Microphone Array Backscatter: An Application-Driven Design for Lightweight Spatial Sound Recording over the Air

Jia Zhao
Simon Fraser University
BC, Canada
zhaojiaz@sfu.ca

Wei Gong
University of Science and Technology of China
Anhui, China
weigong@ustc.edu.cn

Jiangchuan Liu
Simon Fraser University
BC, Canada
jcliu@cs.sfu.ca

ABSTRACT

Modern acoustic wearables with microphone arrays are promising to offer rich experience (e.g., 360° sound and acoustic imaging) to consumers. Realtime multi-track audio streaming with precise synchronization however poses significant challenges to the existing wireless microphone array designs that depend on complex digital synchronization as well as bulky and power-hungry hardware.

This paper presents a novel microphone array sensor architecture that enables synchronous concurrent transmission of multi-track audio signals using analog backscatter communication. We develop novel Pulse Position Modulation (PPM) and Differential Pulse Position Modulation (DPPM) baseband circuits that can generate a spectral-efficient, time-multiplexing, and multi-track-synchronous baseband signal for backscattering. Its lightweight analog synchronization supports parallel multimedia signals without using any ADCs, DSPs, codecs and RF transceivers, hence largely reducing the complexity, latency, and power consumption. To further enhance self-sustainability, we also design an energy harvester that can extract energy from both sound and RF. We have built a microphone array backscatter sensor prototype using an FPGA, discrete components, and analog devices. Our experiments demonstrate a communication range (sensor-to-reader) of up to 28 meters for 8 audio tracks, and an equivalent throughput of up to 6.4 Mbps with a sample rate over 48KHz. Our sensor achieves 87.4μs of streaming latency for 4 tracks, which is 650x improvement as compared with digital solutions. ASIC design results show that it consumes as low as 175.2μW of power. Three sample applications including an acoustic imaging system, a beamform filter, and a voice control system, all built with our phased-array microphone, further demonstrate the applicability of our design.

CCS CONCEPTS

• **Networks** → **Cyber-physical networks**; **Network architectures**; **Sensor networks**.

ACM MobiCom '21, January 31-February 4, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-8342-4/22/01…$15.00
https://doi.org/10.1145/3447993.3483265

KEYWORDS

Microphone Array, Backscatter, Spatial Sound, Internet of Things

ACM Reference Format:
Jia Zhao, Wei Gong, and Jiangchuan Liu. 2022. Microphone Array Backscatter: An Application-Driven Design for Lightweight Spatial Sound Recording over the Air. In *The 27th Annual International Conference on Mobile Computing and Networking (ACM MobiCom '21), January 31-February 4, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3447993.3483265

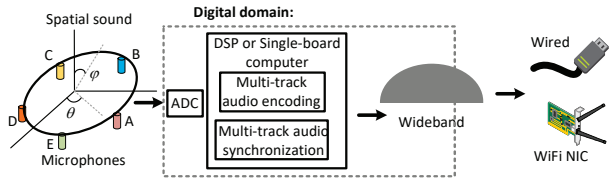
1 INTRODUCTION

Advanced audio sensing and acoustic wearable technologies have been long awaited by mobile consumers [37–45, 51], particularly the spatial sound field technology that offers high-dimension acoustic experience with a broad spectrum of applications (e.g., 3D sound in AR/VR, sound source localization, acoustic camera, acoustic diagnosis, machine condition monitoring, etc.) [52, 53]. Streaming service platforms start supporting spatial sound experience as well. For example, Apple Music has recently added an immersive experience feature with spatial audio in the updated iOS devices [60, 61].

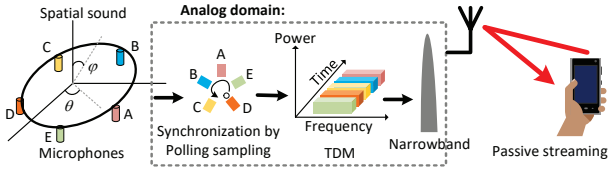
Spatial sound recording typically employs a microphone array to simultaneously record multi-track audio signals in different directions, so as to reconstruct the sound field in a space. It is known that only with highly precise synchronization will the time-domain waveform differences between the microphone-array elements be accurately derived, which is the key information to calculate the sound source position and strength distribution [41, 47, 48]. There have been significant related studies on sound synchronization in the past; yet their evolution towards consumer wearable devices that are of small form factor and low power is full of challenges.

Digital synchronization for multi-track audio. State-of-the-art microphone array products [54, 55] are usually integrated in or connected to PCs or smart devices (e.g., smart speakers), working largely as a sound acquisition frontend only. As shown in Figure 1(a), multi-track synchronization for them is completely in the digital domain, which relies on embedded synchronization algorithms or even a dedicated single-board computer to run a multi-track audio SDK [57–59]. The computational overhead, latency, and transmitted data amount are all very large. Together with such components as ADC, DSP, codec, and amplifier, the hardware complexity and power consumption are generally high.

Synchronization in wireless microphone array. Spatial sound recording over wireless is even more complicated and existing solutions are often heavy weighted. For example, a microphone-array beamforming system in [41] uses multiple smartphones and their WiFi/Cellular interfaces to streaming the audio data to a server



(a) Digital way: digitally synchronizing and wideband active transmission



(b) Our analog parallel streaming suitable for passive transmission

Figure 1: Comparison of spatial sound streaming solutions: (a) the digital way generates large computational overhead and high bandwidth demand; (b) we innovatively combines multi-track sound synchronization with a time-multiplexing modulation in an analog way, hence obtaining narrowband analog data that is suitable for efficient passive transmission.

for synchronization and processing, and large-scale acoustic array sensing in the wild has to use GPS for synchronization[48].

As such, a lightweight wireless microphone array device has to deal with many tradeoffs among digital modules, spectral efficiency, signal quality, latency, and power consumption. State-of-the-art backscatter technologies [1–8, 10–19, 21–25, 29–33] are promising for building ultra low-power wireless sensors. However, advanced applications are still difficult to be built over them for the following reasons:

1) Protocol-oriented backscatter. Most of previous systems, such as [2–8, 12, 13, 16–19, 22, 23], focus on how to design a backscatter baseband to work with ambient excitations using different wireless technologies (e.g., TV/Cellular/WiFi/Bluetooth/ZigBee/LoRa/FM). This is not a complete wireless sensor solution because it does not take the complexity of data acquisition into account and still needs extra power-hungry digital modules (external sensor, ADC, DSP, codec, etc.) to do advanced multimedia sensing.

2) Tradeoff between complex application and lightweight analog modules. As shown in Table 1, uncompressed multi-track audio bitrate can be of multiple Mbps, which is difficult to be supported by a backscatter without dedicated digital modules in Figure 1(a). Although there have been analog backscatter systems for audio/video communication [18, 25, 33], they only deal with single-track streaming and cannot extend to multi-track concurrent streaming, which needs a deliberate synchronization framework.

3) Power consumption. Given the existence of multiple tracks and the higher throughput demands, the overall power consumption may increase as well, challenging a wearable design.

In this paper, we present a novel microphone array sensor architecture that enables synchronous concurrent transmission of multi-track audio signals using analog backscatter communication. To the best of our knowledge, this is the first backscatter system that deals with the synchronization among concurrent high bit rate

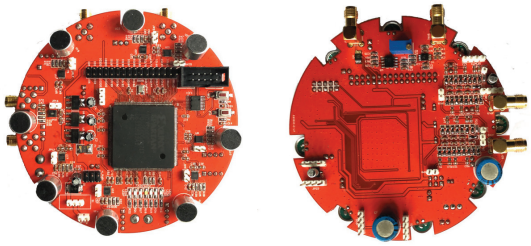


Figure 2: PCB of our microphone array backscatter.

Table 1: Uncompressed multi-track audio bit rate = Sampling rate × Track count × 16 bits (CD bit depth).

Track count	Bit rate for different sampling rates		
	22.05 KHz	44.1 KHz	48 KHz
2	0.71 Mbps	1.41 Mbps	1.54 Mbps
4	1.41 Mbps	2.82 Mbps	3.07 Mbps
6	2.12 Mbps	4.23 Mbps	4.61 Mbps
8	2.82 Mbps	5.64 Mbps	6.14 Mbps

multimedia transmissions, with lightweight design and low power consumption that are suitable for wearables. As shown in Figure 1(b), the key idea is an application-driven design that transforms the microphone array synchronization function entirely from a digital domain into an analog domain, and combines it together with backscatter baseband design. Specifically, our main technical designs can be summarized as follows:

- **Customized PPM and DPPM solutions for multi-track synchronous concurrent transmission.** We design customized pulse position modulation (PPM) and differential PPM (DPPM) to generate an analog baseband, in which each pulse position represents a realtime analog captured from an audio track. It enables a time-division multiplexing that employs roll polling access for synchronization among multiple tracks.
- **Passive transmission.** We design and implement a phased-array microphone backscatter system to support high-quality multi-track sound streaming. The backscatter sensor generates a narrowband baseband signal that fits for phase modulation on excitation signals and carries necessary information of optional sampling rates for sound waveform recovering on every single track.
- **Power analysis.** Our backscatter sensor largely reduces power consumption by eliminating digital signal processing related components and RF transceivers. We have used ASIC design to analyze the power consumption and discuss its opportunities for miniaturizing our design in future work.

