BatComm: Enabling Inaudible Acoustic Communication with High-throughput for Mobile Devices

Yang Bai
WINLAB, Rutgers University
yang.bai.ece@rutgers.edu

Jian Liu
University of Tennessee, Knoxville
jliu@utk.edu

Li Lu*
Shanghai Jiao Tong University
luli_jtu@sjtu.edu.cn

Yilin Yang
WINLAB, Rutgers University
yy450@scarletmail.rutgers.edu

Yingying Chen
WINLAB, Rutgers University
yingche@scarletmail.rutgers.edu

Jiadi Yu
Shanghai Jiao Tong University
jiadiyu@sjtu.edu.cn

ABSTRACT

Acoustic communication is an increasingly popular alternative to existing short-range wireless communication technologies for mobile devices, such as NFC and QR codes. Unlike the current standards, there are no requirements for extra hardware, lighting conditions, or Internet connection. However, the audibility and limited throughput of existing studies hinder their deployment on a wide range of applications. In this paper, we aim to redesign acoustic communication mechanism to push the boundary of potential throughput while keeping the inaudibility. Specifically, we propose BatComm, a high-throughput and inaudible acoustic communication system for mobile devices capable of throughput rates 12× higher than contemporary state-of-the-art acoustic communication for mobile devices. We theoretically model the non-linearity of microphone and use orthogonal frequency division multiplexing (OFDM) to transmit data bits over multiple orthogonal channels with an ultrasound frequency carrier. We also design a series of techniques to mitigate interference caused by sources such as the signal's unbalanced frequency response, ambient noise, and unrelated residual signals created through OFDM, amplitude modulation (AM), and related processes. Extensive evaluations under multiple realistic settings demonstrate that our inaudible acoustic communication system can achieve over $47kbps$ within a $10cm$ communication range. We also show the possibility of increasing the communication range to room scale (i.e., around $2m$) while maintaining high-throughput and inaudibility. Our findings offer a new direction for future inaudible acoustic communication techniques to pursue in emerging mobile and IoT applications.

CCS CONCEPTS

• **Networks** → **Wireless access networks**; **Mobile ad hoc networks**.

*Also with Zhejiang University. Email: li.lu@zju.edu.cn.

SenSys '20, November 16–19, 2020, Virtual Event, Japan

ACM ISBN 978-1-4503-7590-0/20/11…$15.00
https://doi.org/10.1145/3384419.3430773

KEYWORDS

Acoustic Communication, High-throughput, Inaudible, Non-linearity

ACM Reference Format:
Yang Bai, Jian Liu, Li Lu, Yilin Yang, Yingying Chen, and Jiadi Yu. 2020. BatComm: Enabling Inaudible Acoustic Communication with High-throughput for Mobile Devices. In *The 18th ACM Conference on Embedded Networked Sensor Systems (SenSys '20), November 16–19, 2020, Virtual Event, Japan.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3384419.3430773

1 INTRODUCTION

Short-range wireless communication for mobile devices has become increasingly popular in recent years. It has been widely applied in advertisement [25], financial transactions [10], and user authentication [26, 47]. Bluetooth, QR codes, and near field communication (NFC) are among the most common choices, all with their own unique advantages and limitations. Bluetooth is already available in many mobile devices and covers a communication range up to $10m$. It requires a pairing process taking up to seconds and may drain battery quickly if continuously turned-on [34]. QR codes and NFC are developed for shorter communication ranges (around $1m$ and $10cm$, respectively). Internet connections are usually needed for QR codes because they can only carry a short *url* due to the limited data storage capability. Cameras and proper lighting conditions are also required to exchange QR data. Even worse, the QR code usually requires an update whenever the transmitted contents change for new *url* generation. NFC requires specific hardware components (i.e., NFC chips) at the user side. Inexpensive phones less than $400 normally are incompatible with NFC infrastructures [8]. The price of NFC reader is $50-$250 [40], also making it a prohibitively expensive option for wide deployment. Among these, Bluetooth and QR codes are relatively economical for business owners [2, 36]. In addition, Wi-Fi Direct can establish peer-to-peer connections for sharing heavy content (e.g., streaming videos). However, it requires a device-pairing process of a few seconds every time, resulting in a great inconvenience for many "stop-and-go" application scenarios.

In comparison, acoustic communication is one of the most affordable options with a charge of $10-$25 [12], depending on the volume and frequency of signals. Besides, acoustic communication can be enabled on any mobile and wearable devices without Internet connection necessity and environmental constraints. Different from Bluetooth and Wi-Fi Direct which establish a peer-to-peer connection via a several-second device pairing, acoustic communication does not need a pairing process. Microphones can record and process the signals as soon as they are transmitted from the

Table 1: Comparison of existing communication solutions.

Method	Specialized Hardware	Internet Dependent	Environment Dependent	Pairing Process
NFC	✓	✗	✗	✗
QR code	✗	✓	✓	✗
Bluetooth	✗	✗	✗	✓
Wi-Fi Direct	✗	✗	✗	✓
Acoustic	✗	✗	✗	✗

speakers (in a short distance). We compare these short-range communication technologies with more details in Table 1. Acoustic communication thus would be another desirable complementary option meeting various practical requirements. For commercial examples, it can assist merchants looking to transact with anyone who walks into their store in a low-cost manner, or for movie theaters hoping to send advertisements in dark environments. Acoustic communication can also improve the quality of services for users expecting fast and smooth device pairing, mobile payments, and for customers wishing to only receive nearby broadcasts. Additionally, for developing countries without high-end NFC-compatible smartphones popularized, acoustic communication could become a full replacement of traditional NFC. By simply holding their mobile device close to a supported transmitter, users can receive all sorts of information without the concern for Internet, complex user operation, hardware, or environmental constraints. Moreover, in vacation destinations without Wi-Fi or cellular network provided, such as national parks and beaches, ultrasound speakers can be deployed to transmit maps and guidances to travelers. Acoustic communication thus exhibits great potential to provide a lightweight solution for large demographics of mobile device users anytime and anywhere.

The benefits of acoustic communication have compelled many companies to invest in developing acoustic communication techniques for various applications (e.g., highly-secure proximity payments, customer engagement services) [15, 35, 41, 44, 46]. For such systems, high throughput and inaudibility are two key factors to success. High-throughput communication can enable the delivery of large digital files (e.g., audio, image, PDF file) instead of the limited text message or *url* link. It can also save the time cost for transmission and enable the deployment of more robust security protocols requiring additional throughput budget. Inaudibility is necessary to make the system unobtrusive to users.

Many applications could be facilitated if both high-throughput and inaudibility were achieved simultaneously. Figure 1 illustrates several representative ones. Acoustic communication is useful and convenient especially when wireless network connection is unavailable and cellular services are intermittent. For example, in many national parks, high-throughput acoustic communication technology could enable fast and convenient exchanges of dynamic location-dependent information, such as local news, travel itineraries, maps, and hotel reservations, etc. Retail stores, as another example, can distribute detailed coupons and advertisements to customers in a timely fashion. Additionally, high-throughput communication channel in the physical layer also provides a chance to add secure communication protocols, enabling various applications that require higher security levels such as sensitive information transmission, mobile payment, and device pairing [16]. Moreover, BatComm provides a powerful complementary communication channel for the devices equipped with microphones, making it a feasible substitute for many NFC applications.

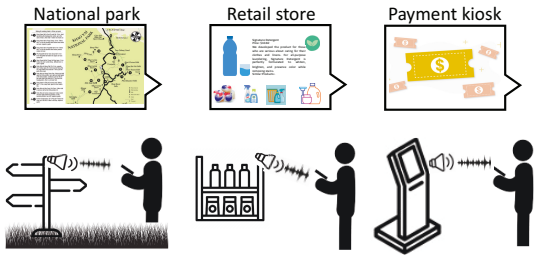


Figure 1: Application scenarios of high-throughput and inaudible near-field acoustic communication.

In this paper, our objective is to develop a robust acoustic communication system with both high-throughput (i.e., >47$kbps$) and inaudibility, which provides a powerful complementary communication channel for the devices equipped with microphones. For the proposed system, as shown in Figure 2, an ultrasound speaker transmits an ultrasound frequency signal (e.g., >40kHz) carrying the entire audio frequency band acoustic signal (i.e., 0-24kHz). By incorporating non-linearity of microphones (i.e., described with details in Section 3) into our technique, a nearby mobile or IoT device can detect and receive signals that are modulated onto this ultrasound frequency carrier. We also explore the opportunity of increasing the communication range to room-scale (e.g., around 2m) to improve user experience (i.e., more easily position the phone within the communication range) and satisfy the needs of some applications requiring longer range.

Studies achieved a relatively high throughput (e.g., 1$kbps$ [27, 55], 2.4$kbps$ [33]) by using the audible acoustic frequency band (e.g., 0-18kHz [27], 0-22kHz [33], 8-10kHz [55]). The audibility of these techniques may disturb users and degrade the user experience. The recordable frequency band is limited to *half of the sampling rate* of audio hardware according to the Nyquist theorem [23]. Based on this principle, existing efforts used near-ultrasound frequency band (i.e., approximately 17-22kHz) [21, 24, 39] to achieve the inaudible communication with 44.1kHz sampling rate, and the achieved throughput is up to 2.76$kbps$. BackDoor [37] verified that ultrasound signals can be designed to be recordable with the non-linearity of microphones and used for data transmission. The communication throughput achieves up to 4$kbps$ within 1m. Unlike BackDoor, we design a robust near-field (i.e., 10cm) inaudible communication solution that can achieve 12× higher throughput (i.e., >47$kbps$) to support another part of applications requiring either less time cost (e.g., device pairing, mobile payment) or transmitting large amounts of data (e.g., mobile advertisement). Our solution is also scalable such that it can support meter-level range (i.e., around 2m), while maintaining up to 17.76$kbps$ throughput.

