Acoustic Eavesdropping Attack Using Self-Mixing Laser Interferometer

Kohei Doi
tokushi.comp@gmail.com
The University of Electro-Communications
Chofu, Tokyo, Japan

Kunihiko Ooi
kunihiko.ooi@uec.ac.jp
The University of Electro-Communications
Chofu, Tokyo, Japan

Takeshi Sugawara
sugawara@uec.ac.jp
The University of Electro-Communications
Chofu, Tokyo, Japan

Abstract

Laser microphone is a remote eavesdropping device that illuminates a laser beam on a target object and reconstructs acoustic vibration by analyzing reflected light. In 2022, Laser Meager Listener showed the advantage of a laser microphone using a laser Doppler vibrometer (LDV), but the need for a scientific-grade LDV increases the attack cost and limits the flexibility. This paper addresses the issues by using self-mixing interferometry (SMI), which can realize an LDV with a single laser diode and simple electronics. Our proof-of-concept SMI-based laser microphone is evaluated in an end-to-end attack scenario considering an attacker eavesdropping sound from a target object 3 meters away through a glass window. We evaluate the intelligibility of sound measured from five target objects under three propagation conditions, and the proposed method achieves a performance comparable to the previous attack using scientific-grade LDV. The proposed method can be extended to an invisible attack using an infrared laser, and the attack cost can be further reduced with our open circuit designs. We finally discuss the physical parameter that limits the attack distance.

CCS Concepts

• **Security and privacy → Hardware attacks and countermeasures**.

Keywords

Acoustic Eavesdropping, Laser Doppler Vibrometer, Self-Mixing Interferometer

ACM Reference Format:
Kohei Doi, Kunihiko Ooi, and Takeshi Sugawara. 2024. Acoustic Eavesdropping Attack Using Self-Mixing Laser Interferometer. In *Proceedings of the 2024 Workshop on Attacks and Solutions in Hardware Security (ASHES '24), October 14–18, 2024, Salt Lake City, UT, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3689939.3695786

1 Introduction

Voice and sound have been the crucial means of human communication. Acoustic eavesdropping has been a known threat and is even more important today with the spread of teleconferencing and voice assistants. Consequently, researchers are studying eavesdropping attacks exploiting acoustic, electromagnetic, and optical side-channels [6, 12, 21, 24, 30, 46].

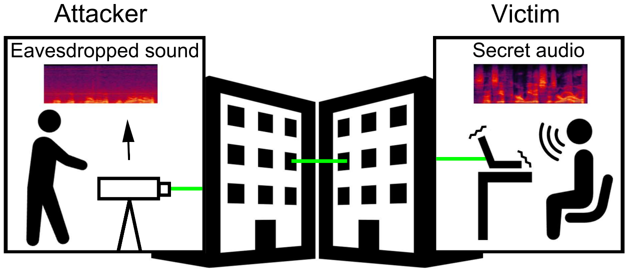

Figure 1: Illustration of an attack scenario of laser microphones. An attacker remotely measures the vibration of a target object with a laser and steals the confidential audio

A laser microphone is a remote eavesdropping device that illuminates a laser beam on a target object and obtains acoustic vibration by analyzing reflected light. A laser microphone takes advantage of a laser beam that can travel a longer distance and penetrate transparent objects (e.g., a glass window) compared to acoustic vibration in the air, as shown in Figure 1. Using invisible light, the attacker can stealthily hear sound in a remote location. The technique has been known for many years but has not been openly discussed in the research community until recently. Knowledge was monopolized by military and intelligence agencies, raising concern for possible abuse, for example, Government Communications Headquarters (GCHQ) demanded journalists to discard classified documents concerning laser microphones [42].

Addressing the issue, researchers began to openly study the feasibility of eavesdropping attacks using laser microphones [26, 46]. In 2022, Laser Meager Listener [46] studied the feasibility of eavesdropping attacks using an instrument called the laser Doppler vibrometer (LDV). When a laser beam emitted from an LDV is reflected at a target object in motion, the frequency of the reflected light is shifted by the Doppler effect. The LDV recovers the frequency shift, which represents the target object's vibration, from the reflected light. Using an LDV, Laser Meager Listener [46] recovered sound from environmental objects, such as a glass window and a cup.

The commercial LDVs are designed for scientific measurement in a laboratory, and Laser Meager Listener [46] used one of such scientific-grade LDVs for verifying the attack. However, the use of such a scientific-grade instrument increases the attack costs and reduces the attack flexibility. First, commercial LDVs are expensive and controlled by export laws. Second, commercial LDVs use

ASHES '24, October 14–18, 2024, Salt Lake City, UT, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1235-7/24/10
https://doi.org/10.1145/3689939.3695786

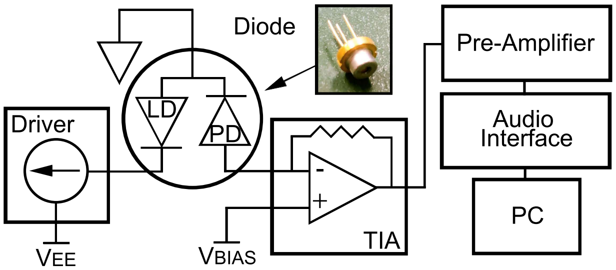

Figure 2: Block diagram of our SMI-LDV setup. A laser current driver drives the anode-grounded laser diode. The photocurrent from the photodiode in the same laser diode package is converted to a voltage signal using a transimpedance amplifier and finally captured by a PC.

low-power and visible lasers for manual aiming, and they are unconfigurable. Third, commercial LDVs are large and heavy, limiting the attack scenario that requires portability.

Addressing the above limitations, this paper tackles the following research questions. *Can an attacker improvise a cheap and compact LDV? If so, does the cheap LDV achieve a performance comparable to that of commercial LDVs in eavesdropping attacks?* Commercial LDVs commonly use the principle called optical heterodyne detection (OHD), and building a cheap OHD-based LDV (OHD-LDV) seems to be challenging because OHD requires expensive optical components, such as beam splitter cubes, prisms, and a frequency shifter precisely assembled on an optical base plate.

We approach the problem using an LDV with the difference detection principle called self-mixing interferometry (SMI) [8, 9, 11, 26, 27, 41]. An SMI-based LDV (SMI-LDV) can be realized with a single laser diode (LD) and simple electronics, as shown in Figure 2; expensive and heavy optical components are unnecessary.

In an SMI-LDV, a LD inside the package emits a laser beam toward a target object and receives reflected light containing the Doppler shift, in the same way as an OHD-LDV. Then, the reflected light is fed back to the LD package through the same optical path. The emitting and reflected lights cause constructive and destructive interference inside the LD package, demodulating the Doppler shift as light intensity. The changes in the light intensity can be measured using a photodiode (PD) typically present inside an LD package.

The electronics necessary for SMI-LDV are simple, as shown in Figure 2: a power supply (a laser current driver) to drive the LD and a transimpedance amplifier (TIA) to convert the photocurrent from the PD into a voltage signal. SMI-LDV is also flexible. We can easily increase the laser power through the current driver. Also, an attacker can choose the color of light, including invisible infrared light, by buying an appropriate LD from an online store.

1.1 Contributions

We study the feasibility of the cheap laser microphone attack by building a cheap SMI-LDV and verifying the attack feasibility. This paper includes the following key contributions.

Attack Feasibility and Capability Characterization (Sections 3 and 4). We build a proof-of-concept SMI-LDV setup from a single LD. We verify the attack feasibility and then characterize the impacts of (i) laser power, (ii) sound volume, and (iii) distance on the sound quality.

End-to-End Evaluation (Section 5). SMI-LDV's performance is evaluated in an end-to-end attack scenario considering an attacker eavesdropping sound from a target object 3 meters away through a glass window. We examine five targets (a beverage can, a charger, a smartphone case, an LCD panel, and a laptop) with three propagation conditions (adjoining, same surface, and aerial). The audio files recovered from our end-to-end experiment are available in the repository[1] The intelligibility and the word error rate after speech-to-text conversion of the audio files are evaluated and compared with those of the previous attack using OHD-LDV.

Attack Extensions, Limitations, and Mitigation (Section 6). The feasibility of the stealthy attack is experimentally verified by replacing an LD with another one emitting infrared light in our SMI-LDV. Moreover, custom circuit boards are designed to reduce the attack cost even further. We also discuss the physical parameter that limits the attack distance and possible countermeasures that prevent or mitigate the proposed attack.

2 Preliminary

We briefly summarize laser microphones and the other optical eavesdropping attacks.

