correlation amplitudes compared to the primary peaks. Importantly, this *mod 4* ambiguity does not compromise *LiTEfoot*'s ability to accurately estimate the PCI, as our evaluation confirms that the highest correlation values consistently align with the correct PCI.

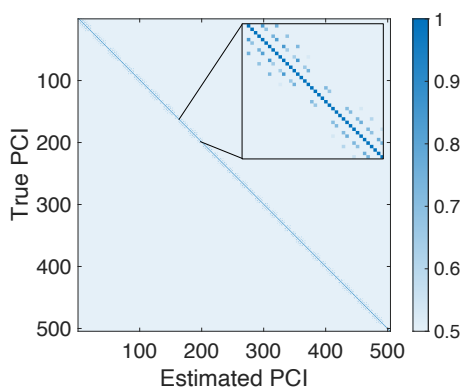


Figure 5.4: Confusion matrix for PCI estimation using SSS^2 (i.e., the SSS signal after non-linearity).

C. Dealing with Folding Interference

Despite *LiTEfoot*'s efficient downconversion capabilities through passive intermodulation, its use of an envelope detector introduces notable challenges in frequency selectivity. Unlike active radios, which have the ability to distinguish and selectively filter frequencies before downconversion, the envelope detector indiscriminately processes every frequency that reaches the antenna. This broad, non-selective approach allows for intermodulation where multiple subcarriers mix with synchronization signals, as detailed in Equation 5.4. This intermodulation introduces unwanted signal components which degrade the overall signal quality. Such interference is particularly detrimental as it can significantly deteriorate SSS correlation. In this section, we outline the three primary sources of interference— inter-subcarrier, inter-band, and inter-synchronization signal interference—and propose specific strategies to mitigate these challenges and enhance system performance.

(1) Inter-subcarrier interference: In LTE networks, which have bandwidth options ranging from 1.4-20 MHz, a challenge emerges with the increase in bandwidth. Despite the constant bandwidth allocated for PSS and SSS, the use of higher bandwidth configurations by cell towers—aimed at supporting an increased user base and higher data rates—introduces inter-subcarrier interference.

As the bandwidth increases, the presence of data subcarriers also increases, leading to a scenario where these subcarriers begin to overshadow the PSS and SSS signals. The interference becomes problematic after non-linear transformation, when the data subcarriers fold on top of synchronization signals aggressively compromising the detectability of synchronization signals. For example, the correlation accuracy drops from 99.7% in a 1.4 MHz LTE frame to 1.5% in a 10

MHz frame.

To solve this challenge, we leverage the inherent periodicity of PSS and SSS (10 ms in LTE). Considering the variability of data subcarriers over frames, we model their amplitude distribution as Gaussian. This is our core intuition for frame stacking. As shown in Figure 5.5, we exploit the sync signal's periodic stability against the background of fluctuating data subcarrier noise. By stacking just 100 frames—a span of 1 second—the Signal-to-Interference-and-Noise Ratio (SINR) of the sync signal is significantly enhanced, improving the accuracy of 95%.

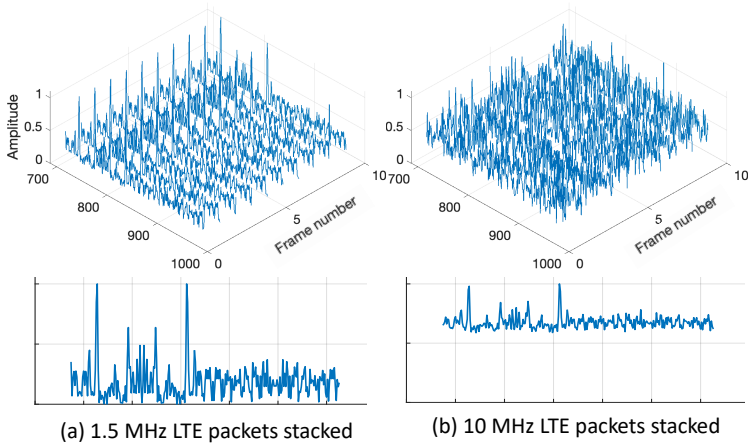


Figure 5.5: LTE packets before and after stacking (a) 1.4 MHz bandwidth (b) 10 MHz bandwidth. After stacking the PSS and SSS strength increases in the 10 MHz case whereas the data fades into a DC bias.

(2) Inter-band interference: Consider two cell towers operating at different bands with center frequencies f_{c1} and f_{c2} and bandwidths Δf_1 and Δf_2. According to Equation 5.4, the output signal will consist of frequency differences from inter-band subcarriers, causing interference.

In LTE, PSS and SSS occupy the central 62 subcarriers, spanning a bandwidth of 1.4 MHz. Except for the 1.4 MHz bandwidth option, which includes a 22.85% guard band, LTE bands typically incorporate a 10% guard band to mitigate interference. Given that the minimum frequency difference possible between sync signals from different cell towers will always exceed

1.4 MHz, we simply employ a low-pass filter with a cutoff frequency of $f_{cutoff} = 1.4$ MHz to remove all inter-band interference from the received signal.

(3) Inter-synchronization signal interference: PSS and SSS are designed with robust correlation characteristics, a feature that is preserved even after non-linear squaring operations. This resilience ensures that, despite the superposition of multiple synchronization signals, the accuracy of PCI estimation remains high. The squaring process, while introducing additional components into the signal, does not significantly disrupt the distinct correlation patterns of PSS and SSS. We also extensively evaluate *LiTEfoot* in the presence of multiple cell towers in the real world, which we discuss in the evaluation section.

D. Recovering Base Station Signal Strength

(1) Blind Separation of Weighted Superposition: A key challenge in multi-cell environments is to independently extract each cell's Received Signal Strength Indicator (RSSI) from the composite received signal, where individual cell contributions are overlapped and intermodulated.

To address this challenge, we develop a blind source separation algorithm. Our approach iteratively isolates each cell's contribution to the composite signal. The core idea is to reconstruct and subtract the time-domain PSS-SSS signal, for each identified PCI, from the received composite signal. We formulate this as an optimization problem:

$$E(A_i) = \frac{1}{N} \sum_{n=1}^{N} \left(x(n) - \alpha_i \cdot SyncSig_{PCI=i}(n - \delta t_i)\right)^2$$

$$\text{Minimize} \quad E(A_i)$$

$$\text{subject to} \quad 0 \leq A_i \leq A_{\text{max}}$$

Here, $x(n)$ is the received signal, $SyncSig_{PCI=i}(n)$ is the PSS-SSS signal for the i-th PCI, α_i is the amplitude factor, δt_i is the time delay in samples, N is the total number of samples, and A_{max} is the upper amplitude limit. Our objective is to minimize the residual energy $E(A_i)$ by adjusting α_i. The minimum residual energy indicates the most accurate subtraction, corresponding to the best estimate of the cell's RSSI, as shown in Figure 5.6(a).

The accuracy of this separation largely depends on correctly estimating both the amplitude factor α_i and the time delay δt_i. Initially, we assume δt_i is known from the correlation process used for PCI identification, though we next refine this with a super-resolution algorithm for phase-level synchronization.

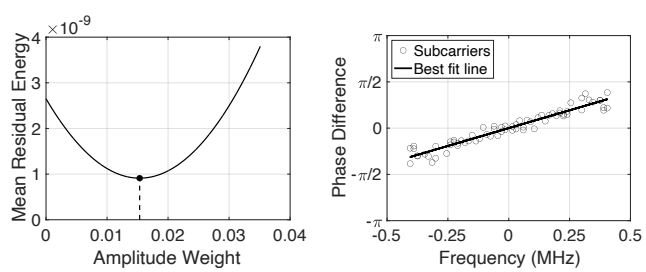


Figure 5.6: (a) Blind Separation of weighted superposition (b) Linear phase change introduced due to sub-sample offset in the OFDM subcarriers.

(2) Sub-sample Offset Estimation: Correlation-based synchronization can only provide sample level time-resolution [81, 222] which naturally depends on the sampling rate or the bandwidth of the receiver. Naturally, any sub-sample delay is missed out and results in poor amplitude estimation. We aim to resolve this sub-sample offsets to estimate the amplitudes accurately.

Fortunately, the cyclic prefix, with a duration of 4.7 microseconds for Frequency Division Duplexing (FDD) LTE, enables us to detect the sub-sample offsets by introducing a predictable phase change in the OFDM subcarriers of the signal. For a signal captured with $N = 62$ subcarriers, the frequency domain of the signal reveals this linear phase change across subcarriers. Figure 5.6(b) shows an example of linear phase changes in the subcarriers of PSS and SSS of a real-world LTE signal.

Let's say our wireless channel is $h = e^{-j2\pi f\tau}$, where τ represents the sub-sample delay, and f corresponds to the frequency of each subcarrier. Given a received signal $X(f)$ and a known reference signal $S(f)$ in the frequency domain, the phase difference $\Delta\phi$ attributable to the sub-sample offset can be calculated as:

$$\Delta\phi = \angle\left(\frac{X(f)}{S(f)}\right)$$

where $\angle$ denotes the phase angle of the complex quotient, indicating the linear phase shift across subcarriers due to sub-sample offsets. This phase difference directly relates to the sub-sample delay $\tau = \frac{\Delta\phi}{2\pi\Delta f}$, where Δf represents the frequency spacing between adjacent subcarriers. We correct for τ to accurately align the PSS and SSS in time before performing the amplitude estimation.

E. Localization Algorithm

(1) Multilateration with Relative Signal Strengths: Our localization approach combines open-source databases (OpenCellID [325], Mozilla Location Services [326], and CellMapper [327]) for PCI-to-coordinate mapping of cell towers with a trilateration technique based on relative signal strengths. We formulate an optimization problem to determine the tag's position (x, y) by

minimizing the discrepancy between observed and expected signal strength ratios from multiple cell towers.

Given n cell towers with known positions (x_i, y_i) and estimated amplitudes A_i, we define our objective function as:

$$f(x, y) = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \left(\frac{A_i}{A_j} - \frac{\sqrt{(x - x_i)^2 + (y - y_i)^2}}{\sqrt{(x - x_j)^2 + (y - y_j)^2}} \right)^2$$

This function compares the ratio of amplitudes $\frac{A_i}{A_j}$ to the ratio of Euclidean distances for each unique pair of cell towers, summing the squared differences. By using amplitude ratios rather than absolute values, we account for real-world signal propagation losses. We initialize the optimization with the centroid of all observed cell towers as our starting point.

Eliminating False Positives. Our system achieves near-perfect recall in PCI identification, leveraging the strong auto-correlation properties of PSS-SSS signals. However, this high sensitivity can lead to false positives. To address this challenge, we implement a refinement strategy that analyzes estimated PCIs along with their corresponding signal delays. We've observed that spurious peaks often cluster around delays associated with true cells, due to low-level correlations with other cells' signals. To filter these false positives, we group and sort signals by their estimated amplitude factors, selecting only those with non-zero amplitudes. This approach effectively balances high recall with improved precision in our PCI estimations, enhancing the overall accuracy of our localization system.

(2) Optimization for On-Road Applications: For applications such as vehicle tracking and urban mobility studies, we enhance our location estimation by incorporating road constraints. Our improved algorithm includes a road snapping technique [328], which aligns estimated positions

with nearby drivable roads using OpenStreetMap [329] data as a sparse graph. We employ a sliding window optimization process that starts with an initial point mapped to the k-nearest road points. For each subsequent estimated point, we identify k candidate points on nearby roads and apply a distance constraint based on the maximum possible travel. This constraint is calculated as the product of the time interval between points and the road's speed limit ($\Delta t \times maxspeed$). We retain only the candidate points that satisfy this distance constraint, defaulting to the nearest point on nearby roads if no candidates meet the criteria. This optimization produces more realistic trajectories for on-road applications by leveraging the constraint that tags predominantly travel on known road networks.

F. Extending *LiTEfoot* to Upcoming Standards

Dual sequence synchronization is a consistent feature in cellular networks that serves the basic two stage frame and slot synchronization used in standards starting from beginning (2G GSM Training Sequence [330], 3G UMTS P-SCH/S-SCH sequence [331], and PSS/SSS sequence for 4G [332] and 5G [332]). *LiTEfoot* relies on the uniqueness of these code sequences after the non-linear transformation, as described in Equation 5.4 (Section 5.3). We have shown that *LiTEfoot* architecture holds correct with other cellular standards and benchmarked its performance for 5G-NR and NB-IoT in the Evaluation section.

5G-NR expands the PCI range to 1008 values, using 3 PSS values (m-sequences of length 127) and 336 SSS values (gold sequences of length 127) [333]. Despite introducing diverse frame structures, 5G maintains key similarities with LTE. Firstly, the frame and sub-frame durations remain consistent [333]. Secondly, while 5G allows dynamic placement of synchronization signals,

these signals maintain consistent location within subsequent frames and occupy 144 subcarriers with a 10ms repetition. After the non-linear transformation, the resulting baseband signal remains independent of the central frequency, requiring only an adjustment of *LiTEfoot*'s low-pass filter cutoff to accommodate 5G's subcarrier spacing.

NB-IoT, with its 15 kHz subcarrier spacing and 10ms frame length [334], is also supported by *LiTEfoot* by adjusting the filter cutoff and the wait-time for signal acquisition. The correlation properties of the code sequences naturally improve with each new standard to support denser deployment of the base stations. It leads to better post-transformation distinguishability of the sequences showing *LiTEfoot*'s potential as a universal, low-power localization system across current and NextG cellular standards.

5.4 *LiTEfoot* Prototype Implementation

Figure 5.7 illustrates the system overview, and Figure 5.8 presents the PCB prototype of *LiTEfoot*'s RF frontend. We've implemented a PCB prototype using low-cost COTS components and simulated a CMOS design to showcase *LiTEfoot*'s potential for further power reduction.

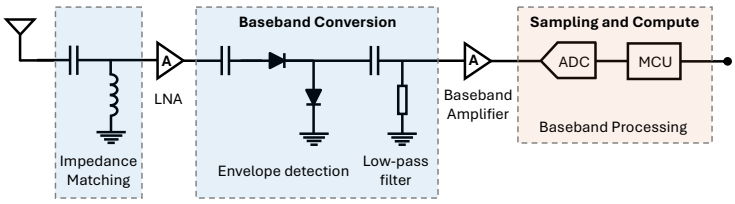


Figure 5.7: The high-level circuit schematic of *LiTEfoot*.

RF Frontend: We design a flexible patch antenna with a center frequency of 1.94 GHz covering LTE bands. Our RF frontend utilizes a tuned RF architecture [335], comprising of a low-noise amplifier (LNA), an envelope detector, and a high-impedance amplifier.

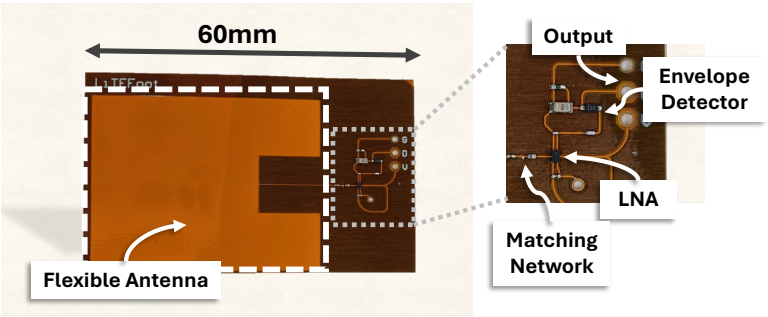
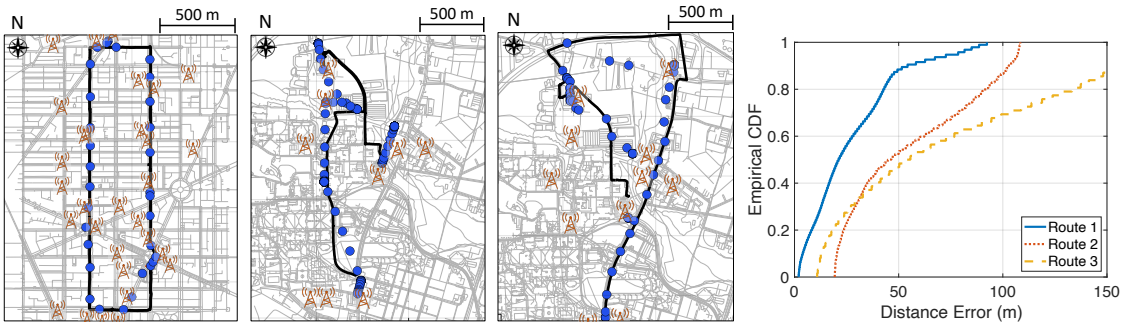


Figure 5.8: *LiTEfoot* PCB tag prototype showing the low-power RF frontend.

Envelope Detector: We implement a Greinacher voltage doubler circuit [336] using a Skyworks Schottky diode [239] for RF-to-baseband conversion. A π matching network ensures 50Ω impedance matching. High Q-value capacitors and inductors minimize losses. The detector's output passes through a passive low-pass filter with a 1.4 MHz cutoff.

Baseband Amplification: The tuned RF 20dB amplifier boosts sensitivity to -70dBm. The high-impedance sub-MHz baseband amplifier uses a 10-stage cascaded common-source configuration, achieving 68 dB voltage gain.



(a) Route1: Urban (b) Route2: Rural (c) Route3: Rural (d) Localization accuracy

Figure 5.9: *LiTEfoot*'s estimated trajectories and localization errors for in (a) urban and (b), (c) rural environments. The black lines denote the GPS trajectory and blue markers denote *LiTEfoot*'s estimated trajectory. The cell towers that are detected during the route are shown in red. (d) The empirical CDF of the localization errors.

ADC: We use an ultra-low-power successive approximation register (SAR) ADC by Texas Instruments [337]. This 12-bit ADC with SPI interface requires an external serial clock and

consumes 585.7 μW offering the power efficiency required for our prototype.

Baseband Processing: We implement *LiTEfoot* using the Apollo3 Blue ultra-low-power ARM Cortex-M4 based MCU SoC [338]. It interfaces with the ADC via SPI and uses Direct Memory Access for efficient signal transfer. The SoC's flash stores the code with zero fetch overhead, while data resides in SRAM and Flash. We power-gate the unused SRAM banks during idle times. Overall, the SoC consumes less than 6μA/MHz at 3V during active processing and approximately 1μA/MHz in deep sleep mode. While we optimize for power efficiency using the COTS MCU, future ASIC implementations could further help in reducing the power consumption [226, 314, 339].

Cost of Prototype: We design the prototype using COTS components. It includes an Skyworks SMS7630 diodes [340] ($0.60), an Infineon BGA7M1N6 LNA [341] ($1.11), a TI ADS7042 ADC [337] ($2.79), and an Ambiq Apollo3 [338] ($3.98). Including additional passive components, the total material cost for a *LiTEfoot* prototype is $9.48. This low-cost design, combined with *LiTEfoot*'s long battery life and minimal maintenance requirements, enables economically viable large-scale deployments for wide-area asset tracking and various IoT applications.

CMOS Simulation: We designed and simulated an Application Specific Integrated Circuit (ASIC) using TSMC-45nm CMOS technology to explore further power efficiency. The ASIC layout, created using Electric VLSI software and simulated with PathWave Advanced Design System (ADS), demonstrates a potential power consumption reduction to 112 μW. For the LNA, we implemented a self-biased inverter architecture [342] optimized for 45nm CMOS, achieving 15 dB gain while consuming only 89 μW. The transistors are biased in the low-inversion region at 0.26V with a 25 $k\Omega$ self-biasing resistor. We simulated a 10-bit charge-redistribution ADC [343]

capable of 2 Msps at 3.8 μW, significantly outperforming the COTS ADC (ADS7042) used in our PCB prototype. These improvements could enable years of battery life for small tags.

5.5 Evaluation

A. Localization Evaluation

We evaluate *LiTEfoot*'s localization performance through real-world experiments across urban and rural environments, comparing it against GPS ground truth. Our experiments cover multiple routes totaling 54km, with *LiTEfoot*'s tag mounted in a vehicle alongside an iPhone 15 running the Matlab mobile application for GPS data collection. Table 5.1 summarizes the key characteristics of our test routes: an urban route (Route 1) covering 1.80 km^2 with 35 cell towers (19.44 cells/km^2), and two rural routes (Routes 2 and 3) spanning 3.60 km^2 and 4.05 km^2 with 10 and 11 towers respectively (2.78 and 2.72 cells/km^2).

Route	**Area (km^2)**	**Number of Towers**	**Cell Density (cells/km^2)**
Route 1	1.80	35	19.44
Route 2	3.60	10	2.78
Route 3	4.05	11	2.72

Table 5.1: Key characteristics of evaluation routes.

Figure 5.9 shows *LiTEfoot*'s estimated trajectories (blue) alongside GPS trajectories (black) for each route and presents the CDF of absolute location errors. Our results show a median error of 22 meters in the urban area and 50 meters in rural areas. These results correlates strongly with cell tower density, with the urban route consisting approx 7 times higher density compared to the rural routes, directly impacting the localization accuracy.

We note that existing baselines were not suitable for direct comparison, as many require prior

information about cell towers (e.g., LTE Band, MNC, MCC) or are incompatible with real-world cell tower setups due to high latency demands, as demonstrated in next section. This evaluation validates *LiTEfoot*'s ability to provide accurate localization in the wild across varied environments.

B. Latency Evaluation

We evaluate *LiTEfoot*'s sensing latency across LTE bands (spanning 450 MHz to 3 GHz) and compare its performance with existing localization techniques. The experiment was conducted on a predefined route in a challenging environment of rural area where we observed 8 unique LTE bands (2, 4, 5, 12, 13, 14, 41, and 71) with at least one serving cell each (see Fig. 5.10).

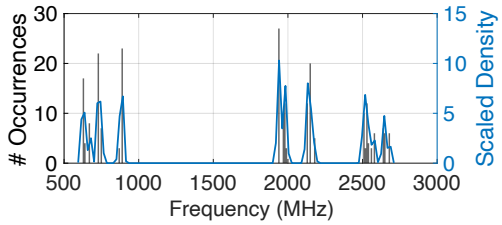


Figure 5.10: Diversity in downlink center frequencies used by the cells in a 200-meter radius.

Baselines: In this experiment we compared *LiTEfoot* against four baseline methods: •GPS: Using a COTS module [315]. •Wideband SDR Search: Scanning the entire LTE band using frequency hopping. •Targeted Band SDR Search: Pre-selecting six frequencies based on prior knowledge. •RSSI Method: Implementing Crescendo [344] on targeted frequencies. Each baseline presents distinct trade-offs. GPS offers high accuracy but with relatively high latency and power consumption. SDR-based methods provide flexibility but suffer from long sensing times due to sequential frequency hopping. The RSSI method improves on SDR approaches but still requires pre-selection of frequencies.

Results and Analysis: Table 5.2 shows the comparison of latency, accuracy, and power

consumption. The results show the fundamental challenge in cellular-based localization: the vast number of possible configurations (65,535 in LTE) necessitates either frequency hopping or prior knowledge of active bands, creating a significant latency bottleneck. *LiTEfoot* overcomes this bottleneck through its spectrum folding technique, enabling simultaneous sensing of all LTE bands. This approach eliminates the need for sequential frequency hopping, resulting in exponential reduction in latency.

Method	**Latency (s)**	**Error (m)**	**Energy (mJ)**
SDR (Wideband search)	66100	38	21×10^4
SDR (Targeted search)	0.66	38	210 [345, 346]
Crescendo [344]	0.66	24	210
GPS [315]	1	2	25
LiTEfoot	0.01	19	0.039

Table 5.2: Comparison between latency, accuracy, and power consumption.

C. Power Evaluation

Measurements Methodology: We implemented *LiTEfoot* on a custom PCB using low-power COTS components. Power consumption of individual modules was measured using a Keysight power monitoring unit [68] for high-precision readings. Additionally, we performed CMOS simulations to estimate potential power savings. The RF peripheral in our PCB prototype interfaces with an Apollo3 MCU.

Power Consumption Analysis: Tables 5.3, 5.4, and 5.5 show the overall power and energy per inference, along with a breakdown of power consumption for the RF chain (PCB and 45nm CMOS) and baseband computation, respectively. In our PCB prototype, the RF frontend consumes 8738.8 μW, with the LNA being the most power-intensive component at 8035.2 μW. The total energy consumption per inference is 902.37 μJ, distributed between the RF frontend (873.88 μJ)

and baseband compute (28.49 μJ). The baseband compute consumes 2035 μW of power and takes 14.52 ms per inference, with SSS correlations dominating the time consumption at 13.02 ms. Our CMOS simulation results demonstrate significant potential for power reduction. The simulated RF frontend consumes only 111.8 μW, representing a 98.7% reduction compared to the PCB prototype. Total energy consumption per inference in the CMOS simulation is 39.67 μJ, a 95.6% reduction from the PCB prototype. The simulated RF frontend achieves high sensitivity (-70 dBm), while the ADC simulation shows a power consumption of 3.8 μW at 2 Msps, further contributing to the overall system efficiency.

Battery Life: We project that the CMOS implementation of *LiTEfoot* can operate continuously for 11 years on a standard CR20we32 coin cell battery (220 mAh) with location updates every 5 seconds.

Module	**Power (μW)**	**Time (s)**	**Energy (μJ)**
RF frontend (PCB)	8738.8	0.10 (10 frames)	873.88
RF frontend (CMOS)	111.8	0.10 (10 frames)	11.18
Baseband Compute	2035	0.0145	28.49
Total (PCB)		**0.1145**	**902.37**
Total (CMOS)		**0.1145**	**39.67**

Table 5.3: Overall power and energy per inference.

Component	**PCB prototype (μW)**	**CMOS sim (μW)**
LNA	8035.2	89
Envelope Detector	0	0
High Imp. PA	162.9	19
ADC	585.7	3.8
Total RF	**8738.8**	**111.8**

Table 5.4: RF frontend power consumption.

Module	Power (μW)	Time (ms)
PSS Search (correlations)	2035	0.66
SSS Search (correlations)	2035	13.02
Amplitude Estimation	2035	0.84
Multilateration	2035	2×10^{-6}
Total Baseband	**2035**	**14.52**

Table 5.5: Baseband power and time per inference.

D. Micro-Benchmarks

Comparison of 5G-NR, NB-IoT, and LTE Standards: We evaluate *LiTEfoot*'s performance across LTE, 5G-NR, and NB-IoT, each employing distinct PCI encoding methods. LTE, our baseline, uses 504 unique PCI values from both PSS and SSS. NB-IoT simplifies this approach, using only the Narrowband SSS (NSSS) to encode 504 NCellID values. In contrast, 5G-NR expands to 1008 unique PCI values while retaining LTE's PSS and SSS structure. As we show in Figure 5.11, NB-IoT achieves the highest PCI estimation accuracy, exceeding 95%, due to its narrower bandwidth and simplified signal structure. 5G-NR initially shows lower accuracy but matches NB-IoT's performance with increased frame stacking, due to its broader bandwidths and fluctuating data subcarriers. LTE maintains consistent performance across frame stack above 20. These results demonstarte *LiTEfoot*'s adaptability across cellular standards, with frame stacking proving particularly effective in managing 5G-NR's dynamic resource allocation.

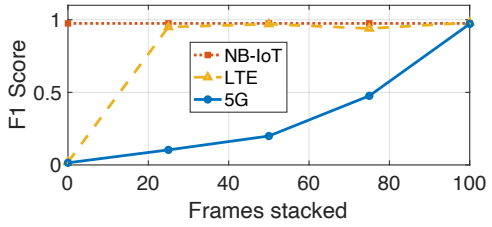
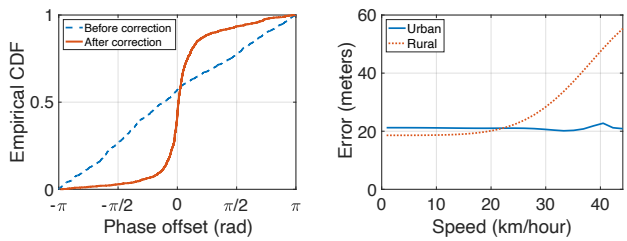


Figure 5.11: PCI estimation accuracy (F1 score) vs. number of stacked frames for LTE, 5G-NR, and NB-IoT.

Sub-sample offset correction: We evaluate the effect of sub-sample offsets due to clock drift or synchronization errors. Since we do not have any way to synchronize the tag with the base stations, there could be ground truth timing errors. Figure 5.12a shows the CDF of phase-offsets for all subcarriers before and after the sub-sample correction algorithm. As we see, before the offset correction, sub-sample shifts within ± 0.5 sample can result in phase shifts of $\pm\pi$ radians. Whereas, after applying our sub-sample correction, most phase differences are reduced to 0. This benchmark uses 1000 real-world synchronization signals with varying signal strengths. We observe there are a few outliers in the CDF when the phase noise is large compared to the signal strength and the best-fit slope of the phase is incorrectly estimated.