To achieve the high throughput and inaudible communication for general mobile devices with limited audio sampling rates (e.g., 48kHz), BatComm (1) applies OFDM and AM to transmit data bits on multiple ultrasound subcarriers concurrently; (2) reconstructs the entire audio frequency band from the received ultrasound signal; and (3) employs a frequency band separation method to better utilize frequency bandwidth. We theoretically and empirically demonstrate that the unrelated residual signals produced by AM under the non-linearity of microphones would interfere with the signals transmitting on other OFDM subcarriers, which induces errors in the received data. To address this issue, we propose an

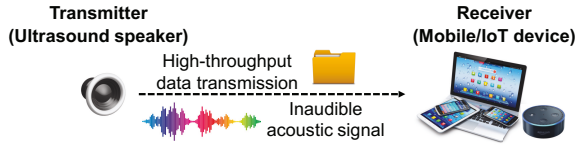


Figure 2: Illustration of inaudible acoustic communication with off-the-shelf mobile devices.

elimination scheme, which elaborately modifies the OFDM waveform before AM, to eliminate the unrelated residual signals in the recorded signals. To make BatComm robust in realistic settings such as outdoor or noisy environments, we also apply a series of techniques (e.g., differential phase shift keying (DPSK), preamble, cyclic prefix, channel estimation, BCH code, interleaving) to handle various interference and fading problems, including frequency selective fading, time selective fading, inter-symbol interference, and practical ambient noises. We also investigate the means of developing a room-scale range communication using multiple speakers. The residual signals are eliminated at the receiver side by carefully choosing the frequency band for robust communication. Our main contributions are summarized as follows:

- To the best of our knowledge, we develop the first high-throughput inaudible near-field acoustic communication system, BatComm, applicable to general mobile devices. The achieved throughput (i.e., over $47kbps$) is 12× higher than existing acoustic communication solutions.
- In order to maximize the throughput while keeping inaudibility, we theoretically model the non-linearity of the device's inbuilt microphone and innovatively use OFDM and AM to transmit data over multiple narrow-band channels in an ultrasound frequency band. The non-linearity of microphones can enable mobile devices to recover modulated data on the entire audio frequency band (i.e., $<24kHz$). We also propose a residual-signal elimination scheme to mitigate the effect caused by the unrelated residual signals produced by AM.
- We test the possibility of adapting our communication range to serve room scale applications by leveraging two speakers to play different segments of the same signal, eliminating audible leakage while increasing the power of transmission.
- Extensive experiments in various realistic settings demonstrate that BatComm can achieve high throughput and low BER, e.g., $47.49kbps$ with 4.3% BER and $25.37kbps$ with 0.5% BER, at centimeter level range and comparable throughput, i.e., $17.76kbps$, with 0.9% BER at up to $2.3m$ range.

2 BACKGROUND

2.1 Microphone System

A microphone on a COTS mobile device mainly consists of four components, i.e., transducer, pre-amplifier, inbuilt low-pass filter and analog-to-digital converter (ADC) [13]. When an acoustic signal carries the energy towards a microphone, the transducer of the microphone first transforms the mechanical sound waves to electric signals through the electromagnetic induction [49]. Then, the pre-amplifier enhances the electric signal to improve the signal-to-noise ratio (SNR) of the signal transformed from sound waves. Next, the inbuilt low-pass filter eliminates the high-frequency harmonic components from the electric signals to match the sampling rate of ADC in the microphone. The inbuilt low-pass filter usually sets the cut-off frequency as $24kHz$ due to the maximum sampling rate of $48kHz$ in the ADC of most widely-deployed mobile devices. After that, ADC samples the electric signals and stores the values as digital signals, which represents the recorded acoustic signals.

2.2 Non-linearity of Microphone

Due to the limited sampling rate (i.e., $48kHz$) of COTS microphones, most mobile devices can only record acoustic signals within a confined frequency range (i.e., $< 24kHz$). However, when receiving an ultrasound signal, the pre-amplifier of the microphone exhibits *non-linearity* in the ultrasound frequency range [1, 9], enabling the capture of ultrasound signals by inbuilt microphones. The non-linearity of microphones can be modeled theoretically. Assume that the microphone receives an acoustic signal s_{in}. After the sound is picked up and amplified by the microphone's transducer and pre-amplifier, the recorded signal s_{out} can be represented as:

$$\begin{aligned} s_{out} &= A_1 s_{in} + \sum_{i=2}^{\infty} \delta(f) A_i s_{in}^i \\ &\approx A_1 s_{in} + \delta(f) A_2 s_{in}^2, \end{aligned} \tag{1}$$

where A_i is the energy gain for the i^{th} order term and $\delta(f)$ is an indicator function. The value of $\delta(f)$ is 1 when f is greater than f_0, where f_0 is the critical frequency of the non-linearity. Otherwise, $\delta(f)$ equals to 0. We empirically find the critical frequency $f_0 \approx 18kHz$ in most commercial mobile devices. Although the non-linear output is an infinite power series, the value of A_i decreases with the increase of i and the third and higher order terms are extremely small. Therefore, we only consider the linear and quadratic terms. This indicates that the pre-amplifier generates additional frequency components, i.e., the quadratic term of Equation 1, other than the original frequency component when the microphone receives an ultrasound signal. The exhibited non-linearity of microphones demonstrates the feasibility of utilizing ultrasound for communication. By selecting an appropriate modulation technique, the modulated data bits carried on an ultrasound carrier can be transmitted inconspicuously and recovered by commercial microphones.

2.3 Speaker System and its Non-linearity

Sound playback is the reverse process of recording. Given a digital signal as input, the digital-to-analog converter (DAC) produces the analog signal and sends it to the amplifier. After enlarging the signal, the speaker's diaphragm oscillates according to the applied analog signal to produce sound in the air. Similar to microphones, if the transmit power at the speaker exceeds a threshold, non-linearities at the speaker's own diaphragm and amplifier can be triggered, resulting in an audible leakage at the output of the speaker [38]. The audible leakage violates the inaudible requirement, which may be obtrusive to users. To avoid audible leakage and maintain good user experience, we operate BatComm at a low transmission power.

3 ACHIEVING HIGH THROUGHPUT WHILE KEEPING INAUDIBILITY

3.1 Challenges

High-throughput and Inaudible Communication for General Mobile Devices. The high-throughput of acoustic communication

requires a wide frequency band for data transmission. However, for most mobile devices, the inbuilt microphone only has a limited ADC sampling rate (i.e., $48kHz$), thus the device can only record acoustic signals within $24kHz$ according to Nyquist theorem [23]. Conversely, in order to achieve inaudibility, a narrow frequency band (i.e., 18-$24kHz$) is necessary for communication, which significantly reduces the possible throughput. Therefore, it is essential to find a way to increase the frequency bandwidth for communication to improve the throughput, while remaining within the confines of inaudibility.

Robust Communication Under Various Environmental Factors. There are many environmental factors affecting the robustness of acoustic communication systems. For instance, acoustic frequency channels may be easily affected by other sound sources in the environment, such as HVAC (Heating, Ventilation, Air Conditioning) noises, people talking, etc. Moreover, due to the omnidirectional propagation of acoustic signals, acoustic communication suffers from significant multipath effects, which may induce time- and frequency-selective fading. Therefore, our system needs to transmit data via acoustic channels in a way that is robust to these noisy environments and signal interference factors.

3.2 Integrating Non-linearity with Signal Multiplexing and Modulation Techniques

Achieving High Throughput based on OFDM. To achieve high-throughput communication, we utilize OFDM technique to modulate the data signals on multiple subcarriers to convey multiple data bits concurrently. Furthermore, the communication capability of an OFDM system increases proportionally to the frequency bandwidth. Therefore, our system aims at using the whole acoustic frequency band, including both audible and inaudible ranges, to achieve high throughput.

Enabling Inaudibility via AM Modulation and the Non-linearity of Microphone. Due to the utilization of audible frequency band in OFDM, directly transmitting the OFDM-multiplexed signals remains audible in ambient areas. To make the acoustic communication inaudible, we further integrate AM technique to transmit audible OFDM-multiplexed signal via an ultrasonic carrier. Specifically, we define the OFDM-multiplexed signal on the m^{th} OFDM subcarrier (subcarrier frequency is f_m) is: $m(t) = \cos(2\pi f_m t)$. Then $m(t)$ can be modulated with an ultrasound carrier signal $\cos(2\pi f_c t)$ ($f_c \geq 2f_m$) through AM. The modulated signal is:

$$s_{in} = \cos(2\pi f_c t) \cdot (1 + \cos(2\pi f_m t)). \tag{2}$$

Combined with the non-linearity of microphone, i.e., Equation 1, we derive the signal s_{out} that is recorded by the device's inbuilt microphone as:

$$\begin{aligned} s_{out} =& A_1 \cos(2\pi f_c t) \\ &+ \frac{A_1}{2}(\cos(2\pi(f_c + f_m)t) + \cos(2\pi(f_c - f_m)t)) \\ &+ \frac{A_2}{4}(4\cos(2\pi f_m t) + 3cos(4\pi f_c t) + \cos(4\pi f_m t)) \\ &+ \frac{A_2}{8}(\cos(4\pi(f_c + f_m)t) + \cos(4\pi(f_c - f_m)t)) \\ &+ \frac{A_2}{2}(\cos(2\pi(2f_c + f_m)t) + \cos(2\pi(2f_c - f_m)t)) \\ &+ \frac{3A_2}{4}. \end{aligned} \tag{3}$$

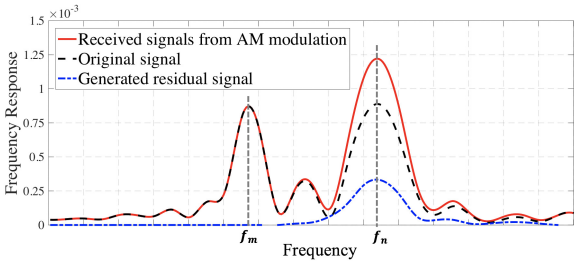


Figure 3: Illustration of the signals on the subcarrier f_n affected by the residual signal (e.g., $f_n = 2f_m$).

