Abstract—Recent research has shown that the integrity of sensor measurements can be violated through out-of-band signal injection attacks. These attacks target the conversion process from a physical quantity to an analog property—a process that fundamentally cannot be authenticated. Out-of-band signal injection attacks thus pose previously-unexplored security risks by exploiting hardware imperfections in the sensors themselves, or in their interfaces to microcontrollers. In response to the growing-yet-disjointed literature in the subject, this article presents the first survey of out-of-band signal injection attacks. It focuses on unifying their terminology and identifying commonalities in their causes and effects through a chronological, evolutionary, and thematic taxonomy of attacks. By highlighting cross-influences between different types of out-of-band signal injections, this paper underscores the need for a common language irrespective of the attack method. By placing attack and defense mechanisms in the wider context of their dual counterparts of side-channel leakage and electromagnetic interference, this study identifies common threads and gaps that can help guide and inform future research. Overall, the ever-increasing reliance on sensors in the wider context of their dual counterparts of side-channel leakage and electromagnetic interference, this study identifies common threads and gaps that can help guide and inform future research. Overall, the ever-increasing reliance on sensors embedded in everyday commodity devices necessitates that a stronger focus be placed on improving the security of such systems against out-of-band signal injection attacks.

Index Terms—Out-of-Band, Signal Injections, Hardware Imperfections, Mixed-Signal Systems, Survey, Attacks and Defenses

I. INTRODUCTION

MATHEMATICALLY secure algorithms can be broken in practice due to a mismatch between the high-level system model used for analysis and the real-world environment on which code runs. For example, data-dependent electromagnetic, optical, and acoustic emanations, as well as variations in power consumption can reveal the information processed by a device [1], with or without the help of intentional faults [2]–[4]. However, attacks exploiting hardware imperfections are not limited to side-channel leakage of confidential data: recent research has shown that it is possible to target the integrity of sensor measurements in a similar out-of-band fashion.

These out-of-band signal injection attacks can be performed using electromagnetic radiation exploiting circuits unintentionally acting as receiver antennas [5]–[7], as well as optical [8], [9] and acoustic [10]–[12] emissions targeting flaws in the conversion process from physical properties into electrical ones. The systems attacked have been equally diverse, and include medical devices [5], drones [10], [13], hard drives [12] and cameras [14], among others. However, despite the wide range of attack methods and devices targeted, research in the field thus far has been disjointed.

The ad-hoc nature of this type of research might stem, in part, from the fact that out-of-band signal injection attacks have so far only been (openly) conducted in a lab environment. However, out-of-band attacks have still garnered the interest of technological and mainstream news publications outside of the academic community [15]–[20]. They have even prompted national agencies to issue Computer Emergency Readiness Team (CERT) advisories [21]. In some cases, the effects of out-of-band attacks can be fatal: for instance, Foo Kune et al. have demonstrated that low-power attacker signals can trick cardiac implantable electrical devices into causing pacing inhibition and defibrillation shocks [5]. Although the techniques used are not identical to those of mass-produced sonic repellents [22] or commercial [23] and military [24] jammers, out-of-band signal injection attacks share the same potential for weaponization. As a result, to bring attention to these potentially severe issues, and to help designers better protect future hardware devices, this study conducts the first comprehensive survey of out-of-band signal injection attacks.

A. Survey Scope

This article focuses on out-of-band signal injection attacks, which target the connections between sensors, actuators, and microcontrollers, or exploit imperfections in the hardware itself. Although the term is defined precisely in Section II, we note here a few key features of the attacks investigated in this survey. The first property of out-of-band signal injection attacks is that they aim to change values processed by a system, rather than infer them. This fact distinguishes them from the side-channel [1] and fault-injection attacks [2]–[4] mentioned in the introduction.

The second feature is that out-of-band attacks do not change the measured quantity itself. For instance, using electromagnetic signals to change the audio recorded by a microphone [5] is an example of an out-of-band signal injection attack, but heating a temperature sensor with an open flame is not.

The final characteristic highlights the physical aspect of out-of-band signal injection attacks. In other words, the attacks studied in this paper alter sensor measurements or actuator inputs at the hardware layer instead of the protocol layer. As a result, spoofing attacks of unauthenticated, digital communication interfaces are out-of-scope. For example, wireless transmissions can be used to spoof the pressure of car tires and trigger warning lights [25], alter the flow of insulin injections [26], or change pacemaker settings to deliver shock commands with implantable cardiac defibrillators [27]. However, because these interfaces can easily be protected with
cryptography, they are not considered in this survey. For similar reasons, jamming, e.g., train signal controls [28], or over-powering legitimate GPS signals [29] are out-of-scope because they rely on intentional communication interfaces. As a result, they do not exploit hardware imperfections, instead relying on high-power, in-band electromagnetic transmissions. Finally, relay attacks on LiDARs, radars, and sonars [9], [14], [30]–[32] are also not out-of-band signal injection attacks. This is because, in order to spooof the distance between the attacker and the victim, they depend on winning a race between the adversarial signal and the true reflected pulses.