(a) (b)

Figure 5.12: (a) CDF of phase offsets in measurements before and after sub-sample offset correction. (b) Localization error for varying speed of vehicle.

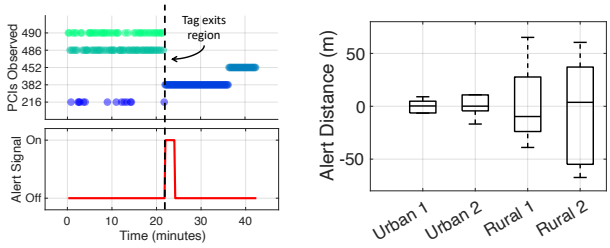
Impact of speed: We evaluate *LiTEfoot*'s performance at varying vehicle speeds in both urban and rural environments, using GPS-measured ground truth speeds. As shown in Figure 5.12b, urban settings demonstrate the minimal impact of speed on localization accuracy. In contrast, rural environments have increased localization errors at higher speeds. Urban environments typically feature a higher density of smaller cells and micro-cells, which change every few meters, contributing to more robust localization across various speeds.

E. Ultra-low-power Geofencing

We also evaluate *LiTEfoot* for geofencing application, which can enable ultra-low-power tracking for elderly care monitoring and route enforcement. Our approach focuses on minimizing power consumption and evaluating continuous PCI monitoring for timely and accurate alert generations.

To reduce computational load, we implement a hierarchical search algorithm that leverages the reduced number of unique PSS signals post non-linear transformation and the periodic nature of synchronization signals. This algorithm first performs a PSS search to identify potential time indices, followed by selective SSS correlation, significantly narrowing the search scope. We further optimize by focusing on known base stations rather than continuously searching for new ones, reducing the search space from thousands to tens of samples in subsequent frames.

We evaluate the alert generation performance in four different scenarios: two for entry restrictions and two for exit restrictions. Geofencing boundaries are marked on maps, and the relevant PCI database is filtered onto the tag. The tag continuously monitors observed PCI values and generates alerts when entering or exiting marked locations. Figure 5.13a illustrates the PCIs detected over time and marks the event when the tag exits a designated region. Figure 5.13b shows the overall accuracy of alert generation distance from marked boundaries across urban and rural settings. Our results show alert generation within 24 meters of the defined boundaries in urban environments and 106 meters in rural environments due to lower cell tower density.



(a) (b)

Figure 5.13: (a) An alert is generated when the tag exits the boundary of the marked region. (b) Alert generation distance from the map boundaries.

5.6 Related Work

Asset tracking: Asset tracking applications span various domains, from supply chain management to healthcare [347–352]. GPS and satellite-based systems [298, 299] offer global coverage but face power and indoor limitations. Cellular IoT technologies like NB-IoT and LTE-M [353, 354] improve on GPS's battery life and coverage for wide-area tracking. RFID [355, 356], BLE beacons [303, 357], and acoustics [231, 232, 358–362] offer low-cost, short-range tracking solutions for controlled indoor environments, each with distinct advantages and limitations. *LiTEfoot* on the other hand leverages ambient cellular signals for low-power, wide-area asset tracking without custom infrastructure.

Cellular localization: Ambient cellular signals have been explored for localization and tracking [309, 363–367]. RSSI-based approaches [368–370] achieve 50-200 meter accuracy. Fingerprinting techniques [371–375] improve accuracy to sub-50 meters but require extensive wardriving. Notably, Crescendo [344] utilizes RSSI distributions within grid cells; and DeepLoc [376] and DeepFeat [377] use RSSI to train a deep learning model to achieve sub-20 meter accuracy. TransparentLoc [378] leverages crowdsourced data for both indoor and outdoor location estimation. However, these methods typically require power-intensive downconversion and

frequency hopping, limiting IoT applicability. ISLA [258] offers cellular self-localization but faces form factor constraints due to multi-antenna requirements. In contrast, *LiTEfoot* enables wideband sensing and localization without power-intensive components, striking a balance between accuracy and energy efficiency for low-power IoT applications.

Non-linearity based receivers: Recent advancements in low-power RF receivers [226, 314, 316, 318, 379–381] have focused on replacing active components with non-linear envelope detectors to reduce power consumption in communication [382–388]. MIXIQ [226] introduced a low-power 802.11ax receiver by modifying OFDM packets to include a helper tone and preserve IQ data during downconversion. These systems primarily rely on the use of a secondary helper signal or changes in the transmitter. ReMix [389] leveraged non-linear backscatter circuits for in-body communication and localization. LivingIoT [176] and Sirius [379] use envelope detectors for direction of arrival estimation. In contrast to all these approaches, *LiTEfoot* operates on unmodified cellular networks, leveraging non-linear transformations of cellular synchronization signals to enable wideband sensing and PCI estimation without requiring helper tones or any protocol modifications. This approach allows *LiTEfoot* to achieve ultra-low power consumption while maintaining compatibility with existing cellular infrastructure.

5.7 Discussion

Scalable and Sustainable IoT: This chapter is a step toward a broader vision of sustainable ambient intelligence and scalable IoT. By leveraging existing cellular infrastructure for self-localization, *LiTEfoot* enables country-wide scalability without dedicated anchors. The system's ultra-low-power design allows tags to operate for years without battery replacement or maintenance,

minimizing electronic waste and environmental impact. This substantially reduces the tag's embodied carbon footprint [390]. *LiTEfoot*'s wide-bandwidth sensing and low-latency inference capabilities enable tracking millions of objects across smart cities and supply chains. We believe this could revolutionize urban sustainability through real-time air quality monitoring or tracking personal belongings across vast distances. *LiTEfoot* paves the way for sustainable, large-scale tracking solutions that align technological progress with environmental consciousness.

Communicating locations: Localization trackers often need a communication module to periodically update location to a server or nearby hub. Standalone GPS trackers (e.g., [391] and [315]) often use cellular or satellite communications to do so. *LiTEfoot* can incorporate an NB-IoT module for relaying location data. Using a low-power NB-IoT transmitter [392], which consumes 76 mW to transmit a 140 kb packet in eDRX mode, the system can intermittently send updates hourly, daily, or monthly depending on the application. This enables *LiTEfoot* to operate for approximately 10 years on a coin cell. The same antenna frontend can be shared between *LiTEfoot* and the communication unit, maintaining a compact form factor while providing flexible communication capabilities.

PCI Distribution: 4G-LTE defines 504 unique PCI numbers to be used in base stations per network. Cellular networks are designed to ensure neighboring base stations get unique PCI numbers [393], while it may repeat in a non-adjacent base station. So, in theory, a PCI-based localization system can face ambiguity when a combination of PCI values from observable base station transmission repeats for two different geographic locations. However, the probability of such repeating combinations is significantly small. Moreover, *LiTEfoot* often senses multiple PCI values from each service provider operating in a region which makes the chances of location ambiguity negligible.

PCI database updates: In LTE networks, the PCI value typically changes infrequently under normal operating conditions. PCI values remain constant unless there is a network reconfiguration such as cell outages, recoveries, maintenance, or upgrades [322, 367, 393, 394]. However, *LiTEfoot* can maintain good accuracy even with infrequent database changes, as it relies on multiple PCI values per location. Moreover, the tags can be updated using the same NB-IoT module used for communication, ensuring long-term reliability with minimal overhead.

Application scenarios of *LiTEfoot*: The ultra-low-power design of the system enables a wide range of long-term tracking applications. In logistics, *LiTEfoot* can provide wide-area asset tracking to combat cargo pilfering, a major supply chain challenge. For smart cities, *LiTEfoot* can enable efficient management of shared e-scooters and bikes by logging usage and location data, transmitting only when docked. In agriculture, it facilitates long-term livestock monitoring and geofencing of large industrial vehicles with minimal communication needs. Urban services can benefit from *LiTEfoot*-equipped smart waste bins or trucks that log positions and communicate only when necessary. The system's 20-meter accuracy suffices for these applications. With its ability to operate for years on a single cell, *LiTEfoot* opens up possibilities for persistent, wide-area monitoring across diverse domains, from supply chain management to environmental sensing.

Limitations and Future Work: Our work reveals opportunities for further refinement. Like GPS and other RF-based systems, *LiTEfoot*'s performance may be affected in metal-enclosed spaces due to signal attenuation. The system's accuracy, while sufficient for many applications, could be enhanced in areas with sparse LTE coverage. Our current PCB prototype, while functional, leaves room for optimization; VNA-tuned impedance matching could significantly improve sensitivity and power efficiency, especially for the LNA. The localization algorithm currently assumes uniform transmitting power from cell towers, which may not always hold true, affecting

precision. Implementing time-of-flight based multilateration could potentially improve accuracy by reducing reliance on signal strength assumptions. Subtle effects of tag orientation and placement on signal reception could be mitigated with improved antenna design. Future ASIC implementations could further reduce power consumption and form factor.

5.8 Conclusion

LiTEfoot is a novel low-power cellular localization system leveraging non-linear spectrum transformation for efficient baseband signal superimposition. It senses a 3 GHz wideband LTE spectrum in 10 ms using non-linear intermodulation, while operating in the μJ energy regime and achieving median localization errors of 22 m.

Chapter 6: Large Network UWB Localization: Algorithms and Implementation

6.1 Introduction

A swarm of connected nodes is the underlying architecture for many emerging applications. Miniaturized sensing modules are scattered like seeds [395,396] or carried by insects [231,360,397] to scale a vast region for fine-grained sensor networks toward border protection [398], animal migration tracking [399], or precision agriculture [379,380,400]. Flocks of drones can localize each other to fly in formation for charting inaccessible regions [38,401,402] or for airshows [403]. A smart network of tags can enable city-scale tracking of deliveries or missing objects in real-time [404]. Future of the cellular [405,406] and vehicular networks [407] can pave the path toward autonomy [408,409] and road safety [410]. Localization of the nodes in such large networks is an essential requirement and at the same time a challenging problem when the number of nodes is large. In this chapter, we present *Locate3D*, a system that is grounded on both theoretical and practical foundations, providing a reliable framework for fast peer-to-peer localization as well as orientation of nodes on a large network.

Multidimensional scaling (MDS) has been the essence of nearly all large-scale localization algorithms for wireless networks. The primary reason MDS is adapted so widely is its ability to reconstruct the relative map of the network nodes with little infrastructural support, even in the form of the anchor nodes with known locations. MDS-based algorithms are fundamentally

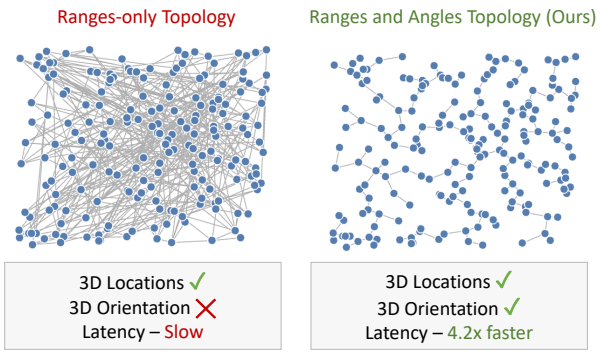


Figure 6.1: *Locate3D*'s approach to include both angle and edge constraints for faster and efficient localization.

centralized ranging-based systems that consider inter-node distances in the Euclidean space to optimize for sensor locations. It, however, suited the capacity of the mobile nodes that can use the time of flight of the signal or RSSI-based model to estimate ranges without additional hardware requirements or computational complexity. However, recent mobile nodes have evolved not only to accurately sense the ranges but are also equipped with multi-dimensional antenna arrays to estimate reliable angles of the nodes. For instance, around 20 years after being released for commercial applications in 2002 [411], off-the-shelf UWB sensors can now sense the angle of the received signal [412, 413]. Unfortunately, the entire class of the MDS-based network localization algorithms can not take advantage of this new-found capability of the mobile nodes. Intuitively the angle information of the peer nodes can serve as additional constraints of the network topology leading to faster convergence to the location estimation. Moreover, the range and angles are estimated simultaneously per exchange of signals between nodes without incurring additional measurement latency. This convenient information is wasted as the MDS objective function can not jointly optimize on the Euclidean distance and angle plane. We develop a new network localization algorithm with a redefined objective function to include range and angle together for optimization to bridge this gap.

The scalability of a network localization solution depends on several practical factors beyond the correctness of the theoretical formulations. The level of dependency on the infrastructure is the most crucial of them. In addition, a dynamic topology of mobile nodes requires short update latency of location estimations. Like any peer-to-peer localization system, *Locate3D* requires at least four anchor nodes for the unambiguous global 3D location of the nodes. However, the relative locations of the entire topology are correct with no assumption on the anchors. We enable our algorithms to incorporate any number of available anchor nodes and other infrastructural opportunities to create virtual anchors that enhance overall localization accuracy. In the proposed system, the joint range and angle-based optimization reduce the measurement and initial topology estimation latency, then a spanning-tree-based optimal edge selection expedites updates on locations after initial estimation, and finally, a graph rigidity-based solution makes the estimation robust to local occlusions or poor peer-to-peer connectivity.

A class of solutions for network localization resort to multimodal data to improve accuracy and reduce the update latency of the system. While effective in a small number of nodes and within restricted environmental conditions, the multimodality assumption limits the scalability of the system. It is infeasible to maintain homogeneous data quality with thousands of nodes spread across a large geographical region. For instance, some recent papers [3] use Visual Inertial Odometry (VIO) - a camera-based solution to track orientation - to improve localization accuracy. We argue that an unimodal solution is ideal for large networks in terms of consistency and ease of practical deployment and maintenance. It makes our solution equally applicable to the networks of resource-constrained low-power nodes. The core localization algorithm is, however, applicable to any modality that can measure the peer-to-peer range and relative angles, and naturally, the location accuracy is defined by the accuracy of the measurements. For instantiating the algorithm

in a prototype and realistic large-scale simulations, we consider off-the-shelf Ultra-WideBand (UWB) wireless sensing nodes.

This chapter strives to improve the accuracy of 3D localization of UWB-enabled nodes on a large single-modality network as shown in Figure 6.1. To this end, we have made the following three specific contributions:

- Developed a novel 3D network localization algorithm that jointly incorporates range and angle information for topology estimation. This algorithm achieves a 75% latency improvement and a median localization accuracy of 12.09 meters for a 100,000-node network spanning several kilometers, using only 15 static anchors. In a 32-node, three-floor building-scale experiment, the algorithm demonstrates a localization accuracy of 0.86 meters.

- Designed a system with a supporting algorithm that addresses challenges of occlusions and limited sensor field of view. The system includes an optimal spanning tree estimation algorithm and a decomposition technique that segments non-rigid graphs into smaller rigid subgraphs, considering both range and angle constraints.

- Implemented a working prototype with UWB-enabled nodes and evaluated its performance in real-world scenarios, including room-scale and multi-floor building experiments. Demonstrated the system's scalability through a city-scale simulation using real-world UWB measurement traces.

6.2 System Design

Our system, *Locate3D*, uses pair-wise Ultra-Wideband (UWB) RF measurements, specifically range and angle-of-arrival (AoA) information, to determine the precise 3D positions of interconnected nodes in relation to one another. The system is designed using four key components:

Joint range-angle localization: We develop an optimization method that uses range and angle measurements as constraints to build a network of localized nodes. Incorporating angles ensures an accurate location estimation with fewer number of edges compared to a range-only method, effectively reducing the latency up to $4.2\times$.

Reference frame transformation: Each node's angle measurement is a composite value of both the relative orientation (determined by the antenna angle with respect to the peer's antenna) and the AoA. *Locate3D* efficiently decomposes this combined measurement to transform all node axes to a common frame of reference.

Scaling to large networks: We develop an edge selection strategy that ensures the formulation of the optimal spanning tree, facilitating the interconnection of large topologies. Large networks often contain flexible or unconstrained edges, which can result in isolated but structurally rigid subgraphs. We aim to identify and resolve these subgraphs separately.

Opportunistic integration with infrastructure: *Locate3D* can leverage any pre-existing infrastructure, incorporating it as anchor points to refine localization accuracy. We further extend our system's capability by introducing the concept of virtual anchors. This mechanism temporarily designates specific mobile nodes as anchors and adjusts edge weights, disseminating the accuracy throughout the network. Below, we discuss the individual algorithmic contributions in detail.

6.2.1 Joint Range-Angle Localization

We formulate the localization problem as an optimization problem that aims to minimize the difference between measured pairwise distances and the distances corresponding to the estimated coordinates. There is a plethora of work done in network localization using range measurements only. SMACOF [414,415], non-metric MDS etc [416–418]. Some recent works also account for missing edges [419,420] and NLOS cases [421–423]. However, none of the works have jointly incorporated AoA information in the network.

Problem formulation: Consider a topology of N nodes (mobile devices), with unknown locations $X = [(x_1, y_1, z_1), \cdots (x_N, y_N, z_N)]$. Suppose the nodes measure the range between each other, $\hat{r_{ij}}$ where (i, j) is the pair nodes i and j in the topology. We aim to solve for X while minimizing the cost function:

$$\min_X \sum_{i,j} w_{ij}(\hat{r_{ij}} - r_{ij}(X))^2 \tag{6.1}$$

where, w_{ij} is the weight assigned to edge between nodes i and j and $r_{ij}(X)$ is the euclidean distance between them given by Equation 6.2. We set $w_{ij} = 0$ for missing edges.

$$r_{ij}(X) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \tag{6.2}$$

Adding angles in the topology: Incorporating angles into the network topology offers the benefit of reducing the necessary number of edges, which subsequently decreases latency. This approach, however, is not as straightforward due to the highly non-convex and discontinuous nature of angular loss functions when simply added with the range-based loss function.

Existing works [424–426] which use L1 and L2 losses directly on this loss function show that the objective function has many local minima.

We can compute the angular loss using the function $f(X)$, expressed as:

$$f(X) = \left[\hat{\theta_{ij}} - \text{arctan}\left[\frac{y_i - y_j}{x_i - x_j}\right]\right]^2 \tag{6.3}$$

This function encapsulates the difference between the measured angles, denoted by $\hat{\theta_{ij}}$, and the angles calculated from the estimated coordinates. The non-convex nature of $f(X)$ occurs due to the arctangent and the least square operation on angles. This results in multiple local minimums, making it highly prone to generating inaccurate topologies. The primary contributor to this issue is the restrictive interval $[-\pi/2, \pi/2]$ that the arctangent function operates in, failing to account for points in the left quadrants of the plane, thereby resulting in the same angles for coordinates (x, y) and $(-x, -y)$. To address this, we first use the 2-argument arctangent, a variant of the arctangent function that considers both x and y inputs and adjusts for the signs, hence returning angles within the interval $[-\pi, \pi]$.

Unfortunately, the transformation is not enough, as the gradient consists of non-differential making it prone to getting ensnared in local minimums. To overcome this, we apply another transformation to the loss function. We take the negative cosine of the angles, creating a smoother, continuous, and differentiable function restricted within the $[0, 1]$ range. This transformation of the new loss function, is defined as:

$$f(X) = 1 - cos(\hat{\theta_{ij}} - arctan2(y_i - y_j, x_i - x_j)) \tag{6.4}$$

Finally, we combine the range and angular loss functions to efficiently integrate angles in the network topology optimization formulating a joint optimization problem.

$$\min_X \Big[\frac{\sum_{i,j} w_{ij}^r (\hat{r_{ij}} - r_{ij}(X))^2}{\sum_{i,j} \hat{r_{ij}}} + \sum_{i,j} w_{ij}^\theta (1 - cos(\hat{\theta_{ij}} - \theta_{ij}(X))) \Big] \tag{6.5}$$

where, $\theta_{ij}(X) = arctan2(y_i - y_j, x_i - x_j)$. Note that we scale the range loss function with $(\sum_i \sum_j \hat{r_{ij}})$. The range loss can be much higher compared to the angular loss, which is in the range $[0, 1]$, and can dominate the overall gradient/overshadow the angular loss. Hence, we normalize the range loss function before adding the angular loss part.

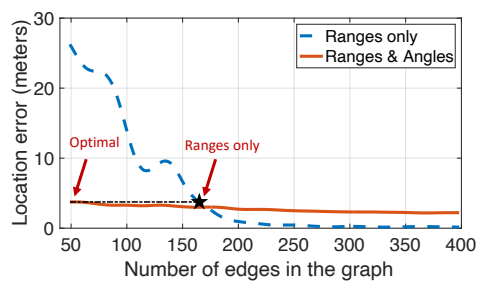


Figure 6.2: Comparative analysis of constraints: Incorporating angles reduces the number of edges required to attain the same level of accuracy as the "Ranges only" approach.

We conducted a simulation to evaluate the effectiveness of integrating angular measurements with distance constraints in topology optimization. The simulation involved a 50-node network within a $100m \times 100m$ area, incrementally adding edges under two constraint scenarios: ranges only, and both ranges and angles. Figure 6.2 shows that using only ranges required more than $3\times$ the number of edges to achieve accuracy comparable to our method combining ranges and angles. However, as more edges were added, the range-only approach eventually surpassed the range-angle method in accuracy. This can be attributed to the use of raw angles, which may include noisy measurements, and the random selection of edges rather than optimal spanning tree

edges. Despite this, the method promises latency improvement over range-only methods without compromising accuracy even for a sub-optimal set of edges. We address the challenges related to noisy angle measurements and edge selection in Section 6.2.2, where we present techniques for filtering erroneous angles and constructing optimal spanning trees.

6.2.2 Scaling to Large networks

In this section, we detail how *Locate3D* can localize large node networks efficiently. To adapt to city-scale networks, we address three significant challenges linked with large networks.

• *Optimal edge selection*: Identifying the optimal edges to sample in the graph topology is vital to forming a unique, rigid structure while optimizing latency and accuracy.

• *Handling erroneous angles*: While in theory, all measured edges should have angle data, the limited FOV of commercial antenna arrays can sometimes result in inaccurate angle information. Correcting these errors is vital as they can adversely affect the graph's rigidity constraints.

• *Rigidity decomposition*: Large networks frequently consist of flexible or unconstrained edges which leads to smaller but rigid subgraphs. We need to identify and solve these subgraphs individually.

Next, we will elaborate on how *Locate3D* mitigates these challenges to scale to larger networks, introducing techniques that could reduce *Locate3D*'s latency by up to $4.2\times$.

6.2.2.1 Optimal Edge Selection

During any measurement time iteration, *Locate3D* compiles a list of optimal edges that are to be measured within the graph. These edges are chosen based on three criteria - (1) They form a rigid and unique topology, eliminating flexibility or ambiguity. (2) They are minimal in number to decrease overall latency. (3) They are feasible, meaning the edges fall within the necessary range and angle FoV.

In a topology comprising n nodes, a maximum of $n(n-1)$ edges are potentially available. However, acquiring data for each possible edge isn't viable as it is time-consuming and increases exponentially as n increases. Hence, rather than overconstraining the topology, we can create the topology with significantly fewer edges. According to Laman's condition [427], for a system that doesn't utilize angles, the minimum number of edges equals to $2n-3$ and $3n-4$ for 2D and 3D setups respectively. Interestingly, our approach necessitates only a *single* edge per node to constrain it, amounting to only $n-1$ edges in total. This efficiency stems from the fact that a single edge contains three constraints - range, azimuth angle, and elevation angle. Moreover, any random $n-1$ edges may not always be enough, as some edges could be redundant. Rather we need *well-distributed* $n-1$ edges to make the topology rigid and accurate. This renders our problem closely analogous to the *Minimum Spanning Tree problem* found in graph theory [428].

Minimum spanning tree: The minimum spanning tree (MST) is a subset of the graph's edges connecting all nodes with the least total edge weight. In our system, we utilize Kruskal's algorithm, a greedy algorithm that arranges the graph's edges in increasing order of their weights and adds edges to the MST as long as they do not form a cycle, thus guaranteeing the minimum combined edge weights.

Edge weight calculation: Our weight assignment algorithm minimizes localization error in the MST by utilizing prior knowledge of the network topology from the previous iteration, including estimated range (r) and angle (θ) values between nodes. The *area of uncertainty*, denoted as A, is calculated using sector area, $A = \pi(\alpha\sigma_r)^2 \frac{\alpha\sigma_\theta}{2\pi}$, where σ_r and σ_θ are the standard deviations of the range and angle estimates, respectively, and α is a factor that scales the standard deviations based on LOS ($\alpha = 1$) or NLOS ($\alpha > 1$) conditions. The edge weight, w, is then computed as $w = A$, representing the potential location uncertainty based on the estimated range, angle, and their associated error distributions. Edges with shorter distances and smaller standard deviations inherently have a smaller area of uncertainty, and edges with available angle information are favored over those restricted to range data only. Figure 6.3a demonstrates that our algorithm's optimal spanning tree ranks within the top 1% of all possible spanning trees in terms of localization error. Algorithm 1 uses these estimated range and angle values to compute edge weights and form the MST, iteratively selecting the lowest-weight edges without introducing cycles. Actual range and angle measurements are taken for the selected edges once the MST is formed. During the cold start phase, when no prior knowledge is available, all nodes measure their neighbors, similar to traditional systems like [3, 5], resulting in increased latency for the first iteration. Subsequent iterations efficiently select edges for measurement by leveraging the prior knowledge of the network topology.

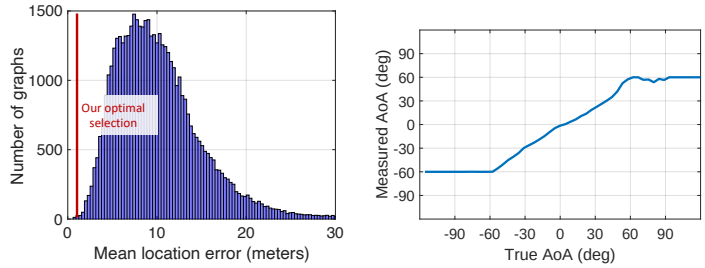
6.2.2.2 Handling Erroneous Angles

In practical settings, nodes may not consistently measure each other's ranges or angles due to limited sensing range or angular FOV. This is due to the limited FOV of antenna arrays present in

Algorithm 1 Optimal Edge Selection Algorithm

Input: $G(V, E)$, $\hat{r}$, $\hat{\theta}$, σ_r, σ_θ, α **Output:** $M(V_{MST}, E_{MST})$

```
procedure
    for each E in G do
        Retrieve r̂, θ̂, σ_r, σ_θ, and α for current E
        Assign weight w(E) = π(ασ_r)^2 ασ_θ/2π
    end for
    G_sorted(V, E) = Sort(G(V, E), w(E))
    Initialize M(V_MST, E_MST) as empty
    for each V, E in G_sorted do
        if adding V, E to M does not form a cycle then
            Add V, E to M
        end if
    end for
    return M(V_MST, E_MST)
end procedure
```



(a) (b)

Figure 6.3: (a) Histogram of localization errors for all spanning trees. (b) Reported angles by COTS UWB sensor.

commonly used UWB modules like Decawave [412,413] and NXP [429]. These modules typically restrict the AoA field of view (FOV) to between $90°$ and $120°$ to maintain a respectable accuracy in measured AoA, primarily because the AoA estimation algorithm, which depends on the phase difference of the incoming signal, can result in significant errors when the AoA approaches the broadside of the antenna. Figure 6.3b shows the reported angles by an off-the-shelf UWB sensor [429]. To address this, *Locate3D* employs two strategies for handling erroneous angles. First, we filter and flag angles as erroneous if they fall outside the range of $|FOV_{angle} - tolerance|$, effectively removing angles close to the broadside. Second, when available, *Locate3D* tracks the

rotation of each node using its inertial sensors. Utilizing this rotation data, we can determine if any two nodes are within each other's FOV and flag the angles as valid or erroneous accordingly. In scenarios where some nodes lack inertial sensors or AoA estimation capability, *Locate3D* can still perform localization using the filtered angle data from the nodes equipped with these capabilities.

6.2.2.3 Rigidity Decomposition

While previous sections operate under the presumption of network topology rigidity, ensuring a unique realization of topology with given constraints, this isn't always the case. Notably, the optimal spanning tree derived earlier doesn't inherently ensure rigidity. Even fully connected graphs can remain non-rigid due to absent constraints, as illustrated in Figure 6.4 where spanning trees lack rigidity due to missing angle constraints. To address this, we develop a rigidity-based graph decomposition technique that segments the graph into smaller *rigid subgraphs*. These subgraphs can then be solved with refined edge weights ensuring that the topology is both rigid and uniquely realizable.

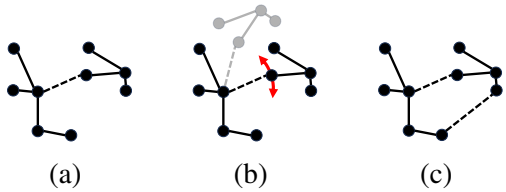


Figure 6.4: Different spanning trees representing rigid and non-rigid graphs. Solid lines indicate both range+angle edges, and dashed lines indicate range-only edges. (a) A connected but non-rigid graph due to missing angle information in an edge. (b) The subgraph is free to rotate. (c) Adding a range measurement makes the graph rigid.