From Equation 3, we find that the frequency components contain $f_c, f_c - f_m, f_c + f_m, f_m, 2f_c, 2f_m, 2(f_c + f_m), 2(f_c - f_m), 2f_c + f_m$, and $2f_c - f_m$. Since f_c ($f_c \geq 2f_m$) is the ultrasound carrier frequency, components with frequency f_c can be eliminated with the low-pass filter in the microphone. After that, the signal becomes:

$$s_{out} = A_2 \cos(2\pi f_m t) + \frac{A_2}{4}\cos(4\pi f_m t) + \frac{3A_2}{4}, \tag{4}$$

which only contains the frequency components f_m and $2f_m$. In addition to the transmitted frequency component f_m, we note that the induced component $\frac{A_2}{4}\cos(4\pi f_m t)$ may impact the OFDM signals on other specific subcarriers (i.e., subcarriers with frequency $2f_m$). It would lead to significant errors when the system recovers the original signal $m(t)$.

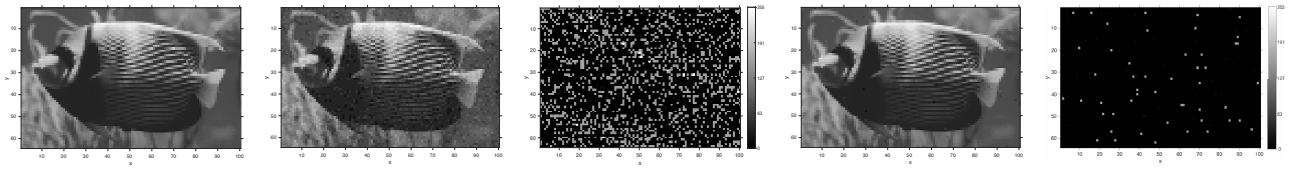
3.3 Eliminating Unrelated Residual Signals Induced by AM Modulation

As shown in Equation 4, the unrelated frequency component $2f_m$, which we call *unrelated residual signals*, may greatly impact the system's capability to recover original signals as well as interfere with OFDM signals transmitting over subcarriers of frequency $2f_m$. We simulate a signal, consisting of two frequency components (i.e., f_m and $f_n = 2f_m$), that is transmitted with an ultrasound carrier through AM modulation. As shown in Figure 3, we observe that the received demodulated signal has great interference on the frequency f_n due to the induced unrelated residual signal.

To eliminate the residual signal, we elaborately modify the OFDM-multiplexed signal before the AM modulation in the transmitter side. According to Equations 1 and 2, the linear term (i.e., $A_1 s_{in}$) of the non-linearity model contains ultrasound frequency f_c and thus would be filtered out by the inbuilt low-pass filter of the microphone. Hence, we only analyze the quadratic term (i.e., $A_2 s_{in}^2$) for the elimination scheme design. Specifically, we define the analog OFDM symbol waveform as $s(t)$ that carried the data bits to be transmitted. According to Equation 1, the quadratic term s_q can be represented as:

$$\begin{aligned} s_q =& A_2(\cos(2\pi f_c t) \cdot (1 + s(t)))^2 \\ =& \frac{A_2}{2}(1 + \cos(4\pi f_c t)) + \frac{A_2}{2}(1 + \cos(4\pi f_c t))(2s(t) + s^2(t)). \end{aligned} \tag{5}$$

We observe that the first term (i.e., $\frac{A_2}{2}(1+\cos(4\pi f_c t))$) of Equation 5 only contains the ultrasound frequency component $2f_c$, which can be discarded due to the low-pass filter in the microphone. As $s(t)$ contains frequency components crossing all the OFDM subcarriers, the second term (i.e., $\frac{A_2}{2}(1+\cos(4\pi f_c t))(2s(t)+s^2(t))$) of Equation 5 would produce unrelated residual signals, as per Equation 3. In order



(a) Transmitted image (b) Received image w/o elimination scheme (c) Error comparision w/o elimination scheme (d) Received image w/ elimination scheme (e) Error comparision w/ elimination scheme

Figure 4: Image comparison between transmitted and received images without and with unrelated residual signal elimination scheme (BER = 10.2% vs. BER = 0.3%).

to eliminate these residual signals, we use the following signal $s_e(t)$ to replace the OFDM symbol waveform $s(t)$ before AM modulation:

$$s_e(t) = \sqrt{(s(t)+1)} - 1. \tag{6}$$

Therefore, we have $2s_e(t) + (s_e(t))^2 = s(t)$, and Equation 5 would be changed to:

$$s_q = \frac{A_2}{2}(1 + \cos(4\pi f_c t)) + \frac{A_2}{2}(1 + \cos(4\pi f_c t))s(t). \tag{7}$$

As a result, by removing components containing ultrasound frequency f_c, only the component $\frac{A_2}{2}s(t)$ can be preserved through the microphone's low-pass filter. By further modulating the modified OFDM signal (i.e., $s_e(t)$) with the ultrasound carrier through AM, unrelated residual signals can be removed from the recorded signal on the receiver side.

To validate whether the proposed scheme could eliminate the unrelated residual signals, we conduct a simulation experiment in which a software-defined transmitter (including OFDM, elimination scheme, and AM) and software-defined receiver (including the microphone's pre-amplifier with non-linearity, inbuilt low-pass filter, and OFDM demodulation process) are implemented. We use an additive white gaussian noise (AWGN) model in the communication channel and set the SNR to 36. A $6.4kB$ image[1], shown in Figure 4(a), is sent from the transmitter to the receiver. Figures 4(b) and 4(d) are the received images without and with the proposed elimination scheme (i.e., using $s(t)$ as OFDM symbol waveform), respectively. Since it is difficult to visually differentiate them, we show the error comparisons without and with the elimination scheme in Figures 4(c) (bit error rate (BER)=10.2%) and 4(e) (BER=0.3%). This result demonstrates that the proposed elimination scheme is effective at mitigating errors introduced by unrelated residual signals. In addition, we carry out in-the-air experiments with our proposed elimination scheme. When transmitting the same image, we find the BER to be higher than the simulation results. This is likely attributed to the ultrasound speaker's frequency response, which would dramatically decreases as transmission frequency increases, making SNR lower in the high frequency band [5]. Such a frequency-response deterioration is particularly obvious in our system as the recovered signal at the receiver side would be scaled down to $\frac{A_2}{2}s(t)$ with respect to the transmitted signal $s(t)$, as shown in Equation 7. Consequentially, with a low frequency response in the high frequency band, the signal strength of the received signal can be greatly diminished, introducing significant errors if these behaviors are not properly accounted for.

[1]Each pixel in the grayscale image contains 8 bits.

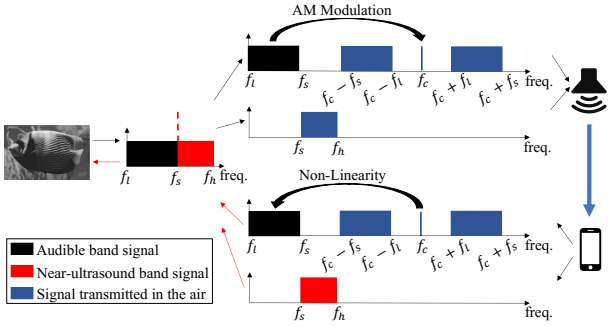


Figure 5: Illustration of the bandwidth separation.

3.4 Separating the Frequency Band of Acoustic Signals for AM Modulation

To mitigate the effect caused by the aforementioned frequency-response deterioration while keeping communication inconspicuous, we only modulate the OFDM signal in the audible band (i.e., $<f_s$) on the ultrasound carrier signal through AM modulation, whereas the OFDM signal in the near-ultrasound band (i.e., $>f_s$) would be directly transmitted. We leave the details of how to choose f_s in Section 5.4.

More specifically, as illustrated in Figure 5, the frequency band for data transmission (i.e., $f_l \sim f_h$) is separated by a breaking point (i.e., f_s) into two parts, which are $f_l \sim f_s$ (i.e., audible band) and $f_s \sim f_h$ (i.e., near-ultrasound band), respectively. For the audible band, the signal is modulated on the ultrasound carrier signal (i.e., with the frequency of f_c) using AM modulation. After that, the frequencies of the modulated signal are $(f_c - f_s) \sim (f_c - f_l), f_c, (f_c + f_l) \sim (f_c + f_s)$. For the near-ultrasound band, the signal can be directly transmitted. The two signals are played by a speaker simultaneously and captured by the receiver to interpret. At the receiver side, the signal in the audible band can be automatically demodulated through the non-linearity of microphone, while the signal in the near-ultrasound band can be captured by the microphone directly. Therefore, the signal strength in the near-ultrasound band will not be scaled down, and the required frequency of ultrasound carrier is lower than before, both leading to a much higher overall SNR.

4 SYSTEM OVERVIEW

Transmitter Design. The transmitter is responsible for modulating data bits to an ultrasound signal for the high-throughput and inaudible acoustic communication. The data bits are first encoded with BCH error correction code [7] and further re-ordered through an interleaving technique to counteract unpredicted errors during signal propagation. Then, the encoded data is converted to phase values through the digital modulation technique, i.e., DPSK.

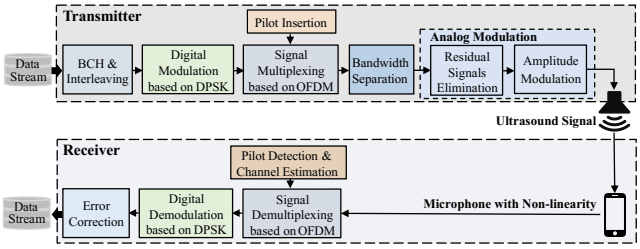


Figure 6: Architecture of BatComm.

To fully utilize the scarce frequency band for communication, the OFDM technique is further applied to modulate the phase values to multiple subcarriers for concurrent data transmission. During OFDM, a pilot is inserted in the OFDM signals for channel estimation so as to mitigate the multipath effect on the received signals. We then partition the OFDM signals into audible and near-ultrasound components. For the audible frequency band signal, the OFDM symbol waveform $s(t)$ is modified to $s_e(t)$ for elimination of unrelated residual signals, which are generated through AM under the non-linearity of microphones. Then the modified OFDM symbol waveform $s_e(t)$ is modulated onto the ultrasound carrier through AM for inaudible communication. For the near-ultrasound frequency band signal, directly applying OFDM is sufficient to keep the transmission below common hearing thresholds.