To verify our design, we have built a microphone array backscatter sensor prototype using an FPGA, discrete components, and analog devices. In the realworld experiments, the prototype demonstrates a communication range (sensor-to-sink) of up to 28m for 8 audio tracks, and an equivalent throughput of up to 6.4Mbps with over 48KHz sample rate. Our sensor achieves 87.4μs of streaming latency for 4 tracks, which is 650x improvement as compared with a digital solution (using sound card, streaming software, and WiFi/Bluetooth adapter). We also optimize the power consumption

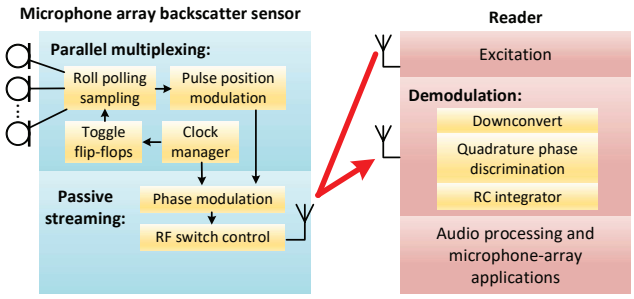


Figure 3: Overview of our microphone array backscatter communication system using passive parallel streaming for synchronous multi-track sound.

using ASIC design. Our results show that the ASIC design of 8-track 48KHz backscatter sensor consumes as low as 175.2μW of power. Finally, three sample applications including an acoustic imaging system, a beamform filter, and a voice control system, all built with our phased-array microphone, further demonstrate the applicability of our design. We also discuss the limitations of modulating multi-track audio in an analog way, including: 1) limited communication range because decoding analog audio data with an RC integrator can impact the fidelity of the recovered audio; 2) pulse position modulation needs synchronization for each sampling cycle, which can potentially increase power consumption on the actively generated synchronization clock; 3) we need a specialized reader to generate tone-excitation and demodulate backscatter signals, which cannot be replaced with a commodity radio in our current implementation.

Our main contributions in this paper are as follows:

- We design, implement and evaluate an application-driven backscatter to demonstrate that it is possible to transform some large-overhead application-related processing tasks (e.g., microphone array synchronization) from a digital domain into a lightweight analog domain, and even combines them with backscatter baseband design. This not only improves efficiency (e.g., reduced data amount, low latency, and low power) but also provides a new thinking that differs from previous protocol-oriented backscatter.
- We explore possible sensor solutions for fully passive multi-track multimedia streaming, potentially to empower ubiquitous self-sustainable 3D sensing and immersive applications.

2 DESIGN

We aim to have spatial sound data acquisition, necessary signal processing, and wireless transmission together in a lightweight architecture for integration. To achieve this, we combine multi-track audio multiplexing with backscatter baseband, and make them completely in the analog domain for accurate multi-track synchronization and low streaming latency. Figure 3 depicts an architectural view of our microphone array backscatter sensor. In this section, we introduce the details of its key modules, including our PPM and DPPM circuits, multipath interference reduction scheme, signal reflection, and demodulation.

2.1 Pulse Position Modulation

2.1.1 Channel Multiplexing for Synchronous Concurrent Transmission. As shown in Table 1, multi-track audio signals have a high aggregated bitrate in the uncompressed condition. Since we do not use digital compression components, we design the concurrent transmission in an analog way, so as to reduce transmitted data amount but retain necessary information for sound waveform recovering.

Our concurrent transmission solution does not use parallel wireless channels because it is difficult to perform multi-channel synchronization. From spectral efficiency perspective, this also avoids over-occupying spectrum resource. When using a single channel, we need a multiplexing mechanism to reduce channel contention.

Our channel-multiplexing solution is inspired by the remote control of model aircrafts. A aeromodelling controller has multiple dimension control commands to be sent, which are all time-variant analog information, such as height, direction, speed, etc. The controller uses the pulse position to multiplex multiple command signals into a single wireless channel. In the pulse position modulation [49, 50], the analog information is presented by a narrow width pulse position relative to the start position of each sampling period or to its previous pulse position. To deal with synchronous concurrent transmission, we need a customized baseband to perform pulse position modulation.

Our PPM can be illustrated with an example circuit in Figure 4. The circuit first uses an analog multiplexer for the multiple access to a single channel, and then uses a comparator and a periodic triangle waveform to perform analog sampling. During each sampling period, the comparator outputs a pulse width that represents instantaneous analog information of the current access track.

At the end of each sampling period, the circuit uses an inverter and a monostable multivibrator to generate the pulse position. Monostable multivibrator is a well studied circuit component. In Figures 4 and 5, we use a simplified component sketch of the monostable multivibrator and let it represent the functionality that converts a pulse width to a pulse position. Its detailed implementation is introduced in Section 3.1. The pulse position is relative to the start position of its associated period (e.g., the pulse position of track 2 in Figure 4). The monostable multivibrator outputs a fixed-length narrow pulse that marks the end position of each input pulse width.

2.1.2 Roll Polling Control. We use roll polling access to synchronize multiple tracks for multiplexing. The polling is controlled by a synchronization clock signal, which generates both multiplexer input selection signals and sampling signals. As illustrated in Figure 4, the selection signals are generated from the synchronization signal with a truth table.

For each roll polling cycle, the first sampling period is reserved as a synchronization blank without generating a pulse. The synchronization clock frequency f_{syn} is related to the sound quality (sample rate f_a) and the number of tracks N. We have the following equation to calculate the frequency of the synchronization clock.

$$f_{syn} = f_a \times (N + 1) \quad (1)$$

Take three typical sample rates 22.05KHz, 44.1KHz and 48KHz as examples. 22.05KHz is used in FM radio broadcasting, for which our synchronization clock frequency increases from 66.15KHz to

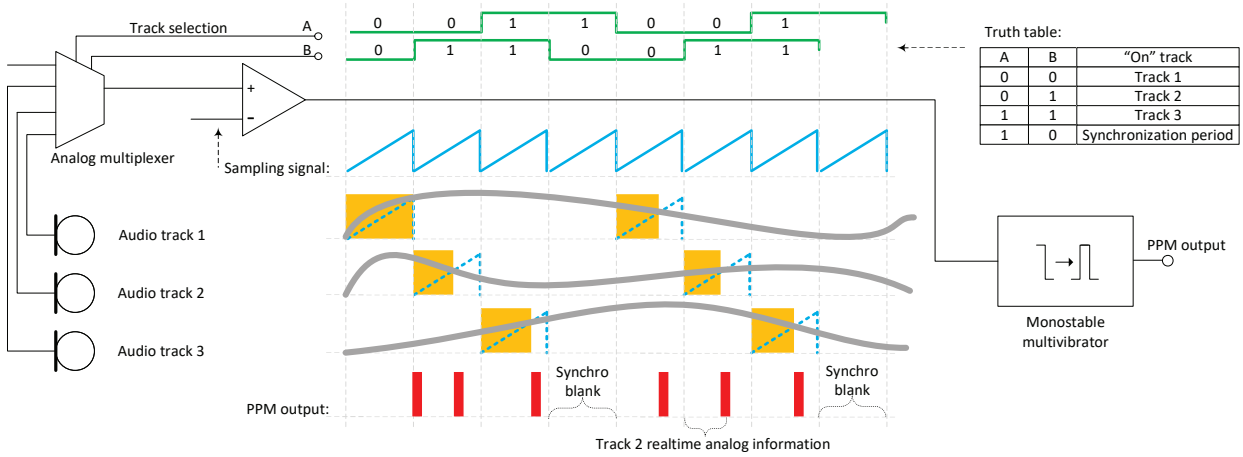


Figure 4: Our pulse position modulator that allows multi-track audio signals to multiplex a single channel for concurrent synchronous transmission.

198.45KHz when the track number increases from 2 to 8. CD and MP3 audio formats commonly use a sample rate of 44.1KHz, for which we need a 396.9KHz synchronization signal to transmit 8 tracks. Finally, 48KHz is used for high-definition audio/video playback, which, if having 8 tracks, needs a 432KHz signal.

2.2 Differential Pulse Position Modulation

Our baseband circuit for differential pulse position modulation uses pulse intervals to represent the analog information of each track in its sampling duration. As discussed in Subsection 2.1, PPM has a synchronization signal at the sensor side, which generates the multiplexer input selection signals and decides the start position of each sampling period; At the reader side, the demodulation of PPM also needs the synchronization signal to recover the original analog information. In contrast, our DPPM circuit has no synchronization control and uses only analog components, cascade connections, and feedbacks. In DPPM, the end of the current track sampling will trigger the start of the next track sampling. For each roll polling cycle, there is a feedback from the last track to the first track, which triggers the next roll polling cycle.

Figure 5 shows an example of a 3-stage DPPM baseband circuit. The circuitry uses comparators, RC calculation circuits, inverters, monostable multivibrators, switches, and an OR gate. For each track, we use two different monostable multivibrators. The output monostable multivibrator can generate a narrow pulse given a negative edge. The output monostable multivibrator marks the end position of a sampling duration for each track. The sampling monostable multivibrator can generate a trigger pulse of a fixed width that represents the maximum sampling duration for each track. A trigger pulse generates a triangular pulse after passing an RC calculation circuit, and also switches on the audio track in the same stage. When sampling ends, the comparator will output a negative edge, which is used to trigger the output monostable multivibrator. The sampling duration t_s and the sampling monostable multivibrator pulse width T_{smm} satisfy $t_s \leq T_{smm}$.

T_{smm} depends on its discrete component parameters. T_{smm} decides the lower bound of the sample rate for each track. Give N tracks, the sample rate f_a for each track satisfies the following equation:

$$f_a \geq \frac{1}{N \times T_{smm}}. \tag{2}$$

For 8 tracks audio quality higher than FM radio (sample rate 22.05KHz), we have $T_{smm} < 5.67\mu s$. To achieve 8 tracks audio quality higher than CD (sample rate 44.1KHz), we have $T_{smm} < 2.83\mu s$. For 8 tracks audio quality higher than DVD (sample rate 48KHz), we have $T_{smm} < 2.60\mu s$.