That said, these related research areas can offer invaluable insight into novel injection techniques and possible countermeasures. As a result, cross-disciplinary connections are made throughout this work, but with a clear focus on how they impact out-of-band signal injection attacks and defenses.

B. Contributions & Organization

Despite the growing literature in the area and extensive parallels between prior work in hardware security research and out-of-band signal injection attacks and defenses, no other article has made these connections explicit, or traced their evolution through time, theme, or approach. Our survey fills these gaps through the following contributions:

1) It unifies the diverse terminology used by different works (Section II) and summarizes the threat model (Section III) to create a common language through which to discuss attack and defense mechanisms.

2) It proposes the first chronological and thematic evolution of out-of-band signal injection attacks (Section VIII), which highlights cross-influences between electromagnetic (Section IV), conducted (Section V), acoustic (Section VI), and other (Section VII) attacks.

3) It creates a taxonomy of countermeasures introduced to prevent and detect out-of-band attacks (Section IX).

4) It places attacks and defenses in the wider context of side-channel leakage and electromagnetic interference attacks (Section X). Using these insights, this study identifies gaps in the experimental approach of published research, and proposes concrete steps to overcome these challenges in the future (Section XI).

II. CHOICE OF TERMINOLOGY

The terms used to describe the numerous acoustic, electromagnetic, and optical attacks on sensor-to-microcontroller and microcontroller-to-actuator interfaces have so far been inconsistent, with some works not even naming the attacks at all [33], [34]. This section sets out to identify and unify the nomenclature used as a first step towards providing a common language through which to compare the various works. As the threat model (Section III) and the causes of vulnerability (Section VIII) will reveal, the commonalities in the attack techniques highlight a need for an all-encompassing term irrespective of the method of injection. In other words, although some of the more restrictive terms are appropriate for describing specific attacks, we find that doing so can hide potential insights that arise from considering different types of attacks jointly. The term we have chosen for this unification is out-of-band signal injection attacks:

**Definition 1 (Out-of-Band Signal Injection Attacks).** Out-of-band signal injection attacks are adversarial manipulations of interfaces not intended for communication involving sensors/actuators that cause a mismatch between the true physical property being measured/acted upon and its digitized version.

To motivate our definition, the term injection was chosen because it captures the fact that values reported by a system are altered; it is not channel-specific; and it has already been adopted by different works [5], [11], [13], [35]–[44]. The out-of-band qualifier is necessary to distinguish the attacks studied in this survey from signal injection attacks on sensors using pulse reflections such as LiDARs [9], [14], [31], signal injection attacks on the physical layer of communication protocols [45], and false data injection attacks [28], [46]. As explained in Section I-A, these attacks are out-of-scope, as they do not depend on hardware vulnerabilities, but instead use external communication interfaces.

By contrast, our term captures attacks which target interfaces using signals outside of their intended frequency of operation. It includes ultraviolet or infrared light against cameras which should only be recording the visible part of the spectrum, and ultrasonic injections against microphones meant to be recording only audible sounds: these attacks transmit signals that are literally outside the operational band. It also includes electromagnetic signals against systems without any (intentional) antennas, and acoustic attacks against gyroscopes and accelerometers: these are also out-of-band, since they inject signals through channels other than the ones used by the sensor to measure the physical property. Our use of the out-of-band modifier is therefore consistent with the definition for out-of-band covert communication [47]. It has also recently been used by Tu et al. [43] to describe acoustic attacks on inertial sensors, further motivating its choice in this survey.

It should be noted that earlier work [31], [48] has proposed a subdivision of signal injections attacks into regular-channel attacks, which target the sensor structure itself by “using the same type of physical quantity sensed”, transmission-channel attacks, which target the connection between the sensor output and the measurement setup, and side-channel attacks, where the sensors themselves are targeted, but “by physical stimuli other than those they are supposed to sense”. We do not adopt this categorization, as it generally follows the medium of injection (optical, electromagnetic, and acoustic respectively). Moreover, regular-channel attacks are usually in-band, while side-channel attacks have an overloaded meaning.