Quantifying rigidity: The Rigidity of a topology determines whether it is possible to uniquely determine the location of all nodes in the topology with respect to other nodes. This leaves us with some trivial transformations of topology like translations and rotations in the space.

The rigidity can be quantified using the degree of freedom (DoF) of the graph. Each node has three DoFs, representing movements in the x, y, and z directions. Thus, a graph with n nodes inherently has $3n$ DoFs. As constraints are imposed, the available DoFs decrease. A graph is deemed rigid if its DoFs are ≤ 3, with these residual DoFs accounting for the whole-graph translational motions.

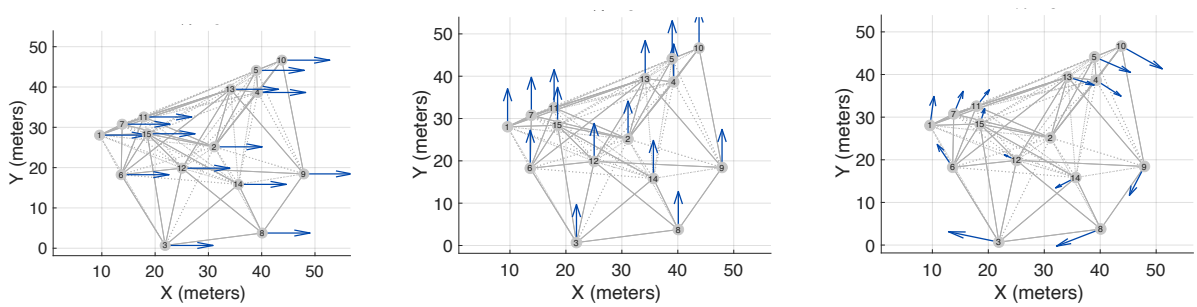


Figure 6.5: The displacements corresponding to zero eigenvalues represent the translational and rotational motions that the nodes can undergo without violating any constraint.

Eigenanalysis of rigidity matrix: To identify independent and unconstrained motions that the topology can perform without violating the constraints, we perform the eigenvector analysis of the topology. To capture the constraints in a compact matrix form, we first introduce the rigidity matrix, R. Every row of R denotes a unique constraint equation, which could be a distance or angle constraint between nodes. Each column of R corresponds to the x, y and z coordinates for each node in the graph. Thus, for a graph having n nodes, the matrix expands to have $3n$ columns.

Each entry r_{ij} in this matrix is determined by the partial derivatives of the constraint equations 6.2 and 6.3, as given by:

$$r_{ij} = \begin{cases} \frac{\partial d_{ij}}{\partial x_m} & \text{if edge } ij \text{ is a distance constraint} \\ \frac{\partial \theta_{ij}}{\partial y_m} & \text{if edge } ij \text{ is an angle constraint} \\ 0 & \text{otherwise} \end{cases}$$

$$R = \begin{bmatrix} \frac{\partial d_{1,2}}{\partial x_1} & \frac{\partial d_{1,2}}{\partial y_1} & \cdots & \frac{\partial d_{1,2}}{\partial x_n} & \frac{\partial d_{1,2}}{\partial y_n} \\ \frac{\partial d_{2,3}}{\partial x_1} & \frac{\partial d_{2,3}}{\partial y_1} & \cdots & \frac{\partial d_{2,3}}{\partial x_n} & \frac{\partial d_{2,3}}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial \theta_{1,2}}{\partial x_1} & \frac{\partial \theta_{1,2}}{\partial y_1} & \cdots & \frac{\partial \theta_{1,2}}{\partial x_n} & \frac{\partial \theta_{1,2}}{\partial y_n} \\ \frac{\partial \theta_{2,3}}{\partial x_1} & \frac{\partial \theta_{2,3}}{\partial y_1} & \cdots & \frac{\partial \theta_{2,3}}{\partial x_n} & \frac{\partial \theta_{2,3}}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \end{bmatrix}$$

Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be the eigenvalues of the matrix $M = RR^T$, where M is the symmetric and positive semi-definite matrix of R. Topology is uniquely rigid if the number of zero eigenvalues of M is equal to $3n-3$. A number of zero eigenvalues, or the degree of freedom, can be represented as:

$$DoF = \sum_{i=1}^{m} \delta(\lambda_i), \quad \text{where } \delta(\lambda) = \begin{cases} 1 & \text{if } |\lambda| < \epsilon \\ 0 & \text{otherwise} \end{cases} \tag{6.6}$$

Here, ϵ is a small positive number close to zero, chosen to account for numerical inaccuracies (e.g., due to floating-point representation limits in computers). If $|\lambda_i| < \epsilon$, it is considered a zero eigenvalue. Figure 6.5 illustrates the eigenvectors corresponding to the three zero eigenvalues $\lambda_1, \lambda_2,$ and λ_3 showing the motion of each node.

Constructing subgraphs: By examining the zero eigenvalues, we gain insights into the *displacement vectors* of each node. For a fully rigid topology, these vectors maintain consistent magnitudes and directions across nodes. In contrast, non-rigid graphs display varied displacements. Our core intuition is that displacement vectors can be used for decomposition: nodes that have identical displacement vectors in terms of magnitude and direction inherently form a rigid subset.

This means they can only undergo collective motion. To extract and group these subsets, we first reshape the eigenvector into a 3×n matrix to extract the displacement vectors of each node. Each row within this matrix signifies the displacement of a node across the three axes. Next, we identify the unique displacement vectors in this matrix where every unique vector signals the presence of a distinct ”rigid subgraph” within the primary graph. By grouping the nodes corresponding to unique displacement vectors, we find the vertices belonging to the associated ”rigid subgraph”.

Critical edges list: Beyond recognizing rigid subgraphs, our approach also pinpoints the edges connecting the separate subgraphs. We term these edges as ’critical edges’, as these edges are crucial for rigidity, serving as the connectors between independent rigid substructures. For enhanced analysis in subsequent iterations, we also prioritize these critical edges based on their latest computed locations. This prioritized list of edges underlines those edges that are more likely to come into proximity with one another, thus enhancing the overall scalability and rigidity of the graph.

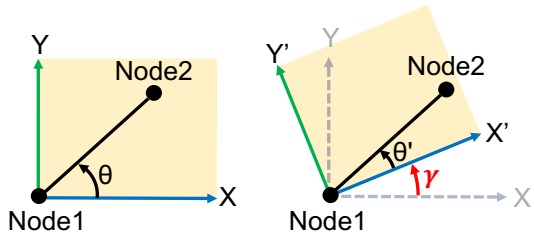


Figure 6.6: (a) When Node 1 is aligned with the global frame of reference, it reports that Node 2 is positioned at angle θ. (b) When Node 1 is rotated by γ relative to the global frame of reference in a 3D space, it reports a distinctly different angle θ', for the same node.

6.2.3 Reference Frame Transformation

In the above section, we assume that all nodes share a common frame of reference. Here, we relax this assumption and address the challenges arising from differing local frames of reference

for each node. Every node measures angles based on two factors: its relative orientation (defined by its antenna angle concerning another node's antenna) and the AoA. The challenge lies in the fact that these angle measurements are in the node's local frame of reference, which moves according to the node's orientation within a broader, global context. Consider the 2D example in Figure 6.6. In scenario (a), Node 1 identifies Node 2 at an angle θ. But in scenario (b), after Node 1 turns by an angle γ, it sees Node 2 at a different angle θ'. This change in reported angles is due to Node 1's rotation, even though the positions of the nodes didn't change. This issue gets complex in 3D, where nodes can rotate in any direction. Our goal is to deconstruct these angle measurements from orientation offsets and align all nodes to a singular, shared frame of reference.

6.2.3.1 Estimating Orientation Offsets

Every node's orientation in a 3D space is defined by a set of three rotation angles. Specifically, they depict the node's rotations about its X, Y, and Z axes and are recognized as the Roll(α), Pitch(β) and Yaw(γ) [430].

To estimate a node's orientation, we leverage a fundamental observation: When the orientations of nodes are in the same global frame, the measured azimuth AoA (θ) and elevation AoA (ϕ) are complementary, implying they obey the condition: $\theta_{ij} - \theta_{ji} = \pi$ and $\phi_{ij} = -\phi_{ji}$ (Figure 6.7). Using this insight, we formulate an optimization problem with the goal of estimating the rotation offsets (α, β, γ). The objective is to minimize the deviation from the above-mentioned constraints:

$$\min_{\alpha,\beta,\gamma} \sum_{i,j} (\hat{\theta}_{ij} - \hat{\theta}_{ji} - \pi) + (\hat{\phi}_{ij} + \hat{\phi}_{ji})$$

In this equation, $\hat{\theta}_{ij}$ and $\hat{\phi}_{ij}$ denote the rotated azimuth and elevation angles after the node has

undergone rotation by α_i, β_i, and γ_i. Solving for this optimization provides us with the required orientation offsets to align all nodes in the system with the global frame of reference. The residual angles, $\hat{\theta}$ and $\hat{\phi}$, thus correspond to the AoAs in the global frame.

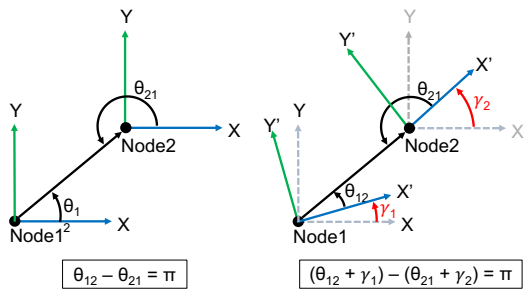


Figure 6.7: The relation between AoAs reported by two nodes (a) when nodes are perfectly aligned, and (b) when nodes are oriented with angle γ relative to the global frame of reference.

Next, we break down the steps in the optimization problem, detailing how we compute $\hat{\theta}$ and $\hat{\phi}$ through the rotation of local frames. The rotation in a 3D context can be decomposed into three unique rotations around the x, y, and z basis axes. *Step1* - Determine the unit vector based on the estimated azimuth and elevation angles:

$$U = \begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{bmatrix} = \begin{bmatrix} cos(\phi) * cos(\theta) \\ cos(\phi) * sin(\theta) \\ sin(\phi) \end{bmatrix}$$

Step2 - Apply the rotation matrix [431] to adjust the unit vector using the computed roll, pitch, and yaw angles (α, β, and γ). This results in a new unit vector, V, which reflects the node's orientation post-adjustment.

$$V = R_z(\gamma)R_y(\beta)R_x(\alpha)U$$

$$V = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & -\cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & -\cos\alpha \end{bmatrix} U$$

Step3 - Compute the azimuth and elevation angles from the rotated vector V to obtain the corrected AoA measurements.

$$\begin{bmatrix} \hat{\theta} \\ \hat{\phi} \end{bmatrix} = \begin{bmatrix} tan^{-1}(y/z) \\ tan^{-1}(z/\sqrt{x^2+y^2+z^2}) \end{bmatrix}$$

If we already have roll and pitch data from another sensor (e.g., IMU), the error function can become simpler. Instead of considering all three rotation angles, we can solve for yaw (γ). When the optimization is initiated with $\alpha_1 = \beta_1 = \gamma_1 = 0$, the first node sets the reference frame, and thus the orientations of all subsequent nodes are calculated relative to this frame.

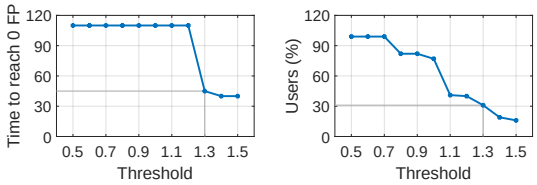
6.3 Opportunistic Anchor Integration

Locate3D is fundamentally designed to function without the necessity for anchors. However, when anchors are available in the infrastructure, *Locate3D* can seamlessly integrate them to refine location and orientation results and transition from relative to global coordinates. This capability offers an edge over range-only systems, which strictly require at least three anchors for global coordinates, which is a challenging requirement in highly mobile environments. We categorize anchors into the following two types:

(a) Static anchors: They are conventional infrastructure anchors with known locations. These anchors have higher edge weightage, which increases their likelihood of being selected in MST over mobile volatile edges. Their impact on the accuracy is evaluated in a later section with

varying anchor densities.

(b) Virtual anchors: As existing infrastructure cameras are available in many applications, we explore the potential to perform accurate localization in the vicinity and assign users within the camera's field-of-view as *virtual anchors*. A virtual anchor is a user with a high-confidence location obtained by an external modality, such as an infrastructure camera. It is important to note that while cameras can provide valuable localization information, they have limitations, such as reduced accuracy in low-light conditions.



(a) Time taken (in seconds) (b) Percent of users

Figure 6.8: (a) Time to reach '0 False positives'. (b) Percentage of users registered for varying trajectory matching thresholds.

User registration: Our registration technique uses correlation in user trajectories to incorporate human motion dynamics, such as varying walking speeds and stationary periods. The system benefits from *heading direction*, *cosine similarity of motion*, and speed analysis. We observe a scenario where some users are equipped with our node, while others are not, all within the camera's FoV. Our primary objective is to accurately register the system users in the camera's FoV while avoiding misclassifications. To address this, we adopt a stringent approach to trajectory similarity, aiming for zero false positives (FP). As depicted in Figure 6.8, keeping a higher threshold for registration effectively achieves zero FPs. By windowing 45 seconds of motion data, we register approximately 30% of visible users. It is important to note that while the availability of anchors can improve the accuracy, *Locate3D* does not rely on them for operation.

We evaluate the system's performance under various anchor densities, to show its adaptability to different infrastructure setups.

6.4 Implementation

To this end, we implement a prototype of *Locate3D* and perform experiments in room-level, building-level, and city-level scenarios.

Node prototype: We developed prototype mobile nodes of *Locate3D*, as shown in Figure 6.11a to evaluate the performance. Each node consists of a Raspberry Pi 3 [432] for computation, an Intel Realsense T261 [433] to collect visual inertial odometry data. We use the NXP SR150 UWB evaluation boards [429] to collect the range, azimuth, and elevation AoA data. We collect the UWB measurements at a rate of 20Hz. All the nodes are time-synced using NTP.

Software: All the prototype nodes are part of a network and communicate with each other using a UDP socket. We collect and store the UWB range, AoA, and VIO data from all nodes using Python scripts running on Raspberry Pi. We then process all measurements offline using Matlab where our algorithms are implemented.

6.5 Evaluation

We evaluate *Locate3D*'s performance in three key scenarios: room-scale, building-scale, and city-scale. In each scenario, we compare *Locate3D* with state-of-the-art baselines and benchmark localization accuracy, system latency, and the impact of various conditions such as line-of-sight (LOS) and non-line-of-sight (NLOS), stationary and moving nodes, lighting conditions, and varying numbers of infrastructure anchors. To assess *Locate3D*'s performance,

we have collected over 4 hours of real-world data using 32 nodes and employed a motion capture system [434] as a ground truth setup to evaluate the centimeter-level localization accuracy and sub-degree-level orientation accuracy. We compare *Locate3D* with two distinct baseline systems: an infrastructure-free system, Capella [3], that uses UWB+VIO for peer-to-peer 3D localization, and an infrastructure-dependent system, ULoc [5], that uses UWB anchors and angular information for parallel node triangulation. Additionally, we simulate a large city-scale network by stitching real-world measurements to evaluate *Locate3D*'s scalability for networks with up to 100,000 nodes.

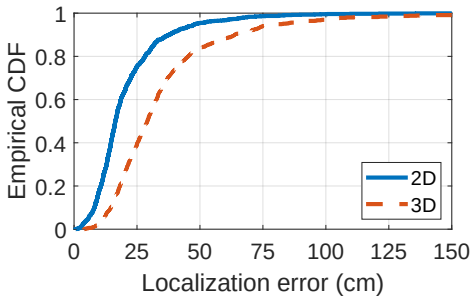


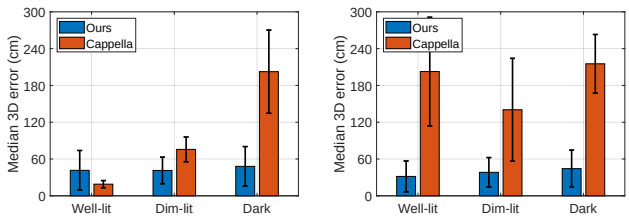
Figure 6.9: Room-scale evaluation: CDF of localization errors.

6.5.1 Localization Evaluation

6.5.1.1 Room-Scale Evaluation

We evaluate *Locate3D*'s performance in room-scale scenarios by conducting experiments in indoor rooms with dimensions of 6m $\times$ 8m and 4m $\times$ 20m, containing various furniture and equipment. We compare *Locate3D*'s performance with Cappella [3], a collaborative localization system that uses VIO for 3D tracking and inter-node UWB ranges to stitch individual trajectories using RBPF [435]. This evaluation involves 6 users walking naturally along random paths,

emulating an AR/VR application, with additional users not participating in the system present to create a crowded and dynamic environment (Figure 6.11b). No anchors are used in these experiments as both *Locate3D* and Cappella are infrastructure-free. We use a Vicon motion capture system [436] as the ground truth, which provides sub-millimeter accuracy at a frame rate of 200Hz. The Vicon system consists of 12 Vantage V8 cameras mounted on the ceiling to track the ground truth locations and orientations of all nodes using active IR light and reflective markers attached to the nodes. We use 9 markers per node to ensure reliable tracking in all orientations. Figure 6.9 shows the CDF of the absolute localization error in 2D and 3D estimated by *Locate3D*, with median errors $18cm$ and $30cm$, respectively, and 90th percentile errors of $38cm$ and $68cm$.



(a) (b)

Figure 6.10: 3D Localization performance of *Locate3D* compared to baseline system - Cappella [3] - which uses visual odometry along with UWB. Results show performance in different lighting conditions for (a) Moving nodes and (b) Static nodes.

Effect of lighting: We evaluate *Locate3D*'s performance under varying lighting conditions to demonstrate its robustness to changes in illumination. As shown in Figure 6.10a, *Locate3D*'s performance remains consistent across different lighting conditions, while Cappella's performance degrades as indoor illumination decreases. The light intensity values for Well-lit, Dim-lit, and Dark conditions were 400, 150, 50 lux respectively which were controlled by analog light dimmers. This highlights *Locate3D*'s advantage of relying on RF-based peer-to-peer measurements, which are unaffected by visual conditions, in contrast to Cappella's reliance on VIO and visual features

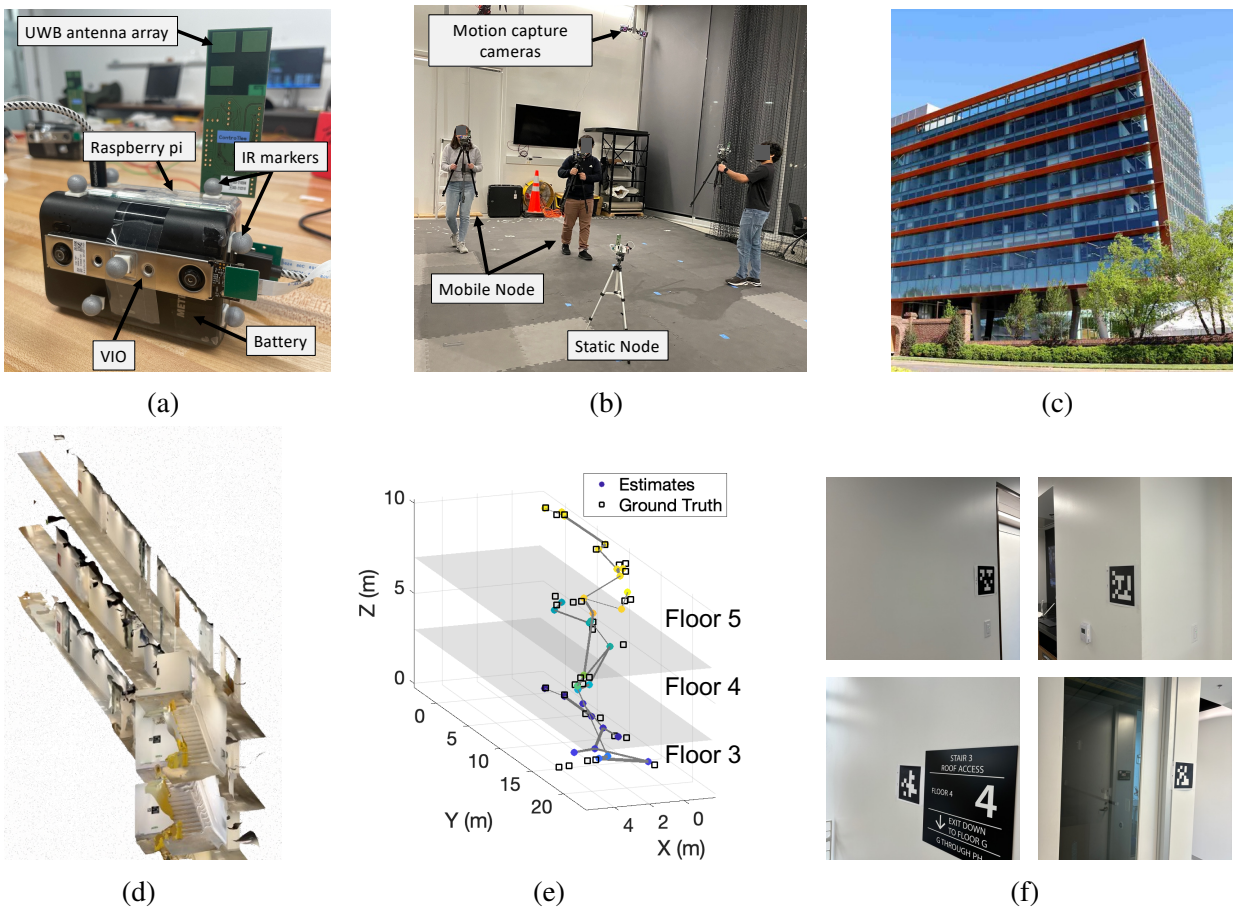


Figure 6.11: (a) Prototype built using Raspberry Pi, UWB sensor, Intel Realsense, IR markers for ground truth and a battery pack. (b) Room-scale evaluation. (c) Building-scale evaluation (d) 3D Lidar scan of the building for reference (not used in computation) (e) Snapshot of estimated locations and MST. (f) AprilTags [4] captured by nodes for ground truth (not used in computation).

of the environment.

Effect of motion: Figure 6.10b shows that when nodes are static and the baseline system lacks a VIO trajectory tail, it fails to estimate 3D locations accurately. In contrast, *Locate3D* does not rely on odometry and can estimate the complete 6DOF location and orientation for all nodes, regardless of whether they are static or mobile.

6.5.1.2 Building-Scale Evaluation

We evaluate *Locate3D*'s performance in a multi-level building scenario using 32 nodes across multiple floors, as shown in Figure 6.11c, 6.11d. The experiments include a combination of LOS and NLOS conditions, with users walking in the building while holding nodes that capture images of calibrated AprilTags [4] placed on the walls for ground truth reference (Figure 6.11f). *Locate3D* achieves a median 3D localization error of 0.86 meters and an average orientation error of 4.5 degrees without using any infrastructure anchors. Figure 6.11e shows a snapshot of the estimated topology. To validate the accuracy of the AprilTag-based ground truth system, we compared it against the Motion capture setup and found a 3D localization error of 0.11 meters. These results show *Locate3D*'s ability to provide accurate localization and orientation estimation in larger-scale, multi-level building environments without relying on infrastructure anchors. In section 6.5.2, we also do an ablation study to understand the contribution of each submodule to the system's latency and accuracy improvements.

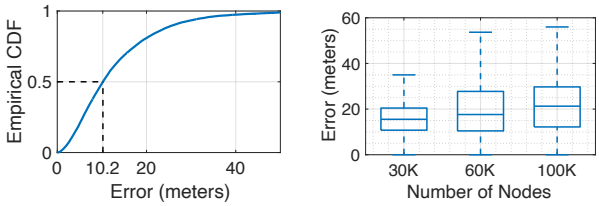


Figure 6.12: City-scale analysis: (a) CDF of localization errors of 30k nodes, $3800m \times 3800m$ area, and 15 anchors. (b) Errors for 30k, 60k, and 100k nodes in $1500m \times 1500m$ area and 1 anchor.

6.5.1.3 City-Scale Evaluation (Simulation)

We evaluate the performance of *Locate3D* in city-scale scenarios by simulating an environment with real-world collected data stitched in the simulation. The simulation incorporates real-world

peer-to-peer measurements collected using 32 nodes in various environments, including indoor, outdoor, LOS, and NLOS scenarios. The dataset consists of 20,000 measurements covering ranges from 0.1 to 50 meters and AoA spanning the full 360 degrees from both static and mobile users. To emulate large-scale city networks with up to 100,000 nodes, we employ a data-driven approach by stitching together real-world measurements to preserve the noise properties. The simulation environment includes obstacles that mimic real-world urban settings, and the corresponding measurements are selected from the appropriate subset of the dataset (e.g., nodes 20m apart, 30° and 200° azimuth and elevation AoAs and NLOS edge) to maintain the integrity of the simulation. During the simulation process, we strictly adhere to the real-world measurements, ensuring that the nodes experience realistic noise and variations in range and AoA based on the specific distance and angle of the interaction. This enables us to evaluate *Locate3D*'s performance in large-scale networks while maintaining a strong connection to real-world conditions. Figure 6.12 shows the CDF of localization errors for 30k nodes using 15 anchors in a $3800m \times 3800m$ area and the localization errors for 30k, 60k, and 100k nodes in a $1500m \times 1500m$ area using 1 anchor (for global translation only). *Locate3D* achieves a median localization error of 12.09 meters for large-scale networks (100,000 nodes with 15 anchors) and 21 meters with 1 anchor.

We further evaluate *Locate3D* in a wide-area simulation of New York City, spanning approximately $22000m \times 3200m$. We use 100k nodes with ground truth locations from the NYC transportation dataset [437], representing a Tuesday evening scenario. Figure 6.13 shows the results for configurations with 1 and 5 anchors. *Locate3D* achieved median localization errors of 82.31m and 34.19m for 1 and 5 anchors, respectively. The improved accuracy with more anchors demonstrates the system's ability to propagate location information effectively across wide-area environments.

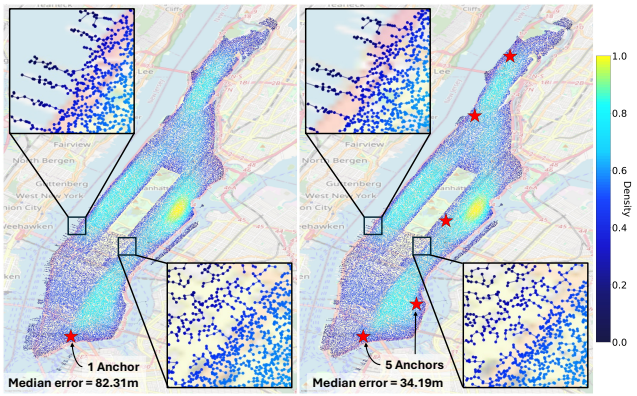
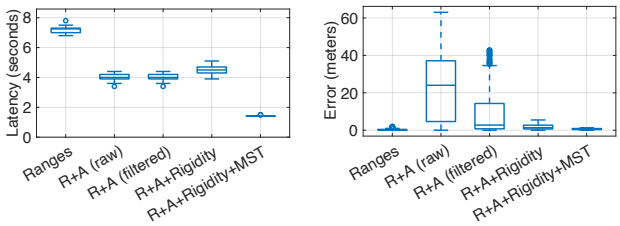


Figure 6.13: New York City wide-area simulation results: CDF of localization errors for 100,000 nodes using 1 (left figure) and 5 (right figure) anchors in a $22000m \times 3200m$ area.

6.5.2 Latency Evaluation

To evaluate the effectiveness of each submodule in latency improvement, we conduct an experimental ablation study in the building-scale scenario. The results are shown in Figure 6.14. Starting with a baseline system that uses only range information, we observe that the introduction of angle information (Range+Angle) significantly reduces the latency. However, the raw angle data leads to an increase in localization error due to the erroneous angles, as discussed in the section 6.2.2.2. To address this issue, we apply our filtering technique to fix erroneous angles which improves the accuracy while maintaining the latency reduction achieved by the introduction of angle information. The incorporation of rigidity constraints (R+A+Rigidity) further enhances the localization accuracy at a slight increase in latency compared to the filtered range and angle system. This is due to the addition of the critical edges that ensure a unique and unambiguous realization of the graph, as discussed in section 6.2.2.3. Finally, the optimal edge selection algorithm (R+A+MST) yields the most significant latency reduction while achieving localization accuracy comparable to the baseline system. This algorithm selects the optimal set of edges to minimize the overall latency while maintaining the graph's rigidity and uniqueness.