Receiver Design. The receiver in our system is a commercial mobile/IoT device (e.g., a smartphone) with an inbuilt microphone, which records and demodulates the received ultrasound signal to receive data. The audible band signal can be automatically demodulated, taking advantage of the non-linearity of the microphone, while the near-ultrasound frequency band signal can be directly recorded. To recover the transmitted signal in the full frequency band, the receiver first performs demultiplexing on the recorded OFDM waveform in order to extract phase values. The pre-inserted pilot in the OFDM signals is used for channel estimation to eliminate multipath interference during this procedure. Afterwards, the extracted phase values are mapped to digital data bits through DPSK demodulation. Finally, the receiver performs error correction on the digital data bits with the pre-inserted BCH code and the interleaving matrix to mitigate the unpredicted errors.

5 TRANSMITTER DESIGN

5.1 Error Correction via BCH Codes and Interleaving

The proposed acoustic communication may have unpredicted errors (i.e., randomly distributed errors, burst errors) induced in the propagation channel. To mitigate randomly distributed errors, we choose BCH code [32], one of the most commonly used multi-bit error correction code, to encode the digital data. Other error correction mechanisms (e.g., Reed-Solomon code [53], Hamming code [52]) are also feasible. Specifically, the digital data is encoded with (N, K)-BCH code, where N is the length of the encoded data and K is the length of original digital data. Moreover, errors usually burst in particular domains of the signal, due to intensive noises appearing in some specific frequency bands or time periods. To address this problem, we further apply a matrix-based interleaving approach [28] to interleave the data stream in a particular known order, which could convert bursts of errors into random-like errors. Such an interleaving approach could greatly improve error correction based on BCH codes when burst errors occur.

5.2 Digital Modulation based on DPSK

In order to transmit data in the air, the digital data bits should first be modulated, which is necessary for digital-to-analog conversion. The most commonly used digital modulation techniques are amplitude shift keying (ASK), frequency shift keying (FSK) and phase shift keying (PSK). ASK and FSK utilize the amplitude and frequency of carrier signals to modulate the digital data bits, respectively. However, due to the vulnerability to noises and the requirement of wide bandwidth, they are impractical for acoustic communication. PSK modulates data bits on several absolute phase values, which is efficient to utilize the scarce acoustic spectrum. However, in practical communication scenarios, the multipath propagation of acoustic signals and ambient noise may induce unpredictable phase shift on the signals, which leads to errors in the absolute phase values. To solve this problem, we use DPSK [51] for the digital modulation in our system, which shows satisfactory performance in low SNR environments [31]. This is because DPSK facilitates non-coherent demodulation, which does not need to keep track of a reference wave for being more robust to the phase deviations in practice.

5.3 Signal Multiplexing based on OFDM

Adding Preamble and Cyclic Prefix for Synchronization. To ensure that the receiver can recover the complete data from acoustic signals, it is necessary for the receiver to precisely find the beginning of each frame. To achieve synchronization, we divide the data stream to be transmitted into a set of frames and insert a preamble designed to follow the protocol of IEEE 802.11a [19] at the beginning of each frame. Moreover, the multipath effect introduces inter-symbol interference [11] between OFDM symbols. To eliminate this interference, the transmitter adds a cyclic prefix in the beginning of each symbol, which is designed as the last quarter of the symbol. With the preamble and cyclic prefix, the receiver can find the beginning of each frame and obtain OFDM symbols precisely. The length of an OFDM symbol is 1, 280 digital samples.

Inserting Pilot for Channel Estimation. Due to the multipath propagation of omni-directional acoustic signals, there exist time and frequency selective fading in the received signals [30]. To mitigate the time selective fading induced by multipath propagation, we insert pre-defined phase values, *comb-type pilot*, on one subcarrier (i.e., subcarrier #502 corresponding to $23.53kHz$)[2]. The pilot symbol, *block-type pilot*, is used to mitigate the frequency selective fading at the receiver end, which is discussed in Section 6.1.

Frequency Band Selection to Resist Interference. Due to the interference of ambient noises in the environment, parts of the available acoustic frequency bands introduce significant errors in the communication. Common sound sources (e.g., human speaking) usually generate acoustic signals lying in the frequency of less than $8kHz$ [48]. On the other hand, commercial smartphones can only record acoustic signals with frequency up to $24kHz$, due to the limited sampling rate. To achieve high-throughput communication

[2] 1024-point IFFT/FFT in OFDM could have 512 orthogonal subcarriers corresponding to the bandwidth of 0-$24kHz$.

while keeping our system resilient to daily noises, the operation band for the OFDM subcarriers is chosen as 8.06-23.53kHz, corresponding to the OFDM subcarrier #172 to #502 in our system.

5.4 Selection of Frequency Threshold in Band Separation

As described in Section 3.4, we only modulate the OFDM waveform in the audible band (i.e., $f_l \sim f_s$) on the ultrasound carrier signal through AM modulation while transmitting the OFDM waveform in the near-ultrasound band (i.e., $f_s \sim f_h$) directly. Hence, how to appropriately choose the partitioning frequency of the spectrum f_s is critical. To keep the communication process inaudible, the frequency band $f_s \sim f_h$ should be beyond human hearing perception. Moreover, f_s should be as small as possible to avoid signal strength deterioration induced by non-linearity of microphone at the receiver side. An existing study [3] shows the average sound pressure level at which humans may begin to hear a 16kHz signal is 41.8$dBSPL$. We empirically demonstrate that the measured sound pressures nearby our implemented transmitters are all below 40$dBSPL$, which is discussed in Section 8.2. As a result, we set the partitioning of the spectrum, f_s, as 16kHz. Specifically, the audible band is 8.06-15.98kHz (i.e., the subcarrier #172 to #341), and the near-ultrasound band is 16.03-23.53kHz (i.e., the subcarrier #342 to #502). The audible band signal is then modulated using AM towards inaudibility, which is discussed in Section 5.5.

5.5 Analog Modulation based on AM Towards Inaudibility

The analog OFDM symbol waveform $s(t)$ needs to be modulated onto an ultrasound carrier, so as to ensure communication is beyond human perception. Particularly, as mentioned in Section 3, the combination of OFDM and AM techniques would produce unrelated residual signals, which would largely interfere with our system and produce significant transmission errors. Hence, we use the modified OFDM symbol waveform $s_e(t)$, as per Equation 6, before AM modulation to eliminate the interference from the residual signals. Then we use AM modulation to modulate $s_e(t)$ onto the ultrasound carrier with frequency f_c, according to Equation 2.

Furthermore, the selection of the frequency f_c is critical. This is because a low carrier frequency induces overlapping between the AM-modulated ultrasound signal and OFDM-multiplexed signal, while a high carrier frequency leads to low-power transmitted signals, i.e. low SNR. As mentioned in Section 2.2, after the ultrasound is recorded by the microphone, the frequency components (e.g., $f_c - f_m$, $f_c + f_m$, f_c) including the carrier frequency f_c should be filtered out by the inbuilt low-pass filter. It indicates that the lowest frequency component (i.e., $f_c - f_m$) should be larger than the cut-off frequency of the filter f_{cut} (i.e., 24kHz). Thus, we have $f_c \geq f_{cut} + f_m$. According to Section 5.4, the operation bandwidth of audible acoustic signal is around 8.06-15.98kHz. Therefore, the carrier frequency should meet the requirement of $f_c \geq 40kHz$. Unless mentioned otherwise, we set $f_c = 40kHz$ in our system.

6 RECEIVER DESIGN

6.1 Signal Demultiplexing based on OFDM

Due to the non-linearity of the microphone and the design of our residual signal elimination scheme, the OFDM waveform modulated on the ultrasound carrier could be automatically picked up by the microphone, as mentioned in Section 3.2 and Section 3.3. Moreover, the OFDM waveform transmitted in the near-ultrasound band (i.e., 16.08-23.53kHz) can be captured by the microphone directly. In BatComm, the receiver first detects the preamble in the received signal to synchronize the OFDM frames, then demodulates the signal through OFDM and finally performs channel estimation to mitigate the multipath effect based on the pre-inserted pilot signal.

Preamble Detection and Synchronization. The receiver first needs to detect the preamble of each OFDM frame so as to synchronize the signal frames. Since the preamble is known to both the transmitter and receiver, we apply correlation to detect the preamble and find the beginning of the transmission [50]. Based on the correlation values, the beginning of the preamble can be found through the maximum value of correlation. After we detect the beginning of each OFDM symbol, the receiver further removes the cyclic prefix to extract the data from the signal. Since the cyclic prefix serves as a guard between two successive OFDM symbols, the inter-symbol interference can be eliminated by removing the cyclic prefix. After that, the receiver demodulates the OFDM symbols through FFT operation to derive the phase values for further processing.

Pilot Detection and Channel Estimation. Due to the omnidirectional propagation property of sound, the unpredicted propagation would introduce unexpected time and frequency selective fading errors in received signals. To address this, the receiver performs least square channel estimation [43] based on the pre-inserted pilot, which is discussed in Section 5.3. Through such a channel estimation, the multipath effect can be mitigated in the received signals, which improves the robustness of communication.

6.2 Digital Demodulation & Error Correction

After demodulating the OFDM signals, the receiver then performs digital demodulation based on DPSK to obtain the digital data bits. Specifically, the receiver first derives the differential phase values and extracts the digital data bits through the constellation mapping scheme. After that, the receiver recovers the digital data bits through interleaving. Since the dimension of the matrix for interleaving is known to the receiver, the receiver performs reverse interleaving to obtain the corrected-order data. Then, the receiver utilizes the BCH code to correct the unpredicted errors in the digital data bits.

7 INCREASING COMMUNICATION RANGE

Although BatComm concentrates on inaudible acoustic communication in a near field, the capability of room-scale (e.g., several meters) high-throughput inaudible acoustic communication, would greatly extend its possible applications. We thus explore how to increase the communication range while maintaining high-throughput and inaudibility.

7.1 Challenges of Designing a Meter-Level Communication

Increasing Transmission Power while Keeping Inaudibility. To increase the communication range, one straightforward way is to increase the power of transmission at the transmitter side (i.e., ultrasound speaker). However, enlarging the power can cause audible leakage with the non-linearity of speakers [38]. When the power

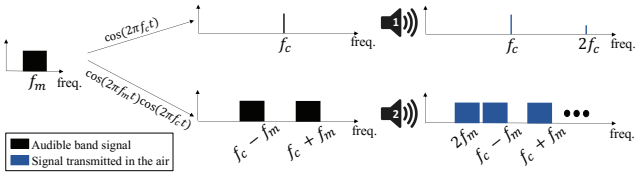


Figure 7: Illustration of the transmission process of room-scale communication.