We similarly find (intentional) interference [5], [10], [12], [36], [38], [40], [42], [49] to be unsuitable as a term because: (a) it does not make it clear that the attackers can in some cases inject waveforms of their choosing; and (b) Intentional Electromagnetic Interference (IEMI) has an established meaning in Electromagnetic Compatibility (EMC) literature [50], [51]. As Section X indicates, IEMI attacks often use high-power, destructive transmissions, and therefore have a different aim than out-of-band signal injection attacks.

The term (sensor) spoofing [8], [9], [14], [43] was also
TABLE I: Terminology used by different works to describe out-of-band signal injection attacks.

<table>
<thead>
<tr>
<th>Terminology</th>
<th>Example References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Injection</td>
<td>[5], [13], [35], [38], [41]–[44]</td>
</tr>
<tr>
<td>Intentional Interference</td>
<td>[5], [10], [12], [36], [38], [40], [42], [49]</td>
</tr>
<tr>
<td>Non-Linearity</td>
<td>[11], [54], [55], [68]</td>
</tr>
<tr>
<td>Spoofing</td>
<td>[8], [9], [14], [43]</td>
</tr>
<tr>
<td>Other (See Text)</td>
<td>[7], [53], [67]</td>
</tr>
</tbody>
</table>

avoided for similar reasons: it has an overloaded meaning in authentication contexts and with in-band signal injection attacks [6], [52]. Moreover, it does not capture the physical aspect of injections, and does not accurately describe coarse-grained attacks which lead to saturation of a sensor.

Other terms used have been specific to the particular channel which is being exploited, including induction attacks [7], acoustic resonance [53], and (acoustic) non-linearity [11], [54], [55]. Such terms were not selected because they are channel-specific, and focus on the mechanism of the attack, rather than the effect. Similarly, methodology-inspired terms which have been avoided include Radio Frequency Injection (RFI), Direct Power Injection (DPI), and other terminology that arises in immunity or susceptibility literature against (non-adversarial) electromagnetic interference (EMI) [56]–[66].

Finally, the term transduction attacks, proposed by Fu and Xu [67] to mean attacks which “exploit a vulnerability in the physics of a sensor to manipulate its output or induce intentional errors” has not yet received mainstream recognition. It also does not necessarily make it clear that the attack may target the interface between the sensor and the rest of the system, instead of just the sensor itself. The various terms which have been used to describe out-of-band signal injection attacks are shown in Table I, along with example references.

Since the majority of attacks in the literature are on sensors rather than actuators, we will refer to both of them collectively as sensors for brevity, and will distinguish between the two only when it is necessary to do so, i.e., when there is a divergence in the attack methodology.

III. SYSTEM AND ADVERSARY MODEL
Systems depend on sensors and actuators to interface with their external environment. They therefore require a conversion of a physical property (e.g., temperature or speed) to or from an electrical quantity (such as voltage or resistance). This electrical measurement is typically analog in nature, and is digitized by an Analog-to-Digital Converter (ADC) before it is processed. Although modern cryptography has mostly solved the problem of secure communication between digital interfaces, there is no way to authenticate the measurement itself, or the analog component of the connection between the sensor or actuator and a microcontroller. This lack of authentication, coupled with hardware imperfections, can be exploited for out-of-band signal injection attacks.

Conceptually, the sensor, the ADC, and the microcontroller perform logically distinct functions, but all three can be fully encapsulated into the same Integrated Circuit (IC) chip. Although this chip presents a digital interface to third parties, which can be protected by cryptographic protocols, the sensor itself can still be vulnerable to out-of-band signal injection attacks. For example, acoustic attacks targeting the resonant frequencies of gyroscopes and accelerometers have proven to be effective even against digital ICs [10], [13], [43].

As explained in Section II, attackers are not allowed to manipulate the property being measured itself (e.g., the radiation being measured by a Geiger counter): in the language of Shoukry et al. [69], the property measured itself is trusted, although the measurement itself is not. Nonetheless, attackers are allowed to transmit signals outside of the limits being sensed, which are still interpreted as valid measurements. For example, an attacker can produce ultrasound waves which are picked up by a microphone recording human speech [11], [54], [55], or shine infrared (IR) light into a camera capturing the visible part of the spectrum [14]. However, attacking a microphone with audible sound, or a camera with visible light is not allowed under this threat model, since the attacker is manipulating the property being sensed in-band. A secondary goal for some attacks is therefore undetectability or concealment [69]. Attacks also need to be non-invasive [69], and preclude direct physical access to the system under attack. It should be noted that different attack techniques have different distance requirements, with electromagnetic attacks theoretically having a longer range compared to optical and acoustic attacks. We defer the discussion of distance and related considerations to Section XI.