(a) (b)

Figure 6.14: Impact of submodules on (a) latency and (b) accuracy.

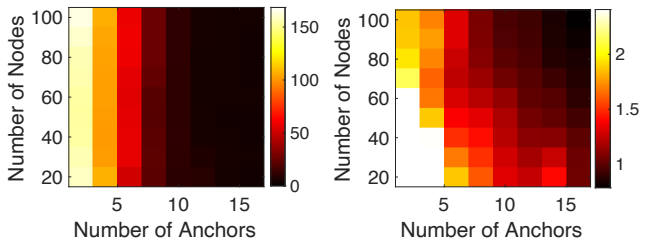
6.5.3 Micro-Benchmarks

6.5.3.1 Comparison with Infrastructure-Based Systems

To compare *Locate3D* with infrastructure-based systems, we evaluate the accuracy and scalability of *Locate3D* with a baseline infrastructure-based system ULoc [5], while varying the area, number of nodes, and number of anchors Our results show that *Locate3D* outperforms infrastructure-based systems in two key scenarios:

Sparse anchor deployments: *Locate3D*'s peer-to-peer network enables localization even when nodes are beyond the range of anchors, while ULoc fails in such cases. Figure 6.15 shows that both ULoc and *Locate3D* achieve meter-level errors when the number of anchors is high. However, as the number of anchors decreases, ULoc's average localization error increases due to nodes being out of range. Figure 6.16a demonstrates that with a fixed number of anchors and increasing area space, *Locate3D* can localize almost all nodes using its spanning tree, while the infrastructure-dependent technique loses connections to nodes making them unreachable.

NLOS edges: *Locate3D*'s edge selection algorithm prioritizes LOS measurements over NLOS, optimizing for overall noise reduction in the graph. In contrast, ULoc, being an infrastructure-based system, is constrained to using anchor-to-node measurements regardless of whether they



(a) Baseline (Infrastructure based)

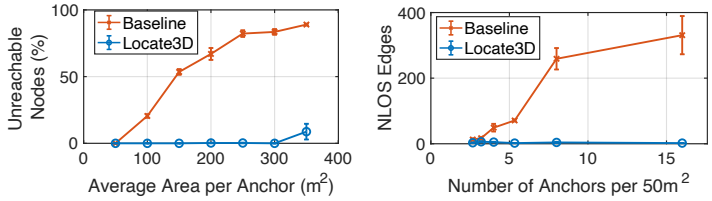
(b) *Locate3D*

Figure 6.15: Localization errors (in meters) for infrastructure baseline [5] and *Locate3D* for varying nodes and anchors.

are LOS or NLOS. This limitation leads to the inclusion of noisy edges in the graph, affecting the overall localization performance. As shown in Figure 6.16b, *Locate3D* selectively includes LOS edges and minimizes the number of NLOS edges resulting in improved accuracy in realistic environments. By reducing the overall noise in the topology, *Locate3D* is able to provide more reliable and precise location estimates compared to ULoc, which is susceptible to errors introduced by NLOS measurements between nodes and anchors. This highlights the benefits of a peer-to-peer approach, which allows for greater flexibility in edge selection and noise reduction compared to infrastructure-based systems. However, infrastructure-based systems outperforms *Locate3D* in the following scenarios:

High anchor density and small-area experiments: As shown in Figure 6.15, infrastructure-based systems outperform in high anchor density cases where the error decreases due to close proximity with multiple anchors. Such small-scale scenarios with clear LOS and at least three anchors are where ULoc achieves better accuracy.

Locate3D's scope is to scale to a large number of nodes, while infrastructure-based systems are well-suited for small-scale deployments with sufficient anchor density.



(a) (b)

Figure 6.16: (a) Number of unreachable nodes and (b) Total number of NLOS measurements made with varying anchor area density.

6.5.3.2 LOS vs NLOS nodes

Figure 6.17(a) presents the CDF of the 2D and 3D localization errors in both line-of-sight and non-line-of-sight scenarios. In the NLOS scenario, there is a slight increase in localization error compared to the LOS scenario. This increase is primarily attributed to the higher noise levels in the received signals at the nodes due to obstructed signal paths and reflections. Specifically, in the 2D scenario, the median localization error rises from $13cm$ in LOS to $28cm$ in NLOS. Similarly, in the 3D scenario, the median localization error increases from $31cm$ in LOS to $39cm$ in NLOS.

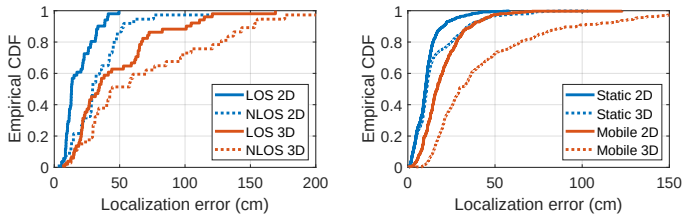
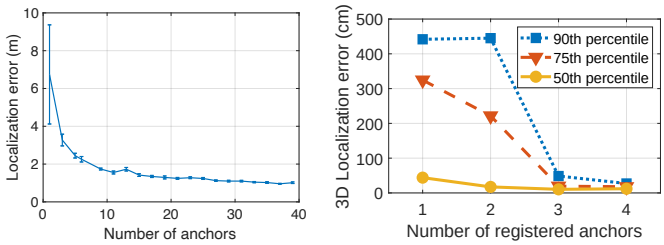


Figure 6.17: (a) LOS and NLOS localization errors. (b) Static vs Mobile nodes localization errors

6.5.3.3 Static vs Mobile nodes

To assess the impact of node mobility on the performance of *Locate3D*, we conduct experiments to compare static and mobile node scenarios. Figure 6.17(b) presents the CDF

of 2D and 3D localization errors in both static and mobile node configurations. In the mobile scenario, the system experiences a slight degradation in performance due to varying environments and node vibrations. In the 2D scenario, the median localization error increases from a low of $6cm$ in the static setup to $13cm$ in the mobile scenario. Similarly, in the 3D scenario, the median localization error rises from $6.5cm$ in the static scenario to $26cm$ in the mobile scenario.



(a) (b)

Figure 6.18: (a) City-scale results for a 1000 node topology simulation in a $200 \times 200 \times 50$ meter 3D space for a varying number of static anchors. (b) Room-scale localization error for real-world 20-node experiments with varying numbers of anchors registered.

6.5.3.4 Impact of Anchor Density

We evaluate the impact of anchor density on localization performance in both large-scale simulations and real-world room-scale experiments. While *Locate3D* does not require anchors for operation, the presence of anchors can improve absolute localization accuracy in relative global coordinates. In a large-scale simulation with 1,000 nodes distributed within a 200m $\times$ 200m $\times$ 50m 3D space, *Locate3D* achieves a localization error below 2 meters with just 5 anchors (Figure 6.18a). In a real-world room-scale experiment, we evaluate the localization error while varying the number of registered anchors. *Locate3D* can peg the entire topology using only two anchors by leveraging angle information alongside location estimates. Figure 6.18b shows that the localization error decreases as more anchors are registered, with the main source of error being

the rotation of the topology when only one anchor is present.

6.5.3.5 Ranging and AoA Performance

The localization errors inherently depend on the raging and AoA accuracies. We report the ranging and AoA estimation performance of the UWB sensor used in our implementation.

While our system can be applied to any technology, here in the implementation we have used the UWB radio to get the distance and AoA error. So in this section, we report the accuracy of the ranging and angle estimation accuracy.

Figure 6.19 shows the ranging and AoA performance of our node for varying distances and incoming angles. We see that there are biases in certain angles as it is likely due to multipath. But still the 90th percentile error remain within $\pm 5°$.

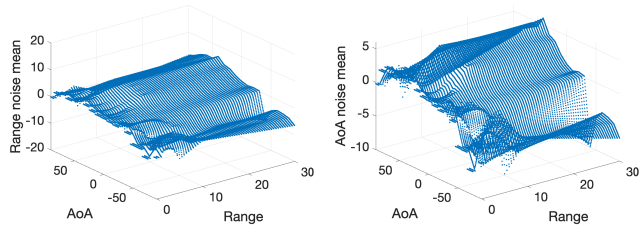


Figure 6.19: Ranging and AoA errors for various range-angles.

6.5.3.6 Orientation Estimation Accuracy

Figure 6.20(a) presents the CDF of the error in orientation estimation results. It is highly important for our system as these orientations play a critical role in transforming AoAs from the local frame of reference of the nodes to the global frame of reference, which is then used localization of nodes within a 3D space. Figure 6.20(a) shows that *Locate3D* consistently estimates accurate node's orientation with a median error of 4 degrees which is competitive with an inertial

sensor.

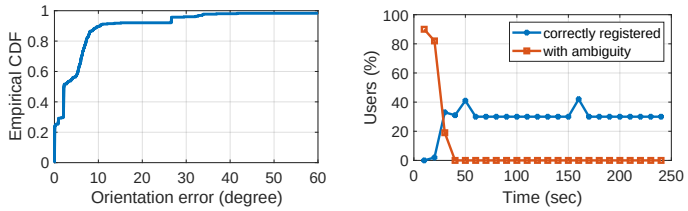


Figure 6.20: (a) CDF orientation errors. (b) Percent of users registered as Virtual anchors over time.

6.5.3.7 Virtual Anchor Registration

We evaluate the virtual anchor registration technique through a building-scale experiment with a network-connected camera. This setup included four node users and three non-participants walking within the camera's FoV, simulating typical crowd dynamics. Figure 6.20(b) shows the results that within the initial 40 seconds, the system registered 40% of the users as virtual anchors, while distinguishing between system users and non-participants.

6.6 Discussion

Scalability in real-time implementation: An extension of the proposed system with a semi-distributed approach can further improve its worst-case latency and robustness. In this case, the network is divided into sub-graphs managed by a local leader node. The leader nodes compute MSTs per sub-graphs in parallel and later aggregate to form the complete graph. This distributed architecture can offer better handling of high mobility scenarios subjected to future research.

Bootstrapping: *Locate3D*'s initial bootstrapping phase establishes the first set of measurements and constructs the initial graph without assumptions about node locations or network topology. Each node measures all neighboring nodes, building the graph incrementally based on responses.

While this cold start phase naturally adds to the latency, it does not compromise localization accuracy. The duration of this phase depends on factors such as the number of nodes and application requirements. In subsequent iterations, *Locate3D* uses previous iteration information to efficiently select edges, reducing latency. While cold start is a common issue in state-of-the-art peer-to-peer localization systems, optimizing the bootstrapping process can be a direction for future work.

Applications: *Locate3D* caters to a wide range of applications requiring accurate and fast localization in infrastructure-free networks. Inspired by first responders scenarios, real-world experiments in multi-level buildings demonstrate the system's potential in emergency response scenarios. *Locate3D*'s ability to operate in GPS-denied environments and adapt to changes in network topology makes it suitable for disaster management and rescue operations. Its scalability, demonstrated through our city-scale simulations, shows its potential for smart city applications such as asset tracking, crowd monitoring, and urban planning. *Locate3D*'s framework can be easily adapted to incorporate application-specific parameters and constraints.

Advancements in angle field-of-view: Future sensor advancements, such as increased antenna counts, may enable 360° field-of-view coverage. In such cases, the spanning tree algorithm can be adapted to consider either range+angle edges or no edges, and account for varying noise distributions across different incoming angles. *Locate3D*'s framework is flexible to accommodate these advancements and generate optimal spanning trees without significant modifications.

Subgraphs without proximity for extended periods: Our critical edges technique addresses cases where subgraphs are out of vicinity for long durations. However, prolonged reliance can lead to drifting. Reinitializing the tree or cold-start can mitigate this issue. Presence of anchors in subgraphs can also provide reference points to align relative topology.

Other RF modalities: Our framework is adaptable to other modalities beyond UWB, including RF and acoustic sensors that can measure pairwise ranges and AoA [438–441].

6.7 Related Work

Infrastructure-based systems: Motion tracking cameras (OptiTrack [442], Vicon [434]), passive markers (ARTags [443, 444], AprilTags [4, 445, 446]), and beacon-based solutions (UWB [447–449], Bluetooth [450–452], ultrasound [452–454]) are used for localization. These methods often combine ranging with odometry from IMU or VIO [448, 455–460]. However, infrastructure-based systems are limited to prepared environments. *Locate3D* leverages peer-to-peer UWB ranges and angles, eliminating the need for infrastructure.

Infrastructure-free systems: Infrastructure-free localization determines relative positions between users without global references. This has been explored in sensor networks for stationary [461, 462] and mobile [182, 463–465] nodes, and in robotics using visual object detection [466–468] or odometry and distance measurements [469–475]. Recent research [3, 474, 475] employs visual-inertial features and UWB ranges for localization, but have limited scalability.

UWB localization: UWB-based systems have shown decimeter-level accuracy [351, 476–479]. However, these systems often require a high number of infrastructure anchors, limiting their scalability. ULoc [5] uses AoA for triangulation but requires dense anchor deployments. XRLoc [480] employs a single large specialized anchor with multiple UWB nodes, focusing on small-scale, meter-level applications. SALMA [476] uses only one anchor but is limited to 2D and supports a small number of concurrent tags. In contrast, *Locate3D* is an infrastructure-free system that leverages both range and angle information for 6DoF localization. Unlike Cappella [3],

which relies on VIO and UWB measurements, *Locate3D*'s performance is independent of lighting conditions. Moreover, *Locate3D* achieves localization without depending on odometry, preventing error accumulation.

6.8 Conclusion

This chapter presents *Locate3D*, a peer-to-peer 3D localization and orientation estimation system for large networks. Our novel optimization and edge selection algorithm integrates both range and angle-of-arrival measurements for enhanced network topology estimation, demanding fewer edge measurements. As a result, *Locate3D* improves the latency by 75% without sacrificing accuracy, outperforming conventional range-only solutions.

Chapter 7: Low-Cost and Dynamic Food Quality Sensing at the Pallet-Level

7.1 Introduction

Food wastage is a pervasive issue that affects millions of people worldwide. Approximately 40% of produced food goes to waste after harvest [481] while 10% of the world's population goes to bed hungry each night [482]. The food industry is also known to have some of the highest emissions, and the environmental impact from produce that gets wasted post-harvest is a significant overhead to the industry's carbon footprint. On these accounts, there is an imminent need to detect and reduce food loss.

In an effort to address this problem, in this work, we ask the question: *Can we use wireless technology to track the change in food quality and nutrition, post-harvest, as it travels through the global supply chain?* Current methods for monitoring food quality are sparse and prone to manual error. Literature shows that moisture content in produce is a good proxy for quality and nutrition [483]. However, the inability to accurately and efficiently track moisture content at scale leads to substantial loss.

To address this challenge, we propose **FreshSense**, a novel wireless sensing system that enables real-time, non-invasive monitoring of moisture content in produce as it moves through the supply chain.

One of the major challenges leading to loss in the food supply chain is the unclear labeling of

shelf life. The current labeling system is static and does not accurately reflect the dynamic nature of food quality as it travels globally. There is a need for a technology that can cost-effectively integrate into the global food supply chain and provide real-time monitoring of food quality.

Decades of optimization in large-scale transport have led to the widespread use of pallets with boxes designed to support air circulation. These pallets can be lifted with forklifts and are easy to store and transport in planes, ships, and trucks. They are usually packed in the early stages of the supply chain, then stored and transported, and finally opened at distribution centers or retail stores.

While individual fruit sensing exists during the pre-packaging stages, post-packaging is where produce spends most of its time — potentially years in cold storage and distribution. Manual monitoring methods are inadequate for the vast quantities involved; for instance, over 9 million metric tonnes of avocados alone are produced annually [484]. This outlines the need for a pallet-level sensing technology that can dynamically track the freshness and nutrition of produce throughout the global supply chain.

Existing technologies for sensing food quality include chemical indicators like ethylene gas sensors, or direct invasive measurement of dry matter using methods like Karl Fischer moisture analysis [485]. Camera-based techniques using computer vision and deep learning are also employed, but they primarily analyze surface colors and textures, not the core of the pallets where most ripening occurs [486].

These existing technologies are not scalable. Although some non-invasive and highly accurate techniques exist, such as portable Near-Infrared (NIR) spectrometers [487], they are extremely expensive and can only support sparse sampling.

In this chapter, we design and implement **FreshSense**, a physics-informed wireless sensing

system that leverages dispersion in radio waves to accurately detect moisture content in produce at the pallet level. By seamlessly integrating into the food supply chain, FreshSense enables cost-effective, real-time monitoring of food quality across multiple stages, from post-harvest to distribution.

7.2 Understanding Food Quality Metrics

In the food industry, two primary metrics are used to assess food quality: *Dry Matter (DM)* and *Brix*. Dry matter represents the percentage of a fruit's weight that is not water, while Brix is a measure of the fruit's sugar content. These metrics are crucial indicators of a fruit's ripeness, lifecycle, nutrition, and overall freshness. The FDA has established regulations [488] to maintain ideal DM and Brix levels in various food products at the time of harvest and consumption, ensuring that consumers receive high-quality, nutritious produce.

As fruit ripens, the DM and Brix values change due to factors such as water loss through osmosis and the conversion of starch into sugars. For example, when an avocado is packaged, its dry matter content is typically around 20%–25%. As it reaches the consumer, the ideal dry matter content for consumption is between 25%–30%. However, if the dry matter content exceeds 35%, the avocado is considered overripe or rotten. Similarly, in apples, the optimal Brix level at harvest is between 12%–14%, indicating the ideal balance of sweetness and ripeness.

Our goal is to estimate these metrics, which serve as measures of maturity and remaining shelf life, enabling more informed decision-making throughout the food supply chain.

In this work, we focus on estimating dry matter, as it is the preferred index for assessing fruit quality and shelf life. An increase in dry matter content inherently enhances the Brix value

as well, due to their correlation. As dry matter content rises, the total soluble solids (TSS) also increase [489].

Dry matter content is a better indicator of TSS because it encompasses all solid components, including soluble sugars and insoluble starches. This is crucial because starch, which eventually converts into sugar, is included, making dry matter an effective predictor of freshness in fruits.

Starting with a certain DM percentage, the value changes as food begins to ripen, and our goal is to sense these changes accurately. Extensive research has been conducted on mapping DM and Brix values to nutrition levels for different types of food items [490].

By sensing water content changes and estimating DM at various stages in the food supply chain, our system has the potential to provide a detailed breakdown of the dynamic nutrition levels in food. This information can be invaluable for consumers, retailers, and food processors alike, enabling them to make more informed decisions about food quality, shelf life, and nutritional value.

7.3 Physics and Core Intuition

Detecting food quality at the pallet level presents unique challenges compared to individual fruit sensing solutions. In the supply chain, produce is frequently shuffled during transportation and handling, leading to dynamic movement and variations in signal paths caused by scattering, reflection, and absorption within the pallet. Pallet-level sensing must contend with these practical challenges, which differ significantly from the controlled environments of conveyor belt systems.

To address these challenges, we employ radio frequencies that can penetrate packaging materials and the pallet itself. Unlike higher frequencies such as millimeter-wave or terahertz,

which are predominantly reflected or absorbed, sub-6 GHz frequencies offer a balance between penetration depth and sensitivity to water content changes. However, microwaves are also less sensitive to changes in water content, necessitating the development of sophisticated sensing techniques to extract meaningful data.

At the core of our approach lies the phenomenon of *dispersion* [491], where different frequencies of a wireless signal travel at different speeds through water-rich environments like fruits and vegetables. In traditional wireless communication, all frequencies are assumed to propagate at the same speed through air. However, in water-rich environments, this assumption no longer holds.

By analyzing the frequency-dependent delay caused by dispersion, we can infer the water content in the pallet and gain insights into the produce's freshness and quality.

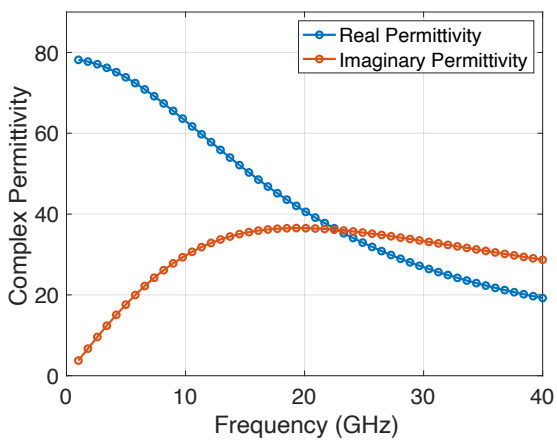


Figure 7.1: Complex permittivity of electromagnetic waves in water.

To understand the physics behind our method, we focus on the complex permittivity of water, denoted as

$$\varepsilon(f) = \varepsilon'(f) - j\varepsilon''(f),$$

where $\varepsilon'(f)$ is the real part (dielectric constant) that determines the signal slowdown in the material,

and $\varepsilon''(f)$ is the imaginary part representing the loss factor [492], as shown in Figure 7.1.

The speed of a signal $c(f)$ through water is given by:

$$c(f) = \frac{c}{\sqrt{\varepsilon'(f)}},$$

where c is the speed of light in a vacuum.

As the dielectric constant of water increases with frequency, higher frequency signals experience a greater slowdown when traveling through produce (Figure 7.2).

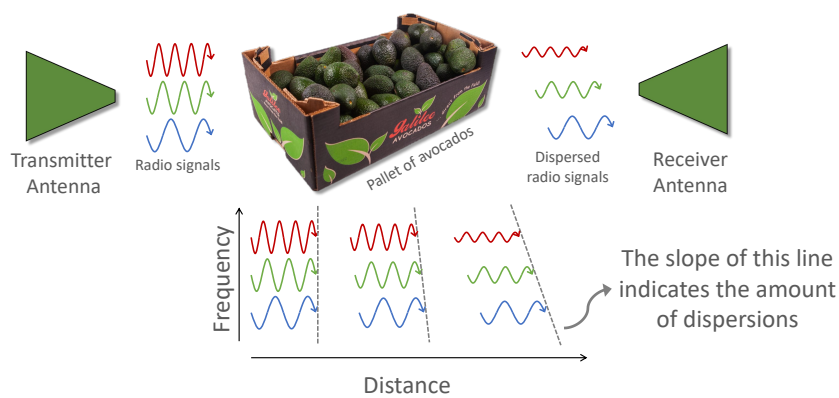


Figure 7.2: FreshSense: Different frequencies travel at different speeds through water resulting in an additional frequency dependent delay.

To validate our approach, we conducted a controlled experiment using a pallet box filled with water bottles, simulating different levels of water content (20% to 90% of the total volume). The bottles were shuffled, and the external environment was varied to capture multipath effects. Figure 7.3 shows the extracted dispersion levels for frequencies between 1–5 GHz, demonstrating a clear correlation between frequency-dependent delay and water content.

While this experiment is an exaggerated representation of real-world scenarios, it validates the fundamental principle that different frequencies travel at different speeds in water-rich

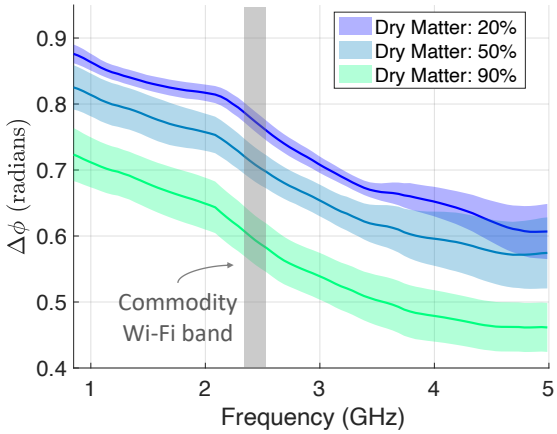


Figure 7.3: Estimated dispersion for varying percentage of water in the box.

environments, manifesting as accumulated phase in the received signal.

7.4 Design and Implementation

Building upon the foundations of dispersion, we proceed to develop a practical system to demonstrate the assessment of DM content in avocados, a high-value crop with significant post-harvest losses. As avocados ripen, the DM content changes, altering the amount of water molecules present.

However, fruits are complex mixtures of water, sugar, fiber, and other components, making it challenging to develop a precise physical model of how each frequency interacts with avocados. To address this, we approach the problem using machine learning.

We build a non-invasive, pallet-level dispersion sensor capable of predicting DM content in produce. Our system transmits and receives 30 MHz chirps, continuously scanning from 800 MHz to 6 GHz by varying the center frequencies of software-defined radios (SDRs). Our experimental setup is shown in Figure 7.4.

To identify the most informative frequencies for DM prediction in avocados, we first perform

Figure 7.4: Evaluation setup with a box of avocados.

a principal component analysis (PCA) on the collected data. This helps determine the frequencies that exhibit the highest correlation with DM content. By focusing on these top frequencies, we can narrow down the range and optimize the system.

In our prototype, we use two separate SDRs for transmitting and receiving the chirps. To synchronize the transmitter and receiver, we employ a reference antenna connected to the receiver SDR, enabling us to determine the relative time difference between the signals received by the reference and main receiving antennas.

Directly feeding the received signals into a machine learning model would require extensive training data and would struggle to generalize. To overcome this, we leverage dispersion physics and extract the extra delay, the unwrapped accumulated phase, across frequencies.

While amplitude-based approaches are highly sensitive to fading and interference, phase-based techniques offer more robustness. We do not rely on raw phase; rather, we focus on extracting the differential time delay caused by slower propagation through water-rich produce. Using a wideband signal with sparse frequency steps, we capture the dispersion characteristics as a more agnostic feature.

To extract the dispersion delays - First, we *dechirp* the received signal by multiplying it with the transmitted signal. Then, we apply a Fast Fourier Transform (FFT) to extract the subcarrier phase at each frequency. Finally, we filter and smooth the stacked subcarrier phases to identify the general trend across the wideband signal.

To further mitigate environmental impacts, we perform two scans at the same antenna locations: (1) One without the pallet (free air line-of-sight), (2) Another with the pallet between the antennas. This approach captures the environmental channel characteristics and isolates the additional phase delay caused by the produce.

We envision integrating such a system into forklifts, commonly used to move and load pallets at various checkpoints. A single reference scan can be performed once, followed by multiple pallet scans. For our machine learning model, we employ a regression approach and ensure diversity in training by shuffling the pallet and slightly varying the quantity of produce. Ground truth measurements are obtained using an industry-standard Felix Instruments NIR Food Quality Sensor, a portable spectroscope meter that uses near-infrared spectrum to estimate DM and Brix values of fruits.

7.5 Performance Evaluation

To evaluate the performance of the **FreshSense** system, we conducted a study on avocados, tracking their ripening process over a span of 14 days. We labeled 5 avocados within the batch and measured the ground truth dry matter (DM) content using the Felix Instruments NIR Food Quality Sensor. Simultaneously, we collected RF measurements twice daily from shuffled pallets using our dispersion-based sensing method.

Figure 7.5: Images of avocados captured across 14 days.

Figure 7.5 shows visual images of the same 5 avocados throughout the 14-day period. While the change in green saturation from Day 1 to Day 14 is clearly visible, this technique has limitations. Color-based assessment cannot provide fine-grained measurements and requires a clear line-of-sight and calibrated imaging setup, making it impractical for pallet-level monitoring.

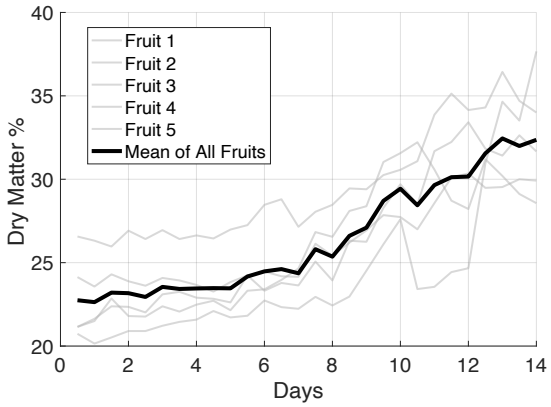


Figure 7.6: Ground truth measured for 5 individual fruits and their average trend.

Figure 7.6 illustrates the ground truth DM content over time for all 5 avocados and their average value. As the fruit ripens, the DM content begins to rise, with a notable acceleration after Day 5. By Day 14, the avocados are considered overripe. This trend highlights the importance of accurately monitoring DM content to optimize shelf life and quality of avocados in the supply chain.

To evaluate the performance of our system, we calculated the R-squared correlation between the true DM values and the estimated values obtained from our model.