It is worth noting that DPPM and PPM have their respective pros and cons, and hence different application contexts. In particular, PPM has accurate sampling rate control. According to Equation (1), given the synchronization signal frequency f_{syn} and N tracks, each track's sample rate is fixed to $f_{syn}/(N+1)$. In addition, the synchronization blanks of PPM provide an accurate synchronization for each roll polling cycle. In contrast, DPPM cannot accurately control the sample rate for each track, but needs no synchronization signal from both the sensor side and the reader side.

2.3 RF Signal Reflection

Let $f_s = 10$MHz and $\omega_s = 2\pi f_s$. A square wave with frequency f_s can be represented as $M(t) = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{sin((2n-1)\omega_s t)}{2n-1}$. The PPM or DPPM baseband signal modulates the square wave signal $M(t)$ to obtain a frequency shift signal $S(t)$. $S(t)$ is an on-off signal that can change antenna impedance to reflect or absorb an excitation signal. We use the baseband signal to generate a time-variant phase $\varphi(t)$ that can change the phase of $M(t)$. Let t_i be the ith pulse position. $\varphi(t)$ can be formulated as

$$\varphi(t) = \frac{\pi}{2} + (i-1)\pi, \quad (t_i < t < t_{i+1}). \tag{3}$$

The phase modulation process from the baseband to the frequency shift signal can be expressed as follows

$$S(t) = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{sin\Big((2n-1)\omega_s t + \varphi(t)\Big)}{2n-1}. \tag{4}$$

$S(t)$ is used to control the antenna impedance for backscattering an RF excitation signal. Subsection 3.2.2 implements it with digital

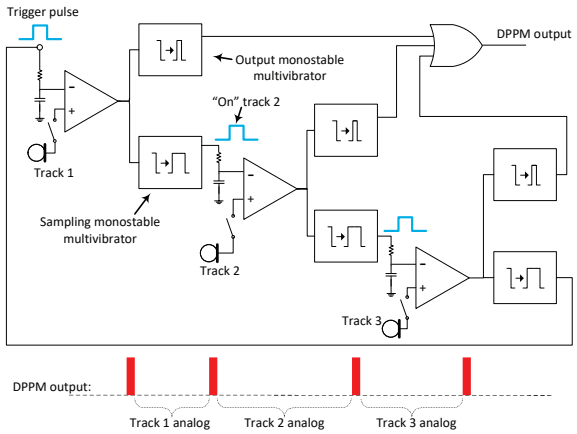
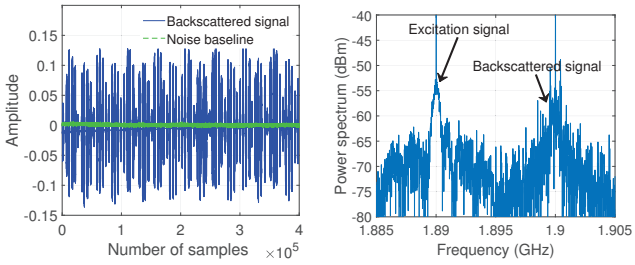


Figure 5: Design of differential pulse position modulator. It uses the delay between two pulses to represent a track's analog information, without using any synchronization clock.



(a) Time domain (b) Frequency domain

Figure 6: Backscattered signal in time and frequency domains.

logic circuits. Figure 6 shows the backscattered signal in time and frequency domains.

2.4 Demodulation

As introduced in Subsection 2.1, 2.2 and 2.4, the sensor side first generates a time-multiplexing PPM or DPPM baseband for multi-track audio signals, then modulates the baseband on a 10MHz frequency shift signal that changes antenna impedance, and finally uses RF reflection to carry the modulated baseband on a high-frequency excitation signal. Therefore, the reader side also needs three corresponding steps to demodulate each track's signal, as illustrated in Figure 7.

The first step is to obtain the modulated 10MHz frequency shift signal by removing the high-frequency excitation carrier. Since the backscattering is a linear modulation, we can use coherent demodulation to remove the carrier. As shown in Figure 7, the backscattered signal is received at the frequency with a 10MHz offset from the carrier central frequency using a band pass filter. We multiply the band pass output by a coherent signal with the same frequency and phase as the carrier. We then use a low pass filter with a cutoff frequency larger than 10MHz to extract the low-frequencies that represent the frequency shift signal.

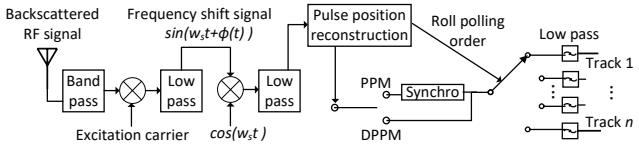


Figure 7: Demodulation operations.

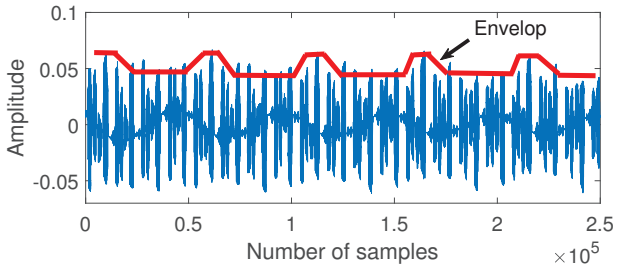


Figure 8: Envelope of demodulated frequency shift singal, which is a approximation to the signal $b(t)$.

The second step is to recover the PPM or DPPM baseband from the frequency shift signal. Equations (3) and (4) indicate that a frequency shift signal $S(t)$ conveys the pulse position information by changing the phase of a 10MHz square wave in real time. Since $S(t)$ is a phase modulation signal, we use the quadrature phase discrimination to demodulate the baseband. The associated operations are also included in Figure 7. The received frequency shift signal $S'(t)$ is the the fundamental component of $S(t)$, and expressed as

$$S'(t) = sin(\omega_s t + \varphi(t)) \tag{5}$$

where $\varphi(t) = \frac{\pi}{2} + (i-1)\pi$, $t_i < t < t_{i+1}$ and t_i is the ith pulse position. To demodulate $S'(t)$, we multiply $S'(t)$ with $cos(\omega_s t)$, i.e., a 10MHz signal with a $\frac{\pi}{2}$ phase difference. The output of the multiplier is passed through a low pass filter to remove the 10MHz sine signal. Let the low pass output be $b(t)$, which is equivalent to let $\omega_s = 0$ in Equation (6), and hence we have

$$b(t) = sin(\varphi(t)). \tag{6}$$

Since the function range of $b(t)$ equals to $\{-1, +1\}$, the envelope of $S'(t)$ is also a approximation to the waveform of $b(t)$, as shown in Figure 8. We obtain the PPM or DPPM baseband by taking all the positive and negative edges of $b(t)$.

The last step is to recover the waveform of each audio track. We assign the pulses in a PPM or DPPM signal to multiple tracks by following the roll polling order. Note that for PPM signal we need a synchronization signal to find the start position of each sampling period. Finally, we use a low pass filter (RC integrator) to recover the original sound waveform of each track. The principle and trace of using RC integrator are shown in Figure 9 and Figure 10, respectively.

3 IMPLEMENTATION

To verify our design, we have built a backscatter sensor prototype (PCB shown in Figure 2) using an FPGA, discrete components, and analog devices. To further optimize the power consumption with

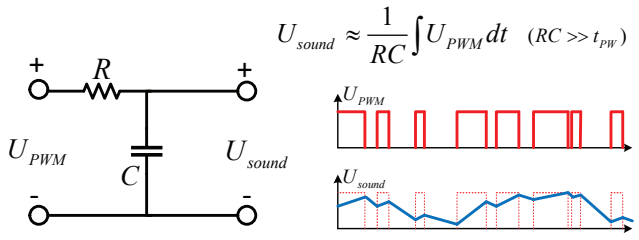


Figure 9: Principle of using RC integrator to convert PWM signal into time-domain waveform.

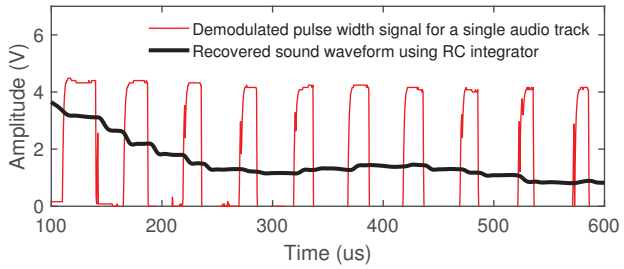


Figure 10: A trace of audio waveform recovering process using RC integrator.

advanced IC manufacturing technology, we conduct an Application-Specific Integrated Circuit (ASIC) design. In this section, we introduce these implementation details, as well as those for our reader and energy harvester, followed by evaluation in the next section.

3.1 Analog Modulation Circuits

The analog modulation circuits for PPM and DPPM follow Figure 4 and Figure 5, respectively. We introduce the implementation of key components as follows.