Figure 1 summarizes the channels that have successfully been exploited in the literature thus far. Some of the out-of-band signal injection attacks use electromagnetic (EM) waves to penetrate the wires connecting sensors and microcontrollers (Section IV), or the power circuit of the device (Section V). The same effects can sometimes also be achieved through shared power lines (conducted attacks). Finally, other attacks may target the sensors themselves through sound (Section VI), and alternative means such as infrared light (Section VII).

The effectiveness of different attacks, however, has mostly been evaluated in a qualitative fashion so far (e.g., whether the system was tricked into performing an action or not). Recent research, however, has attempted to mathematically define security against out-of-band signal injection attacks, and therefore quantify their success [44]. Specifically, Giechaskiel...
et al. [44] introduced probabilistic definitions which attempt to capture the fidelity with which an adversary can make a target waveform appear at the input of the microcontroller of Figure 1. These definitions address both coarse-grained attacks which merely disrupt sensor measurements and fine-grained attacks with precise waveform injections. Although the definitions abstract away from specific hardware considerations, they still model the discrepancy between an adversary’s target waveforms and the actual signal transmitted, which is necessary due to the behavior of the underlying circuits.

The specific circuit properties that allow adversarial signals to be injected into a device vary with the injection method and module targeted. For example, EM attacks typically depend on unintentional antennas in Printed Circuit Board (PCB) traces [5], while acoustic attacks exploit resonance in gyroscopes [10] and accelerometers [13]. As each attack exploits unique properties of the target device, the details of specific hardware imperfections in sensors and other modules are only expanded upon in subsequent sections. This section instead discusses common features of different attacks at a high level. One of these commonalities is that the vulnerability of systems to out-of-band signal injection attacks depends on both: (a) how adversarial signals are received by the devices under attack; and (b) how these signals are digitized.

Giechaskiel et al. [44] recently proposed a general circuit model which uses two transfer functions to separate these two aspects of vulnerability. The first transfer function describes circuit-specific transformations that an adversarial signal undergoes. For example, for EM attacks, these transformations include the (unintentional) low-power, low-gain antenna-like behavior of PCB traces connecting sensors to ADCs [5], [44]. The second transfer function, on the other hand, is ADC-specific, and summarizes the artifacts of the digitization process. These two transfer functions dictate that for a successful injection, attacker signals typically need to be transmitted over high-frequency carriers, and be demodulated into low-frequency, meaningful waveforms. According to the work of Giechaskiel et al., components within ADCs are the culprits for these demodulation effects [44]. The main constituents of an ADC that contribute to its demodulation characteristics are therefore summarized in Figure 2.

According to Pelgrom, an ADC uses three basic components to convert analog signals into digital ones: a ‘sample- or track-and-hold circuit where the sampling takes place, the digital-to-analog converter and a level-comparison mechanism’ [70]. Level-comparison amplifiers contribute to the demodulation properties of ADCs [44] due to non-linear distortions, including *harmonics* and *intermodulation products* [71].

Harmonics are responsible for producing “spectral components at multiples of the fundamental [input] frequency” [71]. As an example, for a sinusoidal of angular frequency \( \omega = 2\pi f \), harmonics transform the input \( v_{in} = \tilde{v} \cdot \sin(\omega t) \) into:

\[
    v_{out} = \left( \frac{a_2 \tilde{v}^2}{2} + \frac{3a_4 \tilde{v}^4}{8} + \cdots \right) + \left( a_1 \tilde{v} + \frac{3a_3 \tilde{v}^3}{4} + \cdots \right) \sin(\omega t) \]

\[
    - \left( \frac{a_2 \tilde{v}^2}{2} + \frac{a_4 \tilde{v}^4}{2} + \cdots \right) \cos(2\omega t) + \cdots
\]

As Equation (1) shows, the output also contains a Direct Current (DC) component, which depends solely on the “even-order nonlinear behavior” [71] of the system.

DC shifts can also be the result of reverse-biased diodes at the input of an ADC. These diodes protect the circuit from Electrostatic Discharge (ESD) by clamping negative voltages to ground, and inputs exceeding the maximum allowed voltage to \( V_{CC} \). This behavior of ESD diodes can cause a (non-linear) DC shift [71], which attackers can also exploit [7].