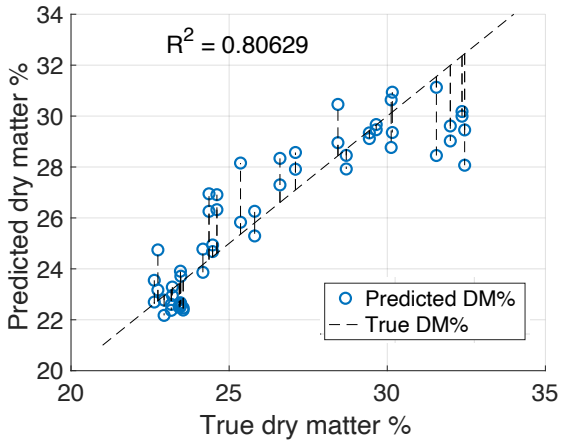


Figure 7.7: Correlation between estimated and true DM% for the testing data.

As shown in Figure 7.7, our dispersion sensing approach achieves an R-squared value of 0.8, indicating a strong correlation between the predicted and true DM content. In comparison, an amplitude-based model yielded a lower R-squared value of 0.64, highlighting the advantages of the dispersion-based approach in capturing ripening.

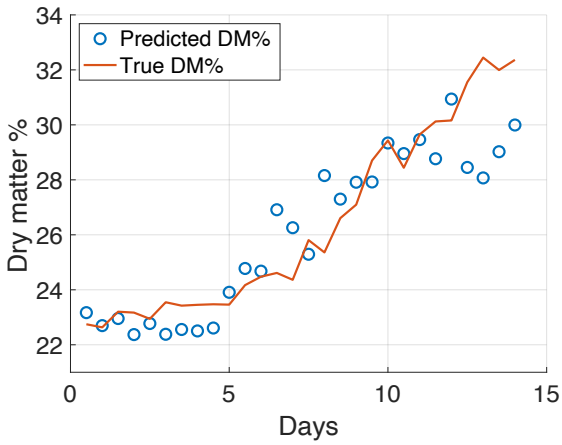


Figure 7.8: Estimated and true DM% across 14 days.

Figure 7.8 presents the estimated DM content over time. Note that **FreshSense** precisely

identifies the critical tipping point at Day 5, when the ripening process accelerates. This ability to detect such pivotal moments demonstrates the potential of our approach to enable proactive decision-making regarding storage and distribution, ultimately optimizing shelf life and reducing food waste.

7.6 Conclusion

FreshSense is a system designed to non-invasively monitor the quality and shelf life of produce at the pallet level in the post-harvest supply chain. It uses dispersion-based wireless sensing and AI to demonstrate efficient quality assessment, which can reduce food loss in the supply chain. This work serves as a foundation for developing more ubiquitous sensing devices and a generalized model for tracking quality metrics of produce. By leveraging advancements in AI and hardware innovations, we aim to transform the way we produce, distribute, and consume food, paving the way for a more sustainable future.

Chapter 8: Conclusion and Future Directions

This thesis explored a new approach to embed intelligence in everyday objects through the development of sensing systems that can operate within extreme resource constraints. By designing sensing and computing architectures from first principles, we have shown how to bridge the gap between conventional intelligent systems and the vision of ambient intelligence at scale. Our work shows that by combining bio-inspired designs, physics-informed machine learning, and scalable architectures, we can create intelligent systems that achieve orders of magnitude improvements in energy efficiency, size, and scalability. In this final chapter, we summarize the key contributions, discusses broader implications, and outlines promising directions for future research.

8.1 Summary of Contributions

This thesis makes the following contributions to the field of sustainable ambient intelligence:

- **Structure-Assisted Spatial Intelligence:** We demonstrated that intelligent passive structures can transform single sensors into spatial perception systems without requiring expensive sensor arrays or power-hungry processing. Our approach draws from nature's solutions to similar perception problems, creating systems that match the capabilities of much more complex devices while using orders of magnitude less power. With *Owlet*, we developed a

direction-finding system that uses 3D-printed metamaterials to achieve 3.6° angular accuracy with a single microphone, matching a 9-microphone array while using $100\times$ less power. Through *SPiDR*, we created a single-sensor depth imaging system for micro-robotics that achieves centimeter-level accuracy while consuming $400\times$ less energy than traditional approaches like LiDAR. Our *Lyra* system provides passive spectral analysis inspired by the human cochlea that performs FFT-like processing with microwatt power consumption. Finally, with *Sirius*, we developed an RF self-localization system that uses reconfigurable antennas to achieve 7° angular accuracy with $1000\times$ less energy than array-based systems.

- **Scalable nextG Systems:** We developed localization for resource-constrained devices by leveraging existing infrastructure and innovative peer-to-peer algorithms. Our systems enable global-scale tracking with minimal energy consumption, opening new possibilities for applications in supply chains, healthcare, and environmental monitoring. *LiTEfoot* provides a cellular-based localization system that achieves GPS-like accuracy with RFID-like form factor, consuming only 40μJ per location fix, a $625\times$ reduction compared to GPS. With *Locate3D*, we created an infrastructure-free 6DoF tracking system that enables coordinating up to 100,000 devices across urban environments with significantly reduced latency.

- **Systems for Sustainability:** We developed novel sensing approaches for monitoring food quality throughout global supply chains, addressing a critical global challenge. *FreshSense* provides a non-invasive system that monitors dry matter content in produce at the pallet level, enabling data-driven decisions to reduce food waste in distribution networks.

8.2 Broader Impact

The systems developed in this thesis have the potential for significant societal impact across multiple domains:

- **Healthcare and Assistive Technologies:** The ultra-low-power sensing approaches we've developed could transform medical implants, wearable health monitors, and assistive technologies for elderly or disabled individuals. By enabling long-term, maintenance-free operation, these systems could provide continuous health monitoring without requiring frequent battery replacements or complex infrastructure.

- **Disaster Response and Recovery:** Our scalable tracking systems could significantly improve coordination in disaster scenarios, enabling first responders to navigate complex environments more effectively and maintain awareness of team members' locations even in infrastructure-compromised settings.

- **Supply Chain Efficiency:** The combination of low-power tracking and quality monitoring systems could transform global supply chains, reducing losses, improving traceability, and enabling more efficient distribution of goods worldwide.

- **Global Food Security:** Our food quality monitoring systems address a critical challenge in global food distribution, where approximately 40% of food is lost to waste. By enabling continuous, non-invasive monitoring throughout supply chains, these systems could help ensure that food reaches consumers in optimal condition, reducing waste and improving access to nutrition worldwide.

8.3 Future Directions

While we have made significant strides toward sustainable ambient intelligence, various exciting research opportunities remain open. We outline some promising directions here:

- **Spatial Intelligence for Large AI Models:** While modern AI systems excel at processing and understanding content, they fundamentally lack the ability to reason about spatial context. Building on our work in direction-aware sensing, future research could develop techniques for integrating spatial awareness into large models. This could involve creating physics-informed neural architectures that process and reason about spatial signals, developing efficient spatiotemporal attention mechanisms, and designing novel multi-modal fusion techniques. These advances would enable transformative applications, from intelligent meeting assistants that track and attribute speech to specific participants, to context-aware AR systems that understand and respond to spatial queries about the environment.

- **Next-Generation Communication and Sensing:** Traditional wireless systems treat sensing and communication as independent functions, leading to inefficient spectrum and energy use. Future research could develop unified architectures that jointly optimize sensing and communication, embedding sensing capabilities within communication signals to achieve simultaneous ranging, imaging, and data transfer.

- **Neuromorphic 3D Perception:** Creating event-driven sensing architectures that achieve efficient 3D reconstruction through biologically-inspired processing. By combining ultra-low power analog computing with adaptive sampling, these systems could enable real-time 3D perception in micro-scale devices.

- **In-body and Wearable Health Monitoring:** Current medical devices face fundamental challenges in continuous monitoring, from power constraints in implants to signal quality in wearables. Future research could explore acoustic backscatter for ultra-low power communication from in-body sensors, enabling continuous glucose monitoring and early disease detection. For wearables, direction-aware earables could enable selective hearing aids that isolate conversations while monitoring vital signs. By combining physics-aware neural architectures with novel backscatter techniques, we could achieve orders of magnitude improvements in battery life while maintaining clinical-grade sensing accuracy.

- **Precision Agriculture and Food Security:** Global agriculture needs to increase food production by 70% by 2050 while using fewer resources. Future work could repurpose existing wireless infrastructure, like cellular networks and satellite constellations, for agricultural sensing. By extracting meaningful information from ambient signals' interaction with crops, we could enable continuous monitoring of soil moisture and plant health at unprecedented scales with minimal infrastructure cost, making data-driven agriculture accessible to farmers worldwide.

8.4 Closing Thoughts

This thesis establishes a new paradigm for intelligent systems, demonstrating that sensing and computation can be designed to operate intrinsically within extreme resource constraints while maintaining intelligent capabilities. Our approach combines nature-inspired design, physics-informed machine learning, and scalable architectures to realize ambient intelligence at unprecedented scale and efficiency. The principles and systems developed in this Ph.D. lay the foundation

for ubiquitous, sustainable, and seamlessly embedded intelligence in everyday objects, where computation becomes as pervasive and energy-efficient as the natural world itself.

Bibliography

[1] Texas Instruments. Texas instruments msp430fr5969. https://www.ti.com/tool/MSP-EXP430FR5969, 2021.

[2] Keysight. Keysight data acquisition system u2331a. https://www.keysight.com/us/en/assets/7018-03510/data-sheets/5991-0566.pdf, 2021.

[3] John Miller, Elahe Soltanaghai, Raewyn Duvall, Jeff Chen, Vikram Bhat, Nuno Pereira, and Anthony Rowe. Cappella: Establishing multi-user augmented reality sessions using inertial estimates and peer-to-peer ranging. In *2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 428–440. IEEE, 2022.

[4] John Wang and Edwin Olson. Apriltag 2: Efficient and robust fiducial detection. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4193–4198. IEEE, 2016.

[5] Minghui Zhao, Tyler Chang, Aditya Arun, Roshan Ayyalasomayajula, Chi Zhang, and Dinesh Bharadia. Uloc: Low-power, scalable and cm-accurate uwb-tag localization and tracking for indoor applications. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(3):1–31, 2021.

[6] Mark Weiser. The computer for the 21 st century. *Scientific american*, 265(3):94–105, 1991.

[7] Joseph M Kahn, Randy H Katz, and Kristofer SJ Pister. Next century challenges: mobile networking for "smart dust". In *Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking*, pages 271–278, 1999.

[8] Philip Sparks. The route to a trillion devices. *White Paper, ARM*, 2017.

[9] David Owald, Suewei Lin, and Scott Waddell. Light, heat, action: neural control of fruit fly behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1677):20140211, 2015.

[10] Rudiger Wehner. Desert ant navigation: how miniature brains solve complex tasks. *Journal of Comparative Physiology A*, 189:579–588, 2003.

[11] Ethan Fast and Eric Horvitz. Long-term trends in the public perception of artificial intelligence. In *Proceedings of the AAAI conference on artificial intelligence*, volume 31, 2017.

[12] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35:24824–24837, 2022.

[13] Haochen Hua, Yutong Li, Tonghe Wang, Nanqing Dong, Wei Li, and Junwei Cao. Edge computing with artificial intelligence: A machine learning perspective. *ACM Computing Surveys*, 55(9):1–35, 2023.

[14] HJ Landau. Sampling, data transmission, and the nyquist rate. *Proceedings of the IEEE*, 55(10):1701–1706, 1967.

[15] Harry Nyquist. Certain topics in telegraph transmission theory. *Proceedings of the IEEE*, 90(2):280–305, 2002.

[16] Meng-Lin Ku, Wei Li, Yan Chen, and KJ Ray Liu. Advances in energy harvesting communications: Past, present, and future challenges. *IEEE Communications Surveys & Tutorials*, 18(2):1384–1412, 2015.

[17] Nakul Garg, Yang Bai, and Nirupam Roy. Owlet: Enabling spatial information in ubiquitous acoustic devices. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, pages 255–268, 2021.

[18] Yang Bai, Nakul Garg, and Nirupam Roy. Spidr: Ultra-low-power acoustic spatial sensing for micro-robot navigation. In *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*, pages 99–113, 2022.

[19] Nakul Garg, Harshvardhan Takawale, Yang Bai, Irtaza Shahid, and Nirupam Roy. Structure assisted spectrum sensing for low-power acoustic event detection. In *Proceedings of Cyber-Physical Systems and Internet of Things Week 2023*, pages 278–284. 2023.

[20] Nakul Garg and Nirupam Roy. Sirius: A self-localization system for resource-constrained iot sensors. In *Proceedings of the 21st annual international conference on mobile systems, applications and services*, pages 289–302, 2023.

[21] Nakul Garg, Aritrik Ghosh, and Nirupam Roy. Litefoot: Ultra-low-power localization using ambient cellular signals. In *Proceedings of the 22nd ACM Conference on Embedded Networked Sensor Systems*, pages 535–548, 2024.

[22] Nakul Garg, Irtaza Shahid, Ramanujan K Sheshadri, Karthikeyan Sundaresan, and Nirupam Roy. Large network uwb localization: Algorithms and implementation. In *22nd USENIX Symposium on Networked Systems Design and Implementation (NSDI 25)*, 2025.

[23] Nakul Garg, Nirupam Roy, Ranveer Chandra, and Vaishnavi Ranganathan. Freshsense: Transforming supply chains with low-cost and dynamic food quality sensing at the pallet-level. *GetMobile: Mobile Computing and Communications*, 28(3):9–13, 2024.

[24] Nakul Garg and Nirupam Roy. Enabling self-defense in small drones. In *Proceedings of the 21st International Workshop on Mobile Computing Systems and Applications*, pages 15–20, 2020.

[25] Nakul Garg, Irtaza Shahid, Erin Avllazagaj, Jennie Hill, Jun Han, and Nirupam Roy. Thermware: Toward side-channel defense for tiny iot devices. In *Proceedings of the 24th International Workshop on Mobile Computing Systems and Applications*, pages 81–88, 2023.

[26] Irtaza Shahid, Yang Bai, Nakul Garg, and Nirupam Roy. Voicefind: Noise-resilient speech recovery in commodity headphones. In *Proceedings of the 1st ACM International Workshop on Intelligent Acoustic Systems and Applications*, pages 13–18, 2022.

[27] Toma Telembici and Lacrimioara Grama. Detecting indoor sound events. *Acta Technica Napocensis*, 59(2):13–17, 2018.

[28] Bing Zhou, Mohammed Elbadry, Ruipeng Gao, and Fan Ye. Batmapper: Acoustic sensing based indoor floor plan construction using smartphones. In *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*, pages 42–55, 2017.

[29] Zhaotai Pan, Yi Ge, Yu Chen Zhou, Jing Chang Huang, Yu Ling Zheng, Ning Zhang, Xiao Xing Liang, Peng Gao, Guan Qun Zhang, Qingyan Wang, et al. Cognitive acoustic analytics service for internet of things. In *2017 IEEE International Conference on Cognitive Computing (ICCC)*, pages 96–103. IEEE, 2017.

[30] Anran Wang, Jacob E Sunshine, and Shyamnath Gollakota. Contactless infant monitoring using white noise. In *The 25th Annual International Conference on Mobile Computing and Networking*, pages 1–16, 2019.

[31] Justin Chan, Thomas Rea, Shyamnath Gollakota, and Jacob E Sunshine. Contactless cardiac arrest detection using smart devices. *NPJ digital medicine*, 2(1):1–8, 2019.

[32] Chiara Semenzin, Lisa Hamrick, Amanda Seidl, Bridgette Kelleher, and Alejandrina Cristia. Towards large-scale data annotation of audio from wearables: validating zooniverse annotations of infant vocalization types. In *2021 IEEE Spoken Language Technology Workshop (SLT)*, pages 1079–1085. IEEE, 2021.

[33] Hillary Ganek and Alice Eriks-Brophy. Language environment analysis (lena) system investigation of day long recordings in children: A literature review. *Journal of Communication Disorders*, 72:77–85, 2018.

[34] Marcin Bernas, Bartłomiej Płaczek, Wojciech Korski, Piotr Loska, Jarosław Smyła, and Piotr Szymała. A survey and comparison of low-cost sensing technologies for road traffic monitoring. *Sensors*, 18(10):3243, 2018.

[35] Anuntapat Anuntachai and Kittituch Pavaranchanakul. Analyze traffic conditions and events with sound processing. In *2020 20th International Conference on Control, Automation and Systems (ICCAS)*, pages 344–349. IEEE, 2020.

[36] Vincent Liu, Aaron Parks, Vamsi Talla, Shyamnath Gollakota, David Wetherall, and Joshua R Smith. Ambient backscatter: Wireless communication out of thin air. *ACM SIGCOMM Computer Communication Review*, 43(4):39–50, 2013.

[37] Mohammad Rostami, Karthik Sundaresan, Eugene Chai, Sampath Rangarajan, and Deepak Ganesan. Redefining passive in backscattering with commodity devices. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, pages 1–13, 2020.

[38] Nakul Garg and Nirupam Roy. Enabling self-defense in small drones. In *Proceedings of the 21st International Workshop on Mobile Computing Systems and Applications. ACM*, 2020.

[39] Nakul Garg and Nirupam Roy. Acoustic sensing for detecting projectile attacks on small drones. In *Proceedings of the 21st International Workshop on Mobile Computing Systems and Applications*, pages 104–104, 2020.

[40] João Andrade, Pedro Santana, and Alexandre Almeida. Motion-induced acoustic noise awareness for socially-aware robot navigation. In *2018 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, pages 24–29. IEEE, 2018.

[41] Jules S Jaffe, Peter JS Franks, Paul LD Roberts, Diba Mirza, Curt Schurgers, Ryan Kastner, and Adrien Boch. A swarm of autonomous miniature underwater robot drifters for exploring submesoscale ocean dynamics. *Nature communications*, 8(1):1–8, 2017.

[42] Federico Mason, Federico Chiariotri, Filippo Campagnaro, Andrea Zanella, and Michele Zorzi. Low-cost auv swarm localization through multimodal underwater acoustic networks. In *Global Oceans 2020: Singapore–US Gulf Coast*, pages 1–7. IEEE, 2020.

[43] Andrew Brughera, Jason Mikiel-Hunter, Mathias Dietz, and David McAlpine. Brainstem biophysics contribute to sound-source localisation in reverberant scenes. *bioRxiv*, page 694356, 2019.

[44] British Trust for Ornithology. Owl hearing. `https://www.bto.org/our-science/projects/project-owl/learn-about-owls/owl-hearing`. Last accessed 19 August 2020.

[45] Heiner Römer. Directional hearing in insects: biophysical, physiological and ecological challenges. *Journal of Experimental Biology*, 223(14), 2020.

[46] Ashutosh Saxena and Andrew Y Ng. Learning sound location from a single microphone. In *2009 IEEE International Conference on Robotics and Automation*, pages 1737–1742. IEEE, 2009.

[47] Dalia El Badawy, Ivan Dokmanić, and Martin Vetterli. Acoustic doa estimation by one unsophisticated sensor. In *International Conference on Latent Variable Analysis and Signal Separation*, pages 89–98. Springer, 2017.

[48] Jill Gilkerson and Jeffrey A Richards. The lena natural language study. *Boulder, CO: LENA Foundation. Retrieved March*, 3:2009, 2008.

[49] Adriel John Orena, Krista Byers-Heinlein, and Linda Polka. Reliability of the language environment analysis recording system in analyzing french–english bilingual speech. *Journal of Speech, Language, and Hearing Research*, 62(7):2491–2500, 2019.

[50] Sadaoki Furui. Speech recognition technology in the ubiquitous/wearable computing environment. In *Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on*, volume 6, pages 3735–3738. IEEE, 2000.

[51] Tiantian Feng, Amrutha Nadarajan, Colin Vaz, Brandon Booth, and Shrikanth Narayanan. Tiles audio recorder: an unobtrusive wearable solution to track audio activity. In *Proceedings of the 4th ACM Workshop on Wearable Systems and Applications*, pages 33–38, 2018.

[52] KN McGuire, Christophe De Wagter, Karl Tuyls, HJ Kappen, and Guido CHE de Croon. Minimal navigation solution for a swarm of tiny flying robots to explore an unknown environment. *Science Robotics*, 4(35), 2019.

[53] Mario Coppola, Kimberly N McGuire, Christophe De Wagter, and Guido CHE de Croon. A survey on swarming with micro air vehicles: Fundamental challenges and constraints. *Frontiers in Robotics and AI*, 7:18, 2020.

[54] Allan D Pierce. Diffraction of sound around corners and over wide barriers. *The Journal of the Acoustical Society of America*, 55(5):941–955, 1974.

[55] KS Peat. A first aproximation to the effects of mean flow on sound propagation through cylindrical capillary tubes. *Journal of sound and vibration*, 175(4):475–489, 1994.

[56] PK Tang and WA Sirignano. Theory of a generalized helmholtz resonator. *Journal of Sound and Vibration*, 26(2):247–262, 1973.

[57] Shadab Ahmad and Tazeem Ahmad. Implementation of recursive least squares (rls) adaptive filter for noise cancellation. *International Journal of Scientific Engineering and Technology*, 1(4):46–48, 2012.

[58] Karol J Piczak. Environmental sound classification with convolutional neural networks. In *2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP)*, pages 1–6. IEEE, 2015.

[59] Francis M Wiener. Sound diffraction by rigid spheres and circular cylinders. *The Journal of the Acoustical Society of America*, 19(3):444–451, 1947.

[60] Robert Bridson. Fast poisson disk sampling in arbitrary dimensions. *SIGGRAPH sketches*, 10:1, 2007.

[61] Autodesk. https://www.autodesk.com/products/fusion-360/overview, 2020.

[62] Autodesk Python API. https://autodeskfusion360.github.io/, 2020.

[63] Elegoo. https://www.elegoo.com/product/elegoo-mars-uv-photocuring-lcd-3d-printer/, 2020.

[64] Tomasz G Zieliński, Kamil C Opiela, Piotr Pawłowski, Nicolas Dauchez, Thomas Boutin, John Kennedy, Daniel Trimble, Henry Rice, Bart Van Damme, Gwenael Hannema, et al. Reproducibility of sound-absorbing periodic porous materials using additive manufacturing technologies: Round robin study. *Additive Manufacturing*, 36:101564, 2020.

[65] Keysight. https://www.keysight.com/us/en/products/waveform-and-function-generators.html, 2020.

[66] Arduino. https://www.arduino.cc/, 2020.

[67] Analog Devices. Admp401: Omnidirectional microphone with bottom port and analog output, 2013.

[68] Keysight. Keysight e6313a power supply. https://www.keysight.com/us/en/assets/9018-04576/user-manuals/9018-04576.pdf, 2021.

[69] Mathworks. Mathworks matlab coder. https://www.mathworks.com/products/matlab-coder.html, 2021.

[70] Kazuhiro Nakadai, Hirofumi Nakajima, Kentaro Yamada, Yuji Hasegawa, Takahiro Nakamura, and Hiroshi Tsujino. Sound source tracking with directivity pattern estimation using a 64 ch microphone array. *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1690–1696, 2005.

[71] Hirofumi Nakajima, Keiko Kikuchi, Toru Daigo, Yutaka Kaneda, Kazuhiro Nakadai, and Yuji Hasegawa. Real-time sound source orientation estimation using a 96 channel microphone array. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 676–683. IEEE, 2009.

[72] Kenta Niwa, Yusuke Hioka, Sumitaka Sakauchi, Ken'ichi Furuya, and Yoichi Haneda. Estimation of sound source orientation using eigenspace of spatial correlation matrix. In *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 129–132. IEEE, 2010.

[73] Yang Bai, Li Lu, Jerry Cheng, Jian Liu, Yingying Chen, and Jiadi Yu. Acoustic-based sensing and applications: A survey. *Computer Networks*, 181:107447, 2020.

[74] António LL Ramos, Sverre Holm, Sigmund Gudvangen, and Ragnvald Otterlei. Delay-and-sum beamforming for direction of arrival estimation applied to gunshot acoustics. In *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense X*, volume 8019, page 80190U. International Society for Optics and Photonics, 2011.

[75] V Krishnaveni, T Kesavamurthy, and B Aparna. Beamforming for direction-of-arrival (doa) estimation-a survey. *International Journal of Computer Applications*, 61(11), 2013.

[76] Gary W Elko. Microphone array systems for hands-free telecommunication. *Speech communication*, 20(3-4):229–240, 1996.

[77] Ralph Schmidt. Multiple emitter location and signal parameter estimation. *IEEE transactions on antennas and propagation*, 34(3):276–280, 1986.

[78] Despoina Pavlidi, Anthony Griffin, Matthieu Puigt, and Athanasios Mouchtaris. Real-time multiple sound source localization and counting using a circular microphone array. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(10):2193–2206, 2013.

[79] Romit Roy Choudhury. Earable computing: A new area to think about. In *Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications*, pages 147–153, 2021.

[80] Sheng Shen, Daguan Chen, Yu-Lin Wei, Zhijian Yang, and Romit Roy Choudhury. Voice localization using nearby wall reflections. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, pages 1–14, 2020.

[81] Rajalakshmi Nandakumar, Vikram Iyer, Desney Tan, and Shyamnath Gollakota. Fingerio: Using active sonar for fine-grained finger tracking. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 1515–1525, 2016.

[82] Wenguang Mao, Mei Wang, and Lili Qiu. Aim: Acoustic imaging on a mobile. In *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, pages 468–481, 2018.

[83] Wenguang Mao and Lili Qiu. Dronetrack: An indoor follow-me system using acoustic signals. *GetMobile: Mobile Computing and Communications*, 21(4):22–24, 2018.

[84] Nirupam Roy, Sheng Shen, Haitham Hassanieh, and Romit Roy Choudhury. Inaudible voice commands: The long-range attack and defense. In *15th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, pages 547–560. USENIX Association, 2018.

[85] Nirupam Roy, Haitham Hassanieh, and Romit Roy Choudhury. Backdoor: Sounds that a microphone can record, but that humans can't hear. *GetMobile: Mobile Computing and Communications*, 21(4):25–29, 2018.

[86] Sheng Shen, Nirupam Roy, Junfeng Guan, Haitham Hassanieh, and Romit Roy Choudhury. Mute: bringing iot to noise cancellation. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*, pages 282–296. ACM, 2018.

[87] Nirupam Roy. *Inaudible acoustics: techniques and applications*. PhD thesis, University of Illinois at Urbana-Champaign, 2018.

[88] Cecilia Casarini, Benjamin Tiller, Carmelo Mineo, Charles N MacLeod, James FC Windmill, and Joseph C Jackson. Enhancing the sound absorption of small-scale 3-d printed acoustic metamaterials based on helmholtz resonators. *IEEE Sensors Journal*, 18(19):7949–7955, 2018.

[89] Gianluca Memoli, Mihai Caleap, Michihiro Asakawa, Deepak R Sahoo, Bruce W Drinkwater, and Sriram Subramanian. Metamaterial bricks and quantization of meta-surfaces. *Nature communications*, 8(1):1–8, 2017.

[90] Dingzeyu Li, David IW Levin, Wojciech Matusik, and Changxi Zheng. Acoustic voxels: computational optimization of modular acoustic filters. *ACM Transactions on Graphics (TOG)*, 35(4):1–12, 2016.

[91] Chris Harrison, Robert Xiao, and Scott Hudson. Acoustic barcodes: passive, durable and inexpensive notched identification tags. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*, pages 563–568, 2012.

[92] Gierad Laput, Eric Brockmeyer, Scott E Hudson, and Chris Harrison. Acoustruments: Passive, acoustically-driven, interactive controls for handheld devices. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2161–2170, 2015.

[93] Valkyrie Savage, Andrew Head, Björn Hartmann, Dan B Goldman, Gautham Mysore, and Wilmot Li. Lamello: Passive acoustic sensing for tangible input components. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1277–1280, 2015.

[94] Ivan Dokmanić and Martin Vetterli. Room helps: Acoustic localization with finite elements. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2617–2620. Ieee, 2012.

[95] Dalia El Badawy and Ivan Dokmanić. Direction of arrival with one microphone, a few legos, and non-negative matrix factorization. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(12):2436–2446, 2018.

[96] Eleftheria Georganti, Tobias May, Steven van de Par, Aki Harma, and John Mourjopoulos. Speaker distance detection using a single microphone. *IEEE transactions on audio, speech, and language processing*, 19(7):1949–1961, 2011.

[97] Nakul Garg, Yang Bai, and Nirupam Roy. Demo: Microstructure-guided spatial sensing for low-power iot. In *The 19th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '21)*, 2021.

[98] Noah T Jafferis, E Farrell Helbling, Michael Karpelson, and Robert J Wood. Untethered flight of an insect-sized flapping-wing microscale aerial vehicle. *Nature*, 570(7762):491–495, 2019.

[99] Julien Dupeyroux, Julien R Serres, and Stéphane Viollet. Antbot: A six-legged walking robot able to home like desert ants in outdoor environments. *Science Robotics*, 4(27), 2019.

[100] Ying Tan and Zhong-yang Zheng. Research advance in swarm robotics. *Defence Technology*, 9(1):18–39, 2013.