2.1 Laser Doppler Vibrometer (LDV) and Its Application to Laser Microphone

An LDV remotely measures the vibration of a target object by emitting a laser light toward a target object and measuring the reflected light. When the emitted laser light hits a vibrating object, the frequency (i.e., wavelength) of the reflected light shifts by the Doppler effect. LDV extracts the frequency shift that represents the vibration by comparing the emitted and reflected frequencies using an interferometer.

Commercial LDVs commonly use OHD for demodulating the frequency shift. The block diagram in Figure 3 explains the principle of OHD. The original laser beam is split into the measuring and reference beams with the beam splitter (BS1). The measuring beam is aimed at the target object, and the reflected beam is directed through another beam splitter (BS2). Meanwhile, the reference beam passes through a prism and a frequency shifter (a Bragg cell). When the two beams are recombined in yet another beam splitter (BS3), they cause constructive and destructive interference, and the intensity of the combined light represents the frequency difference between the two beams. Finally, the light intensity is measured at the photo detector. Commercial LDVs are designed for scientific measurements and are used in laboratories to measure the vibration of a target object remotely and precisely.

Since sound is a form of vibration, an LDV can be used as a remote listening device, which has been studied for a long time in nonsecurity contexts [18, 23]. More recently, Laser Meager Listener [46] rigorously studied the threat of an acoustic eavesdropping attack by capturing the vibrations of environmental objects (e.g., paper cup) using one of a scientific-grade OHD-LDV.

[1] https://osf.io/74cbv/

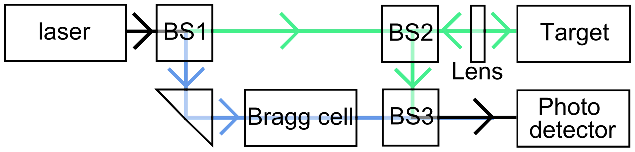

Figure 3: Block diagram of an LDV using optical heterodyne detection (OHD). The original laser beam is split into the measuring and reference beams with the beam splitter BS1. The measuring and reference beams go through the different optical paths and recombined at the beam splitter BS3. The optical interference between the two beams is measured at the photo detector.

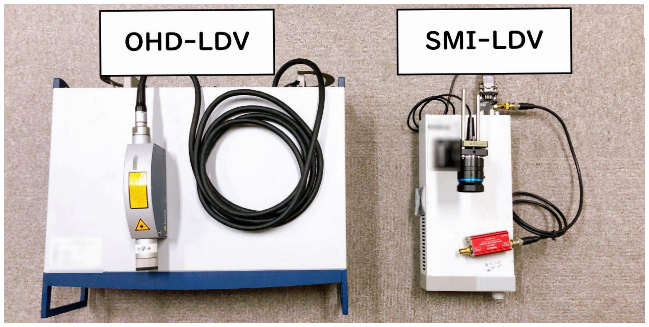

Figure 4: Commercial OHD-LDV (left) and our SMI-LDV setup (right).

Although Laser Meager Listener pioneered the application of LDV as an eavesdropping device, the commercial OHD-LDV applies several restrictions to the attack.

Availability: Commercial LDVs typically cost more than $20,000 [2] and are only available to well-funded attackers. Moreover, they are controlled by export laws in many countries.

Stealthiness: Commercial LDVs use visible lasers because a human operator manually aims a laser beam, and the visibility, i.e., laser wavelengths, is unconfigurable. This limits the stealthiness of the attack because a victim can see the laser spot on a target object[3].

Laser Power: For the same reason as described above, commercial LDVs use a low-power laser considering eye safety. As we will see laser, a higher laser power is necessary for a long-range attack, but such a high-power LDV is unavailable.

Portability: Commercial LDVs are large and heavy because they are designed for a static use case in a laboratory. The system in Figure 3 is realized with discrete optical components assembled on an optical base plate. For example, our OHD-LDV used as a baseline in this paper is 450×360×150 mm and 11.4 kg (see Figure. 4) because laboratory instruments prioritize accuracy over portability. This is in contrast to the eavesdropping scenario in which stealthiness is the priority.

2.2 Laser Microphones by Triangulation

In addition to LDVs, a laser microphone can be realized with triangulation. In this method, an attacker emits a laser beam toward a target object from an angle and receives the reflected light beam. The position of the reflected laser beam represents the displacement of the target object. We can measure the position of the reflected beam with a linear image sensor or its approximation, such as a photodetector with a pinhole filter.

The triangulation method can be inexpensive compared to LDV because it does not involve interferometry. However, as a drawback, the transmitter and receiver must be placed at a right angle, and attack is possible only when such a location is available. Furthermore, this method has a fundamental limitation in a long-range attack. The laser spot in the receiver should be small to distinguish positions, and the method becomes impossible when the diameter of the beam inevitably diverges over a distance. Therefore, this paper follows the LDV-based approach.

2.3 Optical Eavesdropping Attacks

Another approach to an optical eavesdropping attack involves capturing a scene using image sensors. In particular, the visual microphone reconstructs sound by measuring vibrating tiny objects (e.g., foil, tissue) using a high-speed camera [7]. As a drawback, however, the visual microphone requires an expensive high-speed camera, and its recording time is usually limited by the size of an on-camera RAM. Side Eye addresses the issue by exploiting the rolling-shutter effect [25], enabling a camera-based acoustic eavesdropping attack using smartphone cameras. The other papers pursued eavesdropping attacks assuming light-emitting target objects instead of actively injecting a laser. Lamphone [29] listens to sound by measuring the light emitted from a light bulb vibrated by environmental sound. The Little Seal Bug in 2023 [30] extends Lamphone by considering light reflected on lightweight reflective objects, such as an empty beverage can. These attacks can be conducted without a laser setup, but an appropriate light source is necessary in the target room as an additional requirement. Moreover, the attack distance is fundamentally limited by the nature of the incoherent, diffused light source.

3 Proposed Method

3.1 Threat Model

Our threat model follows the previous LDV-based attack, i.e., Laser Meager Listener [46]. The attacker's goal is to eavesdrop on sound in an acoustically isolated location, as illustrated in Figure. 1. The attacker achieves this by measuring the vibration of a reflective target object, induced by a sound source, using an LDV.

To use an LDV, the attacker needs a line of sight towards the target object. The attack is possible as far as the laser light can travel, potentially across transparent obstacles, such as a glass window. The attacker uses the same line of sight to aim the laser spot at the target object while watching the scene with a telescope. The attacker optionally uses an infrared camera to aim an invisible IR laser. The attacker has no control over the target room, and the attacker cannot put a reflective sticker. For effective eavesdropping

[2] Commercial LDVs do not usually disclose the catalog prices, and they are available only upon quotation.

[3] Some commercial LDVs use an IR laser for its advantage in measuring dark and rough surfaces, but they still use a coaxial visible laser for aiming.

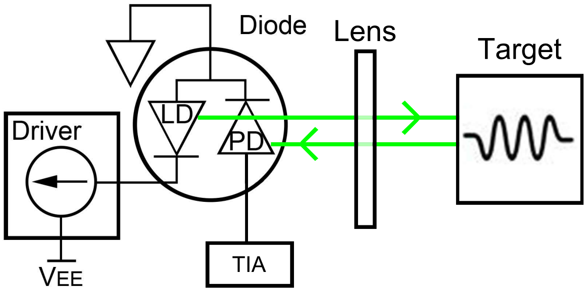

Figure 5: Block diagram for LDV using self-mixing interferometry (SMI). The reflected laser light is fed back to the LD package. A built-in PD meansures constructive and destructive interference happens within the diode package.

attacks, the attacker optionally assumes that a highly reflective and lightweight (natural) object is present in the room.

A typical attack scenario occurs in an office environment with a glass window in which a person or a laptop speaker generates sound with confidential information. The attacker is present in the next room/building separated by a glass window and uses the line of sight to aim the laser toward a reflective target object, such as a beverage can and a metallic backplate of a laptop computer.

3.2 Self-Mixing Interferometer for Acustic Eavesdropping

SMI is a simple interferometer using a single LD [8, 9, 11, 26, 27, 41]. The principle of SMI has been studied since the 1980s [22], and has several sensing applications, including displacement [31], velocity [36], distance [32], and vibration [27]. SMI is based on optical interference between the light emitting and reflected within the laser cavity inside the LD package, as illustrated in Figure 5.

An LD package commonly has an LD and an integrated PD; the PD continuously monitors the light emission from the LD so that an external controller can stabilize the laser power with feedback control. In an SMI-LDV, a laser beam from the LD is aimed at a target object, and the reflected light containing the Doppler shift is fed back to the same LD package. Inside the LD package, the emitting and reflected light causes constructive and destructive interference, and the frequency difference is converted to the light intensity. This principle is called self-mixing. Finally, the built-in PD generates a photocurrent that represents the light intensity.