Monostable multivibrator. Monostable multivibrator is an essential component in our system, which converts pulse width to pulse position. In our analog modulation circuits, the PPM baseband uses one monostable multivibrator, while the DPPM baseband uses multiple monostable multivibrators. The classic discrete-component-based monostable multivibrator involves complicated circuits and consumes high power. As shown in Figure 11, we have two simplified implementations, both of which can serve our goal. Implementation I uses only two resistors, a capacitor, and a NAND gate. We can see that it can generate a narrow pulse immediately following a negative edge of the input pulse. Implementation II is a further simplified approach that only uses an inverter, an RC differential circuit and a diode. It does not generate a rectangular pulse; instead, it can output a positive edge if there is a pulse negative edge. PPM or DPPM can use these pulse edges to modulate data. Implementation II incurs very low power and can be easily realized in both prototype PCB and ASIC design.

Multiplexing and sampling. We use a CMOS analog multiplexer CD4051 to allow multi-track analog signals to multiplex one channel. The binary control inputs of CD4051 are connected to a roll polling control circuit using an FPGA, which will be introduced in Subsection 3.2. The sampling process uses an RC integral circuit and a high-speed voltage Comparator TLV3501.

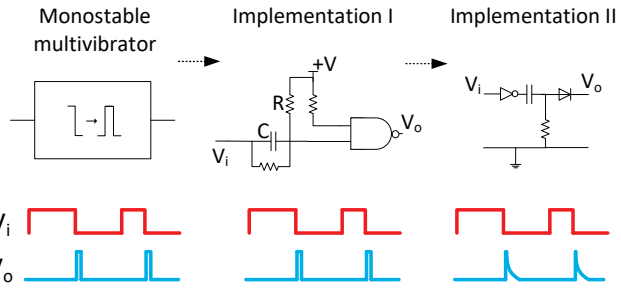


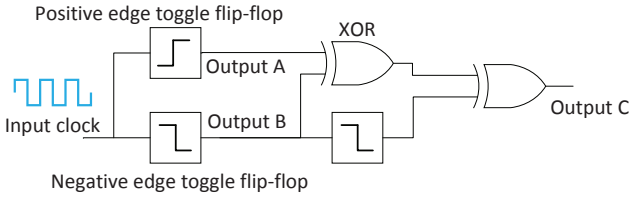
Figure 11: Monostable multivibrator implementation.

3.2 Digital Circuits

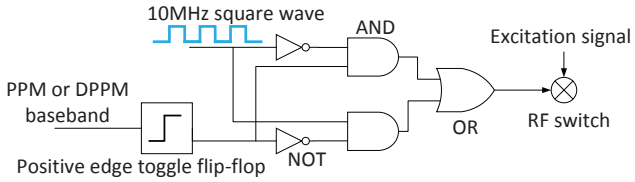
We use a XILINX Spartan XC3S500E-4PQ208 FPGA to implement the key modules that include the synchronization clock, the roll polling logic, and the phase modulation on the frequency shift signal. All the digital modules are written in Verilog. For the synchronization clock, we follow the design in Subsection 2.1 and use the FPGA's clock frequency division to generate synchronization signals for different sample rates and number of tracks. For the other two modules, we have the following detailed implementations.

3.2.1 Roll Polling Control Circuit. Following the design in Figure 4, we need a digital logic to control the analog multiplexer. This control logic should not only enable a fixed "On" period for each multiplexer input, but also allow the inputs to multiplex the output channel in an orderly manner. This roll polling logic is implemented with an FPGA circuit that can use a synchronization clock signal to generate the multiplexer's control input signals. We configure the synchronization clock period T by following the sampling rate analysis in Subsection 2.1. The synchronization clock ensures that all the channels of the multiplexer have the same "On" state period. We then use the circuit in Figure 12(a) to generate the multiplexer control signals. We first take a 2-input 4-channel multiplexer as an example. The Output A and Output B in Figure 12(a) should follow the truth table in Figure 4. From the truth table in Figure 4, we can see that the two control signals have the same period $2T$ but have a 90 degree phase difference. We use two edge toggle flip-flop modules to obtain the two control signals. The positive edge toggle flip-flop uses each positive edge of the input clock to trigger its output switching between 0 and 1, while the negative edge toggle flip-flop uses each negative edge of the input clock to perform the toggle. Therefore, there is a 90 degree phase difference between Output A and Output B, and they can be directly connected to the two control inputs of a 4-channel multiplexer. For a 3-input 8-channel multiplexer, we can still use Output A and Output B as two control inputs, and the last control input is connected to Output C in Figure 12(a). To generate Output C, we connect Output A and Output B to a circuit that consists of two XOR gates and the positive edge toggle flip-flop.

3.2.2 Phase Modulation Circuit. Equations (3) and (4) demonstrate the conversion from a PPM or DPPM baseband to a frequency shift signal. We perform the conversion using phase modulation. The formulated solution is implemented with a practical phase modulation circuit, which is shown in Figure 12(b). The baseband signal is passed through a positive edge toggle flip-flop to generate a time-variant signal that switches between high and low voltage



(a) Roll polling control circuit.



(b) Phase modulation circuit.

Figure 12: FPGA modules that generate (a) roll polling control signal for the analog multiplexer; (b) frequency shift signal using phase modulation.

levels. High voltage corresponds to 0 degree phase, and low voltage corresponds to 180 degree phase. Then we uses a 10MHz square wave and its inversion, respectively, as the inputs of two AND gates. The two AND gates' remaining inputs are connected to the flip-flop output and its inversion, respectively. With these two AND gates, the high output of the flip-flop can switch on the 0 degree delay square wave and switch off the 180 degree delay square wave. The low output of the flip-flop has the opposite effect. The two AND gate outputs are connected to an OR gate. The output of the OR gate is the frequency shift signal that directly controls an ADG902 RF switch.

3.3 Miniaturization: from PCB to ASIC Design

While we have made efforts on reducing the power consumption with the analog components and simplified designs, as discussed above, it is known that further improvement can be made by customized ASIC (Application-Specific Integrated Circuit) design, an integrated circuit designing process that can transform a customized large scale circuit into a tiny working silicon. Compared with PCB-level implementation, ASIC can provide an integration that allows small size, low cost, high speed, reliability and very low power consumption, which are worthy of exploration.

Our backscatter sensor prototype includes the analog circuit modules and the digital circuit modules, both of which can be optimized through the ASIC design. For the analog circuits, we use the IC design software Cadence IC6.17 Virtuoso [64]. We have implemented our design in the Virtuoso analog design environment using a TSMC 65nm CMOS Low Power technology library [65], and performed power simulations for all the analog modules in PPM and DPPM baseband circuits. For the digital circuits (written in Verilog code), we use the synthesis tool XST and power analysis tool XPower [62] in Xilinx Integrated Synthesis Environment. The ASIC design will be closely examined in Section 4.

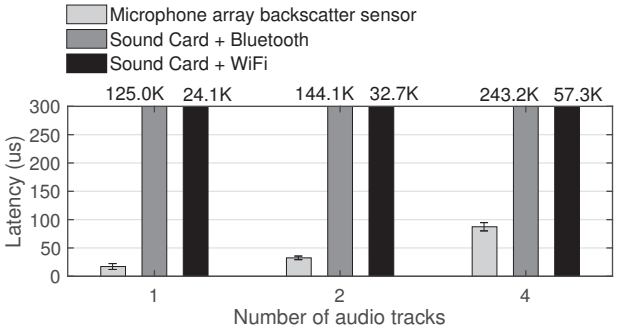


Figure 13: Latency on streaming.

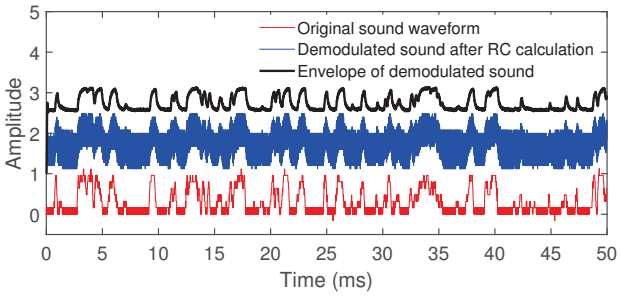


Figure 14: Fidelity analysis by comparing original sound wavefrom and demodualted sound waveform.

3.4 Reader

We implement the demodulation modules on a USRP N210 software-defined radio device, which is used as the reader. We use another USRP N210 device as the RF excitation source that can generate single-tone signals at 895MHz. The excitation uses an HA901I-APC RF amplifier and an RFMAX S9028PCR antenna with an 8.5 dBic gain. The reader uses an MTI MT-242025/TRH/A antenna with a 6.5 dBic gain.

4 EVALUATION

In this section, we evaluate our microphone array backscatter sensor from multiple aspects, including the fidelity of demodulated spatial sound, the backscattered signal quality over different sensor-to-reader and excitation-to-sensor distances, power consumption, and energy harvesting performance. We conduct both line-of-sight and non-line-of-sight experiments to evaluate the signal-to-noise ratio (SNR) and audio distortion. We also optimize the power consumption of our design by simulating its circuit modules in an ASIC design platform.

4.1 Latency

Latency on streaming. We first test the latency for audio data acquisition and transmission in a sensing IoT node. We compare our system with the regular solutions that use sound cards for quantization and general-purpose radios for transmission. We use Conexant 20751 sound cards, Realtek RTL8723BE WLAN NICs, and Realtek Bluetooth 4.0+ High Speed Chips. To measure the sound card latency, we write an audio recording program and track its running time for recording each audio frame. We calculate transmission delay using packet timestamps from traffic analysis. Our

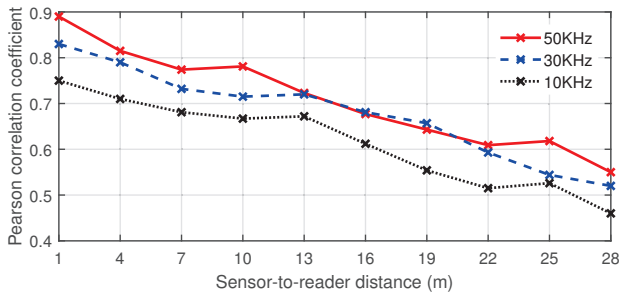


Figure 15: Correlation between demodulated sound waveform and original sound waveform.