[101] David L Christensen, Elliot W Hawkes, Srinivasan A Suresh, Karen Ladenheim, and Mark R Cutkosky. μtugs: Enabling microrobots to deliver macro forces with controllable adhesives. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4048–4055. IEEE, 2015.

[102] Nikita Korobov, Oleg Shipitko, Ivan Konovalenko, Anton Grigoryev, and Marina Chukalina. Swap-c based comparison of onboard computers for unmanned vehicles. In *Proceedings of 14th International Conference on Electromechanics and Robotics "Zavalishin's Readings"*, pages 573–583. Springer, 2020.

[103] Benjamin Langmann, Klaus Hartmann, and Otmar Loffeld. Depth camera technology comparison and performance evaluation. In *ICPRAM (2)*, pages 438–444, 2012.

[104] Jaebum Choi, Simon Ulbrich, Bernd Lichte, and Markus Maurer. Multi-target tracking using a 3d-lidar sensor for autonomous vehicles. In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, pages 881–886. IEEE, 2013.

[105] David K Barton. Modern radar system analysis. *Norwood*, 1988.

[106] Winthrop N Kellogg. Sonar system of the blind. *Science*, 137(3528):399–404, 1962.

[107] Benjamin Goldberg, Raphael Zufferey, Neel Doshi, Elizabeth Farrell Helbling, Griffin Whittredge, Mirko Kovac, and Robert J Wood. Power and control autonomy for high-speed locomotion with an insect-scale legged robot. *IEEE Robotics and Automation Letters*, 3(2):987–993, 2018.

[108] Gilles Caprari, Thomas Estier, and Roland Siegwart. Fascination of down scaling-alice the sugar cube robot. In *IEEE International Conference on Robotics and Automation (ICRA 2000): Workshop on Mobile Micro-Robots*, number CONF, 2000.

[109] E Farrell Helbling, Sawyer B Fuller, and Robert J Wood. Altitude estimation and control of an insect-scale robot with an onboard proximity sensor. In *Robotics Research*, pages 57–69. Springer, 2018.

[110] Vikram Iyer, Ali Najafi, Johannes James, Sawyer Fuller, and Shyamnath Gollakota. Wireless steerable vision for live insects and insect-scale robots. *Science robotics*, 5(44), 2020.

[111] Marco F Duarte, Mark A Davenport, Dharmpal Takhar, Jason N Laska, Ting Sun, Kevin F Kelly, and Richard G Baraniuk. Single-pixel imaging via compressive sampling. *IEEE signal processing magazine*, 25(2):83–91, 2008.

[112] Pieter Kruizinga, Pim van der Meulen, Andrejs Fedjajevs, Frits Mastik, Geert Springeling, Nico de Jong, Johannes G Bosch, and Geert Leus. Compressive 3d ultrasound imaging using a single sensor. *Science advances*, 3(12):e1701423, 2017.

[113] John C Curlander and Robert N McDonough. *Synthetic aperture radar*, volume 11. Wiley, New York, 1991.

[114] Patrick Kapita Mvemba, Simon Kidiamboko Guwa Gua Band, Aime Lay-Ekuakille, and Nicola Ivan Giannoccaro. Advanced acoustic sensing system on a mobile robot: design, construction and measurements. *IEEE Instrumentation & Measurement Magazine*, 21(2):4–9, 2018.

[115] Ingmar Jager, Richard Heusdens, and Nikolay D Gaubitch. Room geometry estimation from acoustic echoes using graph-based echo labeling. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE, 2016.

[116] Ivan Dokmanić, Reza Parhizkar, Andreas Walther, Yue M Lu, and Martin Vetterli. Acoustic echoes reveal room shape. *Proceedings of the National Academy of Sciences*, 110(30):12186–12191, 2013.

[117] Xuyu Wang, Lingjun Gao, Shiwen Mao, and Santosh Pandey. Csi-based fingerprinting for indoor localization: A deep learning approach. *IEEE Transactions on Vehicular Technology*, 66(1):763–776, 2016.

[118] Raspberry Pi Foundation. Raspberry pi 4. `https://www.raspberrypi.org/products/raspberry-pi-4-model-b/`, 2021.

[119] Apollo. Apollo3 blue datasheet. `https://cdn.sparkfun.com/assets/c/1/b/7/6/Apollo3_Blue_MCU_Data_Sheet_v0_10_0.pdf`, 2021.

[120] Adafruit. Adafruit edgebadge datasheet. `https://media.digikey.com/pdf/Data%20Sheets/Adafruit%20PDFs/4400_Web.pdf`, 2021.

[121] ARM Developer. Optimizing c code with neon intrinsics. `https://developer.arm.com/documentation/102467/0100/Matrix-multiplication-example`, 2021.

[122] HiLetgo Inc. Hiletgo tct40-16r/t 16mm split ultrasonic sensor ultrasonic transceiver. `https://www.amazon.com/HiLetgo-Ultrasonic-Sensor-Transceiver-Receiver/`, 2021.

[123] Roberto Di Leonardo, Francesca Ianni, and Giancarlo Ruocco. Computer generation of optimal holograms for optical trap arrays. *Optics Express*, 15(4):1913–1922, 2007.

[124] Jean-Jacques Fuchs. Recovery of exact sparse representations in the presence of bounded noise. *IEEE Transactions on Information Theory*, 51(10):3601–3608, 2005.

[125] Rafael C Gonzalez, Richard E Woods, et al. Digital image processing, 2002.

[126] I Pits and AN Venetsanopoulos. Non linear digital filters, 1990.

[127] Wen-Yan Lin, Siying Liu, Yasuyuki Matsushita, Tian-Tsong Ng, and Loong-Fah Cheong. Smoothly varying affine stitching. In *CVPR 2011*, pages 345–352. IEEE, 2011.

[128] Chunwei Tian, Lunke Fei, Wenxian Zheng, Yong Xu, Wangmeng Zuo, and Chia-Wen Lin. Deep learning on image denoising: An overview. *Neural Networks*, 2020.

[129] Intel. Intel realsense lidar camera l515. `https://www.intelrealsense.com/lidar-camera-l515/`, 2021.

[130] Parallax Inc. Ping ultrasonic distance sensor. `https://www.parallax.com/product/ping-ultrasonic-distance-sensor/`, 2021.

[131] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.

[132] Ivan Dokmanić and Ivan Tashev. Hardware and algorithms for ultrasonic depth imaging. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6702–6706. Ieee, 2014.

[133] Péter Fankhauser, Michael Bloesch, Diego Rodriguez, Ralf Kaestner, Marco Hutter, and Roland Siegwart. Kinect v2 for mobile robot navigation: Evaluation and modeling. In *2015 International Conference on Advanced Robotics (ICAR)*, pages 388–394. IEEE, 2015.

[134] Intel. Intel realsense lidar camera l515. `https://www.intelrealsense.com/lidar-camera-l515/`, 2021.

[135] Mathias Jensen. Coin cells and peak current draw. *White Paper. Texas Instruments*, 2010.

[136] Mems microphone breakout. `http://www.analog.com/media/en/technical-documentation/obsolete-data-sheets/ADMP401.pdf`. Last accessed 24 September 2017.

[137] Texas Instruments. Dual rail-to-rail low-voltage low power operational amplifier. `https://www.ti.com/product/TLV2252`, 2021.

[138] TI. Texas instruments energytrace++. `https://www.ti.com/tool/ENERGYTRACE`, 2022.

[139] David L Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006.

[140] Yoni Sher, Lior Cohen, Daniel Istrati, and Hagai S Eisenberg. Low intensity lidar using compressed sensing and a photon number resolving detector. In *Emerging Digital Micromirror Device Based Systems and Applications X*, volume 10546, page 105460J. International Society for Optics and Photonics, 2018.

[141] Lixiang Li, Guoqian Wen, Zeming Wang, and Yixian Yang. Efficient and secure image communication system based on compressed sensing for iot monitoring applications. *IEEE Transactions on Multimedia*, 22(1):82–95, 2019.

[142] Tomás Mantecón, Ana Mantecón, Carlos R del Blanco, Fernando Jaureguizar, and Narciso García. Hand-gesture-based human-machine interface system using compressive sensing. In *2015 International Symposium on Consumer Electronics (ISCE)*, pages 1–2. IEEE, 2015.

[143] Ehsan Ardestanizadeh, Mahdi Cheraghchi, and Amin Shokrollahi. Bit precision analysis for compressed sensing. In *2009 IEEE International Symposium on Information Theory*, pages 1–5. IEEE, 2009.

[144] Nirupam Roy, Sheng Shen, Haitham Hassanieh, and Romit Roy Choudhury. Inaudible voice commands: The long-range attack and defense. In *15th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 18)*, pages 547–560, 2018.

[145] Nirupam Roy and Romit Roy Choudhury. Ripple ii: faster communication through physical vibration. In *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, pages 671–684, 2016.

[146] Nirupam Roy, Mahanth Gowda, and Romit Roy Choudhury. Ripple: Communicating through physical vibration. In *12th USENIX Symposium on Networked Systems Design and Implementation (NSDI 15)*, pages 265–278, 2015.

[147] Nirupam Roy and Romit Roy Choudhury. Listening through a vibration motor. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*, pages 57–69. ACM, 2016.

[148] Jie Yang, Simon Sidhom, Gayathri Chandrasekaran, Tam Vu, Hongbo Liu, Nicolae Cecan, Yingying Chen, Marco Gruteser, and Richard P Martin. Sensing driver phone use with acoustic ranging through car speakers. *IEEE Transactions on Mobile Computing*, 11(9):1426–1440, 2012.

[149] Zhengxia Zou, Zhenwei Shi, Yuhong Guo, and Jieping Ye. Object detection in 20 years: A survey. *arXiv preprint arXiv:1905.05055*, 2019.

[150] Xinzhu Ma, Zhihui Wang, Haojie Li, Pengbo Zhang, Wanli Ouyang, and Xin Fan. Accurate monocular 3d object detection via color-embedded 3d reconstruction for autonomous driving. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 6851–6860, 2019.

[151] Martin Weinmann, Jens Leitloff, Ludwig Hoegner, Boris Jutzi, Uwe Stilla, and Stefan Hinz. Thermal 3d mapping for object detection in dynamic scenes. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2(1), 2014.

[152] Feryel Zoghlami, Okan Kamil Sen, Harald Heinrich, Germar Schneider, Emec Ercelik, Alois Knoll, and Thomas Villmann. Tof/radar early feature-based fusion system for human detection and tracking. In *2021 22nd IEEE International Conference on Industrial Technology (ICIT)*, volume 1, pages 942–949. IEEE, 2021.

[153] Serdal Ayhan, Mario Pauli, Thorsten Kayser, Steffen Scherr, and Thomas Zwick. Fmcw radar system with additional phase evaluation for high accuracy range detection. In *2011 8th European Radar Conference*, pages 117–120. IEEE, 2011.

[154] Tara Boroushaki, Junshan Leng, Ian Clester, Alberto Rodriguez, and Fadel Adib. Robotic grasping of fully-occluded objects using rf perception. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 923–929. IEEE, 2021.

[155] Fadel Adib, Zach Kabelac, Dina Katabi, and Robert C Miller. 3d tracking via body radio reflections. In *11th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 14)*, pages 317–329, 2014.

[156] Jorge Beltrán, Carlos Guindel, Francisco Miguel Moreno, Daniel Cruzado, Fernando Garcia, and Arturo De La Escalera. Birdnet: a 3d object detection framework from lidar information. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 3517–3523. IEEE, 2018.

[157] Yan Wang, Wei-Lun Chao, Divyansh Garg, Bharath Hariharan, Mark Campbell, and Kilian Q Weinberger. Pseudo-lidar from visual depth estimation: Bridging the gap in 3d object detection for autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8445–8453, 2019.

[158] Sriram Sami, Yimin Dai, Sean Rui Xiang Tan, Nirupam Roy, and Jun Han. Spying with your robot vacuum cleaner: eavesdropping via lidar sensors. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, pages 354–367, 2020.

[159] Dariush Forouher, Marvin Große Besselmann, and Erik Maehle. Sensor fusion of depth camera and ultrasound data for obstacle detection and robot navigation. In *2016 14th international conference on control, automation, robotics and vision (ICARCV)*, pages 1–6. IEEE, 2016.

[160] Donghe Yang, Jinsong Xia, and Liu Xi'ang. The research on intelligent mobile robot's avoiding obstacle by ultrasound. In *2010 International Conference on Artificial Intelligence and Computational Intelligence*, volume 1, pages 77–79. IEEE, 2010.

[161] Dong-Won Jung, Zhong-Soo Lim, Byung-Geuk Kim, and Nak-Ku Kim. Multi-channel ultrasonic sensor system for obstacle detection of the mobile robot. In *2007 International Conference on Control, Automation and Systems*, pages 2347–2351. IEEE, 2007.

[162] Jongil Lim, SeokJu Lee, Girma Tewolde, and Jaerock Kwon. Indoor localization and navigation for a mobile robot equipped with rotating ultrasonic sensors using a smartphone as the robot's brain. In *2015 IEEE International Conference on Electro/Information Technology (EIT)*, pages 621–625. IEEE, 2015.

[163] Anna C Bakenecker, Anselm von Gladiss, Hannes Schwenke, André Behrends, Thomas Friedrich, Kerstin Lüdtke-Buzug, Alexander Neumann, Joerg Barkhausen, Franz Wegner, and Thorsten M Buzug. Navigation of a magnetic micro-robot through a cerebral aneurysm phantom with magnetic particle imaging. *Scientific reports*, 11(1):1–12, 2021.

[164] Fang Wu, Akash Vibhute, Gim Song Soh, Kristin L Wood, and Shaohui Foong. A compact magnetic field-based obstacle detection and avoidance system for miniature spherical robots. *Sensors*, 17(6):1231, 2017.

[165] Eleni Kelasidi, Signe Moe, Kristin Pettersen, Anna M Kohl, Pål Liljebäck, Jan Tommy Gravdahl, et al. Path following, obstacle detection and obstacle avoidance for thrusted underwater snake robots. *Frontiers in Robotics and AI*, 6:57, 2019.

[166] Gabe Cohn, Sidhant Gupta, Tien-Jui Lee, Dan Morris, Joshua R Smith, Matthew S Reynolds, Desney S Tan, and Shwetak N Patel. An ultra-low-power human body motion sensor using static electric field sensing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 99–102, 2012.

[167] Zhongqin Wang, Fu Xiao, Ning Ye, Ruchuan Wang, and Panlong Yang. A see-through-wall system for device-free human motion sensing based on battery-free rfid. *ACM Transactions on Embedded Computing Systems (TECS)*, 17(1):1–21, 2017.

[168] P Offermans, HD Tong, CJM Van Rijn, P Merken, SH Brongersma, and M Crego-Calama. Ultralow-power hydrogen sensing with single palladium nanowires. *Applied Physics Letters*, 94(22):223110, 2009.

[169] I Elmi, S Zampolli, E Cozzani, M Passini, GC Cardinali, and M Severi. Development of ultra low power consumption hotplates for gas sensing applications. In *SENSORS, 2006 IEEE*, pages 243–246. IEEE, 2006.

[170] Mandy Lai Yi Sin, Gary Chun Tak Chow, Gary Man Keung Wong, Wen Jung Li, Philip Heng Wai Leong, and Ka Wai Wong. Ultralow-power alcohol vapor sensors using chemically functionalized multiwalled carbon nanotubes. *IEEE Transactions on Nanotechnology*, 6(5):571–577, 2007.

[171] Josue Pagan, Ramin Fallahzadeh, Mahdi Pedram, Jose L Risco-Martin, Jose M Moya, Jose L Ayala, and Hassan Ghasemzadeh. Toward ultra-low-power remote health monitoring: An optimal and adaptive compressed sensing framework for activity recognition. *IEEE Transactions on Mobile Computing*, 18(3):658–673, 2018.

[172] Yoonmyung Lee, David Blaauw, and Dennis Sylvester. Ultralow power circuit design for wireless sensor nodes for structural health monitoring. *Proceedings of the IEEE*, 104(8):1529–1546, 2016.

[173] Nhat Pham, Tuan Dinh, Zohreh Raghebi, Taeho Kim, Nam Bui, Phuc Nguyen, Hoang Truong, Farnoush Banaei-Kashani, Ann Halbower, Thang Dinh, et al. Wake: a behind-the-ear wearable system for microsleep detection. In *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services*, pages 404–418, 2020.

[174] Nam Bui, Nhat Pham, Hoang Truong, Phuc Nguyen, Jianliang Xiao, Robin Deterding, Thang Dinh, and Tam Vu. ebp: Frequent and comfortable blood pressure monitoring from inside human's ears. *GetMobile: Mobile Computing and Communications*, 23(4):34–38, 2020.

[175] Rajalakshmi Nandakumar, Vikram Iyer, and Shyamnath Gollakota. 3d localization for sub-centimeter sized devices. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*, pages 108–119, 2018.

[176] Vikram Iyer, Rajalakshmi Nandakumar, Anran Wang, Sawyer B Fuller, and Shyamnath Gollakota. Living iot: A flying wireless platform on live insects. In *The 25th Annual International Conference on Mobile Computing and Networking*, pages 1–15, 2019.

[177] Andrea Ricci, Matteo Grisanti, Ilaria De Munari, and Paolo Ciampolini. Improved pervasive sensing with rfid: An ultra-low power baseband processor for uhf tags. *IEEE transactions on very large scale integration (VLSI) systems*, 17(12):1719–1729, 2009.

[178] Nicolas Gay and Wolf-Joachim Fischer. Ultra-low-power rfid-based sensor mote. In *SENSORS, 2010 IEEE*, pages 1293–1298. IEEE, 2010.

[179] Young-Joon Kim, Hansraj S Bhamra, Jithin Joseph, and Pedro P Irazoqui. An ultra-low-power rf energy-harvesting transceiver for multiple-node sensor application. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 62(11):1028–1032, 2015.

[180] Shilong Zhai, Huaijun Chen, Changlin Ding, Fangliang Shen, Chunrong Luo, and Xiaopeng Zhao. Manipulation of transmitted wave front using ultrathin planar acoustic metasurfaces. *Applied Physics A*, 120(4):1283–1289, 2015.

[181] Yangbo Xie, Tsung-Han Tsai, Adam Konneker, Bogdan-Ioan Popa, David J Brady, and Steven A Cummer. Single-sensor multispeaker listening with acoustic metamaterials. *Proceedings of the National Academy of Sciences*, 112(34):10595–10598, 2015.

[182] Nakul Garg, Yang Bai, and Nirupam Roy. Owlet: Enabling spatial information in ubiquitous acoustic devices. In *The 19th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '21), June 24–July 2, 2021, Virtual, WI, USA*. ACM.

[183] Sheng Shen, Daguan Chen, Yu-Lin Wei, Zhijian Yang, and Romit Roy Choudhury. Voice localization using nearby wall reflections.

[184] E Caroli, JB Stephen, G Di Cocco, L Natalucci, and A Spizzichino. Coded aperture imaging in x-and gamma-ray astronomy. *Space Science Reviews*, 45(3):349–403, 1987.

[185] Roummel F Marcia, Zachary T Harmany, and Rebecca M Willett. Compressive coded aperture imaging. In *Computational Imaging VII*, volume 7246, page 72460G. International Society for Optics and Photonics, 2009.

[186] Xiaopeng Peng, Garreth J Ruane, and Grover A Swartzlander Jr. Randomized aperture imaging. *arXiv preprint arXiv:1601.00033*, 2016.

[187] MD Brown, DI Nikitichev, BE Treeby, and BT Cox. Generating arbitrary ultrasound fields with tailored optoacoustic surface profiles. *Applied Physics Letters*, 110(9):094102, 2017.

[188] Kai Melde, Andrew G Mark, Tian Qiu, and Peer Fischer. Holograms for acoustics. *Nature*, 537(7621):518–522, 2016.

[189] Yangbo Xie, Chen Shen, Wenqi Wang, Junfei Li, Dingjie Suo, Bogdan-Ioan Popa, Yun Jing, and Steven A Cummer. Acoustic holographic rendering with two-dimensional metamaterial-based passive phased array. *Scientific reports*, 6(1):1–6, 2016.

[190] Mitsuharu Matsumoto and Shuji Hashimoto. A miniaturized adaptive microphone array under directional constraint utilizing aggregated microphones. *The Journal of the Acoustical Society of America*, 119(1):352–359, 2006.

[191] Mitsuharu Matsumoto and Shuji Hashimoto. Multiple signal classification by aggregated microphones. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 88(7):1701–1707, 2005.

[192] Nan Zou and Arye Nehorai. Circular acoustic vector-sensor array for mode beamforming. *IEEE transactions on signal processing*, 57(8):3041–3052, 2009.

[193] Sina Sadeghpour, Michael Kraft, and Robert Puers. Highly efficient piezoelectric micromachined ultrasound transducer (pmut) for underwater sensor networks. In *2019 20th International Conference on Solid-State Sensors, Actuators and Microsystems & Eurosensors XXXIII (TRANSDUCERS & EUROSENSORS XXXIII)*, pages 162–165. IEEE, 2019.

[194] Flavius V Pop, Bernard Herrera, Cristian Cassella, Guofeng Chen, Emrecan Demirors, Raffaele Guida, Tommaso Melodia, and Matteo Rinaldi. Novel pmut-based acoustic duplexer for underwater and intrabody communication. In *2018 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4. IEEE, 2018.

[195] P-C Eccardt, K Niederer, T Scheiter, and C Hierold. Surface micromachined ultrasound transducers in cmos technology. In *1996 IEEE Ultrasonics Symposium. Proceedings*, volume 2, pages 959–962. IEEE, 1996.

[196] David E Dausch, John B Castellucci, Derrick R Chou, and Olaf T Von Ramm. Theory and operation of 2-d array piezoelectric micromachined ultrasound transducers. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 55(11):2484–2492, 2008.

[197] Nirupam Roy. Owlet: Insect-scale spatial sensing with 3d-printed acoustic structures. *GetMobile: Mobile Computing and Communications*, 25(2):14–20, 2021.

[198] Irtaza Shahid, Yang Bai, Nakul Garg, and Nirupam Roy. In *Proceedings of the 1st ACM International Workshop on Intelligent Acoustic Systems and Applications*, 2022.

[199] Per Hyberg, Magnus Jansson, and Björn Ottersten. Array interpolation and bias reduction. *IEEE Transactions on Signal Processing*, 52(10):2711–2720, 2004.

[200] Hiroki Katahira, Nobutaka Ono, Shigeki Miyabe, Takeshi Yamada, and Shoji Makino. Virtually increasing microphone array elements by interpolation in complex-logarithmic domain. In *21st European Signal Processing Conference (EUSIPCO 2013)*, pages 1–5. IEEE, 2013.

[201] Ryoga Jinzai, Kouei Yamaoka, Mitsuo Matsumoto, Takeshi Yamada, and Shoji Makino. Microphone position realignment by extrapolation of virtual microphone. In *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 367–372. IEEE, 2018.

[202] Efren Fernandez-Grande and Angeliki Xenaki. Compressive sensing with a spherical microphone array. *The Journal of the Acoustical Society of America*, 139(2):EL45–EL49, 2016.

[203] Fangli Ning, Jingang Wei, Lianfang Qiu, Hongbing Shi, and Xiaofan Li. Three-dimensional acoustic imaging with planar microphone arrays and compressive sensing. *Journal of Sound and Vibration*, 380:112–128, 2016.

[204] Xun Huang. Compressive sensing and reconstruction in measurements with an aerospace application. *AIAA journal*, 51(4):1011–1016, 2013.

[205] Dovid Levin, Emanuël AP Habets, and Sharon Gannot. Maximum likelihood estimation of direction of arrival using an acoustic vector-sensor. *The Journal of the Acoustical Society of America*, 131(2):1240–1248, 2012.

[206] Arye Nehorai and Eytan Paldi. Acoustic vector-sensor array processing. *IEEE Transactions on signal processing*, 42(9):2481–2491, 1994.

[207] Joseph Tabrikian, Reuven Shavit, and Dayan Rahamim. An efficient vector sensor configuration for source localization. *IEEE Signal Processing Letters*, 11(8):690–693, 2004.

[208] Michael Rubenstein, Christian Ahler, and Radhika Nagpal. Kilobot: A low cost scalable robot system for collective behaviors. In *2012 IEEE international conference on robotics and automation*, pages 3293–3298. IEEE, 2012.

[209] Stanley S Baek, Fernando L Garcia Bermudez, and Ronald S Fearing. Flight control for target seeking by 13 gram ornithopter. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2674–2681. IEEE, 2011.

[210] Christophe De Wagter, Sjoerd Tijmons, Bart DW Remes, and Guido CHE de Croon. Autonomous flight of a 20-gram flapping wing mav with a 4-gram onboard stereo vision system. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4982–4987. IEEE, 2014.

[211] Remo Brühwiler, Benjamin Goldberg, Neel Doshi, Onur Ozcan, Noah Jafferis, Michael Karpelson, and Robert J Wood. Feedback control of a legged microrobot with on-board sensing. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5727–5733. IEEE, 2015.

[212] Andrew T Baisch, Christian Heimlich, Michael Karpelson, and Robert J Wood. Hamr3: An autonomous 1.7 g ambulatory robot. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5073–5079. IEEE, 2011.

[213] Seth Hollar, Anita Flynn, Colby Bellew, and KSJ Pister. Solar powered 10 mg silicon robot. In *The Sixteenth Annual International Conference on Micro Electro Mechanical Systems, 2003. MEMS-03 Kyoto. IEEE*, pages 706–711. IEEE, 2003.

[214] Gabriel T Sibley, Mohammad H Rahimi, and Gaurav S Sukhatme. Robomote: A tiny mobile robot platform for large-scale ad-hoc sensor networks. In *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292)*, volume 2, pages 1143–1148. IEEE, 2002.

[215] P. Birkmeyer and R. S. Fearing. Dash: A resilient high-speed 15g hexapedal robot. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 418–419, 2009.

[216] Bernhard Klaassen and Karl L Paap. Gmd-snake2: a snake-like robot driven by wheels and a method for motion control. In *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*, volume 4, pages 3014–3019. IEEE, 1999.

[217] Koichi Osuka and Hiroshi Kitajima. Development of mobile inspection robot for rescue activities: Moira. In *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)(Cat. No. 03CH37453)*, volume 4, pages 3373–3377. IEEE, 2003.

[218] Kousuke Suzuki, Atsushi Nakano, Gen Endo, and Shigeo Hirose. Development of multi-wheeled snake-like rescue robots with active elastic trunk. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4602–4607. IEEE, 2012.

[219] M Pytasz, G Granosik, and Krzysztof Tchon. Modelling of wheeeler–hyper mobile robot. *In: Advances in Robotics. Control, Perception, and Communication. K. Tchon (Ed.)*, pages 313–322, 2006.

[220] Philip Taylor, Tara Crewe, Stuart Mackenzie, Denis Lepage, Yves Aubry, Zoe Crysler, George Finney, Charles Francis, Christopher Guglielmo, Diana Hamilton, et al. The motus wildlife tracking system: a collaborative research network to enhance the understanding of wildlife movement. *Avian Conservation and Ecology*, 12(1), 2017.

[221] Sandeep Kamath and Joakim Lindh. Measuring bluetooth low energy power consumption. *Texas instruments application note AN092, Dallas*, 2010.

[222] Deepak Vasisht, Swarun Kumar, and Dina Katabi. {Decimeter-Level} localization with a single {WiFi} access point. In *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, pages 165–178, 2016.

[223] Omid Abari, Hariharan Rahul, Dina Katabi, and Mondira Pant. Airshare: Distributed coherent transmission made seamless. In *2015 IEEE Conference on Computer Communications (INFOCOM)*, pages 1742–1750. IEEE, 2015.

[224] Jie Xiong and Kyle Jamieson. Arraytrack: A fine-grained indoor location system. In *10th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 13)*, pages 71–84, 2013.

[225] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. Spotfi: Decimeter level localization using wifi. In *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, pages 269–282, 2015.

[226] Mohammad Rostami, Xingda Chen, Yuda Feng, Karthikeyan Sundaresan, and Deepak Ganesan. Mixiq: re-thinking ultra-low power receiver design for next-generation on-body applications. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*, pages 364–377, 2021.

[227] Ander Galisteo, Ambuj Varshney, and Domenico Giustiniano. Two to tango: Hybrid light and backscatter networks for next billion devices. In *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services*, pages 80–93, 2020.

[228] Junbo Zhang, Elahe Soltanaghai, Artur Balanuta, Reese Grimsley, Swarun Kumar, and Anthony Rowe. {PLatter}: On the feasibility of building-scale power line backscatter. In *19th USENIX Symposium on Networked Systems Design and Implementation (NSDI 22)*, pages 897–911, 2022.