With SMI-LDV, interferometric detection of the doppler shift is achieved by a single LD only, which is in contrast to OHD in Figure 3. The electronics necessary are also simple: a laser current driver to drive the LD and a TIA to convert the photocurrent to a voltage signal. As a result, our eavesdropping attack using SMI-LDV addresses the problems in the previous OHD-LDV-based attack [46] as follows.

Availability: An attacker can improvise its own SMI-LDV using off-the-shelf electronic components purchased from online resellers. SMI-LDV is also inexpensive. Our setup in Figure 4-(right) reduces the attack cost by an order of magnitude, despite the use of the off-the-shelf laser current driver and TIA. Further cost reduction is possible by replacing them with custom circuit boards, as we will discuss in Section 6.2.

Stealthiness: An LD has standard packages and is easily replaceable. The LDs with different wavelengths are available in the market, including an invisible infrared LD. By choosing a laser with the desired wavelength, the attacker can improve stealthiness with invisible light or penetrate certain obstacles with selective transmittance, such as colored glasses. Since we conduct our experiments with a visible laser for safety requirements, we separately verify the feasibility with an infrared laser in Section 6.1.

Laser Power: An attacker can easily improve the laser power by increasing the driving current. For example, our green LD (Osram PLT5 520B, available for less than $50) can emit up to 110 mW of laser power, which is in contrast to our OHD-LDV (Polytec NLV-2500) limited to less than 1 mW.

Portability: SMI-LDV is smaller, as compared in Figure 4. The off-the-shelf laser driver occupies the most space, and it can be smaller with a custom circuit board (see Section 6.2).

4 Attack Feasibility and Characterization

This section verifies the feasibility of eavesdropping attack using an SMI-LDV, followed by the characterization of physical parameters that impact the attacker capability.

4.1 Attack Feasibility

Setup. We build our SMI-LDV by following the diagram in Figure 2. The picture in Figure 6 shows the optical part of the setup. The laser source is a 520nm green LD (Osram PLT5 520B) mounted on a socket (Thorlabs SR9). With 40 mA of driving current, the LD emits 2.5 mW of optical power (measured with a Thorlabs PM100D laser power meter). We use a C-mount camera lens (Raspberry Pi SC0123) as the focusing optics. The diode and lens are assembled in the optical cage system (Thorlabs CP1LM56/M and CP13/M). The laser socket and the laser current driver have standard DB9 connectors. We design a custom plug-in adapter in Figure 10 to intercept the PD terminal. The PD terminal is accessible through an SMA connector, which is connected to a TIA (Thorlabs AMP102). The PD is biased with 6.6 V. The photocurrent from the PD is converted to a voltage signal with the TIA. In this feasibility experiment, we measured the voltage signal using a spectrum analyzer (SignalHound SA44B) configured with 100kHz resolution and video bandwidths.

Procedure. We verify the feasibility by measuring a vibration speaker[4] with a reflective tape, as shown in Figure 6. The vibration speaker is driven by a 5 V_{pp} linear chirp signal from a function generator (Rigol DG1022Z) that sweeps 0.1–20 kHz in 20 seconds. The sound level emitted is 60 dB (measured with a Rion NL-06 noise meter) placed 0.5 meters away from the speaker. The distance between the speaker and the LDV is 0.5 meters.

Results. Figure 7 shows the spectrogram captured in the spectrum analyzer, where the vertical and horizontal axes are time and frequency, respectively. The spectrogram shows a straight line representing the chirp emitted from the speaker. This result verifies that our simple LDV setup successfully captures the vibration.

[4] https://amzn.asia/d/0VkWlXJ

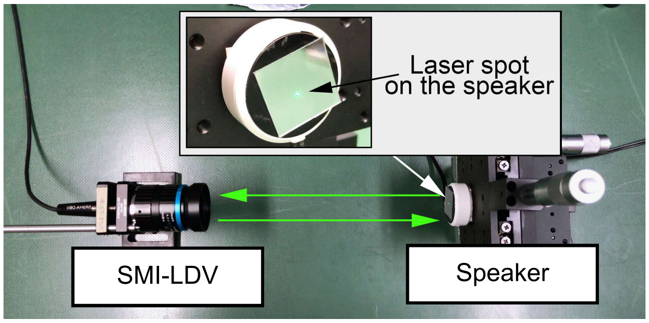

Figure 6: Setup for the SMI-LDV feasibility and characterization experiments. The laser beam from the SMI-LDV is aimed at the reflective tape on the target vibration speaker. The reflected beam comes back to the SMI-LDV on the same optical path.

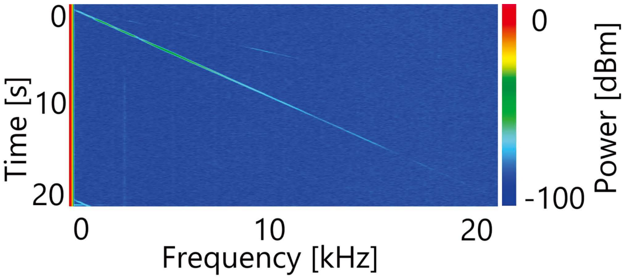

Figure 7: Spectrogram from our SMI-LDV setup aimed at the vibration speaker. The straight line represents the sound from the speaker, the chirp signal swept between 0.1–20 kHz in 20 seconds.

4.2 Attacker Capability Characterization

By following the basic verification in the previous section, we characterize the attacker's capability by repeating the experiments with different physical parameters and evaluating the signal-to-noise ratio (SNR) and the intelligibility of the captured audio.

Setup. We first describe the baseline parameters. The target is the same vibration speaker with a reflective tape. We use the same SMI-LDV in Section 4.1, wherein PD is biased with 6.6 V. For comparison, we make the same measurement using an off-the-shelf OHD-LDV (Polytec NLV-2500), shown in Figure 4. The sensitivity of OHD-LDV is set to 10mm/s/V. The distance between the LDVs and the speaker is 0.5 meters. The setup is contained in a laser enclosure unless otherwise noted.

Signal-to-Noise Ratio (SNR) Evaluation. We evaluate SNR by playing a tone signal from the speaker. The frequency is 940 Hz, which is the most efficient within the frequency range of human voice (0.1–2 kHz) found in the previous experiment. The peak-to-peak voltage driving the speaker is 5V. We measure the TIA output with the spectrum analyzer (see Section 4.1) and obtain the peak-to-peak output voltages with and without the tone, namely V_{signal} and V_{noise}. Then, SNR is obtained as SNR $= 10 \cdot \log_{10} \frac{V_{\text{signal}}}{V_{\text{noise}}}$

Word Error Rate (WER) Evaluation. We also evaluate WER by playing natural spoken language. Following the previous work [46], we use the sound source from the Audio Quality Testing Stimuli [16, 33] based on the Harvard Sentence [1]. We use two audio sources from [33]: (i) "Four hours of steady work faced us." spoken by a male voice and (ii) "The birch canoe slid on the smooth planks." spoken by a female voice. In this section, the LDV measures the vibration speaker while playing the female voice.

The voltage signal from the TIA is converted to an audio file with the setup in Figure 2. We first amplify the voltage signal with a preamplifier (Stanford Research SR560) configured with ×200 amplification gain and 0.03–10 kHz band-pass filtering. The preamplifier output is captured with an audio interface (Focusrite Scarlett 2i2) and recording software (Audacity). We finally apply basic noise reduction using Audacity.

The recorded audio is converted to text using the speech-to-text functionality of the messaging application (Slack). When an output of the speech-to-text includes W_c correct words, W_i inserted words, W_r wrong words, and W_d deleted words, the WER is given by WER $= \frac{W_i+W_r+W_d}{W_c}$.

4.2.1 Laser Power. We first evaluate the impact of laser power by changing it from 1 to 30 mW. Table 1 summarizes SNR and WER for each laser power. The OHD-LDV achieves 16.14 dB of SNR at its fixed laser power of 1 mW, which is better than our SMI-LDV with the same laser power. However, the SMI-LDV's SNR improves as we increase the laser power, and it exceeds that of OHD-LDV beyond 10 mW. As a result, SMI-LDV can be better than the commercial OHD-LDV in SNR by compensating for the low SNR with more laser power. We note that the laser power of 10 mW is conservative because the LD tolerates up to 110 mW.