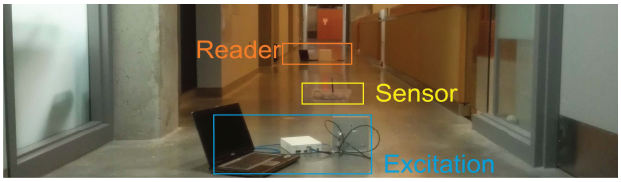


Figure 16: Line-of-sight setup to test communication range.

system's streaming latency is the time taken by a captured audio signal to pass through the PPM modulation circuit and the backscatter modulator. We use an oscilloscope to capture the time difference between audio analog input and baseband output. Figure 13 shows the streaming latency under different number of audio tracks. Our system achieves μs-level streaming latency (e.g., 17.2μs for a single track and 87.4μs for 4 tracks), while the other two solutions have ms-level latencies (e.g., 57.3ms for 4 tracks using sound card and WiFi, 243.2ms for 4 tracks using sound card and Bluetooth).

4.2 Fidelity of Demodulated Sound

Figure 14 plots an audio track's original and demodulated sound waveforms in 50ms. To examine the fidelity of the demodulated multi-track audio signals, we calculate the Pearson correlation coefficient of the original sound waveform and the envelope of the demodulated sound waveform. Figure 15 shows that the similarity between the two waveforms decreases with sensor-to-reader distance. This is due to the downward trend of the SNR over the distance. We also see that the similarity is improved by increasing analog baseband sample rate for each track. In Figure 7, the demodulation uses lowpass filters (RC calculation) to convert a PPM baseband into a time-domain waveform, which also brings distortion. Such a distortion is similar to the quantization error, which can be reduced by oversampling [36]. Therefore, a higher sample rate achieves better approximation to original waveform.

4.3 Sensor-to-Reader Distance

We examine the communication range in an indoor environment (as shown in Figure 16) with both line-of-sight and non-line-of-sight configurations. We use a USRP N210 device together with an HA901I-APC RF amplifier as the excitation source to generate an 895MHz single-tone signal for our backscatter sensor. We then use another USRP N210 device as the reader. At the beginning, the backscatter sensor is placed close to the source, and the reader is located near the sensor. We then move the reader to increase the distance between the sensor and the reader, and measure how

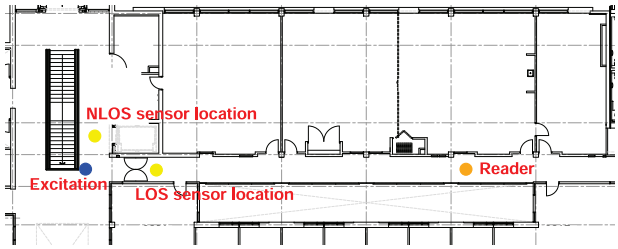
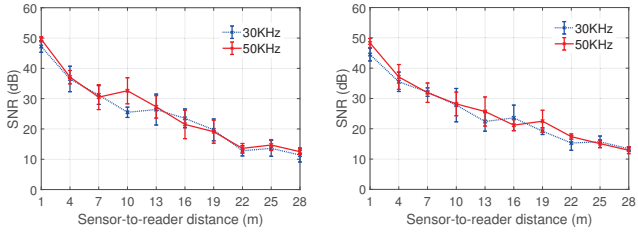


Figure 17: Floorplan of line-of-sight and none-line-of-sight deployments.



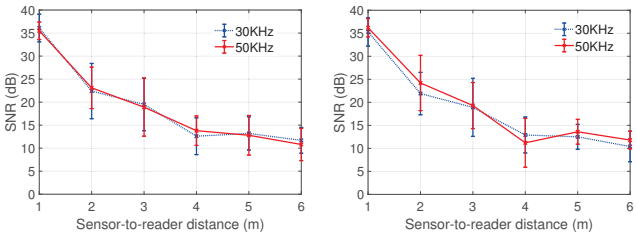
(a) 4 audio tracks (b) 8 audio tracks

Figure 18: Backscattered signal SNR versus sensor-to-reader distance in the line-of-sight experiment.

the backscattered signal's SNR changes with the distance, sample rate, and track count, respectively. While the type of baseband does not necessarily impact the the backscattered signal strength, our statistics include both PPM and DPPM results.

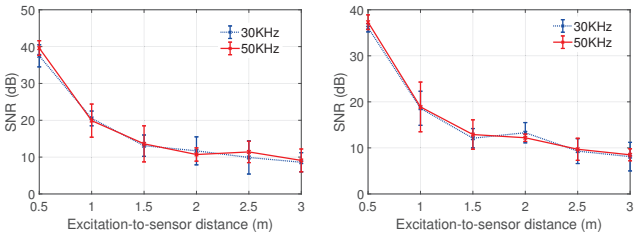
Figure 18 shows our results in the line-of-sight experiment. Figure 18(a) shows that the backscattered signal's SNR decreases with the distance between the sensor and the reader for 4-track audio. For a sample rate of 30KHz, the SNR is 47.2dB at a distance of 1m, and it decreases to 11.4dB when the sensor-to-reader distance increases to 28m. When the sample rate is increased to 50KHz, we can achieve an SNR of more than 45dB at 1m and an SNR of less than 15dB at 28m. When we increase audio track count to 8, the backscattered signal's SNR also shows a general downward trend as the sensor-to-read distance increases, and its dynamic range also falls between 10dB to 50dB, without a significant change as compared with the 4-track case. Since the data rate of an uncompressed multi-track audio equals sample rate × track count × ADC quantization depth, our reader can achieve an equivalent throughput of up to 6.4Mbps using a 16-bit ADC for 50KHz sample rate and 8 audio tracks.

The line-of-sight experiment results show that backscattered signal's SNR decreases with the sensor-to-reader distance, and sample rates and number of tracks do not noticeably impact the SNR during concurrent synchronous transmission. Here, the SNR during wireless transmission is different from the SNR for quantifying the dynamic performance of ADC devices. In digital signal processing, sampling at a higher frequency makes the quantization noise to distribute over a wide range of frequencies, thus reducing the density of noise power spectrum. Given a fixed number of bits for quantization in the ADC, oversampling can improve the signal to quantization noise ratio (SQNR) and effective number of bits (ENOB). Our design is an analog synchronous transmission solution. Therefore, we focus on the impact of analog noise during



(a) 4 audio tracks (b) 8 audio tracks

Figure 19: Backscattered signal SNR versus sensor-to-reader distance in the none-line-of-sight experiment.



(a) 4 audio tracks (b) 8 audio tracks

Figure 20: Received backscatter signal SNR versus excitation-to-sensor distance.

wireless transmission. While the sample rate and the number of tracks do not impact the backscattered signal's SNR, they actually determine the complexity of the PPM and DPPM baseband circuits, and directly impact the power consumption of the circuits, which will be analyzed in Subsection 4.2.

The non-line-of-sight floorplan is shown in Figure 17 and the experiment result is shown in Figure 19. The backscattered signal's SNR also decreases with the distance between the sensor and the reader. The communication range can be up to 6m. For a 30KHz sample rate and four audio tracks, we achieve an SNR over 35dB at a distance of 1m, and it decreases quickly to less than 15dB when the sensor-to-reader distance increases to 6m. For a 50KHz sample rate and four audio tracks, we have observed similar results. When we increase the track count to 8, SNR has a similar dynamic range between 10dB and 35dB.

4.4 Excitation-to-Sensor Distance

We also examine how the excitation-to-sensor distance impacts the backscattered signal quality. We place the reader with a fixed distance of 1m away from the backscatter sensor. Then we move the excitation to increase the excitation-to-reader distance. Figure 20 shows that the backscattered signal quality degrades across the excitation-to-sensor distance without being affected by the sample rate and audio track count. Our result suggests that the excitation source should fall within a range of 3m from the sensor in order to maintain the quality of backscatter communication.

4.5 Power Consumption

We next use an ASIC design to optimize the power consumption of our design. The implementation follows Subsection 3.3: The 10MHz oscillator consumes 142.3μW; The monostable multivibrator consumes 9.5μW; The analog sampling circuit consumes 13.4μW; The digital circuits written in Verilog code consume 10μW of power (Digital circuit power consumption in ASIC design is the Verilog code synthesis result [62], and it is irrelevant to the FPGA chip power consumption of the PCB implementation). Accordingly we can calculate the total power consumption of our backscatter sensor modulation circuits. When using the PPM baseband, the sensor has a analog sampling circuit, a monostable multivibrator, an oscillator and an digital circuit module. As shown in Table 2, the sensor using PPM baseband consumes a total power of 175.2μW. For the DPPM baseband, the sensor's power consumption changes with the number of audio tracks. For example, in Figure 5, the DPPM baseband for three tracks uses three analog sampling circuits, six monostable multivibrators, a oscillator and a digital circuit module. Therefore, the total power consumption of three tracks is 249.5μW. While the DPPM baseband circuit eliminates the synchronization clock as compared with PPM baseband, it is more complex and our results suggest that it indeed consumes more power, e.g., twice higher than PPM for 8 audio tracks. For a practical ASIC of microphone array backscatter, the choice between PPM and DPPM depends on the user requirements of power consumption and hardware complexity. First, the above results show that DPPM consumes more power than PPM and DPPM's power consumption increases with the number of audio tracks. PPM is a good choice if the goal is to enable as many as possible audio tracks with relatively low power. Second, DPPM is simpler than PPM in terms of hardware because it needs no clock to synchronize to every sampling period. Hence, DPPM is appropriate for less digital components.