(volume) of the speaker exceeds a threshold, the non-linearity of the speaker can be triggered. Some leaked low-frequency signals caused by the non-linearity thus can be induced, thereby making the outcoming signal audible.

Interference after Enlarging the Received Signal Strength. As discussed in Section 3.1, we use the signal $\sqrt{s(t)+1}-1$ to replace the OFDM multiplexed waveform $s(t)$ to eliminate the residual signals. This approach can effectively eliminate the residual signal while using the full frequency band and OFDM for communication. However, this significantly downgrades the strength of the usable signal to $\frac{A_2}{2}$, which largely limits the communication range. A different approach that can eliminate the residual signals without sacrificing the signal strength is required.

7.2 Keeping Inaudibility while Enlarging Power of Transmission

When we enlarge the power of transmission to lengthen the range of communication, the audible leakage induced by the non-linearity of speaker would also be triggered, as described in Section 2.3. To achieve inaudibility with the non-linearity of the speaker, we separate the AM modulated signal into two parts and use two speakers to play them separately. The AM modulated signal s_{in} (i.e., $cos(2\pi f_c t)(1+s(t))$) is separated into $s_1 = cos(2\pi f_c t)$ and $s_2 = s(t)cos(2\pi f_c t)$, which are played by two ultrasound speakers respectively. As shown in Figure 7, the frequency component of s_1 is f_c. For simplicity, we define the frequency of $s(t)$ as f_m, i.e., $s_2 = cos(2\pi f_m t)cos(2\pi f_c t)$. The frequency components of s_2 include $f_c - f_m$ and $f_c + f_m$. Due to the non-linearity of speakers, the frequency components of the signal played by the first speaker (i.e., s_{out1}) are f_c and $2f_c$. Since they are both in the inaudible range, s_{out1} is inaudible to humans. The output of the second speaker is:

$$\begin{aligned} s_{out2} =& A_1 s_2 + A_2 s_2^2 \\ =& A_1 cos(2\pi f_m t)cos(2\pi f_c t) \\ &+ A_2 (cos(2\pi f_m t)cos(2\pi f_c t))^2. \end{aligned} \tag{8}$$

From Equation 8, we find that the frequency components are $f_c - f_m$, $f_c + f_m$, $2f_c - 2f_m$, $2f_c + 2f_m$, $2f_m$, and $2f_c$. Since f_c ($f_c \geq 2f_m$) is the frequency of ultrasound carrier, $f_c - f_m$, $f_c + f_m$, $2f_c - 2f_m$, $2f_c + 2f_m$, and $2f_c$ are all inaudible to humans. To guarantee $2f_m$ is inaudible, we set $2f_m$ within the near-ultrasound band.

7.3 Residual Signal Elimination

To extend the communication range of the proposed system, it is necessary to carefully design the techniques without reducing the signal strength for data transmission. However, the aforementioned residual signal elimination scheme (as mentioned in Section 3.3) would downgrade the strength of the signal carrying data. Hence, we turn to another direction, i.e., carefully choosing data-carrying frequency band and ultrasound carrier frequency, for ensuring the transmission signal strength while eliminating the interference of residual signals.

In particular, at the receiver side, a microphone records the outputs of the two ultrasound speakers, i.e., $s_{out1} + s_{out2}$. Based on the non-linearity of the microphone, the frequency components of the received signal include an informative signal (i.e., f_m), high-frequency residual signals (i.e., $f_c + f_m$, $f_c + 2f_m$, and $f_c + 3f_m$), and low-frequency residual signals (i.e., $2f_m$, $f_c - f_m$, $f_c - 2f_m$, and $f_c - 3f_m$). Among them, high-frequency residual signal can be automatically filtered out by the microphone's low-pass filter. On the other hand, the impacts of low-frequency residual signals can be eliminated by distinctly separate them from the informative signal, without interfering with each other. The key technique is to choose a separating frequency f_s and the carrier frequency f_c for independent data transmission on AM-modulated band and directly transmitted band. Specifically, assume $f_m \in [f_l, f_s]$. To separate $2f_m$ from the informative signal, f_s should satisfy:

$$f_s \leqslant 2f_l. \tag{9}$$

Moreover, other low-frequency residual signals, i.e., $f_c - f_m$, $f_c - 2f_m$, and $f_c - 3f_m$, can be eliminated through carefully choosing f_c to avoid overlapping with f_m, i.e.,

$$\begin{cases} f_c - f_m \geqslant f_s, \\ f_c - 2f_m \leqslant f_l, \\ f_c - 3f_m \leqslant f_l. \end{cases} \tag{10}$$

By solving Equation 10, we can obtain $2f_s \leqslant f_c \leqslant 3f_l$. Combined with the aforementioned analysis, we could choose the f_c satisfying $2f_s \leqslant f_c \leqslant 3f_l$, and the f_s meeting $f_s \leqslant 2f_l$. With these choices, the informative signal would not be overlapped by other residual signals, and can be recovered at the receiver side.

8 PERFORMANCE EVALUATION

8.1 Experimental Setup & Methodology

Device and Setting. To evaluate the performance of BatComm, we implement three settings of the transmitter to meet various application requirements. In the frontend of these settings, we use three types of ultrasound speakers (i.e., a commercial ultrasound speaker, Vifa [5], a customized ultrasound speaker with 44-transducer array, and a customized ultrasound speaker with 2 transducers) and two types of amplifiers to transmit ultrasound signals, as shown in Figure 8(a)-(c). To evaluate the impact of power, we use different kinds of amplifiers for these setups. The output power of Vifa is $10W$, while the other two customized speakers are amplified by an Adafruit amplifier module [14] with $3.6W$ output. In addition, we include two types of transducer arrays with different numbers of transducers to study their impact on our system. We use commercial off-the-shelf mobile devices (i.e., a Galaxy S6, a Galaxy Note 4, and a Samsung Tab P7510) as receivers to evaluate BatComm. We use the primary microphone, which is located at the bottom of the devices, to record the acoustic signal. Unless otherwise mentioned, the operation band for OFDM subcarriers is 8.06-23.53kHz, the digital modulation scheme is 16DPSK, the carrier frequency f_c for AM is $40kHz$ and error correction is based on $(63, 45)$-BCH code. We use the COTS ultrasound speaker and a Galaxy S6 phone as our

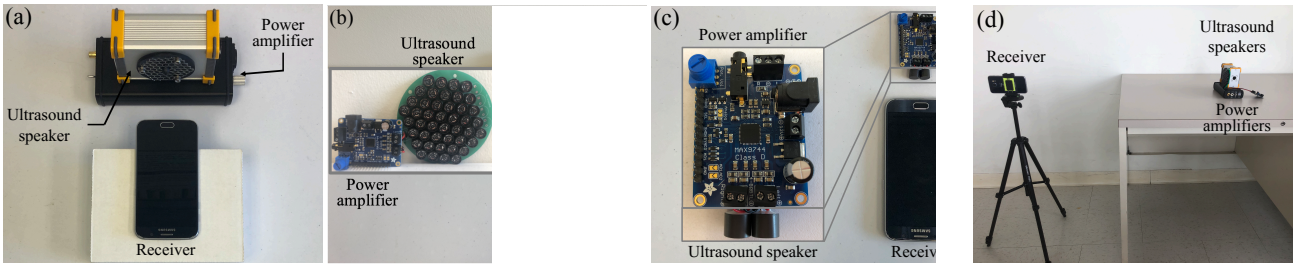


Figure 8: Illustration of four experimental settings with different types and numbers of ultrasound speakers (power cables are not included). (a) COTS ultrasound speaker (Vifa); (b) Customized 44-transducer array; (c) Customized 2-transducer array; and (d) Two ultrasound speakers for room-scale communication.

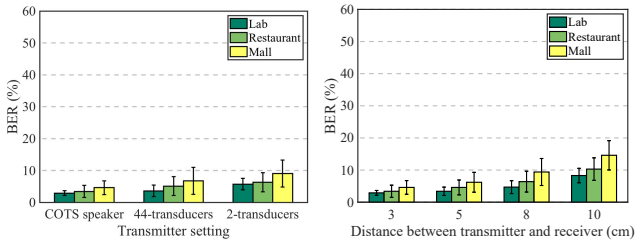


Figure 9: BER of BatComm under different (a) transmitter settings and environments (b) distances between transmitter and receiver (Throughput = 34.13$kbps$**).**

default transmitter-receiver setting, unless mentioned otherwise. The distance between the ultrasound speaker and the receiver (i.e., smartphone) is 3cm, which is natural and appropriate for a near-field communication application. We also evaluate longer distances (e.g., 5, 8, and 10cm), which is discussed in Section 8.2. In addition, we evaluate the performance of room-scale communication using two ultrasound speakers as the transmitter, i.e., Vifa and a customized ultrasound speaker with 44-transducer array, as shown in Figure 8(d).

Data Collection. In order to evaluate BatComm and visually check the occurrence of errors, we transmit various grayscale images (i.e., 6.4kB for each) from the transmitter to the receiver. In each round of transmission, the transmitter transmits the modulated data through ultrasound signals, and we repeatedly conduct 20 rounds of transmissions for each environmental setup with various communication parameter settings. To test the impact of environmental noises on BatComm, we perform the experiments in three representative environments: lab, restaurant, and mall. In the lab environment, there are HVAC noises with an average noise level of 39.7$dBSPL$. In the restaurant environment, sources of noise include people making orders and eating, averaging at a noise level of 57.3$dBSPL$. The noise level of the mall environment is even higher, with people walking and chatting over background music. Experiments held in the mall lobby recorded an average noise level of 80.2$dBSPL$.