4.2.2 Sound Volume. We then assess the impact of the target speaker's sound volume by changing it between 40–70 dB measured with the noise meter placed 1 meter away from the speaker. The laser power in SMI-LDV is 30 mW. Table 2 summarizes the SNR and WER results for each sound volume.

We obtain a higher SNR with a louder sound. The result is natural because a louder sound volume means a large displacement on the target object, which results in a larger frequency shift. The results are aligned with the previous work [30, 46]. The SMI-LDV and OHD-LDV behave similarly to different sound volumes. WER is robust against the changes in sound volume.

4.2.3 Distance. We evaluate the impact of the distance by changing the LDV-to-target distance between 2–10 meters. The volume of the audio source is 70 dB. This experiment is conducted outside the laser enclosure, and 10 meters is the maximum length available in our laboratory environment. To meet the laser safety requirement for open-air laser emission, we limit the laser power to 5 mW (cf. 30 mW), which is the power of commercial laser pointers [34].

The experimental results summarized in Table 3 show that the distance has a minor impact on the SNR in the examined range of 2–10 meters, and SMI has 0% WER up to 10 meters.

Table 1: SNR and WER with various laser power. The laser power of SMI-LDV is changed from 1 to 30 mW. The laser power of the commercial OHD-LDV is fixed to 1 mW.

LDV Type	Laser Power	SNR	WER
SMI	1 mW	9.607 dB	0%
SMI	5 mW	13.37 dB	0%
SMI	10 mW	16.26 dB	0%
SMI	15 mW	17.24 dB	0%
SMI	20 mW	17.24 dB	0%
SMI	25 mW	17.32 dB	0%
SMI	30 mW	17.32 dB	0%
OHD	1 mW	16.14 dB	0%

Table 2: SNR and WER with various sound volume at the target vibration speaker. The sound volume in dB is measured with a noise meter at 1 meter away.

LDV Type	Sound Volume	SNR	WER
SMI	40 dB	15.12 dB	0%
SMI	50 dB	16.96 dB	0%
SMI	60 dB	17.31 dB	0%
SMI	70 dB	17.87 dB	0%
OHD	40 dB	14.15 dB	0%
OHD	50 dB	15.76 dB	0%
OHD	60 dB	16.08 dB	0%
OHD	70 dB	16.16 dB	0%

Table 3: SNR and WER for different LDV-to-target distances between 2–10 meters.The laser power of the SMI-LDV is fixed to 5 mW.

LDV Type	Distance	SNR	WER
SMI	2 m	15.19 dB	0.0%
SMI	4 m	14.82 dB	0.0%
SMI	6 m	14.79 dB	0.0%
SMI	8 m	14.58 dB	0.0%
SMI	10 m	14.79 dB	0.0%
OHD	2 m	16.00 dB	0.0%
OHD	4 m	16.03 dB	0.0%
OHD	6 m	16.02 dB	0.0%
OHD	8 m	16.02 dB	0.0%
OHD	10 m	16.02 dB	0.0%

5 End-to-End Evaluation

Based on the feasibility and attacker capability verified in the previous experiments, we perform an end-to-end evaluation following the attack scenario in the previous work [46].

5.1 Experimental Setup

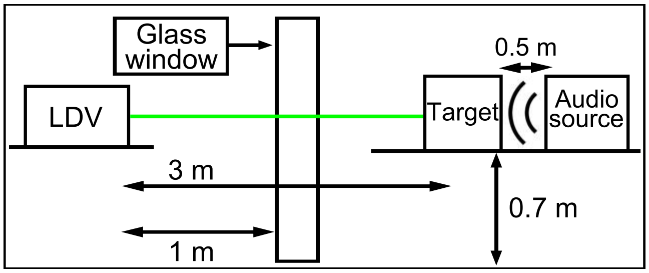

Figure 8: SMI-LDV setup for the end-to-end evaluation. A target object is placed 3m away from the SMI-LDV with a glass windows in between. We examine three propagation conditions, and the figure shows the Same Surface case.

Figure 8 shows the setup representing an attack scenario in which an attacker measures the target object near the audio source across a glass window. The LDV, the target object, and the audio source are placed on tables at the same height. The distance between the LDV and the target object is 3 meters, which is classified as *far* in the previous work [46]. There is a glass between the LDV and the target.

We use the same SMI-LDV setup as in the previous section, except for the focusing optics replaced with a single lens (Thorlabs C240TMD-A) optimized for the target distance. The laser power is limited to 5 mW [34] for open-space laser emission.

The audio source is a speaker [40] driven by an audio amplifier (ELEGIANT F900). The speaker plays the male and female voices described in section 4.2 The volume of sound measured with the noise meter placed 1 meter away is 60 dB, which is a normal range for human conversations [45]. The experiments are carried out in a normal office with an environmental noise level of 45 dB.

We examine the five objects with reflective surfaces in Figure 9: (1) an empty beverage can, (2) a charger, (3) a smartphone case, (4) an display panel of laptop, and (5) a back panel of laptop[5]. The following three conditions are considered regarding the connection between the speaker and the target object.

Adjoining: The sound source and the target object have physical contact. An example includes the chassis of a laptop generating sound from its speakers. We evaluate this condition by placing a vibration speaker on the target object.

Same Surface: The sound source and the target object do not have physical contact but are placed on the same surface. An example is a target object and a loudspeaker placed on the same desk. The speaker and target object are 0.5 meters away from each other.

Aerial: The sound source and the target object are isolated and connected only through the air. This is the most challenging setting for an attacker. The speaker and target object are 0.5 meters away from each other.

In addition to WER (see Section 4.2), we also evaluate the intelligibility score with Taal et al.'s method [39] using its open-source implementation [28]. The intelligibility score has a value between 0

[5] From the targets evaluated in Laser Meager Listener [46], we exclude less reflective objects, such as a paper cup. The attack success rate with these objects was already low in the previous work; it was only 65.4% even with an OHD-LDV, a large sound volume (over 70 dB), and on the same surface. We focus on objects with a reflective surface as a viable target to assess the differences between SMI-LDV and OHD-LDV.

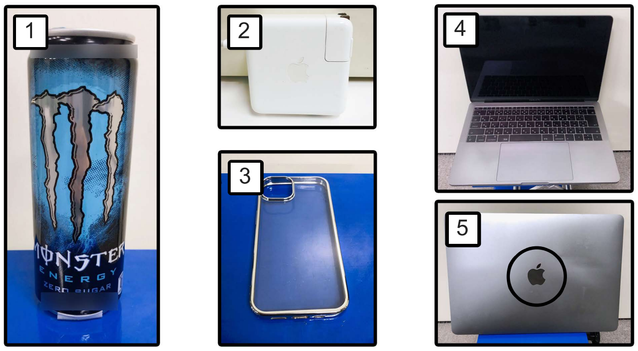

Figure 9: The target objects: (1) an empty beverage can, (2) a charger, (3) a smartphone case, (4) a display panel of a laptop, and (5) a back panel of a laptop.

and 1, and most people successfully recognize 90% of the words [39] with the score > 0.7.

5.2 Results

The recorded audio files are available from the online storage[6] Figure 14 shows the frequency spectrogram obtained from the laptop back panel with the male voice, which clearly shows the patterns correlated with the original audio. Table 4 summarizes the intelligibility and WER. The table highlights the good values: the intelligibility scores > 0.7 and the WER with ≤2 wrong words.

Adjoining. We first compare the OHD-LDV and SMI-LDV under the adjoining condition. Intelligibility scores are high in most cases, and the number of wrong words is below 2 in all cases with both the OHD-LDV and the SMI-LDV. The beverage can is the exceptional case having >0% WER with the SMI-LDV. The difficulty in measuring the beverage can is aligned between the different LDVs and the different conditions.

Same Surface. Although the sound recovery becomes more challenging with the Same Surface condition, both the SMI-LDV and the OHD-LDV achieve high intelligibility scores and low WER. The difference between target objects becomes clear under this condition, and the smartphone case, the charger, and the laptop are particularly efficient, achieving > 0.7 intelligibility score and 0% WER with the SMI-LDV and with the male voice.

Aerial. Aerial is the most challenging condition and the intelligibility score is less than 0.7 in most cases in both SMI-LDV and OHD-LDV, as shown in Table 4. The intelligibility scores with SMI-LDV are generally lower, making critical differences in WER. For a laptop with a female voice, for example, the OHD-LDV and the SMI-LDV have 25.0% and 100.0% of WER, respectively. Meanwhile, SMI-LDV achieves a comparable result in a particular condition, e.g., 28.6% WER in the case with the laptop and the male voice.