[229] Songfan Li, Hui Zheng, Chong Zhang, Yihang Song, Shen Yang, Minghua Chen, Li Lu, and Mo Li. Passive {DSSS}: Empowering the downlink communication for backscatter systems. In *19th USENIX Symposium on Networked Systems Design and Implementation (NSDI 22)*, pages 913–928, 2022.

[230] Warren L Stutzman and Gary A Thiele. *Antenna theory and design*. John Wiley & Sons, 2012.

[231] Yang Bai, Nakul Garg, and Nirupam Roy. Spidr: Ultra-low-power acoustic spatial sensing for micro-robot navigation. In *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*, pages 99–113, 2022.

[232] Nakul Garg, Harshvardhan Takawale, Yang Bai, Irtaza Shahid, and Nirupam Roy. Structure assisted spectrum sensing for low-power acoustic event detection. In *Proceedings of Cyber-Physical Systems and Internet of Things Week 2023*, pages 278–284, 2023.

[233] Nakul Garg, Irtaza Shahid, Erin Avllazagaj, Jennie Hill, Jun Han, and Nirupam Roy. Thermware: Toward side-channel defense for tiny iot devices. In *Proceedings of the 24th International Workshop on Mobile Computing Systems and Applications*, pages 81–88, 2023.

[234] Bashima Islam and Shahriar Nirjon. Zygarde: Time-sensitive on-device deep inference and adaptation on intermittently-powered systems. *arXiv preprint arXiv:1905.03854*, 2019.

[235] Yang Bai, Nakul Garg, and Nirupam Roy. Ultra-low-power acoustic imaging. In *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*, pages 523–524, 2022.

[236] Yaxiong Xie, Jie Xiong, Mo Li, and Kyle Jamieson. md-track: Leveraging multi-dimensionality for passive indoor wi-fi tracking. In *The 25th Annual International Conference on Mobile Computing and Networking*, pages 1–16, 2019.

[237] Ambuj Varshney, Andreas Soleiman, and Thiemo Voigt. Tunnelscatter: Low power communication for sensor tags using tunnel diodes. In *The 25th Annual International Conference on Mobile Computing and Networking*, pages 1–17, 2019.

[238] Ambuj Varshney, Wenqing Yan, and Prabal Dutta. Judo: addressing the energy asymmetry of wireless embedded systems through tunnel diode based wireless transmitters. In *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*, pages 273–286, 2022.

[239] Skyworks. https://store.skyworksinc.com/products/detail/sms7630006-skyworks/150612/, 2022.

[240] Kubra Alemdar, Divashree Varshney, Subhramoy Mohanti, Ufuk Muncuk, and Kaushik Chowdhury. Rfclock: timing, phase and frequency synchronization for distributed wireless networks. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*, pages 15–27, 2021.

[241] Infinion. https://www.infineon.com/dgdl/Infineon-BAR50SERIES-DS-v01_01-en.pdf?fileId=db3a304314dca3890114fea7dd410a92, 2022.

[242] Infinion. https://www.infineon.com/cms/en/product/rf/rf-diode/rf-pin-diode/antenna-switch/bar50-02v/?redirId=192411, 2022.

[243] Texas Instruments. https://www.ti.com/lit/an/slaa720/slaa720.pdf, 2022.

[244] Texas Instruments. https://www.ti.com/tool/MSP-EXP430FR5969, 2022.

[245] Ansys. https://www.ansys.com/products/electronics/ansys-hfss, 2022.

[246] Ettus. https://www.ettus.com/all-products/vert900/, 2022.

[247] Ettus. https://www.ettus.com/all-products/vert2450/, 2022.

[248] Qorvo. https://www.qorvo.com/products/p/RF5110G, 2022.

[249] Ettus. https://www.ettus.com/all-products/un210-kit/, 2022.

[250] Matlab. https://www.mathworks.com/help/supportpkg/usrpradio/ug/qpsk-receiver-with-usrp-tm-hardware.html, 2022.

[251] Mathworks. https://www.mathworks.com/help/antenna/ug/ray-tracing-for-wireless-communications.html, 2023.

[252] Angus CK Mak, Corbett R Rowell, Ross D Murch, and Chi-Lun Mak. Reconfigurable multiband antenna designs for wireless communication devices. *IEEE Transactions on Antennas and Propagation*, 55(7):1919–1928, 2007.

[253] Ajith Kumar MM, Amalendu Patnaik, and Christos G Christodoulou. Design and testing of a multifrequency antenna with a reconfigurable feed. *IEEE Antennas and wireless propagation letters*, 13:730–733, 2014.

[254] Swarun Kumar, Stephanie Gil, Dina Katabi, and Daniela Rus. Accurate indoor localization with zero start-up cost. In *Proceedings of the 20th annual international conference on Mobile computing and networking*, pages 483–494, 2014.

[255] Maurizio Rea, Aymen Fakhreddine, Domenico Giustiniano, and Vincent Lenders. Filtering noisy 802.11 time-of-flight ranging measurements from commoditized wifi radios. *IEEE/ACM Transactions on Networking*, 25(4):2514–2527, 2017.

[256] Jie Xiong, Karthikeyan Sundaresan, and Kyle Jamieson. Tonetrack: Leveraging frequency-agile radios for time-based indoor wireless localization. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, pages 537–549, 2015.

[257] Li Sun, Souvik Sen, Dimitrios Koutsonikolas, and Kyu-Han Kim. Widraw: Enabling hands-free drawing in the air on commodity wifi devices. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, pages 77–89, 2015.

[258] Suraj Jog, Junfeng Guan, Sohrab Madani, Ruochen Lu, Songbin Gong, Deepak Vasisht, and Haitham Hassanieh. Enabling iot self-localization using ambient 5g signals. In *19th USENIX Symposium on Networked Systems Design and Implementation (NSDI 22)*, pages 1011–1026, 2022.

[259] Alex T Mariakakis, Souvik Sen, Jeongkeun Lee, and Kyu-Han Kim. Sail: Single access point-based indoor localization. In *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*, pages 315–328, 2014.

[260] Xiuyuan Zheng, Chen Wang, Yingying Chen, and Jie Yang. Accurate rogue access point localization leveraging fine-grained channel information. In *2014 IEEE conference on communications and network security*, pages 211–219. IEEE, 2014.

[261] Elahe Soltanaghaei, Adwait Dongare, Akarsh Prabhakara, Swarun Kumar, Anthony Rowe, and Kamin Whitehouse. Tagfi: Locating ultra-low power wifi tags using unmodified wifi infrastructure. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(1):1–29, 2021.

[262] Jue Wang, Deepak Vasisht, and Dina Katabi. Rf-idraw: Virtual touch screen in the air using rf signals. *ACM SIGCOMM Computer Communication Review*, 44(4):235–246, 2014.

[263] Yunfei Ma, Nicholas Selby, and Fadel Adib. Minding the billions: Ultra-wideband localization for deployed rfid tags. In *Proceedings of the 23rd annual international conference on mobile computing and networking*, pages 248–260, 2017.

[264] Ricardo Gonçalves, Pedro Pinho, and Nuno B Carvalho. Compact, frequency reconfigurable, printed monopole antenna. *International Journal of Antennas and Propagation*, 2012, 2012.

[265] Liping Han, Caixia Wang, Xinwei Chen, and Wenmei Zhang. Compact frequency-reconfigurable slot antenna for wireless applications. *IEEE Antennas and Wireless Propagation Letters*, 15:1795–1798, 2016.

[266] Zhong-Liang Lu, Xue-Xia Yang, and Guan-Nan Tan. A multidirectional pattern-reconfigurable patch antenna with csrr on the ground. *IEEE Antennas and Wireless Propagation Letters*, 16:416–419, 2016.

[267] Tamer Aboufoul, Clive Parini, Xiaodong Chen, and Akram Alomainy. Pattern-reconfigurable planar circular ultra-wideband monopole antenna. *IEEE Transactions on Antennas and Propagation*, 61(10):4973–4980, 2013.

[268] Anuradha A Palsokar and SL Lahudkar. Frequency and pattern reconfigurable rectangular patch antenna using single pin diode. *AEU-International Journal of Electronics and Communications*, 125:153370, 2020.

[269] Fan Wu and Kwai Man Luk. Wideband tri-polarization reconfigurable magneto-electric dipole antenna. *IEEE transactions on antennas and propagation*, 65(4):1633–1641, 2017.

[270] Yasir IA Al-Yasir, Abdulkareem S Abdullah, Naser Ojaroudi Parchin, Raed A Abd-Alhameed, and James M Noras. A new polarization-reconfigurable antenna for 5g applications. *Electronics*, 7(11):293, 2018.

[271] Arash Valizade, Mouhammad Ojaroudi, and Nasser Ojaroudi. Cpw-fed small slot antenna with reconfigurable circular polarizion and impdeance bandwidth characteristcis for dcs/wimax applications. *Progress In Electromagnetics Research C*, 56:65–72, 2015.

[272] Daniele Piazza, Prathaban Mookiah, Michele d'Amico, and Kapil R Dandekar. Experimental analysis of pattern and polarization reconfigurable circular patch antennas for mimo systems. *IEEE Transactions on Vehicular Technology*, 59(5):2352–2362, 2010.

[273] Naser Ojaroudi Parchin, Haleh Jahanbakhsh Basherlou, Yasir IA Al-Yasir, Ahmed M Abdulkhaleq, and Raed A Abd-Alhameed. Reconfigurable antennas: Switching techniques—a survey. *Electronics*, 9(2):336, 2020.

[274] Jennifer T Bernhard. Reconfigurable antennas. *Synthesis lectures on antennas*, 2(1):1–66, 2007.

[275] Marta Cabedo-Fabres, Eva Antonino-Daviu, Alejandro Valero-Nogueira, and Miguel Ferrando Bataller. The theory of characteristic modes revisited: A contribution to the design of antennas for modern applications. *IEEE Antennas and Propagation Magazine*, 49(5):52–68, 2007.

[276] Aidin Mehdipour, Tayeb A Denidni, Abdel-R Sebak, Christopher W Trueman, Iosif D Rosca, and Suong V Hoa. Mechanically reconfigurable antennas using an anisotropic carbon-fibre composite ground. *IET Microwaves, Antennas & Propagation*, 7(13):1055–1063, 2013.

[277] Daniel Rodrigo, Lluís Jofre, and Bedri A Cetiner. Circular beam-steering reconfigurable antenna with liquid metal parasitics. *IEEE transactions on antennas and propagation*, 60(4):1796–1802, 2012.

[278] Riaz Ahmed Liyakath, Arash Takshi, and Gokhan Mumcu. Multilayer stretchable conductors on polymer substrates for conformal and reconfigurable antennas. *IEEE Antennas and Wireless Propagation Letters*, 12:603–606, 2013.

[279] Joseph Costantine, Youssef Tawk, Jonathan Woodland, Noah Flaum, and Christos G Christodoulou. Reconfigurable antenna system with a movable ground plane for cognitive radio. *IET Microwaves, Antennas & Propagation*, 8(11):858–863, 2014.

[280] Mateusz Groth, Mateusz Rzymowski, Krzysztof Nyka, and Lukasz Kulas. Espar antenna-based wsn node with doa estimation capability. *IEEE Access*, 8:91435–91447, 2020.

[281] Xue-Song Yang, Bing-Zhong Wang, Sai Ho Yeung, Quan Xue, and Kim Fung Man. Circularly polarized reconfigurable crossed-yagi patch antenna. *IEEE Antennas and Propagation Magazine*, 53(5):65–80, 2011.

[282] Abdul Zohur, H Mopidevi, D Rodrigo, M Unlu, L Jofre, and Bedri A Cetiner. Rf mems reconfigurable two-band antenna. *IEEE Antennas and Wireless Propagation Letters*, 12:72–75, 2013.

[283] Bedri A Cetiner, Gemma Roqueta Crusats, Lluís Jofre, and Necmi Biyikli. Rf mems integrated frequency reconfigurable annular slot antenna. *IEEE Transactions on Antennas and Propagation*, 58(3):626–632, 2009.

[284] Alfred Grau, Jordi Romeu, Ming-Jer Lee, Sebastian Blanch, Lluís Jofre, and Franco De Flaviis. A dual-linearly-polarized mems-reconfigurable antenna for narrowband mimo communication systems. *IEEE Transactions on Antennas and Propagation*, 58(1):4–17, 2009.

[285] Amjad Iqbal, Amor Smida, Nazih Khaddaj Mallat, Ridha Ghayoula, Issa Elfergani, Jonathan Rodriguez, and Sunghwan Kim. Frequency and pattern reconfigurable antenna for emerging wireless communication systems. *Electronics*, 8(4):407, 2019.

[286] Muzammil Jusoh, Tamer Aboufoul, Thennarasan Sabapathy, Akram Alomainy, and Muhammad Ramlee Kamarudin. Pattern-reconfigurable microstrip patch antenna with multidirectional beam for wimax application. *IEEE antennas and wireless propagation letters*, 13:860–863, 2014.

[287] Vida Vakilian, Hoang V Nguyen, Samer Abielmona, Sébastien Roy, and Jean-François Frigon. Experimental study of direction-of-arrival estimation using reconfigurable antennas. In *2014 IEEE 27th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–4. IEEE, 2014.

[288] Emrah Kaderli, İsrafil Bahçeci, Kathleen M Kaplan, and Bedri A Cetiner. On the use of reconfigurable antenna arrays for doa estimation of correlated signals. In *2016 IEEE Radar Conference (RadarConf)*, pages 1–5. IEEE, 2016.

[289] Aki Hakkarainen, Janis Werner, Nikhil Gulati, Damiano Patron, Doug Pfeil, Henna Paaso, Aarne Mämmelä, Kapil Dandekar, and Mikko Valkama. Reconfigurable antenna based doa estimation and localization in cognitive radios: Low complexity algorithms and practical measurements. In *2014 9th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, pages 454–459. IEEE, 2014.

[290] Vikas Hassija, Vinay Chamola, Vatsal Gupta, Sarthak Jain, and Nadra Guizani. A survey on supply chain security: Application areas, security threats, and solution architectures. *IEEE Internet of Things Journal*, 8(8):6222–6246, 2020.

[291] Henry H Willis and David Santana Ortiz. Evaluating the security of the global containerized supply chain. 2004.

[292] Kunal Maurya, Mandeep Singh, and Neelu Jain. Real time vehicle tracking system using gsm and gps technology-an anti-theft tracking system. *International Journal of Electronics and Computer Science Engineering*, 1(3):1103–1107, 2012.

[293] Fahim Ahmed, Mark Phillips, Stephen Phillips, and Kyoung-Yun Kim. Comparative study of seamless asset location and tracking technologies. *Procedia Manufacturing*, 51:1138–1145, 2020.

[294] Partha Pratim Ray, Dinesh Dash, and Debashis De. A systematic review and implementation of iot-based pervasive sensor-enabled tracking system for dementia patients. *Journal of medical systems*, 43:1–21, 2019.

[295] Raul Fernández-Garcia and Ignacio Gil. An alternative wearable tracking system based on a low-power wide-area network. *Sensors*, 17(3):592, 2017.

[296] Attila Frankó, Gergely Vida, and Pal Varga. Reliable identification schemes for asset and production tracking in industry 4.0. *Sensors*, 20(13):3709, 2020.

[297] Sooyoung Yoo, Seok Kim, Eunhye Kim, Eunja Jung, Kee-Hyuck Lee, and Hee Hwang. Real-time location system-based asset tracking in the healthcare field: lessons learned from a feasibility study. *BMC medical informatics and decision making*, 18:1–10, 2018.

[298] GlobalStar. Globalstar iot asset tracking. `https://www.globalstar.com/en-us/products/iot`, 2024.

[299] DigitalMatter. Digitalmatter asset tracking. `https://www.digitalmatter.com/our-devices/4g-gps-trackers/`, 2024.

[300] Atul Bansal, Akshay Gadre, Vaibhav Singh, Anthony Rowe, Bob Iannucci, and Swarun Kumar. Owll: Accurate lora localization using the tv whitespaces. In *Proceedings of the 20th International Conference on Information Processing in Sensor Networks (co-located with CPS-IoT Week 2021)*, pages 148–162, 2021.

[301] Willem Bouten, Edwin W Baaij, Judy Shamoun-Baranes, and Kees CJ Camphuysen. A flexible gps tracking system for studying bird behaviour at multiple scales. *Journal of Ornithology*, 154:571–580, 2013.

[302] Youjing Cui and Shuzhi Sam Ge. Autonomous vehicle positioning with gps in urban canyon environments. *IEEE transactions on robotics and automation*, 19(1):15–25, 2003.

[303] Roshan Ayyalasomayajula, Deepak Vasisht, and Dinesh Bharadia. Bloc: Csi-based accurate localization for ble tags. In *Proceedings of the 14th International Conference on emerging Networking EXperiments and Technologies*, pages 126–138, 2018.

[304] Amee Trivedi and Deepak Vasisht. Digital contact tracing: technologies, shortcomings, and the path forward. *ACM SIGCOMM Computer Communication Review*, 50(4):75–81, 2020.

[305] Hazem Ibrahim, Rohail Asim, Matteo Varvello, and Yasir Zaki. I tag, you tag, everybody tags! In *Proceedings of the 2023 ACM on Internet Measurement Conference*, pages 561–568, 2023.

[306] Laura Dodds, Isaac Perper, Aline Eid, and Fadel Adib. A handheld fine-grained rfid localization system with complex-controlled polarization. In *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*, pages 1–15, 2023.

[307] Alexander Heinrich, Niklas Bittner, and Matthias Hollick. Airguard-protecting android users from stalking attacks by apple find my devices. In *Proceedings of the 15th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, pages 26–38, 2022.

[308] FCC. `https://docs.fcc.gov/public/attachments/DA-15-1487A1.pdf`, 2023.

[309] Yuda Feng, Yaxiong Xie, Deepak Ganesan, and Jie Xiong. Lte-based pervasive sensing across indoor and outdoor. In *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*, pages 138–151, 2021.

[310] Nakul Garg, Irtaza Shahid, Ramanujan K Sheshadri, Karthikeyan Sundaresan, and Nirupam Roy. Fast localization and tracking in city-scale uwb networks. *arXiv preprint arXiv:2310.02211*, 2023.

[311] José A del Peral-Rosado, Ronald Raulefs, José A López-Salcedo, and Gonzalo Seco-Granados. Survey of cellular mobile radio localization methods: From 1g to 5g. *IEEE Communications Surveys & Tutorials*, 20(2):1124–1148, 2017.

[312] Kelly R Evenson, Michelle M Goto, and Robert D Furberg. Systematic review of the validity and reliability of consumer-wearable activity trackers. *International Journal of Behavioral Nutrition and Physical Activity*, 12(1):1–22, 2015.

[313] Raghav Subbaraman, Kevin Mills, Aaron Schulman, and Dinesh Bharadia. Crescendo: Towards wideband, real-time, high-fidelity spectrum sensing systems. In *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*, pages 1–14, 2023.

[314] Xiuzhen Guo, Longfei Shangguan, Yuan He, Nan Jing, Jiacheng Zhang, Haotian Jiang, and Yunhao Liu. Saiyan: Design and implementation of a low-power demodulator for {LoRa} backscatter systems. In *19th USENIX Symposium on Networked Systems Design and Implementation (NSDI 22)*, pages 437–451, 2022.

[315] u blox. Ultra-low power gps with the u-blox m10, 2023. Accessed: 2024-07-01.

[316] Carlos Pérez-Penichet, Claro Noda, Ambuj Varshney, and Thiemo Voigt. Battery-free 802.15. 4 receiver. In *2018 17th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 164–175. IEEE, 2018.

[317] Anran Wang, Vikram Iyer, Vamsi Talla, Joshua R Smith, and Shyamnath Gollakota. {FM} backscatter: Enabling connected cities and smart fabrics. In *14th USENIX Symposium on Networked Systems Design and Implementation (NSDI 17)*, pages 243–258, 2017.

[318] Ambuj Varshney and Lorenzo Corneo. Tunnel emitter: Tunnel diode based low-power carrier emitters for backscatter tags. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, pages 1–14, 2020.

[319] Kai Lv, Xiaolong Pan, Junwen Zhang, Kaihui Wang, Dong Guo, Qi Zhang, ao Li Deng, and Xiangjun Xin. Envelope detection of 16qam single carrier signal for im/dd rof system at q-band. *IEICE Electronics Express*, 16(4):20180986–20180986, 2019.

[320] Vivek Mangal and Peter R Kinget. A wake-up receiver with a multi-stage self-mixer and with enhanced sensitivity when using an interferer as local oscillator. *IEEE Journal of Solid-State Circuits*, 54(3):808–820, 2019.

[321] Yi Wu, Hai Jiang, Ye Wu, and Dongmei Zhang. Physical cell identity self-organization for home enodeb deployment in lte. In *2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*, pages 1–6. IEEE, 2010.

[322] Hakan Kavlak and Hakki Ilk. Pci planning strategies for long term evolution networks. In *NETWORKING 2012 Workshops: International IFIP TC 6 Workshops, ETICS, HetsNets, and CompNets, Held at NETWORKING 2012, Prague, Czech Republic, May 25, 2012. Proceedings 11*, pages 151–156. Springer, 2012.

[323] Stefania Sesia, Issam Toufik, and Matthew Baker. *LTE-the UMTS long term evolution: from theory to practice*. John Wiley & Sons, 2011.

[324] Tao Zhang, Shuxing Li, Tao Feng, and Gennian Ge. Some new results on the cross correlation of m-sequences. *IEEE transactions on information theory*, 60(5):3062–3068, 2014.

[325] Open Cell ID. Open cell id. `https://opencellid.org/`, 2024.

[326] Mozilla. Mozilla location services. `https://location.services.mozilla.com`, 2024.

[327] CellMapper. Cellmapper. `https://www.cellmapper.net/`, 2024.

[328] Google. Google snap to roads. `https://developers.google.com/maps/documentation/roads/snap`, 2024.

[329] OpenStreetMap contributors. OpenStreetMap. `https://www.openstreetmap.org`, 2024.

[330] 3rd Generation Partnership Project (3GPP). Gsm/edge physical layer on the radio path. general description. Technical Specification 45.001, Technical Specification Group Radio Access Network, 2024. Available: 3GPP website.

[331] Ajay R Mishra. *Advanced cellular network planning and optimisation: 2G/2.5 G/3G... evolution to 4G*. John Wiley & Sons, 2007.

[332] 3GPP. 3gpp specifications per tsg round. https://www.3gpp.org/specifications-technologies/specifications-by-series/specifications-per-tsg-round, 2024. Accessed: 2024-07-01.

[333] Technical Specification Group Radio Access Network. Nr; physical channels and modulation. Technical Specification 38.211, 3rd Generation Partnership Project (3GPP).

[334] Aya Shimura, Mamoru Sawahashi, Satoshi Nagata, and Yoshihisa Kishiyama. Physical cell id detection performance applying frequency diversity reception to npss and nsss for nb-iot. In *2018 24th Asia-Pacific Conference on Communications (APCC)*, pages 514–519, 2018.

[335] David D Wentzloff, Abdullah Alghaihab, and Jaeho Im. Ultra-low power receivers for iot applications: A review. In *2020 IEEE Custom Integrated Circuits Conference (CICC)*, pages 1–8. IEEE, 2020.

[336] Hewlett-Packards. https://www.hpl.hp.com/hpjournal/95dec/dec95a12.pdf.

[337] Texas Instruments. 12-bit ultra-low-power sar adc with spi interface. https://www.ti.com/product/ADS7042, 2024.

[338] Ambiq. Ultra-low power microcontroller apollo3 blue. https://ambiq.com/apollo3/, 2024.

[339] Xiuzhen Guo, Yuan He, Zihao Yu, Jiacheng Zhang, Yunhao Liu, and Longfei Shangguan. Rf-transformer: a unified backscatter radio hardware abstraction. In *Proceedings of the 28th annual international conference on mobile computing and networking*, pages 446–458, 2022.

[340] Skyworks Solutions Inc. Sms7630-079lf schottky diode - single, 20v, 100ma, 2 pf @ 0v, 30 ohm @ 5ma, sc-79 package. https://www.digikey.com/en/products/detail/skyworks-solutions-inc/SMS7630-079LF/2052135, Accessed 2024. Available at Digi-Key Electronics.

[341] Infineon Technologies. Bga7m1n6e6327xtsa1 - rf amplifier, lte band 7, 2.5ghz to 2.7ghz, 15.6db gain, 0.75db noise figure, 20dbm oip3, tsnp-6 package. https://www.digikey.com/en/products/detail/infineon-technologies/BGA7M1N6E6327XTSA1/5959862, Accessed 2024. Available at Digi-Key Electronics.

[342] T. Taris, JB. Begueret, and Y. Deval. A 60µw lna for 2.4 ghz wireless sensors network applications. In *2011 IEEE Radio Frequency Integrated Circuits Symposium*, pages 1–4, 2011.

[343] Michiel Van Elzakker, Ed Van Tuijl, Paul Geraedts, Daniel Schinkel, Eric Klumperink, and Bram Nauta. A 1.9 μw 4.4 fj/conversion-step 10b 1ms/s charge-redistribution adc. In *2008 IEEE International Solid-State Circuits Conference-Digest of Technical Papers*, pages 244–610. IEEE, 2008.

[344] Rizanne Elbakly and Moustafa Youssef. Crescendo: An infrastructure-free ubiquitous cellular network-based localization system. In *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6. IEEE, 2019.

[345] Qorvo. Rf2052 rf frequency converter, 30 - 2500 mhz, integrated pll/vco, rf mixer. `https://www.rfmw.com/products/detail/rf2052-qorvo/326547/?gad_source=1&gclid=EAIaIQobChMI0LTFvrT_hgMVnWJHAR0uzQKnEAAYBCAAEgL_nfD_BwE`, 2024. Accessed: 2024-06-28.

[346] Infineon Technologies. Bga7m1n6e6327xtsa1 rf mosfet, low noise amplifier. `https://www.win-source.net/products/detail/infineon-technologies/bga7m1n6e6327xtsa1.html`, 2024. Accessed: 2024-06-28.

[347] PowerFleet. Asset tracking solutions, Year. Accessed: Month Day, Year.

[348] Digital Matter. Cold chain monitoring, Year. Accessed: Month Day, Year.

[349] Hanhaa. Cold chain monitoring, Year. Accessed: Month Day, Year.

[350] Asil Oztekin, Foad M Pajouh, Dursun Delen, and Leva K Swim. An rfid network design methodology for asset tracking in healthcare. *Decision support systems*, 49(1):100–109, 2010.

[351] Nakul Garg, Irtaza Shahid, Karthik Sankar, Mallesham Dasari, Ramanujan K Sheshadri, Karthikeyan Sundaresan, and Nirupam Roy. Bringing ar/vr to everyday life-a wireless localization perspective. In *Proceedings of the 24th International Workshop on Mobile Computing Systems and Applications*, pages 142–142, 2023.

[352] Ygal Bendavid, Harold Boeck, and Richard Philippe. Redesigning the replenishment process of medical supplies in hospitals with rfid. *Business Process Management Journal*, 16(6):991–1013, 2010.

[353] Christos Laoudias, Adriano Moreira, Sunwoo Kim, Sangwoo Lee, Lauri Wirola, and Carlo Fischione. A survey of enabling technologies for network localization, tracking, and navigation. *IEEE Communications Surveys & Tutorials*, 20(4):3607–3644, 2018.

[354] Stefano Parrino, Giacomo Peruzzi, and Alessandro Pozzebon. Lopatran: Low power asset tracking by means of narrow band iot (nb-iot) technology. *Sensors*, 21(11):3772, 2021.

[355] Jue Wang and Dina Katabi. Dude, where's my card? rfid positioning that works with multipath and non-line of sight. In *Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM*, pages 51–62, 2013.

[356] Lei Yang, Yekui Chen, Xiang-Yang Li, Chaowei Xiao, Mo Li, and Yunhao Liu. Tagoram: Real-time tracking of mobile rfid tags to high precision using cots devices. In *Proceedings of the 20th annual international conference on Mobile computing and networking*, pages 237–248, 2014.

[357] Ruofeng Liu, Zhimeng Yin, Wenchao Jiang, and Tian He. Wibeacon: Expanding ble location-based services via wifi. In *Proceedings of the 27th annual international conference on mobile computing and networking*, pages 83–96, 2021.

[358] Sangki Yun, Yi-Chao Chen, Huihuang Zheng, Lili Qiu, and Wenguang Mao. Strata: Fine-grained acoustic-based device-free tracking. In *Proceedings of the 15th annual international conference on mobile systems, applications, and services*, pages 15–28, 2017.

[359] Nakul Garg, Yang Bai, and Nirupam Roy. Owlet: Enabling spatial information in ubiquitous acoustic devices. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, pages 255–268, 2021.

[360] Yang Bai, Nakul Garg, and Nirupam Roy. Microstructure-assisted vision: Adding new senses to low-power devices. *GetMobile: Mobile Computing and Communications*, 27(4):15–20, 2024.