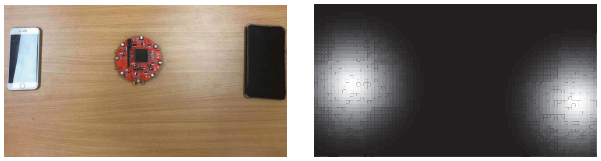
Table 2: Power consumption of our design.

Track count	Baseband circuits	
	PPM	DPPM
2	175.2 μW	217.1 μW
3	175.2 μW	249.5 μW
4	175.2 μW	281.9 μW
5	175.2 μW	314.3 μW
6	175.2 μW	346.7 μW
7	175.2 μW	379.1 μW
8	175.2 μW	411.5 μW

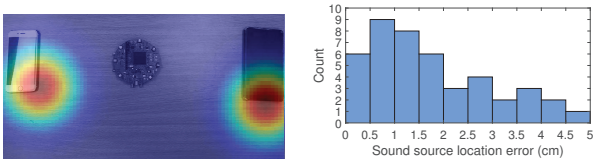
5 APPLICATION CASES

Our microphone array backscatter sensor can work on wearables and mobile devices to provide diverse acoustic and multimedia experience. To demonstrate its applicability, we have built prototypes for three representative scenarios.

Acoustic imaging. We have built an acoustic imaging system that can visually localize sound sources. Figure 21(a) shows the targeted area in which our phased-array microphone collects spatial sound signals from two smartphone speakers. The data is processed by a microphone array location algorithm, and the sound source positions are shown in Figure 21(b). Figure 21(c) visualizes the sound sources in the targeted area with jet color map. Figure 21(d) shows the distribution of location error across 44 trials in which about 66% of the trials fall within a 2cm error range.

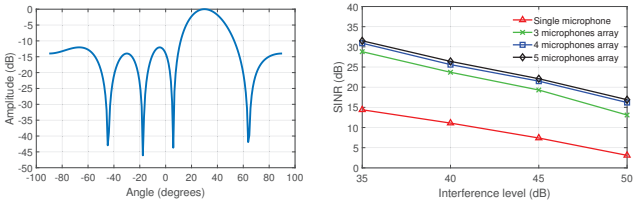


(a) Original data (b) Detected sound source location



(c) Visualization (d) Precision

Figure 21: Acoustic imaging using our system.



(a) 30° beamform filter design (b) Signal to interference-plus-noise

Figure 22: Beamforming and noise reduction for directional audio signal enhancement.

Spatial filtering and speech enhancement. Our phased-array microphone can also work as a directional signal reception device for speech enhancement. For example, we use five microphones and their relative positions to design a beamform filter with a 30° in azimuth angle as shown in Figure 22(a). Given that a directional audio signal comes from 30° in azimuth and 0° in elevation and an interference signal comes from -30° in azimuth and 0° in elevation, the result in Figure 22(b) demonstrates that our phased-array microphone can improve signal-to-interference-plus-noise ratio (SINR) by up to 15dB, as compared to a single microphone. To evaluate the dynamic range of the received audio signals, we introduce peak signal-to-noise ratio (PSNR) and test it across multiple audio tracks. Our results show that the PSNR of our speech system decreases with the distance between our microphone array sensor and the sound source, and it is up to 50dB when the distance is less than 1m. Our system's potential to support immersive applications or work as wearable devices is associated with the quality of user experience. To demonstrate it, we introduce a subjective assessment of single- or multi-track audio quality by following the ITU's Recommendation [56]. We use a unipolar discrete five-grade scale (1: Bad, 2: Poor, 3: Fair, 4: Good, and 5: Excellent) and let the listeners compare the received audio with the reference to value the quality. We test single-track audio with multiple attributes such as transparency, timbre, echoes, harmonic distortions, quantization noise, and background noise. The average quality score is 3.62 for single-track audio. We also examine the multi-track audio quality, which mainly focuses on the subjective impression of stereophonic or

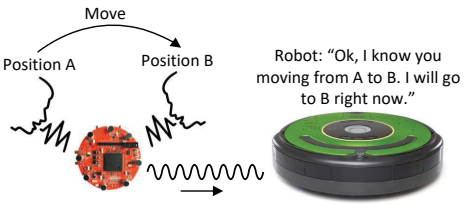
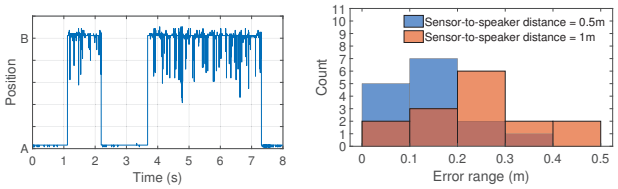


Figure 23: A voice control application. Cleaning robot can identify speaker position with two audio tracks information sent by our microphone array sensor.



(a) Speaker position change (b) Location errors

Figure 24: Speaker position identified by robot.

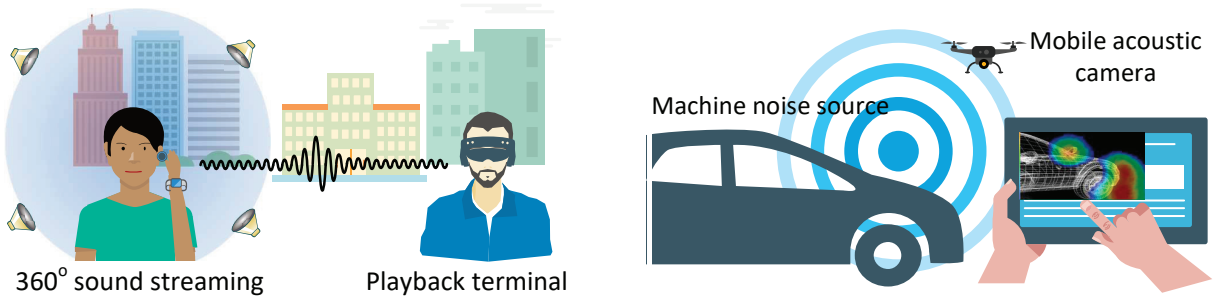
spatial effects (e.g., the relative position of multiple sound sources). The multi-track audio achieves an average score of 3.37.

Localization and voice control. We have built a voice control system that enables a commodity robot to identify and move towards a person's position if the person makes a loud voice at that position. The system includes a microphone array backscatter sensor, two backscatter readers, a PC for audio processing, and an off-the-shelf cleaning robot (iRobot Create 2 with Raspberry Pi). We use two microphones on our backscatter sensor. These two microphones have a fixed direction difference. These two tracks realtime audio signals can be sent to a reader using our proposed analog synchronous concurrent transmission. The reader then demodulates the two-track signals and sends them to a PC. The PC performs audio signal processing, which includes recovering each track's audio waveform, calculating the inter-track difference, and using the microphone array localization algorithm to derive the instantaneous position of the person. We have written and installed a robot movement control terminal on the PC. As shown in Figure 23, the robot can track the person's position. Figure 24(a) shows the position transition when the person moves back and forth between A and B. The glitches are detection errors brought by analog phase detector AD8302. As shown in Figure 24(b), that localization precision degrades when we increase the sensor-to-person distance, but is still within an acceptable distance of 0.3m for 83.3% of the trials.

6 DISCUSSION

We discuss the limitations, potential solutions, and future studies as follows.

Compatibility. While we have demonstrated an equivalent bitrate of 6.4 Mbps for 8 audio tracks at a distance up to 28 meters, high-bitrate multimedia signals are still difficult to be embedded in commodity radios in our current implementation. Our current design only supports single-tone excitation. To allow the communication between our microphone array sensor and commodity radios, we need more accurate synchronization on excitation packets [6, 19, 23].



(a) Wearables for spatial sound recording in AR/VR live streaming.

(b) Sound field imaging for location, monitoring and diagnosis.

Figure 25: Potential applications of microphone array backscatter sensor.

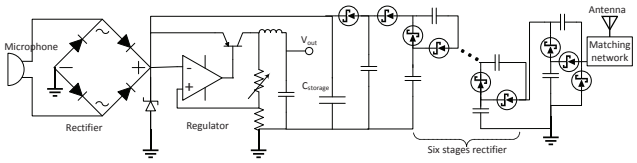


Figure 26: Sound-WiFi energy harvester design.

Eavesdropping. Our current design does not address the security issues and we need a strict access control mechanism to resist to such malicious behaviors as covert listening. Encryption would greatly increase the computing and communication overhead, which may not fit the backscatter design. We may incorporate the acoustic-radio transformation [42] that withstands eavesdropping attack. We also believe that a lightweight source-channel coding scheme that exploits both the multi-track sound signals and the excitation signals would be a potential solution.

Operating range. A backscatter communication system relates its performance to two key variables: the distance between excitation source and backscatter device, and the distance between backscatter device and receiver. As introduced in Section 3, we implement both the excitation and the receiver with USRP N210 devices. Note that we can also combine them as a reader using a single USRP N210 device, with one integrated antenna generating a single-tone excitation signal at 895MHz and the other integrated antenna receiving the backscattered signal at 905MHz. Given that the excitation is separate from the reader, we have shown in our experiment that the sensor-to-reader distance can be up to 28m when the excitation is located nearby the sensor. This communication range is sufficient for most indoor scenarios but limited to support data transfers over a very long distance in outdoor environments. To extend the range, we can increase the transmission power or introduce coding in our backscatter uplink. In the case of using a single reader to perform both excitation and demodulation, the sensor-to-reader distance becomes very limited for our current implementation. In Section 4.4, we have shown that the excitation-to-sensor distance can be up to 3m, which also means that a combined reader can communicate with the backscatter sensor at a distance of less than 3m. To enhance the performance, a potential solution is the full-duplex backscatter reader design [46].