Comparison. Table 5 compares the results with previous work, namely Laser Meager Listener [46] and Little Seal Bug [30]. Compared with Laser Meager Listener, our experiments obtained a better WER both in the Same Surface and in the Aerial conditions. This is because Laser Meager Listener targeted less reflective objects such as a styrofoam cup; the reflective targets in Figure 9 are selected because of those previous results. The results with the OHD-LDV in Table 4 represent those of Laser Meager Listener against the reflective objects. Although OHD-LDV always gets better results, SMI-LDV achieves comparable performance as discussed above.

Compared with Little Seal Bug, the proposed method achieves a higher intelligibility score and a lower WER in the Aerial condition even with a more challenging condition, i.e., a longer distance and a lower sound volume. This is attributed to the directivity of a laser beam compared to diffused light in Little Seal Bug. Moreover, the proposed method does not require a light source near the target, as we discussed in Section 2.3.

6 Discussion

6.1 Invisible Attack

Section 4 and Section 5 verify the proposed method using a visible green laser due to safety requirements regarding the open-air laser emission. This section verifies the feasibility of a stealthy attack using an infrared laser invisible to human eyes.

Setup. We repeat the feasibility experiment in Section 4.2.1 by replacing the light source with an invisible 940nm infrared LD (Rohm Semiconductor, RLD94-PZJ5). We examine the laser power 1, 5, and 10 mW. This experiment is conducted inside a laser enclosure.

Results. Table 6 summarizes the SNR and WER with 1, 5, and 10 mW of laser power. For comparison, the table also shows the ones with the green laser under the same conditions. The table shows that the infrared laser always achieves better performance both in SNR and WER. The results verify that the proposed method can be extended to an invisible attack by simply replacing an LD.

6.2 Attack Cost Reduction

Although our setup in Figure 4-(right) reduces the attack cost by an order of magnitude from commercial OHD-LDVs, the setup can be even cheaper and smaller. We verify the feasibility of the proposed method with an inexpensive setup by replacing the off-the-shelf components with our open circuit design. The schematic of the circuit we designed is described in the appendix A.

Experimental Evaluation. Figure 12 in appendix A shows our cheap setup using the custom-made laser current driver and the TIA discussed above. We verify the feasibility of the attack using this cheap setup by repeating the end-to-end evaluation in Section 5 targeting the laptop in the adjoining condition. Table 7 compares the intelligibility and WER values between the original and cheap setup. The results show the feasibility of retrieving intelligible sound from the target using the cheap setup. The sound quality is degraded from the previous setup; our cheap setup is intended for a proof-of-concept, and it can be improved in many ways.

6.3 Attack Distance and its Limitation

We discuss the physical parameters that limit the attack distance. While the transmitted laser beam travels through the air efficiently, the condition of the reflected light is determined by the target's

[6] https://osf.io/74cbv/

Table 4: Intelligibility and WER in the end-to-end evaluation. The table highlights (i) intelligibility > 0.7 with which most people successfully recognize 90% of the words [39] and (ii) WER values with up to 2 wrong words.

Medium	Target	Voice	Intelligibility		WER	
			OHD	SMI	OHD	SMI
Adjoining	Smartphone Case	Female	**0.8831**	**0.7459**	**0.0**%	**0.0**%
Adjoining	Smartphone Case	Male	**0.8334**	**0.7069**	**0.0**%	**0.0**%
Adjoining	Charger	Female	**0.9390**	**0.8214**	**0.0**%	**0.0**%
Adjoining	Charger	Male	**0.9142**	**0.8599**	**0.0**%	**0.0**%
Adjoining	Laptop	Female	**0.8802**	**0.8147**	**0.0**%	**0.0**%
Adjoining	Laptop	Male	**0.8452**	**0.7167**	**0.0**%	**0.0**%
Adjoining	LCD Panel	Female	**0.8698**	**0.7830**	**0.0**%	**0.0**%
Adjoining	LCD Panel	Male	**0.8373**	**0.7767**	**0.0**%	**0.0**%
Adjoining	Beverage Can	Female	**0.7101**	0.6600	**0.0**%	**25.0**%
Adjoining	Beverage Can	Male	**0.7204**	0.5619	**0.0**%	**14.3**%
Same Surface	Smartphone Case	Female	**0.7554**	**0.7026**	**0.0**%	**12.5**%
Same Surface	Smartphone Case	Male	**0.7100**	**0.7073**	**0.0**%	**0.0**%
Same Surface	Charger	Female	**0.7600**	**0.7183**	**0.0**%	**0.0**%
Same Surface	Charger	Male	**0.7004**	**0.7102**	**0.0**%	**0.0**%
Same Surface	Laptop	Female	**0.7068**	**0.7092**	**0.0**%	**25.0**%
Same Surface	Laptop	Male	**0.7191**	**0.7035**	**0.0**%	**0.0**%
Same Surface	LCD Panel	Female	**0.7010**	0.6184	**12.5**%	37.5%
Same Surface	LCD Panel	Male	**0.7064**	0.5390	**0.0**%	28.6%
Same Surface	Beverage Can	Female	0.6305	0.2370	75.0%	100.0%
Same Surface	Beverage Can	Male	0.5857	0.2443	100.0%	100.0%
Aerial	Smartphone Case	Female	0.5484	0.3725	92.3%	100.0%
Aerial	Smartphone Case	Male	0.5917	0.3336	100.0%	100.0%
Aerial	Charger	Female	0.3686	0.3035	100.0%	100.0%
Aerial	Charger	Male	0.6065	0.3739	**28.6**%	100.0%
Aerial	Laptop	Female	0.6941	0.5312	**25.0**%	100.0%
Aerial	Laptop	Male	0.6061	0.5812	**28.6**%	**28.6**%
Aerial	LCD Panel	Female	0.3009	0.2884	100.0%	100.0%
Aerial	LCD Panel	Male	0.3128	0.3125	100.0%	100.0%
Aerial	Beverage Can	Female	0.5450	0.3829	100.0%	100.0%
Aerial	Beverage Can	Male	0.3855	0.3196	100.0%	100.0%

Table 5: Comparison with Laser Meager Listener [46] and Little Seal Bug [30].

Ref.	Reflective Target?	Medium	Distance†	Sound Volume	Intelligibility (Max.)	WER (Min.)
Little Seal Bug [30] (Table. I)	Yes	Same Surface	(2.5, 0.05)	95 dB	0.79	–
Laser Meager Listener [46] (Table. 1)	No	Same Surface	(3.0, 0.5)	60 dB	–	100%
Laser Meager Listener [46] (Table. 1)	No	Same Surface	(3.0, 0.5)	> 70 dB	–	34.6%
Ours	Yes	Same Surface	(3.0, 0.5)	60 dB	0.72	0.00%
Little Seal Bug [30] (Table. IV)	Yes	Aerial	(2.5, 0.25)	75 dB	0.46	–
Laser Meager Listener [46] (Table. 1)	No	Aerial	(3.0, 0.5)	60 dB	–	92.6%
Laser Meager Listener [46] (Table. 1)	No	Aerial	(3.0, 0.5)	> 70 dB	–	82.8%
Ours	Yes	Aerial	(3.0, 0.5)	60 dB	0.58	28.6%

†The tuple represents the LDV-to-target and the target-to-speaker distances in meters, respectively.

surface and is uncontrollable, which limits the attack distance in both the SMI-LDV and OHD-LDV.

When the transmitted laser hits a smooth and flat surface, it makes specular reflection, and the reflected light travels efficiently

Table 6: SNR and WER results using the invisible, infrared laser with 1, 5, 10 mW laser power. The results with the visible, green laser Section 4.2.1 are also shown for comparison.

Laser Type	Laser Power	SNR	WER
Infrared (940 nm)	1 mW	14.33 dB	0%
Infrared (940 nm)	5 mW	14.70 dB	0%
Infrared (940 nm)	10 mW	15.25 dB	0%
Green (520 nm)	1 mW	8.32 dB	0%
Green (520 nm)	5 mW	12.67 dB	0%
Green (520 nm)	10 mW	13.55 dB	0%

as a laser beam. In this case, the distance has a low impact on sound quality, as we saw in our distance experiment (Section 4.2) wherein the target (a reflective tape, as shown in Figure 6) makes efficient specular reflection. The charger, the smartphone case, and the laptop have smooth and flat surfaces that make specular reflection, and they are relatively easy targets as shown in Table 4. Meanwhile, the LCD panel is more difficult because it has a rough surface that randomly reflects the incoming laser light. The resulting reflected light quickly diffuses over a distance and limits the attack distance. While the beverage can has a smooth surface, the curved surface makes a divergent reflected beam, resulting in lower intelligibility scores and higher WER. As a result, the attack distance heavily depends on the target surface condition, and a target objective that makes an efficient specular reflection is ideal for the attacker.