[361] Dong Li, Jialin Liu, Sunghoon Ivan Lee, and Jie Xiong. Fm-track: pushing the limits of contactless multi-target tracking using acoustic signals. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, pages 150–163, 2020.

[362]

[363] Da Li, Yingke Lei, and Haichuan Zhang. A novel outdoor positioning technique using lte network fingerprints. *Sensors*, 20(6):1691, 2020.

[364] Giovanni Pecoraro, Simone Di Domenico, Ernestina Cianca, and Mauro De Sanctis. Csi-based fingerprinting for indoor localization using lte signals. *EURASIP Journal on Advances in Signal Processing*, 2018:1–18, 2018.

[365] Kimia Shamaei and Zaher M Kassas. Lte receiver design and multipath analysis for navigation in urban environments. *Navigation*, 65(4):655–675, 2018.

[366] Kimia Shamaei, Joe Khalife, and Zaher M Kassas. Exploiting lte signals for navigation: Theory to implementation. *IEEE Transactions on Wireless Communications*, 17(4):2173–2189, 2018.

[367] Yuda Feng, Yaxiong Xie, Deepak Ganesan, and Jie Xiong. Lte-based low-cost and low-power soil moisture sensing. In *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*, pages 421–434, 2022.

[368] Mohamed Ibrahim and Moustafa Youssef. A hidden markov model for localization using low-end gsm cell phones. In *2011 IEEE International Conference on Communications (ICC)*, pages 1–5. IEEE, 2011.

[369] Jie Yang, Alexander Varshavsky, Hongbo Liu, Yingying Chen, and Marco Gruteser. Accuracy characterization of cell tower localization. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pages 223–226, 2010.

[370] Ding-Bing Lin and Rong-Terng Juang. Mobile location estimation based on differences of signal attenuations for gsm systems. *IEEE transactions on vehicular technology*, 54(4):1447–1454, 2005.

[371] Mohamed Ibrahim and Moustafa Youssef. Cellsense: An accurate energy-efficient gsm positioning system. *IEEE Transactions on Vehicular Technology*, 61(1):286–296, 2011.

[372] Heikki Laitinen, Jaakko Lahteenmaki, and Tero Nordstrom. Database correlation method for gsm location. In *IEEE VTS 53rd Vehicular Technology Conference, Spring 2001. Proceedings (Cat. No. 01CH37202)*, volume 4, pages 2504–2508. IEEE, 2001.

[373] Sinem Coleri Ergen, Huseyin Serhat Tetikol, Mehmet Kontik, Raffi Sevlian, Ram Rajagopal, and Pravin Varaiya. Rssi-fingerprinting-based mobile phone localization with route constraints. *IEEE Transactions on Vehicular Technology*, 63(1):423–428, 2013.

[374] Ayon Chakraborty, Luis E Ortiz, and Samir R Das. Network-side positioning of cellular-band devices with minimal effort. In *2015 IEEE Conference on Computer Communications (INFOCOM)*, pages 2767–2775. IEEE, 2015.

[375] Hazem Sallouha, Alessandro Chiumento, and Sofie Pollin. Localization in long-range ultra narrow band iot networks using rssi. In *2017 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2017.

[376] Ahmed Shokry, Marwan Torki, and Moustafa Youssef. Deeploc: A ubiquitous accurate and low-overhead outdoor cellular localization system. In *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 339–348, 2018.

[377] Amel Mohamed, Mohamed Tharwat, Mohamed Magdy, Tarek Abubakr, Omar Nasr, and Moustafa Youssef. Deepfeat: Robust large-scale multi-features outdoor localization in lte networks using deep learning. *IEEE Access*, 10:3400–3414, 2022.

[378] Shuli Zhu, Lingkun Li, Xuyu Wang, Changcheng Liu, Yuqin Jiang, Zengwei Huo, Hua Chai, Jiqiang Liu, Dan Tao, and Ruipeng Gao. Experience: Large-scale cellular localization for pickup position recommendation at black-hole. In *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*, pages 1–15, 2023.

[379] Nakul Garg and Nirupam Roy. Sirius: A self-localization system for resource-constrained iot sensors. In *Proceedings of the 21st Annual International Conference on Mobile Systems, Applications and Services*, pages 289–302, 2023.

[380] Nakul Garg and Nirupam Roy. poster: Ultra-low-power angle-of-arrival estimation using a single antenna. In *Proceedings of the 21st Annual International Conference on Mobile Systems, Applications and Services*, pages 569–570, 2023.

[381] Joshua F Ensworth, Alexander T Hoang, and Matthew S Reynolds. A low power 2.4 ghz superheterodyne receiver architecture with external lo for wirelessly powered backscatter tags and sensors. In *2017 IEEE International Conference on RFID (RFID)*, pages 149–154. IEEE, 2017.

[382] Pengyu Zhang, Colleen Josephson, Dinesh Bharadia, and Sachin Katti. Freerider: Backscatter communication using commodity radios. In *Proceedings of the 13th international conference on emerging networking experiments and technologies*, pages 389–401, 2017.

[383] Pengyu Zhang, Mohammad Rostami, Pan Hu, and Deepak Ganesan. Enabling practical backscatter communication for on-body sensors. In *Proceedings of the 2016 ACM SIGCOMM Conference*, pages 370–383, 2016.

[384] Vamsi Talla, Mehrdad Hessar, Bryce Kellogg, Ali Najafi, Joshua R Smith, and Shyamnath Gollakota. Lora backscatter: Enabling the vision of ubiquitous connectivity. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, 1(3):1–24, 2017.

[385] Mohamed R Abdelhamid, Ruicong Chen, Joonhyuk Cho, Anantha P Chandrakasan, and Fadel Adib. Self-reconfigurable micro-implants for cross-tissue wireless and batteryless connectivity. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, pages 1–14, 2020.

[386] Yao Peng, Longfei Shangguan, Yue Hu, Yujie Qian, Xianshang Lin, Xiaojiang Chen, Dingyi Fang, and Kyle Jamieson. Plora: A passive long-range data network from ambient lora transmissions. In *Proceedings of the 2018 conference of the ACM special interest group on data communication*, pages 147–160, 2018.

[387] Farzan Dehbashi, Ali Abedi, Tim Brecht, and Omid Abari. Verification: can wifi backscatter replace rfid? In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*, pages 97–107, 2021.

[388] Ambuj Varshney, Oliver Harms, Carlos Pérez-Penichet, Christian Rohner, Frederik Hermans, and Thiemo Voigt. Lorea: A backscatter architecture that achieves a long communication range. In *Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems*, pages 1–14, 2017.

[389] Deepak Vasisht, Guo Zhang, Omid Abari, Hsiao-Ming Lu, Jacob Flanz, and Dina Katabi. In-body backscatter communication and localization. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*, pages 132–146, 2018.

[390] Thibault Pirson and David Bol. Assessing the embodied carbon footprint of iot edge devices with a bottom-up life-cycle approach. *Journal of Cleaner Production*, 322:128966, 2021.

[391] Global View. Buy gps tracking devices, 2024. Accessed: 2024-07-01.

[392] u blox. Sara-n310 module. `https://www.u-blox.com/en/product/sara-n310-module`, 2024. Accessed: 2024-07-01.

[393] J Salo, M Nur-Alam, and K Chang. Practical introduction to lte radio planning. *A white paper on basics of radio planning for 3GPP LTE in interference limited and coverage limited scenarios, European Communications Engineering (ECE) Ltd, Espoo, Finland*, 47, 2010.

[394] RFAssurance. The value of pci planning in lte. `https://main.rfassurance.com/?q=node/79`, 2015. Accessed: 2024-07-01.

[395] Vikram Iyer, Hans Gaensbauer, Thomas L Daniel, and Shyamnath Gollakota. Wind dispersal of battery-free wireless devices. *Nature*, 603(7901):427–433, 2022.

[396] Benjamin Thompson and Noah Baker. The floating sensors inspired by seeds. *Nature*, 2021.

[397] Vikram Iyer, Maruchi Kim, Shirley Xue, Anran Wang, and Shyamnath Gollakota. Airdropping sensor networks from drones and insects. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, pages 1–14, 2020.

[398] Neha Bhadwal, Vishu Madaan, Prateek Agrawal, Awadesh Shukla, and Anuj Kakran. Smart border surveillance system using wireless sensor network and computer vision. In *2019 international conference on Automation, Computational and Technology Management (ICACTM)*, pages 183–190. IEEE, 2019.

[399] Inhee Lee, Roger Hsiao, Gordy Carichner, Chin-Wei Hsu, Mingyu Yang, Sara Shoouri, Katherine Ernst, Tess Carichner, Yuyang Li, Jaechan Lim, et al. msail: milligram-scale multi-modal sensor platform for monarch butterfly migration tracking. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*, pages 517–530, 2021.

[400] Deepak Vasisht, Zerina Kapetanovic, Jongho Won, Xinxin Jin, Ranveer Chandra, Sudipta Sinha, Ashish Kapoor, Madhusudhan Sudarshan, and Sean Stratman. {FarmBeats}: an {IoT} platform for {Data-Driven} agriculture. In *14th USENIX Symposium on Networked Systems Design and Implementation (NSDI 17)*, pages 515–529, 2017.

[401] Judy Scott and Carlton Scott. Drone delivery models for healthcare. 2017.

[402] Yong-Gu Han, Youngho Cho, Ohseok Kwon, et al. The use of conservation drones in ecology and wildlife research. *Journal of Ecology and Environment*, 38(1):113–118, 2015.

[403] Technology and systems for drone shows. `https://www.verge.aero/`. Last accessed 21 September 2023.

[404] Daniel Minoli and Benedict Occhiogrosso. Ultrawideband (uwb) technology for smart cities iot applications. In *2018 IEEE international smart cities conference (ISC2)*, pages 1–8. IEEE, 2018.

[405] Fatma Raissi, Sami Yangui, and Frederic Camps. Autonomous cars, 5g mobile networks and smart cities: Beyond the hype. In *2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, pages 180–185. IEEE, 2019.

[406] Nakul Garg, Aritrik Ghosh, and Nirupam Roy. Litefoot: Ultra-low-power localization using ambient cellular signals. In *Proceedings of the 22nd ACM Conference on Embedded Networked Sensor Systems (SenSys 2024)*, New York, NY, USA, 2024. ACM. University of Maryland, College Park.

[407] Nima Alam and Andrew G Dempster. Cooperative positioning for vehicular networks: Facts and future. *IEEE transactions on intelligent transportation systems*, 14(4):1708–1717, 2013.

[408] Jiadai Wang, Jiajia Liu, and Nei Kato. Networking and communications in autonomous driving: A survey. *IEEE Communications Surveys & Tutorials*, 21(2):1243–1274, 2018.

[409] Kan Zheng, Qiang Zheng, Haojun Yang, Long Zhao, Lu Hou, and Periklis Chatzimisios. Reliable and efficient autonomous driving: the need for heterogeneous vehicular networks. *IEEE Communications Magazine*, 53(12):72–79, 2015.

[410] Henk Wymeersch, Gabriel Rodrigues de Campos, Paolo Falcone, Lennart Svensson, and Erik G Ström. Challenges for cooperative its: Improving road safety through the integration of wireless communications, control, and positioning. In *2015 international conference on computing, networking and communications (ICNC)*, pages 573–578. IEEE, 2015.

[411] Fira. Unleashing the potential of uwb: Regulatory considerations. *FiRa Consortium*, August 2022.

[412] Qorvo. `https://www.mouser.com/datasheet/2/412/DWM1000_Data_Sheet-1950396.pdf`, 2016.

[413] Qorvo. `https://www.qorvo.com/products/d/da008334`, 2021.

[414] Jan De Leeuw and Patrick Mair. Multidimensional scaling using majorization: Smacof in r. *Journal of statistical software*, 31:1–30, 2009.

[415] Moshe Hamaoui. Non-iterative mds method for collaborative network localization with sparse range and pointing measurements. *IEEE Transactions on Signal Processing*, 67(3):568–578, 2018.

[416] Zhongmin Pei, Zhidong Deng, Shuo Xu, and Xiao Xu. Anchor-free localization method for mobile targets in coal mine wireless sensor networks. *Sensors*, 9(4):2836–2850, 2009.

[417] Brian Beck, Robert Baxley, and Joseph Kim. Real-time, anchor-free node tracking using ultrawideband range and odometry data. In *2014 IEEE international conference on ultra-wideband (ICUWB)*, pages 286–291. IEEE, 2014.

[418] Francesco Sottile, Maurizio A Spirito, Mauricio A Caceres, and Jaron Samson. Distributed-weighted multidimensional scaling for hybrid peer-to-peer localization. In *2010 Ubiquitous Positioning Indoor Navigation and Location Based Service*, pages 1–6. IEEE, 2010.

[419] Yi Shang, Wheeler Rumi, Ying Zhang, and Markus Fromherz. Localization from connectivity in sensor networks. *IEEE Transactions on parallel and distributed systems*, 15(11):961–974, 2004.

[420] Yi Shang and Wheeler Ruml. Improved mds-based localization. In *IEEE INFOCOM 2004*, volume 4, pages 2640–2651. IEEE, 2004.

[421] Hongyang Chen, Gang Wang, Zizhuo Wang, Hing-Cheung So, and H Vincent Poor. Non-line-of-sight node localization based on semi-definite programming in wireless sensor networks. *IEEE Transactions on Wireless Communications*, 11(1):108–116, 2011.

[422] Saipradeep Venkatraman, James Caffery, and Heung-Ryeol You. A novel toa location algorithm using los range estimation for nlos environments. *IEEE Transactions on Vehicular Technology*, 53(5):1515–1524, 2004.

[423] Carmelo Di Franco, Amanda Prorok, Nikolay Atanasov, Benjamin Kempke, Prabal Dutta, Vijay Kumar, and George J Pappas. Calibration-free network localization using non-line-of-sight ultra-wideband measurements. In *Proceedings of the 16th ACM/IEEE International Conference on Information Processing in Sensor Networks*, pages 235–246, 2017.

[424] Volodymyr Melnykov and Igor Melnykov. Initializing the em algorithm in gaussian mixture models with an unknown number of components. *Computational Statistics & Data Analysis*, 56(6):1381–1395, 2012.

[425] Christophe Biernacki, Gilles Celeux, and Gérard Govaert. Choosing starting values for the em algorithm for getting the highest likelihood in multivariate gaussian mixture models. *Computational Statistics & Data Analysis*, 41(3-4):561–575, 2003.

[426] Yihong Qi. *Wireless geolocation in a non-line-of-sight environment*. Princeton University, 2003.

[427] Gerard Laman. On graphs and rigidity of plane skeletal structures. *Journal of Engineering mathematics*, 4(4):331–340, 1970.

[428] Ronald L Graham and Pavol Hell. On the history of the minimum spanning tree problem. *Annals of the History of Computing*, 7(1):43–57, 1985.

[429] Mouser. `https://www.mouser.com/new/murata/murata-type2bp-uwb-module-eval-kit/`, 2024.

[430] Eric W Weisstein. Euler angles. `https://mathworld.wolfram.com/`, 2009.

[431] Gregory G Slabaugh. Computing euler angles from a rotation matrix. *Retrieved on August*, 6(2000):39–63, 1999.

[432] Raspberry Pi. `https://www.raspberrypi.com/products/raspberry-pi-3-model-b/`, 2023.

[433] Intel. `https://dev.intelrealsense.com/docs/tracking-camera-t265-datasheet`, 2023.

[434] Vicon. `https://www.vicon.com/`. Last accessed 12 March 2023.

[435] Sebastian Thrun. Probabilistic robotics. *Communications of the ACM*, 45(3):52–57, 2002.

[436] Vicon. https://docs.vicon.com/display/Vantage/V8+camera+specifications, 2023.

[437] NYC Metropolitan Transportation Authority. Nyc mta turnstile data, 2024. [Online]. Available: https://data.ny.gov/browse/select_dataset?category=Transportation.

[438] Fen Liu, Jing Liu, Yuqing Yin, Wenhan Wang, Donghai Hu, Pengpeng Chen, and Qiang Niu. Survey on wifi-based indoor positioning techniques. *IET communications*, 14(9):1372–1383, 2020.

[439] Pengfei Wang and Yufeng Luo. Research on wifi indoor location algorithm based on rssi ranging. In *2017 4th International Conference on Information Science and Control Engineering (ICISCE)*, pages 1694–1698. IEEE, 2017.

[440] Hongbo Liu, Yu Gan, Jie Yang, Simon Sidhom, Yan Wang, Yingying Chen, and Fan Ye. Push the limit of wifi based localization for smartphones. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 305–316, 2012.

[441] Chunyi Peng, Guobin Shen, Yongguang Zhang, Yanlin Li, and Kun Tan. Beepbeep: a high accuracy acoustic ranging system using cots mobile devices. In *Proceedings of the 5th international conference on Embedded networked sensor systems*, pages 1–14, 2007.

[442] Optitrack. https://optitrack.com/. Last accessed 12 March 2023.

[443] Manfred Klopschitz and Dieter Schmalstieg. Automatic reconstruction of wide-area fiducial marker models. In *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 71–74. IEEE, 2007.

[444] Alessandro Mulloni, Hartmut Seichter, and Dieter Schmalstieg. Handheld augmented reality indoor navigation with activity-based instructions. In *Proceedings of the 13th international conference on human computer interaction with mobile devices and services*, pages 211–220, 2011.

[445] Jan Kallwies, Bianca Forkel, and Hans-Joachim Wuensche. Determining and improving the localization accuracy of apriltag detection. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 8288–8294. IEEE, 2020.

[446] Boxin Zhao, Zongzhe Li, Jun Jiang, and Xiaolin Zhao. Relative localization for uavs based on april-tags. In *2020 Chinese Control And Decision Conference (CCDC)*, pages 444–449. IEEE, 2020.

[447] Thien Hoang Nguyen, Thien-Minh Nguyen, and Lihua Xie. Range-focused fusion of camera-imu-uwb for accurate and drift-reduced localization. *IEEE Robotics and Automation Letters*, 6(2):1678–1685, 2021.

[448] Fredrik Olsson, Jouni Rantakokko, and Jonas Nygårds. Cooperative localization using a foot-mounted inertial navigation system and ultrawideband ranging. In *2014 International conference on indoor positioning and indoor navigation (IPIN)*, pages 122–131. IEEE, 2014.

[449] Wang Shule, Carmen Martínez Almansa, Jorge Peña Queralta, Zhuo Zou, and Tomi Westerlund. Uwb-based localization for multi-uav systems and collaborative heterogeneous multi-robot systems. *Procedia Computer Science*, 175:357–364, 2020.

[450] Srdjan Čapkun, Maher Hamdi, and Jean-Pierre Hubaux. Gps-free positioning in mobile ad hoc networks. *Cluster Computing*, 5:157–167, 2002.

[451] Chong Shao, Bashima Islam, and Shahriar Nirjon. Marble: Mobile augmented reality using a distributed ble beacon infrastructure. In *2018 IEEE/ACM Third International Conference on Internet-of-Things Design and Implementation (IoTDI)*, pages 60–71. IEEE, 2018.

[452] Patrick Lazik, Niranjini Rajagopal, Oliver Shih, Bruno Sinopoli, and Anthony Rowe. Alps: A bluetooth and ultrasound platform for mapping and localization. In *Proceedings of the 13th ACM conference on embedded networked sensor systems*, pages 73–84, 2015.

[453] David Gómez, Paula Tarrío, Juan Li, Ana M Bernardos, and José R Casar. Indoor augmented reality based on ultrasound localization systems. In *Highlights on Practical Applications of Agents and Multi-Agent Systems: International Workshops of PAAMS 2013, Salamanca, Spain, May 22-24, 2013. Proceedings 11*, pages 202–212. Springer, 2013.

[454] Jinyang Li, Zhiheng Xie, Xiaoshan Sun, Jian Tang, Hengchang Liu, and John A Stankovic. An automatic and accurate localization system for firefighters. In *2018 IEEE/ACM Third International Conference on Internet-of-Things Design and Implementation (IoTDI)*, pages 13–24. IEEE, 2018.

[455] Ashutosh Dhekne, Ayon Chakraborty, Karthikeyan Sundaresan, and Sampath Rangarajan. Trackio: tracking first responders inside-out. In *16th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 19)*, pages 751–764, 2019.

[456] Christian Gentner and Markus Ulmschneider. Simultaneous localization and mapping for pedestrians using low-cost ultra-wideband system and gyroscope. In *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–8. IEEE, 2017.

[457] Ran Liu, Chau Yuen, Tri-Nhut Do, Dewei Jiao, Xiang Liu, and U-Xuan Tan. Cooperative relative positioning of mobile users by fusing imu inertial and uwb ranging information. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5623–5629. IEEE, 2017.

[458] Niranjini Rajagopal, John Miller, Krishna Kumar Reghu Kumar, Anh Luong, and Anthony Rowe. Improving augmented reality relocalization using beacons and magnetic field maps. In *2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–8. IEEE, 2019.

[459] Yang Song, Mingyang Guan, Wee Peng Tay, Choi Look Law, and Changyun Wen. Uwb/lidar fusion for cooperative range-only slam. In *2019 international conference on robotics and automation (ICRA)*, pages 6568–6574. IEEE, 2019.

[460] Chen Wang, Handuo Zhang, Thien-Minh Nguyen, and Lihua Xie. Ultra-wideband aided fast localization and mapping system. In *2017 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pages 1602–1609. IEEE, 2017.

[461] Radhika Nagpal, Howard Shrobe, and Jonathan Bachrach. Organizing a global coordinate system from local information on an ad hoc sensor network. In *Information Processing in Sensor Networks: Second International Workshop, IPSN 2003, Palo Alto, CA, USA, April 22–23, 2003 Proceedings*, pages 333–348. Springer, 2003.

[462] Andreas Savvides, Chih-Chieh Han, and Mani B Strivastava. Dynamic fine-grained localization in ad-hoc networks of sensors. In *Proceedings of the 7th annual international conference on Mobile computing and networking*, pages 166–179, 2001.

[463] Tolga Eren, OK Goldenberg, Walter Whiteley, Yang Richard Yang, A Stephen Morse, Brian DO Anderson, and Peter N Belhumeur. Rigidity, computation, and randomization in network localization. In *IEEE INFOCOM 2004*, volume 4, pages 2673–2684. IEEE, 2004.

[464] David Moore, John Leonard, Daniela Rus, and Seth Teller. Robust distributed network localization with noisy range measurements. In *Proceedings of the 2nd international conference on Embedded networked sensor systems*, pages 50–61, 2004.

[465] Hadi Jamali Rad, Alon Amar, and Geert Leus. Cooperative mobile network localization via subspace tracking. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2612–2615. IEEE, 2011.

[466] Ty Nguyen, Kartik Mohta, Camillo J Taylor, and Vijay Kumar. Vision-based multi-mav localization with anonymous relative measurements using coupled probabilistic data association filter. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3349–3355. IEEE, 2020.

[467] Viktor Walter, Nicolas Staub, Antonio Franchi, and Martin Saska. Uvdar system for visual relative localization with application to leader–follower formations of multirotor uavs. *IEEE Robotics and Automation Letters*, 4(3):2637–2644, 2019.

[468] Thomas Ziegler, Marco Karrer, Patrik Schmuck, and Margarita Chli. Distributed formation estimation via pairwise distance measurements. *IEEE Robotics and Automation Letters*, 6(2):3017–3024, 2021.

[469] Nan Bai, Yuan Tian, Ye Liu, Zhengxi Yuan, Zhuoling Xiao, and Jun Zhou. A high-precision and low-cost imu-based indoor pedestrian positioning technique. *IEEE Sensors Journal*, 20(12):6716–6726, 2020.

[470] Zhiqiang Cao, Ran Liu, Chau Yuen, Achala Athukorala, Benny Kai Kiat Ng, Muraleetharan Mathanraj, and U-Xuan Tan. Relative localization of mobile robots with multiple ultra-wideband ranging measurements. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5857–5863. IEEE, 2021.

[471] Kexin Guo, Xiuxian Li, and Lihua Xie. Ultra-wideband and odometry-based cooperative relative localization with application to multi-uav formation control. *IEEE transactions on cybernetics*, 50(6):2590–2603, 2019.

[472] Kexin Guo, Zhirong Qiu, Wei Meng, Lihua Xie, and Rodney Teo. Ultra-wideband based cooperative relative localization algorithm and experiments for multiple unmanned aerial vehicles in gps denied environments. *International Journal of Micro Air Vehicles*, 9(3):169–186, 2017.

[473] Ming Li, Zhuang Chang, Zhen Zhong, and Yan Gao. Relative localization in multi-robot systems based on dead reckoning and uwb ranging. In *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*, pages 1–7. IEEE, 2020.

[474] Hao Xu, Luqi Wang, Yichen Zhang, Kejie Qiu, and Shaojie Shen. Decentralized visual-inertial-uwb fusion for relative state estimation of aerial swarm. In *2020 IEEE international conference on robotics and automation (ICRA)*, pages 8776–8782. IEEE, 2020.

[475] Hao Xu, Yichen Zhang, Boyu Zhou, Luqi Wang, Xinjie Yao, Guotao Meng, and Shaojie Shen. Omni-swarm: A decentralized omnidirectional visual–inertial–uwb state estimation system for aerial swarms. *IEEE Transactions on Robotics*, 38(6):3374–3394, 2022.

[476] Bernhard Großwindhager, Michael Rath, Josef Kulmer, Mustafa S Bakr, Carlo Alberto Boano, Klaus Witrisal, and Kay Römer. Salma: Uwb-based single-anchor localization system using multipath assistance. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*, pages 132–144, 2018.

[477] Frank Hartmann, Felix Pistorius, Andreas Lauber, Kai Hildenbrand, Jürgen Becker, and Wilhelm Stork. Design of an embedded uwb hardware platform for navigation in gps denied environments. In *2015 IEEE Symposium on Communications and Vehicular Technology in the Benelux (SCVT)*, pages 1–6. IEEE, 2015.

[478] Benjamin Kempke, Pat Pannuto, Bradford Campbell, and Prabal Dutta. Surepoint: Exploiting ultra wideband flooding and diversity to provide robust, scalable, high-fidelity indoor localization. In *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM*, pages 137–149, 2016.

[479] Bernhard Großwindhager, Michael Stocker, Michael Rath, Carlo Alberto Boano, and Kay Römer. Snaploc: An ultra-fast uwb-based indoor localization system for an unlimited number of tags. In *Proceedings of the 18th International Conference on Information Processing in Sensor Networks*, pages 61–72, 2019.

[480] Aditya Arun, Shunsuke Saruwatari, Sureel Shah, and Dinesh Bharadia. Xrloc: Accurate uwb localization to realize xr deployments. In *Proceedings of the 21st ACM Conference on Embedded Networked Sensor Systems*, pages 459–473, 2023.

[481] Dana Gunders and Jonathan Bloom. *Wasted: How America is losing up to 40 percent of its food from farm to fork to landfill*. 2017.

[482] The state of food and agriculture 2019. https://www.fao.org/state-of-food-agriculture/2019/en/, 2019.

[483] John W. Palmer, F. Roger Harker, D. Stuart Tustin, and Jason Johnston. Fruit dry matter concentration: a new quality metric for apples. *Journal of the Science of Food and Agriculture*, 90(15):2586–2594, 2010.

[484] The state of food and agriculture. https://www.fao.org/publications/home/fao-flagship-publications/the-state-of-food-and-agriculture/.

[485] Steven K. MacLeod. Moisture determination using karl fischer titrations. *Analytical Chemistry*, 63(10):557A–566A, 1991.

[486] Matteo Rizzo, Matteo Marcuzzo, Alessandro Zangari, Andrea Gasparetto, and Andrea Albarelli. Fruit ripeness classification: A survey. *Artificial Intelligence in Agriculture*, 7:44–57, 2023.

[487] F-750 produce quality meter. https://felixinstruments.com/.

[488] Cfr - code of federal regulations title 21. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=101&showFR=1.

[489] Alessio Scalisi and Mark Glenn O'Connell. Relationships between soluble solids and dry matter in the flesh of stone fruit at harvest. *Analytica*, 2(1):14–24, 2021.

[490] John S. Toll. Causality and the dispersion relation: logical foundations. *Physical Review*, 104(6):1760, 1956.

[491] Udo Kaatze. Complex permittivity of water as a function of frequency and temperature. *Journal of Chemical and Engineering Data*, 34(4):371–374, 1989.

[492] R. Pedreschi, E. Ponce, I. Hernández, C. Fuentealba, A. Urbina, J.J. González-Fernández, J.I. Hormaza, D. Campos, R. Chirinos, and E. Aguayo. Short vs. long-distance avocado supply chains: Life cycle assessment impact associated to transport and effect of fruit origin and supply conditions chain on primary and secondary metabolites. *Foods*, 11(12):1807, 2022.