Evaluation Metrics. We mainly use three metrics to evaluate the performance of BatComm. (1) *Human Perception*: Sound pressure level measurements and subjective audibility ratings (i.e., user study) are used to measure the human perception, which are presented as $dBSPL$ and a score from 0-10, respectively; (2) *Throughput*: Assume a data stream of D bits is transmitted from the transmitter to the receiver with a time of T seconds. The throughput of acoustic communication is defined as $\frac{D}{T}$ bits per second (bps); and (3) *Bit Error Rate (BER)*: Assume the system transmits n_t bits digital data. Due to noise, interference, distortion or bit synchronization errors, n_e bits data is altered during the communication. The BER is defined as $\frac{n_e}{n_t}$, which is presented as a percentage.

8.2 Overall System Performance

With the operation parameters in Section 8.1, BatComm can achieve a throughput of 34.13$kbps$, which is over 8× higher than the state-of-the-art solutions. Note that different environments and transmitter settings do not affect the throughput of the communication system.

Human Perception. We conduct experiments to verify BatComm is beyond ordinary human hearing perception, both quantitatively and qualitatively, through sound pressure measurements and user studies, respectively. The lower bounds of human hearing thresholds have been documented to be around 40$dBSPL$ [6], at which point most sounds cannot be perceived. In our experiments, we measure the sound pressure in a quiet room using a Holdpeak 882a digital sound level meter [18]. We find the pressure to be 31.7$dBSPL$ if BatComm is not transmitting data. When our system is functioning, we measure sound levels at a distance of 1m from our ultrasound speaker. With three types of implemented ultrasound speakers (i.e., the ones shown in Figure 8), we find the sound levels to be 32.6$dBSPL$, 32.1$dBSPL$, and 31.8$dBSPL$, respectively, all of which are below the accepted threshold for human hearing. Additionally, we recruit 10 volunteers to evaluate the audibility of our system in the same environment. Each volunteer is positioned 1m, 2m, and 3m from the ultrasound speaker and asked to provide subjective audibility ratings for his/her perceived noises, with 0 being silence, 1 being perceived as rustle of leaves and 10 as the perception in noisy factories [4]. In every tested configuration, all participants report silence (i.e., 0) during the communication process, suggesting BatComm is inaudible to human hearing.

Interference with Nearby Microphone-based Applications. Due to the same modality, BatComm may interfere with microphone-based applications nearby (e.g., phone calls, video recordings). Similar to human perception, we quantitatively and qualitatively validate whether BatComm can coexist with other applications. We use a sound level meter to measure the sound pressures at different ranges in the lab. When the range is 1m, the sound levels are 41.2$dBSPL$, 41.1$dBSPL$, and 39.8$dBSPL$ for the three types of sound speakers respectively, which are only a little higher than that of environmental noises (i.e., 39.7$dBSPL$). In a 2m range, the

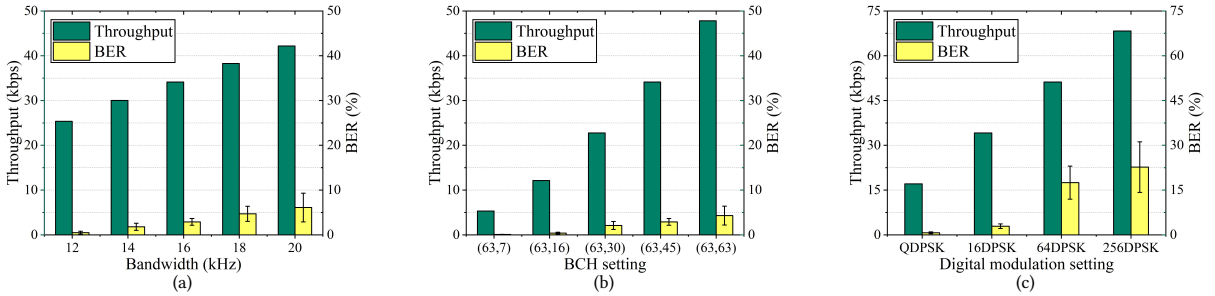
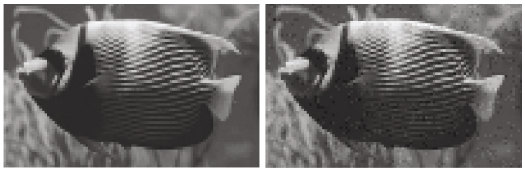


Figure 10: Performance of BatComm with different (a) bandwidth, (b) BCH settings, (c) digital modulation settings.



(a) Fish (transmitted) (b) Fish (received)

Figure 11: Image comparison between transmitted (i.e., (a)) and received images (i.e., (b)), where BERs is 5.7% (Throughput = 34.13$kbps$).

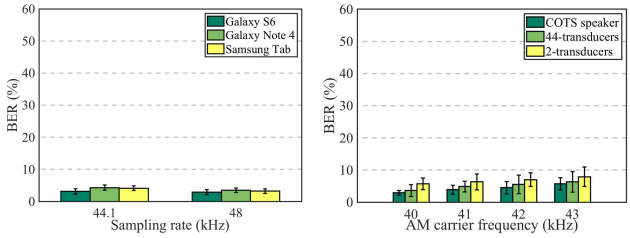


Figure 12: BER of BatComm under different (a) sampling rates and receiver devices (b) AM carrier frequencies.

sound levels of recorded signals, i.e., 39.8$dBSPL$, 39.8$dBSPL$, and 39.7$dBSPL$, are almost the same as the environmental noise level. Additionally, we use Galaxy S6 to record three pieces of music (60 seconds) in 1m when BatComm is on and off, respectively, and ask 10 volunteers to differentiate if the music pieces are recorded during our experiment. Only one volunteer reports that the music pieces are different, but cannot tell where each music source come from (i.e., whether BatComm is on or off). This result indicates that BatComm can hardly cause noticeable interference in most cases.

Environments and Transmitter Settings. We evaluate the BER of BatComm under different environments and transmitter settings, as shown in Figure 9(a). It can be observed that the average BER of BatComm in the quiet lab is 2.9% with standard derivation less than 1%. For the 44-transducer array and 2-transducer array settings, the average BERs are 3.6% and 5.7%, respectively. Figure 11 shows a set of transmitted and received images during the experiments. We find that although the BER for the image transmission is 5.7%, the received image is clear for human recognition. This result indicates that such a BER is satisfactory for acoustic communication. Moreover, all three settings achieve a satisfactory performance (i.e., BER is less than 6%), indicating that BatComm can use various ultrasound speakers at the transmitter end. Additionally, the system performance degrades with the increase of background noise levels. However, BatComm can still achieve an average BER of around 8% in the noisy mall environment (with a noise level around 80$dBSPL$), which indicates our system is robust to different environments. As described in Section 5.3, to keep our system resilient to daily noises whose frequencies are less than 8kHz, we choose the operation band 8.06-23.53kHz, which cannot be interfered by common sound sources (e.g., human speaking, music).

Transmitter-Receiver Distances and Environments. We also evaluate the impact of distance between the transmitter and receiver on BatComm. Figure 9(b) shows the BER of BatComm under different distances and environments. We observe that the BER of our system slightly increases as the distance increases. This is because as the propagation distance increases, SNR of the acoustic signals decreases, which induces more errors in the recorded signals. In the quiet lab environment, the average BER is around 5% with a standard deviation of less than 2% under a distance between the transmitter and receiver less than 8cm. Even for the mall, the BER is 6.2% under the distance of less than 5cm. These results indicate that BatComm can achieve high throughput with an acceptable BER for almost all the near-field applications.

8.3 Impact of OFDM Bandwidth

The bandwidth for OFDM multiplexing directly affects the throughput of the acoustic communication system. Also, since the background noises would impact the audible frequency band used in OFDM, the bandwidth, especially the lower bound of bandwidth, affects the BER of acoustic communication system. Hence, we evaluate the performance of BatComm under different bandwidths (i.e., different lower bound of the bandwidth). Figure 10(a) shows the throughput and BER of BatComm under different bandwidths. We can see that the throughput increases as the OFDM bandwidth increases, which is consistent with the theoretical analysis. Moreover, it can be observed that the BER also slightly increases as the OFDM bandwidth increases. This is because as the bandwidth increases, the lower bound of bandwidth decreases. Hence, the wider bandwidth introduces more errors in recorded signals and would increase the BER of the system. However, even for the 20kHz bandwidth of OFDM operation (4-24kHz), the average BER of BatComm achieves 6.1% with a standard deviation of 3.2%, while the throughput of BatComm can achieve 42.18$kbps$, which indicates satisfactory performance.

8.4 Impact of BCH Code

In our system, BCH code is used to mitigate the unpredicted errors in recorded signals. However, different BCH settings affect the performance of the acoustic communication system. Particularly, the digital data is encoded with (N, K)-BCH code, where N is the length of encoded data bits, K is the length of original digital data bits. Figure 10(b) shows the throughput and BER of BatComm under different BCH settings. We observe that without BCH error correction (i.e., (63, 63) BCH setting), BatComm can achieve a high throughput of $47.49kbps$, and the average BER is only 4.3% with a standard deviation of 2.1%. By adding more BCH coding bits for error correction, both throughput and BER would decrease. To meet specific application requirements, we can use some particular BCH settings to achieve a near-zero BER (e.g., $5.31kbps$ with 0.1% BER; $12.13kbps$ with 0.4% BER).

8.5 Impact of Digital Modulation

In BatComm, DPSK modulates data bits into phase values for the digital modulation. Figure 10(c) shows the throughput and BER of BatComm under different digital modulation settings. We can see that both the throughput and BER increase when the digital modulation setting changes from QDPSK to 256DPSK. Specifically, for QDPSK and 16DPSK, the average BERs of our system are 0.7% and 2.9% with a standard deviation of 0.3% and 0.8%, respectively, which are satisfactory for acoustic communication. However, for 64DPSK and 256DPSK, although the throughputs are quite high (i.e., larger than $50kbps$), the achieved BERs are not quite satisfactory (i.e., higher than 17%). This is because when the digital modulation setting changes from QDPSK to 256DPSK, the number of data bits modulated by each phase increases, which narrows the difference between adjacent phases. This makes the data bits modulated in phase values highly possible to be misjudged.