6.4 Countermeasures

Conventional countermeasures against laser microphones are still effective against SMI-LDV-based eavesdropping attacks. As discussed in the previous section, avoiding reflective objects is a common defense strategy. Choosing objects with less reflective surfaces is an effective administrative control. Also, we can cover a highly reflective part of the target with less reflective material, e.g., putting a sticker on the logo on the laptop (see Figure 9).

Blocking the line of sight between the attacker and the target object is another important defense strategy. Attacks are no longer possible when secret conversation takes place in a room without a glass window. As suggested by the authors of Little Seal Bug [30], blocking the line of sight with opaque obstacles, such as curtains and window shades will make the attack harder or impossible. As a variant, putting an IR filter on a window will prevent the stealthy attack without losing the window's ability to get sunlight.

7 Related Works

7.1 Non-Optical Eavesdropping Attacks

Electromagnetic Eavesdropping. Electromagnetic side-channel has been exploited for eavesdropping attacks, i.e., TEMPEST. Researchers used this method to steal images from computer monitors [2, 3, 15, 20, 35] and keystrokes from keyboards [43]. *Nonstop* is another electromagnetic eavesdropping attack that injects radio wave toward the target object, and recovers secret information by analyzing the modulated reflected wave [3, 44]. This approach is similar to laser microphones in actively emitting signals. More recently, Choi et al. showed that an attacker can eavesdrop on sound by exploiting electromagnetic emission from system-on-chip (SoC) [6]. Meanwhile, MagEar eavesdrops sound by exploiting the electromagnetic field caused by the vibrating headphone diaphragm [24], and Chen et al. proposed the acoustic eavesdropping attack exploiting electromagnetic leakage from audio amplifiers [13].

Oversensing. There is a line of local eavesdropping attacks using a malware installed in a target device. The access to the microphone requires special permissions [45], and there is a line of research works for eavesdropping sound by using the other sensors available with lower privileges. For example, gyroscopes and accelerometers can be used as a listening device [4, 14]. Similarly, an embedded vibration sensor within a hard disk drive can work as a microphone [21].

7.2 Laser-Based Integrity Attacks

Besides eavesdropping, lasers have been used to attack the integrity of target devices. A laser injected on a semiconductor chip generates a photocurrent in the integrated circuit, which changes the circuit behavior that results in transient errors. Since Skorobogatov and Anderson first exploited this phenomenon to attack cryptography implemented on smartcards and microcontrollers [37], the resistance against such laser injection attacks has been studied in the field of secure implementation of cryptography [10, 19].

In addition, lasers have been used to inject false measurements into sensor systems. LiDARs used in autonomous vehicles measure the 3D point cloud of the scene by emitting laser pulses and measuring their echoes They are susceptible to laser injection attacks [5, 17]. Moreover, the energy delivered through a laser beam can interfere with non-optical sensors, e.g., fake signals injected into microphones [38].

8 Conclusion and Future Works

This paper studied the eavesdropping attack using an inexpensive single-diode vibrometer based on self-mixing interferometry. Our proof-of-concept setup successfully recovered intelligible audio from different target objects under different propagation conditions. The proposed method enables flexibility in choosing laser power and wavelength, including the infrared laser for invisible attacks, which is in contrast to commercial laser vibrometers using visible lasers for human operations. The attack becomes less expensive and more portable with our open circuit designs. Finally, we discussed the countermeasures and mitigations against the presented attack.

There are several directions for future research. In particular, we used only simple signal processing (filtering and noise reduction) and applied an existing speech-to-text service to recover the spoken phrase; this part has room for further improvements with advanced machine-learning techniques.

Laser Safety. All of our laser experiments were conducted in closed controlled environments by trained personnel.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP21K1188, JP22H00519 and JST CREST JPMJCR23M4.

Table 7: Intelligibility and WER in the end-to-end evaluation using the cheap setup in Figure 12.

Medium	Target	Voice	Intelligibility			WER		
			OHD	SMI	Cheap	OHD	SMI	Cheap
Adjoining	Laptop	Female	**0.8802**	**0.8147**	**0.7348**	**0.00**%	**0.00**%	37.50%
Adjoining	Laptop	Male	**0.8452**	**0.7167**	0.6595	**0.00**%	**0.00**%	**14.30**%

References

[1] 1969. IEEE Recommended Practice for Speech Quality Measurements. *IEEE No 297-1969* (1969), 1–24.

[2] 1985. Electromagnetic radiation from video display units: An eavesdropping risk? *Computers & Security* 4, 4 (1985), 269–286.

[3] Ross J. Anderson. 2008. Emission Security. In *Security engineering — a guide to building dependable distributed systems (2nd ed.)*. 523–546.

[4] Z. Ba, T. Zheng, X. Zhang, Z. Qin, B. Li, X. Liu, and K. Ren. 2020. Learning-based Practical Smartphone Eavesdropping with Built-in Accelerometer. In *2020 Network and Distributed System Security Symposium (NDSS)*.

[5] Y. Cao, S. Bhupathiraju, P. Naghavi, T. Sugawara, Z. Morley Mao, and Sara Rampazzi. 2023. You Can't See Me: Physical Removal Attacks on LiDAR-based Autonomous Vehicles Driving Frameworks. In *32nd USENIX Security Symposium*.

[6] J. Choi, H. Yang, and D. Cho. 2020. TEMPEST Comeback: A Realistic Audio Eavesdropping Threat on Mixed-signal SoCs. In *2020 ACM SIGSAC Conference on Computer and Communications Security (CCS '20)*. 1085–1101.

[7] A. Davis, M. Rubinstein, N. Wadhwa, G. Mysore, F. Durand, and W. Freeman. 2014. The Visual Microphone: Passive Recovery of Sound from Video. *ACM Transactions on Graphics (Proc. SIGGRAPH)* 33, 4 (2014), 79:1–79:10.

[8] S. Donati. 2012. Developing self-mixing interferometry for instrumentation and measurements. *Laser & Photonics Reviews* 6, 3 (2012), 393–417.

[9] S. Donati and M. Norgia. 2018. Overview of self-mixing interferometer applications to mechanical engineering. *Optical Engineering* 57, 5 (2018), 051506.

[10] J. Dutertre, J. Fournier, A. Mirbaha, D. Naccache, J. Rigaud, B. Robisson, and A. Tria. 2011. Review of fault injection mechanisms and consequences on countermeasures design. In *2011 6th International Conference on Design & Technology of Integrated Systems in Nanoscale Era (DTIS)*. IEEE, 1–6.

[11] G. Giuliani, M. Norgia, S. Donati, and T. Bosch. 2002. Laser diode self-mixing technique for sensing applications. *Journal of Optics A: Pure and Applied Optics* 4, 6 (nov 2002), S283.

[12] M. Guri, Y. Solewicz, A. Daidakulov, and Y. Elovici. 2017. SPEAKE(a)R: Turn Speakers to Microphones for Fun and Profit. In *11th USENIX Workshop on Offensive Technologies (WOOT '17)*.

[13] Y. Hu Z. Ning K. Li Z. Qin M. Duan Y. Xie D. Liu H. Chen, W. Jin and M. Li. 2024. Eavesdropping on Black-box Mobile Devices via Audio Amplifier's EMR. In *31th Annual Network and Distributed System Security Symposium (NDSS)*.

[14] J. Han, A. Chung, and P. Tague. 2017. PitchIn: Eavesdropping via Intelligible Speech Reconstruction Using Non-acoustic Sensor Fusion. In *2017 16th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. 181–192.

[15] Y. Hayashi, N. Homma, M. Miura, T. Aoki, and H. Sone. 2014. A Threat for Tablet PCs in Public Space: Remote Visualization of Screen Images Using EM Emanation. In *ACM SIGSAC Conference on Computer and Communications Security (CCS 2014)*. 954–965.

[16] Hawkins JE Hudgins CV. 1947. The development of recorded auditory tests for measuring hearing loss for speech. *Laryngoscope* 57, 1 (1947), 57–89.

[17] M. Feiri J. Petit, B. Stottelaar and F. Kargl. 2015. Remote attacks on automated vehicles sensors: Experiments on camera and LiDAR. In *Black Hat Europe*.

[18] D. Liu H. Zang J. Shang, Y. He and W. Chen. 2006. Laser doppler vibrometer for real-time speech-signal acquirement. In *Chin. Opt. Lett.* 732–733.

[19] D. Karaklajic, J. Schmidt, and Ingrid Verbauwhede. 2013. Hardware Designer's Guide to Fault Attacks. *IEEE Trans. Very Large Scale Integration (VLSI) Systems* 21, 12 (2013), 2295–2306.