Energy harvesting. To maximize the self-sustainability of our sensor, we also need an energy harvester that can jointly harvest energy from both sound and WiFi signals. As shown in Figure 26, the harvester can be built with the matching network, the rectifier, the regulator, and the storage capacitor. The matching network is used to reduce RF power losses. The rectifier converts alternating current (AC) to direct current (DC). The regulator maintains a stable DC voltage output. The harvester output can be connected to a booster for upgrading voltage. The sound energy harvesting rectifier uses a conventional diode bridge. The WiFI energy harvesting can use a six-stage rectifier by referring to the multi-stage Modified Greinacher Rectifier model in [9]. To discuss how much an energy harvester can contribute to a future battery-free design, we conduct a pilot experiment over a practical harvesting device. It uses electret microphones, 2.4G omnidirectional antennas, a multi-layer chip bandpass filter BF2012-L2R4DART/LF, a six-stage SMS7621 Schottky rectifier, an energy harvesting Power Supply LTC3588, a boost charger BQ25505, 25V 10μF electrolytic capacitors, and LX 0.1F Maxcap double layer storage capacitors. Our preliminary results reveal that the harvested sound energy increases with sound signal strength. We test a wide range of sound strength from 40dB to 90dB. With 40dB, 70dB, and 75 dB sound sources, the harvested powers are up to 19.2μW, 152.5μW, 203.6μW, respectively. 90dB sound can provide almost 406μW power. While a sound above 70dB can be annoying to human beings, it is commonly available in the real world (a busy street can easily reach this level). Other harsh environments, such as underground train or factory, can have a frequent or even constant presence of 90dB sound. When we increase the sound strength from 40dB to 90dB, the charging speed increases from 0.15mV/s to 0.75mV/s. According to Table 2, we can see that: 1) harvesting the energy from 75dB+ sound has the potential to support the PPM backscatter sensor; 2) 80dB sound energy harvesting has the potential to support the DPPM backscatter with less than 4 tracks; 3) 85dB sound energy harvesting has the potential to support the DPPM backscatter with less than 6 tracks; and 4) 90dB has the potential to support at most 7 tracks for the DPPM backscatter. For WiFi energy harvesting, the power decreases with the distance between the harvester and the WiFi transmitter (a Qualcomm AR938x NIC). As it increases from 0.05m to 1.55m, the harvested power changes from 85.65μW to 0.57μW. The charging speed decreases from 4.79mVs to 0.36mV/s when the distance increases from 0.05m to 1.55m. In other words, WiFi energy harvesting alone is not sufficient to power our sensor. Based on our preliminary results, we expect to enhance its efficiency in our future work through incorporating other latest designs for ambient

energy harvesting, e.g., WISP, energy management, beamforming, and piezoelectric transducer [12, 20, 23, 32, 63].

Multipath interference. In order to generate a PPM baseband with robustness to phase noise, we need to refer to the conventional time-domain equalization in wireless communication systems. For example, traditional GSM systems use a training sequence to perform initial synchronization, channel estimation, and equalization [34, 35]. The training sequence is a specific pattern signal that is sent in an empty time slot. The receiver detects the sequence and performs synchronization by delay compensation, which is useful for reducing consecutive bits' mutual interference. While our system transmits analog information rather than data bits, we can also detect the delay between original pulses and echoes, and use it to cancel the echoes. A potential solution is to design a decision feedback equalizer. DPPM uses pulse intervals to convey information and has no synchronization blank. As such, any training patterns can hardly be inserted into a DPPM baseband. It is known that increasing the baseband randomness can increase the robustness of wireless communication. Unfortunately, the conventional pseudorandom code sequence spread spectrum (e.g., direct-sequence spread spectrum) will not work here. This is because the sequence spread spectrum uses fixed-length sequences while DPPM pulse intervals are time-variant. Yet we can use angle modulation to change the patterns of DPPM baseband and improve the resilience to intersymbol interference (ISI). A potential solution is to modulate two consecutive pulse intervals with different phases. When receiving this phase modulated signal, we can use a phase detector and a low pass filter to recover the original pulse intervals. Another option is to use frequency modulation. This is similar to the conventional frequency-hopping spread spectrum. The only difference is that we need to modulate the time-variant pulse intervals rather than fixed-length symbols.

7 RELATED WORK

Acoustic wearable and sensing systems. There has been a significant interest in acoustic wearable and sensing systems [37–45, 51] for their broad use in consumer and industrial application scenarios. LipRead [37] explores ultrasound attack and develops a software-based defense system. Dia [41] designs a microphone-array synchronization algorithm and a microphone-array based adaptive beamforming system. [43] develops a high-precision acoustic motion tracking system as a new interface for VR/AR and smart appliances. [44] develops an acoustic imaging system with a smartphone. These works design and implement acoustic data processing algorithms, while our work focuses on an acoustic sensor's wireless transmission and hardware design.

Protocol-oriented backscatter and analog backscatter. Existing protocol-oriented backscatter devices [2–8, 12, 13, 16–19, 22, 23, 26, 27] can work with ambient radios using different protocols, but it does not include a data acquisition design that previously relies on power-consuming digital modules for multimedia streaming. In contrast, our system is a complete multimedia sensor design that uses analog backscatter technology for lightweight spatial sound parallel streaming. The work on hybrid backscatter [33] demonstrates analog backscatter sensing by continuously varying antenna impedance, and it also builds a battery-free wireless microphone using analog backscatter. Our work differs from the hybrid backscatter in two ways. First, our analog backscatter design focuses on multi-track audio synchronization, while the hybrid backscatter focuses on analog audio signal sensing with a single-track. Second, our PPM and DPPM baseband circuits are different from the audio sensing circuit in the hybrid backscatter. Battery-free cellphone [18] demonstrates an audio communication system using analog backscatter. Our work differs from it, not only in our design goal of analog multi-track synchronization, but also in the audio signal acquisition circuits: batter-free cellphone uses the variation of JFET impedance, whereas we use PPM and DPPM circuits. Video streaming backscatter [25] presents an analog backscatter system that uses pulse width modulation, whereas we use (differential) pulse position modulation, which is suited for multi-track analog signal modulation. Video streaming backscatter does not consider multi-track synchronization, either. Recent work on backscatter sensor mesh [28] enables lightweight sensor data transmission in a multi-hop way, where each backscatter node can use either digital or analog baseband circuits to sense and forward the temperature, light intensity, or single-track audio data. Our work extends backscatter-based sensing applications to an advanced level and significantly increases the communication range compared to [28].

8 CONCLUSION

We have developed a new wireless microphone-array sensor using backscatter technology. Its analog synchronization design brings not only low-power and lightweight hardware but also high throughput for synchronous concurrent transmission of multi-track sound streams. This is the first step towards enabling advanced acoustic technologies on ubiquitous wearables and mobile devices in a convenient and efficient way. In our future work, we will improve its performance for entertainment and industry applications. Lightweight and wireless spatial sound streaming will not only benefit a wide range of acoustic sensing and multimedia applications, but also offer a beyond-listening experience to common users. Such a design will change a complex 3D spatial audio recording system into a simple wearable working anytime and anywhere. Figure 25(a) illustrates a live AR/VR scenario, where one person wears a phased-array microphone device to transmit high-quality 360° sound to another person's AR/VR device that can add the 3D sound effect to live video. With this design, a user can even visualize a spatial sound field showing the location, intensity, and movement of the sound sources, which can be used for acoustic imaging in a variety of monitoring tasks. As shown in Figure 25(b), by incorporating an acoustic camera, mobile sensing devices can be made to move close to a targeted sound source (e.g., machine noise) and wirelessly transmit the multi-track sound to nearby smart devices, serving for a visualization of the sound source location.

ACKNOWLEDGMENTS

We would like to thank the anonymous shepherd and reviewers for their valuable and insightful comments. This work was supported by a Canada Technology Demonstration Program (TDP) grant, and a Canada NSERC Discovery Grant.