8.6 Impact of Sampling Rate

To ensure that BatComm is capable of transmitting data to most mobile devices, we evaluate the performance of our system using three mobile devices (i.e., Galaxy S6, Galaxy Note 4, and a Samsung Tab P7510) as receivers with different sampling rates (i.e., $44.1kHz$ and $48kHz$). According to the Nyquist theorem, different sampling rates lead to different bandwidth for communication, which affects the throughput. Specifically, the throughput of BatComm achieves $34.13kbps$ and $30.01kbps$ under $48kHz$ and $44.1kHz$ sampling rate, respectively. The BERs of our system under different receiver sampling rates are shown in Figure 12(a). We observe that the comparable low BERs can be achieved under different sampling rates and device models. Specifically, the overall BERs for the three receiver models are 3.2% and 3.8% for $48kHz$ and $44.1kHz$ sampling rate, respectively. These results demonstrate that BatComm is suitable for most commercial off-the-shelf mobile devices.

8.7 Impact of AM Carrier Frequency

As mentioned in Section 5.5, due to the limited frequency response of ultrasound speakers, the ultrasound carrier frequency in AM affects the performance of acoustic communication system. Since the AM carrier frequency f_c should satisfy $f_c \geq 40kHz$ to avoid the interference between ultrasound signal and automatically demodulated signals, we evaluate the performance of BatComm under different carrier frequencies larger than $40kHz$. Figure 12(b) shows the BER of BatComm under different AM carrier frequencies. We can see that BER increases proportionally to AM carrier frequency. This is because the response of the ultrasound speaker decreases as the signal frequency increases, which leads to a lower SNR of the received signal, and is consistent with our theoretical analysis.

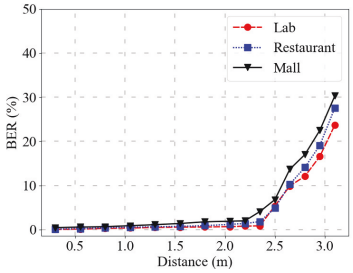


Figure 13: BER of room-scale system under different environments.

8.8 Room-Scale Communication Performance

To evaluate the effectiveness of the proposed scheme, we conduct experiments for the performance evaluation in a room-scale setting. As shown in Figure 8(d), we implement the proposed scheme with two ultrasound speakers (i.e., Vifa and a customized speaker with 44-transducer array) for the experiment. A dual-channel waveform generator (Keysight 33500b series [42]) is used to drive the speakers. We use a commercial mobile device (i.e., Galaxy S6) as receiver, where the primary microphone located at the bottom of the device is employed to record the signal. The operation band for OFDM subcarriers is 12.04-23.53kHz (i.e., subcarrier #257 to #502). We separate the band to AM-modulated band (i.e., 12.04-18.00kHz) and directly transmitted band (i.e., 18.04-23.53kHz), and choose the carrier frequency f_c for AM as $36kHz$, following the scheme as mentioned in Section 7.3. The digital modulation scheme is QDPSK, and error correction is not used. The data collection process is the same as that in Section 8.1.

Figure 13 shows the performance of BatComm under different distances between the transmitter and receiver in different environments. We observe that the communication range can be up to $2.3m$ with $17.76kbps$ throughput without significant BER increase. In lab environment, the BER remains below 1% within $2.3m$. The BER only slightly increases within the certain distance, and follows by a sudden increase to around 5%. This is because as the propagation distance increases, the signal strength decreases, which induces more errors in the recorded signals. We also measure the BERs under two kinds of noises to simulate two representative environments: restaurant and mall. We play the noises from a nearby speaker at the intensities of $55.4dBSPL$ and $72.6dBSPL$ (which are typical sound levels in the two environments) respectively for the environment simulation. In both environments, the BER remains under 2% within $2m$, which shows similar performance with that in lab environment. These indicate satisfactory performance in resisting ambient noises. The reason lies in the difference frequency band of noises and the proposed communication system. Through spectral density analysis of noises, we find the main frequency

components are within $4kHz$, but our operation band lies above $12kHz$.

9 DISCUSSION

Ultrasound Perception by Vulnerable Groups. Certain demographics may be sensitive to ultrasound signals, such as pets, young children, or people afflicted with medical conditions such as hyperacusis. BatComm achieves inaudible acoustic communication using $40kHz$ acoustic signals, which may induce discomfort in such groups. According to existing studies [17, 22, 45], frequencies of $40kHz$ and sound intensity above $40dBSPL$ may be perceived by those with sensitive hearing. However, the sound intensity of BatComm is lower than $40dBSPL$ within $1m$ distance, as demonstrated in Section 8.2, minimizing potential disturbances.

Further Improvement of Performance. While our findings demonstrate BatComm is capable of achieving unprecedented levels of throughput for an acoustic communication system, we acknowledge several aspects of our implementation can be further improved upon. In the context of related acoustic communication systems, the current BER is still a nontrivial concern on performance that warrants further reduction. To do so, we intend to explore techniques such as the use of dual microphones at the receiver side, as well as develop reliable, error-checking methods modeled on existing protocols such as transmission control protocol (TCP). Specifically, most commercial mobile devices are equipped with multiple microphones, which are used for stereo recording and noise cancellation. There is great potential in using these microphones to reduce interference during acoustic communication, as verified by Jeub et al. [20], and recording-quality calibration.

Bidirectional Communication. The aforementioned design details of BatComm concentrate on realizing the physical layer of a unidirectional acoustic communication system, with ultrasound speakers as the transmitter and commercial mobile devices as the receiver, which achieve the high-throughput and inaudibility simultaneously. In addition to the broadcasting manner, users may demand a bidirectional channel, i.e., the data transmission could be forwarded from either side to the other. BatComm can be easily extended to support bidirectional communication between commercial devices, as long as they are both equipped with ultrasound speakers. With the on-going development of mobile devices, we believe ultrasound speakers will become commonplace in off-the-shelf mobile devices in the near future, taking advantage of ultrasound-based communication.

10 RELATED WORK

Audible Acoustic Communication. Early work [27] evaluated the impact of digital modulation techniques, such as ASK and FSK, on human perception to achieve human-pleasant communication with throughput up to $400bps$. Then, PriWhisper [55] and Dhwani [33] used jamming techniques with self-interference cancellation at the receiver to provide a secure acoustic communication channel between devices with throughputs of $1kbps$ and $2.4kbps$. The communication ranges are $15cm$ [55] and $30cm$ [33], respectively. Audible studies may be vulnerable to interference from ambient noises and can distract the user.

Inconspicuous Acoustic Communication. To improve the user experience, other bodies of work disguised transmissions by embedding data in daily sounds for acoustic communications. For instance, existing studies embed data bits into music pieces imperceptibly through OFDM multiplexing [29], audio files by modifying the phases of original signal [54], and daily sounds without arousing human perception [48]. The throughput achieved are around $40bps$ [29], $600bps$ [54], and $500bps$ [48], and the communication ranges are $4m$ [29], $2m$ [54], and $10m$ [48]. These approaches hide data in daily sounds, but may not be applicable in quiet environments.

Inaudible Acoustic Communication. Some existing studies used the near-ultrasound band (i.e., 17-$22kHz$) to achieve inaudible acoustic communication. They employed the chirp binary orthogonal keying [24], chirp quaternary orthogonal keying [21], and Gaussian minimum-shift keying [39] to implement a near-ultrasound acoustic communication for mobile and wearable devices. These studies enable $25m$ [24], $2.7m$ [21], and $1m$ [39] communication range. However, these studies can only achieve low throughputs (i.e., $15bps$ [24], $16bps$ [21], and $2.76kbps$ [39]) due to the limited frequency band for inaudible acoustic communication, which is unsatisfactory for many emerging applications. BackDoor [37] has shown initial success in leveraging the non-linearity of microphones for inaudible acoustic communication. This work achieves up to $1m$ communication range with $4kbps$ throughput. Different from BackDoor, we mainly focus on near-field acoustic communication and achieve a much higher throughput of over $47kbps$. We also manifest the possibility of increasing the range to $2.3m$ with around $17kbps$ throughput.

11 CONCLUSION

We proposed BatComm, a new acoustic method that integrates OFDM and AM techniques with microphone non-linearity to achieve high-throughput and inaudible acoustic communication. BatComm modifies the OFDM symbol waveform before AM to counteract interference from residual signals and improve performance. This allows for much higher throughput rates than existing proposals, opening the possibility for greater integration in practical services such as mobile advertisement and payment. Additionally, to mitigate noise in practical scenarios, a series of interference reduction techniques (e.g., DPSK digital modulation, pilot-based channel estimation, BCH error correction code, and interleaving) are integrated into BatComm to improve robustness. Extensive experiments demonstrated that BatComm achieved a throughput as high as $47.49kbps$, which is $12\times$ higher than contemporary state-of-the-art acoustic communication systems within a $10cm$ range. We also explored the possibility of increasing the range up to $2.3m$ with comparable throughput (i.e., $17.76kbps$). We hope that elevating the possible throughput and range in acoustic communication will inspire other researchers to follow suit, leading to new robust techniques, use cases, and discoveries.

12 ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for the constructive comments to improve the paper. This work was partially supported by the National Science Foundation Grants CNS1820624 and CNS1814590.

REFERENCES

[1] Muhammad Taher Abuelma'atti. 2003. Analysis of the effect of radio frequency interference on the DC performance of bipolar operational amplifiers. IEEE Transactions on Electromagnetic compatibility 45, 2 (2003), 453–458.

[2] Amazon. 2020. Bluetooth Transmitter Receiver. [Online]. Available: https://www.amazon.com/MP3-Player-Bluetooth-Transmitters/.

[3] Kaoru Ashihara. 2007. Hearing thresholds for pure tones above 16 kHz. The Journal of the Acoustical Society of America 122, 3 (2007), EL52–EL57.

[4] BCcampus. 2019. Sound Intensity and Sound Level. [Online]. Available: https://opentextbc.ca/physicstestbook2/chapter/sound-intensity-and-sound-level/.

[5] Avisoft Bioacoustics. 2019. Ultrasonic Dynamic Speaker Vifa. [Online]. Available: http://www.avisoft.com/usg/vifa.htm.

[6] Michael Borgers. 2019. Why 40dB is silence for human. [Online]. Available: https://www.improvestudyhabits.com/best-noise-level-for-studying/.

[7] R.C. Bose and D.K. Ray-Chaudhuri. 1960. On a class of error correcting binary group codes. Information and Control 3, 1 (1960), 68 – 79.