[20] M.G. Kuhn. 2002. Optical time-domain eavesdropping risks of CRT displays. In *Proceedings 2002 IEEE Symposium on Security and Privacy*. 3–18.

[21] A. Kwong, W. Xu, and K. Fu. 2019. Hard Drive of Hearing: Disks that Eavesdrop with a Synthesized Microphone. In *2019 IEEE Symposium on Security and Privacy (SP)*. 905–919.

[22] R. Lang and K. Kobayashi. 1980. External optical feedback effects on semiconductor injection laser properties. *IEEE Journal of Quantum Electronics* 16, 3 (1980), 347–355.

[23] W. Li, M. Liu, Z. Zhu, and T.S. Huang. 2006. LDV Remote Voice Acquisition and Enhancement. In *18th International Conference on Pattern Recognition (ICPR'06)*, Vol. 4. 262–265.

[24] Q. Liao, Y. Huang, Y. Huang, Y. Zhong, H. Jin, and K. Wu. 2022. MagEar: eavesdropping via audio recovery using magnetic side channel. In *20th Annual International Conference on Mobile Systems, Applications and Services (MobiSys '22)*. 371–383.

[25] Y. Long, P. Naghavi, B. Kojusner, K. Butler, S. Rampazzi, and K. Fu. 2023. Side Eye: Characterizing the Limits of POV Acoustic Eavesdropping from Smartphone Cameras with Rolling Shutters and Movable Lenses. In *44th IEEE Symposium on Security and Privacy*. 1857–1874.

[26] C. Mackin, N. Neuenfeldt, M. Hamel, A. Slagel, N. Melena, and C. Smith. 2012. Remote Listening Device: Creation of a Covert IR Laser Microphone.

[27] A. Magnani, A. Pesatori, and M. Norgia. 2014. Real-Time Self-Mixing Interferometer for Long Distances. *IEEE Transactions on Instrumentation and Measurement* 63, 7 (2014), 1804–1809.

[28] Pariente Manuel. [n. d.]. Python implementation of STOI. https://github.com/mpariente/pystoi. Accessed: 2024-02-08.

[29] B. Nassi, Y. Pirutin, R. Swisa, A. Shamir, Y. Elovici, and B. Zadov. 2022. Lamphone: Passive Sound Recovery from a Desk Lamp's Light Bulb Vibrations. In *31st USENIX Security Symposium*.

[30] B. Nassi, R. Swissa, J. Shams, B. Zadov, and Y. Elovici. 2023. The Little Seal Bug: Optical Sound Recovery from Lightweight Reflective Objects. In *2023 IEEE Security and Privacy Workshops (SPW)*. 298–310.

[31] M. Norgia and S. Donati. 2003. A displacement-measuring instrument utilizing self-mixing interferometry. *IEEE Transactions on Instrumentation and Measurement* 52, 6 (2003), 1765–1770.

[32] M. Norgia, G. Giuliani, and S. Donati. 2007. Absolute Distance Measurement With Improved Accuracy Using Laser Diode Self-Mixing Interferometry in a Closed Loop. *IEEE Transactions on Instrumentation and Measurement* 56, 5 (2007), 1894–1900.

[33] Deaf/Hard of Hearing Technology Rehabilitation Engineering Research Center. [n. d.]. Audio Quality Testing Stimuli. https://www.deafhhtech.org/rerc/products/audio-quality-stimuli/. Accessed: 2024-02-29.

[34] Jr. R. James Rockwell, William J. Ertle, and C. Eugene Moss. [n. d.]. Safety Recommendations of Laser Pointers. https://www.rli.com/resources/articles/pointer.aspx. Accessed: 2024-02-09.

[35] P. Rohatgi. 2009. Electromagnetic Attacks and Countermeasures. In *Cryptographic Engineering*. 407–430.

[36] L. Scalise, Yanguang Yu, G. Giuliani, G. Plantier, and T. Bosch. 2004. Self-mixing laser diode velocimetry: application to vibration and velocity measurement. *IEEE Transactions on Instrumentation and Measurement* 53, 1 (2004), 223–232.

[37] S. P. Skorobogatov and R. J. Anderson. 2002. Optical fault induction attacks. In *Conference on Cryptographic Hardware and Embedded Systems (CHES)*.

[38] T. Sugawara, B. Cyr, S. Rampazzi, D. Genkin, and K. Fu. 2020. Light Commands: Laser-Based Audio Injection Attacks on Voice-Controllable Systems. In *29th USENIX Security Symposium*.

[39] C. Taal, R. Hendriks, R. Heusdens, and J. Jensen. 2011. An Algorithm for Intelligibility Prediction of Time–Frequency Weighted Noisy Speech. *IEEE Transactions on Audio, Speech, and Language Processing* 19, 7 (2011), 2125–2136.

[40] TafuOn. [n. d.]. Transmission Vibration Speaker. https://www.amazon.co.jp/dp/B09TNC5CQ8. Accedssed: 2024-03-11.

[41] T. Taimre, M. Nikolić, K. Bertling, Y. Lim, T. Bosch, and A. Rakić. 2015. Laser feedback interferometry: a tutorial on the self-mixing effect for coherent sensing. *Adv. Opt. Photon.* 7, 3 (Sep 2015), 570–631.

[42] The Guardian. [n. d.]. Laser spying: is it really practical? https://www.theguardian.com/world/2013/aug/22/gchq-warned-laser-spying-guardian-offices. Accessed: 2024-02-08.

[43] M. Vuagnoux and S. Pasini. 2009. Compromising Electromagnetic Emanations of Wired and Wireless Keyboards. In *18th USENIX Security Symposium*. 1–16.

[44] S. Wakabayashi, S. Maruyama, T. Mori, S. Goto, M. Kinugawa, and Y. Hayashi. 2018. A Feasibility Study of Radio-frequency Retroreflector Attack. In *12th USENIX Workshop on Offensive Technologies (WOOT 2018)*.

[45] P. Walker and N. Saxena. 2021. SoK: Assessing the Threat Potential of Vibration-Based Attacks against Live Speech Using Mobile Sensors. In *14th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec '21)*. 273–287.

[46] P. Walker and N. Saxena. 2022. Laser Meager Listener: A Scientific Exploration of Laser-based Speech Eavesdropping in Commercial User Space. In *2022 IEEE 7th European Symposium on Security and Privacy (EuroS&P)*. 537–554.

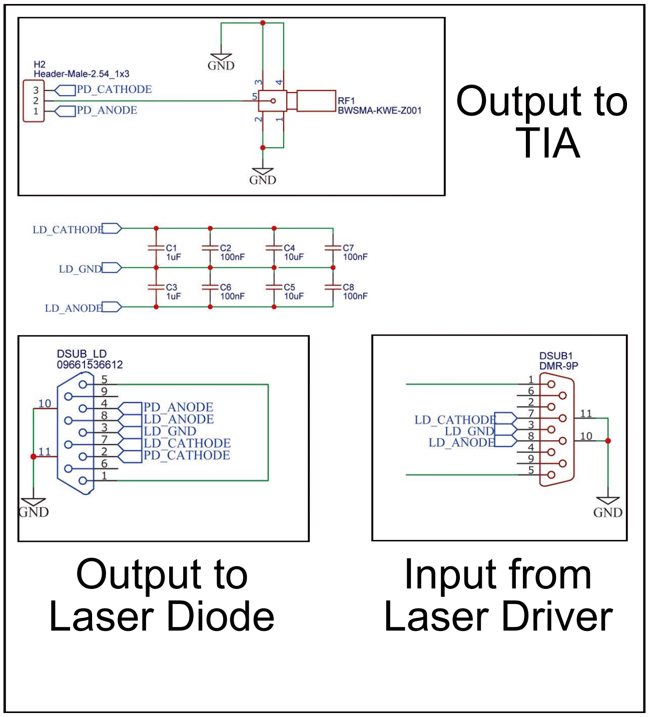

Figure 10: Schematic diagram of an adapter board for interfacing the standard DB9 connector from the laser current driver to a transimpedance amplifier (TIA).

A Cheap Setup

This appendix describes the details of the cheap setup evaluated in Section 6.2. Figure 12 shows the picture of the entire setup comprising three pieces of custom-made PCBs. The black PCB in the middle is a simple interface board for the D-sub-9 connector from the laser diode, as shown in Figure 10. The two green PCBs are the cheap laser driver and TIA. The entire setup works with three standard 9V batteries.