REFERENCES

[1] S. Gollakota, M. S. Reynolds, J. R. Smith, and D. J. Wetherall. The Emergence of RF-Powered Computing. *Computer*, 47(1): 32-39, 2014.
[2] A. Abedi, M. H. Mazaheri, O. Abari, and T. Brecht. WiTAG: Rethinking Backscatter Communication for WiFi Networks. *ACM HotNets*, 2018.
[3] B. Kellogg, A. Parks, S. Gollakota, J. R. Smith, and D. Wetherall. WiFi Backscatter: Internet Connectivity for RF-Powered Devices. *ACM SIGCOMM*, 2014.
[4] D. Bharadia, K. Joshi, M. Kotaru, and S. Katti. BackFi: High Throughput WiFi Backscatter. *ACM SIGCOMM*, 2015.
[5] B. Kellogg, V. Talla, S. Gollakota, and J. R. Smith. Passive Wi-Fi: Bringing Low Power to Wi-Fi Transmissions. *USENIX NSDI*, 2016.
[6] V. Iyer, V. Talla, B. Kellogg, S. Gollakota, and J. R. Smith. Inter-Technology Backscatter: Towards Internet Connectivity for Implanted Devices. *ACM SIGCOMM*, 2016.
[7] P. Zhang, M. Rostami, P. Hu, and D. Ganesan. Enabling Practical Backscatter Communication for On-body Sensors. *ACM SIGCOMM*, 2016.
[8] P. Zhang, D. Bharadia, K. Joshi, and S. Katti. HitchHike: Practical Backscatter Using Commodity WiFi. *ACM SenSys*, 2016.
[9] J. -P. Curty, N. Joehl, F. Krummenacher, C. Dehollain, and M. J. Declercq. Model for μ-Power Rectifier Analysis and Design. *IEEE Trans. Circuits and Systems*, 52(12): 2771-2779, 2005.
[10] P. Hu, P. Zhang, and D. Ganesan. Laissez-faire: Fully Asymmetric Backscatter Communication. *ACM SIGCOMM*, 2015.
[11] P. Hu, P. Zhang, M. Rostami, and D. Ganesan. Braidio: An Integrated Active-Passive Radio for Mobile Devices with Asymmetric Energy Budgets. *ACM SIGCOMM*, 2016.
[12] V. Liu, A. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. R. Smith. Ambient Backscatter: Wireless Communication Out of Thin Air. *ACM SIGCOMM*, 2013.
[13] A. N. Parks, A. Liu, S. Gollakota, and J. R. Smith. Turbocharging Ambient Backscatter Communication. *ACM SIGCOMM*, 2014.
[14] V. Talla, B. Kellogg, B. Ransford, S. Naderiparizi, S. Gollakota, and J. R. Smith. Powering the Next Billion Devices with Wi-Fi. *ACM CoNEXT*, 2015.
[15] P. Zhang and D. Ganesan. Enabling Bit-by-Bit Backscatter Communication in Severe Energy Harvesting Environments. *USENIX NSDI*, 2014.
[16] P. Zhang, P. Hu, V. Pasikanti, and D. Ganesan. Ekhonet: High Speed Ultra Low-Power Backscatter for Next Generation Sensors. *ACM MobiCom*, 2014.
[17] A. Wang, V. Iyer, V. Talla, J. R. Smith, and S. Gollakota. FM Backscatter: Enabling Connected Cities and Smart Fabrics. *USENIX NSDI*, 2017.
[18] V. Talla, B. Kellogg, S. Gollakota, and J. R. Smith. Battery-free Cellphone. *ACM UbiComp*, 2017.
[19] P. Zhang, C. Josephson, D. Bharadia, and S. Katti. FreeRider: Backscatter Communication Using Commodity Radios. *ACM CoNEXT*, 2017.
[20] Y. Ma, Z. Luo, C. Steiger, G. Traverso, and F. Adib. Enabling Deep-Tissue Networking for Miniature Medical Devices. *ACM SIGCOMM*, 2018.
[21] X. Xu, Y. Shen, J. Yang, C. Xu, G. Shen, G. Chen, and Y. Ni. PassiveVLC: Enabling Practical Visible Light Backscatter Communication for Battery-free IoT Applications. *ACM MobiCom*, 2017.
[22] V. Talla, M. Hassar, B. Kellogg, A. Najafi, J. Smith, and S. Gollakota. LoRa Backscatter: Enabling the Vision of Ubiquitous Connectivity. *ACM Ubicomp*, 2017.
[23] Y. Peng, L. Shangguan, Y. Hu, Y. Qian, X. Lin, X. Chen, D. Fang, and K. Jamieson. PLoRa: Passive Long-Range Data Networks from Ambient LoRa Transmissions. *ACM SIGCOMM*, 2018.
[24] D. Vasisht, G. Zhang, O. Abari, D. Katabi, H. -M. Lu, and J. Flanz. In-body Backscatter Communication and Localization. *ACM SIGCOMM*, 2018.
[25] S. Naderiparizi, M. Hessar, V. Talla, S. Gollakota, and J. R. Smith. Towards Battery-Free HD Video Streaming. *USENIX NSDI*, 2018.
[26] J. Zhao, W. Gong, and J. Liu. Spatial Stream Backscatter Using Commodity WiFi. *ACM MobiSys*, 2018.
[27] J. Zhao, W. Gong, and J. Liu. X-Tandem: Towards Multi-hop Backscatter Communication with Commodity WiFi. *ACM MobiCom*, 2018.
[28] J. Zhao, W. Gong, and J. Liu. Towards Scalable Backscatter Sensor Mesh with Decodable Relay and Distributed Excitation. *ACM MobiSys*, 2020.
[29] M. Kotaru, P. Zhang, and S. Katti. Localizing Low-power Backscatter Tags Using Commodity WiFi. *ACM CoNEXT*, 2017.
[30] V. Iyer, R. Nandakumar, A. Wang, S. B. Fuller, and S. Gollakota. Living IoT: A Flying Wireless Platform on Live Insects. *ACM MobiCom*, 2019.
[31] M. Hessar, A. Najafi, and S. Gollakota. NetScatter: Enabling Large-Scale Backscatter Networks. *USENIX NSDI*, 2019.
[32] J. Jang and F. Adib. Underwater Backscatter Networking. *ACM SIGCOMM*, 2019.
[33] V. Talla and J. R. Smith. Hybrid analog-digital backscatter: A new approach for battery-free sensing. *IEEE RFID*, 2013.
[34] C. Drane, M. Macnaughtan, and C. Scott. Positioning GSM telephones. *IEEE Communications magazine*, 36(4): 46-54, 1998.
[35] A. G. Orozco-Lugo, M. M. Lara, and D. C. McLernon. Channel Estimation Using Implicit Training. *IEEE Trans. Signal Processing*, 52(1): 240-254, 2004.
[36] Z. Li, M. S. Drew, and J. Liu. 2014. *Fundamentals of Multimedia* (2nd. ed.). Springer, 145-150.
[37] N. Roy, S. Shen, H. Hassanieh, and R. R. Choudhury. Inaudible Voice Commands: The Long-Range Attack and Defense. *USENIX NSDI*, 2018.
[38] N. Roy, M. Gowda, and R. R. Choudhury. Ripple: Communicating through Physical Vibration. *USENIX NSDI*, 2015.
[39] K. Yatani and K. N. Truong. BodyScope: A Wearable Acoustic Sensor for Activity Recognition. *ACM UbiComp*, 2012.
[40] www.printedelectronicsworld.com/articles/10275/tiny-soft-and-wearable-acoustic-sensor
[41] S. Sur, T. Wei, and X. Zhang. Autodirective Audio Capturing Through a Synchronized Smartphone Array. *ACM MobiSys*, 2014.
[42] T. Wei, S. Wang, A. Zhou, and X. Zhang. Acoustic Eavesdropping through Wireless Vibrometry. *ACM MobiCom*, 2015.
[43] W. Mao, J. He, and L. Qiu. CAT: High-Precision Acoustic Motion Tracking. *ACM MobiCom*, 2016.
[44] W. Mao, M. Wang, and L. Qiu. AIM: Acoustic Imaging on a Mobile. *ACM MobiSys*, 2018.
[45] A. Wang and S. Gollakota. MilliSonic: Pushing the Limits of Acoustic Motion Tracking. *ACM CHI*, 2019.
[46] M. Katanbaf, A. Weinand, and V. Talla. Simplifying Backscatter Deployment: Full-Duplex LoRa Backscatter. *USENIX NSDI*, 2021.
[47] J. Benesty, J. Chen, and Y. Huang. 2008. *Microphone Array Signal Processing*. Springer.
[48] D. J. Mennill, M. Battiston, D. R. Wilson, J. R. Foote, and S. M. Doucet. Field Test of An Affordable, Portable, Wireless Microphone Array for Spatial Monitoring of Animal Ecology and Behaviour. *Methods in Ecology and Evolution*, 3(4): 704-712, 2012.
[49] D. S. Shiu and J. M. Kahn. Differential Pulse-Position Modulation for Power-Efficient Optical Communication. *IEEE Trans. Communications*, 47(8): 1201-1210, 1999.
[50] K. T. Wong. Narrowband PPM Semi-'Blind' Spatial-Rake Receiver & Co-channel Interference Suppression. *European Trans. Telecommunications*, 18(2): 193-197, 2007.
[51] https://physicsworld.com/a/wearable-acoustic-system-monitors-foetal-motion/
[52] https://sonicscoop.com/2018/02/05/audio-mixing-for-vr-the-beginners-guide-to-spatial-audio-3d-sound-and-ambisonics/
[53] www.acoustic-camera.com/fileadmin/acoustic-camera/support/downloads/AC_brochure_2017_EN.pdf
[54] https://www.minidsp.com/applications/usb-mic-array
[55] www.acoustic-camera.com/en/products/microphone-arrays.html
[56] https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1284-2-201901-I!!PDF-E.pdf
[57] https://facebookincubator.github.io/facebook-360-spatial-workstation/
[58] https://superpowered.com/superpowered-audio-sdk-for-ios-android-windows-macos-linux-tvos
[59] https://www.adobe.io/apis/creativecloud/audition.html
[60] https://www.gearpatrol.com/tech/a36932143/apple-spatial-audio-vs-dolby-atmos-whats-the-difference/
[61] https://www.apple.com/ca/newsroom/2021/05/apple-music-announces-spatial-audio-and-lossless-audio/
[62] https://www.xilinx.com/products/design-tools/xst.html
[63] http://wisp5.wispsensor.net
[64] https://www.cadence.com/en_US/home/tools/custom-ic-analog-rf-design/circuit-design.html
[65] https://www.tsmc.com/english/dedicatedFoundry/technology/logic.htm#l_65nm_technology