[8] Fabio Buckell. 2019. Should You Buy a Cheap Low-End Android Phone? [Online]. Available: https://www.maketecheasier.com/cheap-low-end-android-phone/.

[9] Gordon KC Chen and James J Whalen. 1981. Comparative RFI performance of bipolar operational amplifiers. In IEEE International Symposium on Electromagnetic Compatibility. Boulder, CO, USA, 1–5.

[10] Jiajun Jim Chen and Carl Adams. 2004. Short-range wireless technologies with mobile payments systems. In Proceedings of the 6th international conference on Electronic commerce. 649–656.

[11] Shaoping Chen and Cuitao Zhu. 2004. ICI and ISI analysis and mitigation for OFDM systems with insufficient cyclic prefix in time-varying channels. IEEE Transactions on Consumer Electronics 50, 1 (2004), 78–83.

[12] Digi-Key. 2019. Price of Speakers. [Online]. Available: https://www.digikey.com/products/en/audio-products/speakers/156/page/5?k=speaker.

[13] EDN. 2014. Basic principles of MEMS microphones. [Online]. Available: https://www.edn.com/basic-principles-of-mems-microphones/.

[14] Adafruit Electronics. 2019. Adafruit Stereo Audio Amplifier - MAX9744. [Online]. Available: https://learn.adafruit.com/adafruit-20w-stereo-audio-amplifier-class-d-max9744.

[15] Google. 2016. Google Tone. [Online]. Available: https://chrome.google.com/webstore/detail/google-tone/nnckehldicaciogcbchegobnafnjkcne?hl=en.

[16] Mahmoud Reza Hashemi and Elahe Soroush. 2006. A Secure m-Payment Protocol for Mobile Devices. In In Proceedings of Canadian Conference on Electrical and Computer Engineering. IEEE, Ottawa, Ont., Canada, 294 – 297.

[17] H. Heffner. 1983. Hearing in large and small dogs: Absolute thresholds and size of the tympanic membrane. Behavioral Neuroscience 97 (Jan 1983), 310–318.

[18] Holdpeak. 2019. Holdpeak 882A Digital Sound Level Meter. [Online]. Available: https://www.amazon.com/gp/product/B07NYVLCM5/ref=ppx_yo_dt_b_asin_image_o00_s00?ie=UTF8&psc=1.

[19] IEEE. 2016. IEEE Standard for Information technology–Telecommunications and information exchange between systems Local and metropolitan area networks–Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. IEEE Std 802.11-2016 (Revision of IEEE Std 802.11-2012) (2016), 1–3534.

[20] Marco Jeub, Christian Herglotz, Christoph Nelke, Christophe Beaugeant, and Peter Vary. 2012. Noise reduction for dual-microphone mobile phones exploiting power level differences. In Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Kyoto, Japan, 1693–1696.

[21] Soonwon Ka, Tae Hyun Kim, Jae Yeol Ha, Sun Hong Lim, Su Cheol Shin, Jun Won Choi, Chulyoung Kwak, and Sunghyun Choi. 2016. Near-ultrasound communication for TV's 2nd screen services. In Proceedings of Annual International Conference on Mobile Computing and Networking (MobiCom). New York, NY, USA, 42–54.

[22] Udi Katzenell and Samuel Segal. 2001. Hyperacusis: review and clinical guidelines. Otology & neurotology 22, 3 (2001), 321–327.

[23] HJ Landau. 1967. Sampling, data transmission, and the Nyquist rate. Proc. IEEE 55, 10 (1967), 1701–1706.

[24] Hyewon Lee, Tae Hyun Kim, Jun Won Choi, and Sunghyun Choi. 2015. Chirp signal-based aerial acoustic communication for smart devices. In Proceedings of IEEE Conference on Computer Communications (INFOCOM). Hong Kong, China, 2407–2415.

[25] Sheena Leek and George Christodoulides. 2009. Next-generation mobile marketing: how young consumers react to bluetooth-enabled advertising. Journal of advertising research 49, 1 (2009), 44–53.

[26] Jian Liu, Hongbo Liu, Yingying Chen, Yan Wang, and Chen Wang. 2019. Wireless sensing for human activity: A survey. IEEE Communications Surveys & Tutorials (2019).

[27] Cristina Videira Lopes and Pedro MQ Aguiar. 2001. Aerial acoustic communications. In IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics. New Platz, NY, USA, 219–222.

[28] Mathworks. 2019. Interleaving. [Online]. Available: https://www.mathworks.com/help/comm/ug/interleaving.html.

[29] Hosei Matsuoka, Yusuke Nakashima, and Takeshi Yoshimura. 2006. Acoustic communication system using mobile terminal microphones. NTT DoCoMo Tech. J 8, 2 (2006), 2–12.

[30] Heinrich Meyr, Marc Moeneclaey, and Stefan Fechtel. 1997. Digital communication receivers: synchronization, channel estimation, and signal processing. John Wiley & Sons, Inc.

[31] Keiichi Mizutani, Naoto Wakatsuki, and Koichi Mizutani. 2008. Differential phase shift keying acoustic communication in air for low-signal-to-noise ratio environment. Japanese Journal of Applied Physics (2008), 6526.

[32] P.A. Monte, R.M. Tanner, and Santa Cruz. Computer Research Laboratory University of California. 1988. A Table of Efficient Truncated BCH Codes. Computer Research Laboratory, University of California, Santa Cruz. https://books.google.com/books?id=zPofAQAAIAAJ

[33] Rajalakshmi Nandakumar, Krishna Kant Chintalapudi, Venkat Padmanabhan, and Ramarathnam Venkatesan. 2013. Dhwani: secure peer-to-peer acoustic NFC. ACM SIGCOMM Computer Communication Review 43, 4 (2013), 63–74.

[34] Electronics notes. 2019. Bluetooth network connection and pairing. [Online]. Available: https://www.electronics-notes.com/articles/connectivity/bluetooth/network-pairing-connection.php.

[35] Paytm. 2019. Paytm.com. [Online]. Available: https://paytm.com/.

[36] PrintingNews. 2019. Case Study: Case of Pricing QR Codes. [Online]. Available: https://www.printingnews.com/events/article/10257799/.

[37] Nirupam Roy, Haitham Hassanieh, and Romit Roy Choudhury. 2017. Backdoor: Making microphones hear inaudible sounds. In Proceedings of Annual International Conference on Mobile Systems Applications and Services (MobiSys). Niagara Falls, NY, USA, 2–14.

[38] Nirupam Roy, Sheng Shen, Haitham Hassanieh, and Romit Roy Choudhury. 2018. Inaudible voice commands: The long-range attack and defense. In Proceedings of USENIX Symposium on Networked Systems Design and Implementation (NSDI). Renton, WA, USA, 547–560.

[39] G Enrico Santagati and Tommaso Melodia. 2017. A software-defined ultrasonic networking framework for wearable devices. IEEE/ACM Transactions on Networking 25, 2 (2017), 960–973.

[40] ShopNFC. 2019. NFC Readers and Writers. [Online]. Available: https://www.shopnfc.com/en/17-nfc-readers-writers.

[41] Techcrunch. 2013. Alipay Launches Sound Wave Mobile Payments System In Beijing Subway. [Online]. Available: https://techcrunch.com/2013/04/14/alipay-launches-sound-wave-mobile-payments- system-in-beijing-subway/.

[42] Keysight Technologies. 2019. Keysight 33500b Series. [Online]. Available: https://www.keysight.com/us/en/assets/7018-03536/data-sheets/5991-0692.pdf.

[43] Samuli Tiiro, Jari Ylioinas, Markus Myllyla, and Markku Juntti. 2009. Implementation of the least squares channel estimation algorithm for MIMO-OFDM systems. In Proceedings of the International ITG Workshop on Smart Antennas (WSA). Germany, 16–18.

[44] ToneTag. 2019. Cashless & Contactless Payments Solution | Homepage | ToneTag. [Online]. Available: https://www.tonetag.com/.

[45] Sandra E Trehub, Bruce A Schneider, Barbara A Morrongiello, and Leigh A Thorpe. 1989. Developmental changes in high-frequency sensitivity: Original papers. Audiology 28, 5 (1989), 241–249.

[46] Verifone. 2019. Verifone.com. [Online]. Available: https://www.verifone.com/en/us.

[47] Chen Wang, Yan Wang, Yingying Chen, Hongbo Liu, and Jian Liu. 2020. User authentication on mobile devices: Approaches, threats and trends. Computer Networks 170 (2020), 107118.

[48] Qian Wang, Kui Ren, Man Zhou, Tao Lei, Dimitrios Koutsonikolas, and Lu Su. 2016. Messages behind the sound: real-time hidden acoustic signal capture with smartphones. In Proceedings of Annual International Conference on Mobile Computing and Networking (MobiCom). New York, NY, USA, 29–41.

[49] GF West and JC Macnae. 1991. Physics of the electromagnetic induction exploration method. In Electromagnetic Methods in Applied Geophysics: Volume 2, Application, Parts A and B. Society of Exploration Geophysicists, 5–46.

[50] Wikipedia. 2019. Cross-correlation. [Online]. Available: https://en.wikipedia.org/wiki/Cross-correlation.

[51] Wikipedia. 2019. Differential Phase Shift Keying. [Online]. Available: https://en.wikipedia.org/wiki/Phase-shift_keying.

[52] Wikipedia. 2020. Hamming code. [Online]. Available: https://en.wikipedia.org/wiki/Hamming_code.

[53] Wikipedia. 2020. Reed-Solomon error correction. [Online]. Available: https://en.wikipedia.org/wiki/Reed-Solomon_error_correction.

[54] Hwan Sik Yun, Kiho Cho, and Nam Soo Kim. 2010. Acoustic data transmission based on modulated complex lapped transform. IEEE Signal Processing Letters 17, 1 (2010), 67–70.

[55] Bingsheng Zhang, Qin Zhan, Si Chen, Muyuan Li, Kui Ren, Cong Wang, and Di Ma. 2014. PriWhisper: Enabling Keyless Secure Acoustic Communication for Smartphones. IEEE Internet of Things Journal 1, 1 (2014), 33–45.