Laser Current Driver. The off-the-shelf laser current driver (Thorlabs LDC202C, ~$1,000) is the most expensive building block in the previous setup and occupies the largest space (see Figure 4-(right)). The laser current driver is no more than a power supply and can be replaced with a simple circuit in Figure 13. The essential components are the linear voltage regulators operated as a stable current supply with feedback control, and the amount of current can be configured through the resistors (R1, R2, U10, and U11 in Figure 11). The module costs $40 for 5 pieces in a popular PCBA service, including fabrication, assembly, and components.

Transimpedance Amplifier. The off-the-shelf TIA (Thorlabs AMP102, ~$500) is the second most expensive building block and can be replaced with an op-amp circuit in Figure 11. The design follows the regular current-to-voltage converter using a popular and cheap op-amp (Texas Instruments TL082) with a feedback resistor. We can configure the transimpedance gain and the voltage bias applied to the PD by changing the resistor values. The costs are similar to the current driver board, i.e., $40 for 5 pieces.

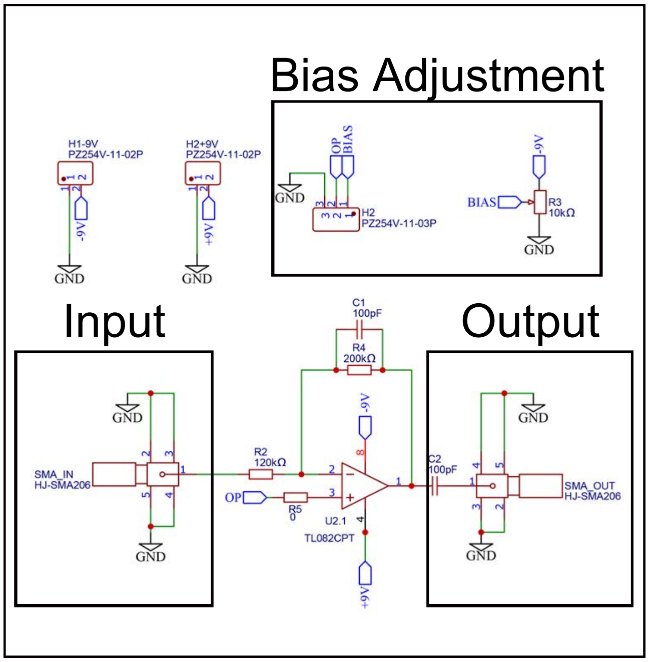

Figure 11: Schematic diagram of the TIA for the cheap attack setup.

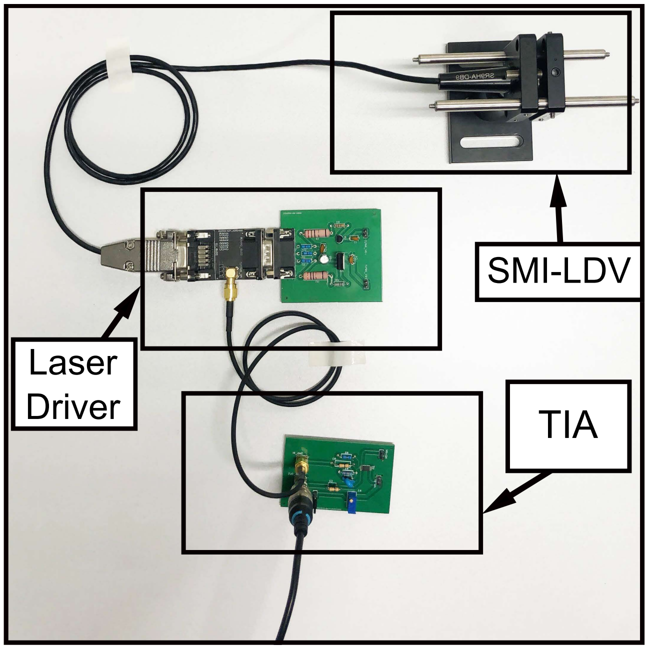

Figure 12: Cheap SMI-LDV setup. The two green PCB boards are the laser driver (Figure 13) and the TIA (Figure 11). They can be operated with three standard 9V batteries.

B Specrtogram of the audio

Figure 14 shows the spectrogram of the acoustic waveform measured during the end-to-end evaluation (see Section 5) from the

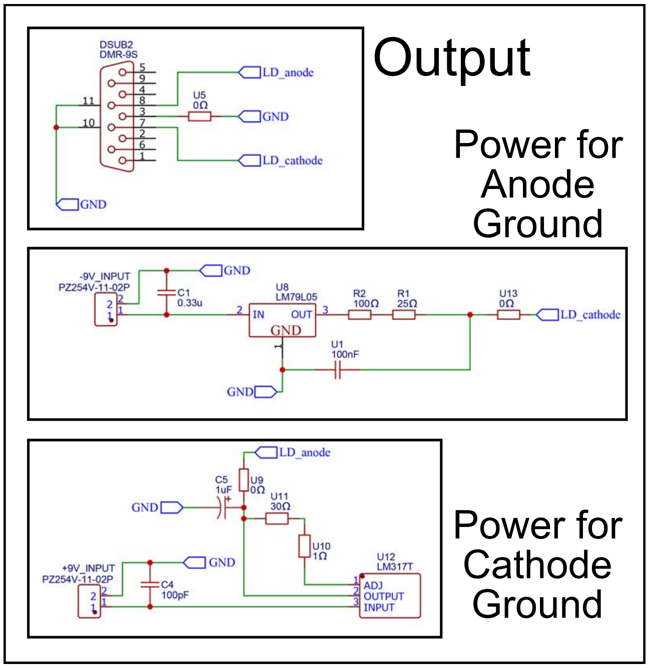

Figure 13: Schematic diagram of the laser current driver for the cheap attack setup.

laptop back panel in Figure 9. Figure 14-(a) is the original audio, and Figures 14-(b)–(d) are the audio measured with the SMI-LDV in the adjoining, same surface, and the aerial conditions, respectively.

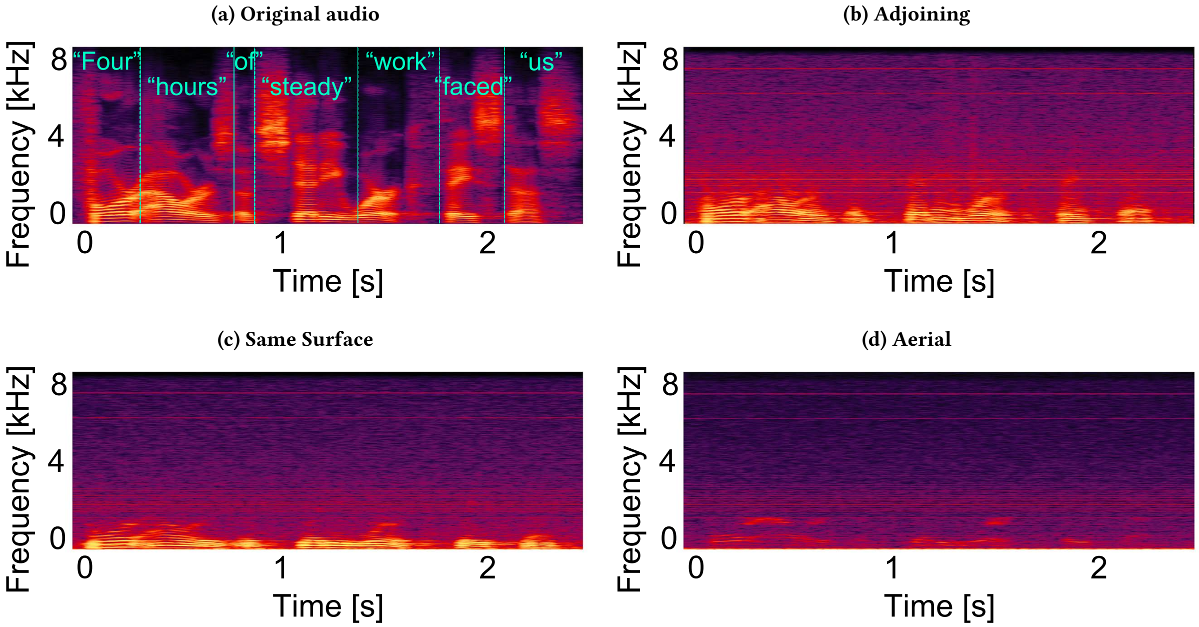

Figure 14: Specrtogram of the audio measured from a laptop back panel (see Figure 9) in the end-to-end evaluation. (a) Original audio. (b)–(d) Recovered audio measured with the SMI-LDV in the different coupling conditions.