Intermodulation distortions, on the other hand, arise when the input signal contains signals of two different frequencies. For example, a sum of two sinusoidals \( v_{in} = \tilde{v}_1 \cdot \sin(\omega_1t) + \tilde{v}_2 \cdot \sin(\omega_2t) \) may represent an adversarial signal laid on top of the legitimate sensor signal [44]. Non-linearities and trigonometric identities would then dictate that the output signal contains frequencies of the form \( n\omega_1 \pm n\omega_2 \) for integers \( n,m \). As Giechaskiel et al. note, both types of “non-linearities demodulate attacker waveforms, even when they are modulated on high-frequency carriers” [44]. This fact makes them crucial for out-of-band signal injection attacks.

The final component of ADCs which is relevant to the adversarial injections studied in this article is the sample- and-hold circuitry shown in Figure 2. In its simplest form, the sample-and-hold mechanism consists of a resistor and a capacitor (RC circuit) connected to the input of the ADC. The transfer function of the voltage across the capacitor is therefore \( H_{S/H}(j\omega) = \frac{1}{1+j\omega RC} \). This dictates that as the angular frequency \( \omega \) increases, the gain \( G_{S/H} = \frac{1}{\sqrt{1+(\omega RC)^2}} \) is reduced [44]. To put it differently, the sample-and-hold mechanism acts as a low-pass filter, counteracting the aliasing effect, which occurs when input signals are faster than half the sampling rate of the ADC (Nyquist frequency). One would expect this filtering behavior to reduce the vulnerability of systems to out-of-band signal injection attacks. However, the filter’s cutoff frequency in practice “is often much higher than the sampling rate of the ADC” [44], necessitating additional anti-aliasing filters before the ADC input [70].

Overall, imperfections in the sensors themselves, the ADCs, or the connections between them can result in high-frequency signals being interpreted as meaningful low-frequency ones. As modulation over high-frequency signals is often necessary to enter the targeted circuit [5], [44], the demodulation properties of ADCs allow remote attackers without physical access to inject signals into a system in an out-of-band fashion.
The subsequent sections discuss these imperfections in greater detail, with a focus on the specific method of injection.

IV. ELECTROMAGNETIC TRANSMISSIONS

The antenna-like behavior of wires and traces is extensively studied in the fields of Electromagnetic Compatibility (EMC) and Electromagnetic Immunity. This is done to ensure interoperability between the various household electrical appliances, by guaranteeing that devices neither cause nor are susceptible to undue interference [72]. Such research has shown an inverse relationship between the length of microstrip PCB traces and the frequencies at which the traces are resonant [73]. However, to better predict the response of PCB traces to external EM fields, many additional parameters are important in practice. For instance, “field incidence, polarization angles, and the magnitude and phase of the impedances loading the microstrip terminations” [74] are all useful in modeling trace behavior.

Although going into the details of such antenna models is outside the scope of this article, the effects of unintentional antennas have been central in the security community. Until now, unintentional transmitting antennas have been key for side-channel analysis: data-dependent emissions can reveal the context of a distance-bounding protocol [33]. This would break protocol properties which depend on the speed of sound constant: adversarial transmissions propagating at the speed of light allow an adversary to operate from a longer distance. This attack is perhaps the first out-of-band electromagnetic signal injection, since it utilized EM emanations to attack an ultrasound-based protocol: unintentional antennas were found on the path “from the reception circuit to the piezo element”, which was effectively “working as a microphone” [33].

Although the attack by Rasmussen et al. was more of a side-note to an otherwise-secure protocol [33], Foo Kune et al.’s seminal 2013 “Ghost Talk” paper [5] made such adversarial injections the focal point of research. It showed that EM emissions could affect Electrocardiogram (ECG) measurements and cause IMDs to deliver fatal defibrillation shocks [5]. Foo Kune et al. succeeded in injecting arbitrary analog measurements, making a marked improvement in the literature compared to coarse replay and jamming attacks on IMDs [26], [27], [96], [97]. Their approach also significantly differed from high-power Intentional Electromagnetic Interference (IEMI) leading to the transient upset or destruction of commercial equipment [50], [76]–[78], [98]–[103].

The attack on IMDs by Foo Kune et al. [5] used low-power (≤10 W), low-frequency (kHz range) EM signals which coupled to the leads of ECGs and Cardiac Implantable Electrical Devices (CIEDs). In the open air, the distance achieved was up to 1.67 m, but when submerging the devices in saline (to approximate the composition of the human body), successful attacks were limited to less than 10 cm. The signals emitted targeted the baseband (i.e., the frequency of operation) of the IMDs directly, and thus did not make use of the non-linearities identified above and in Section III. However, this attack should still be considered out-of-band, as the IMD leads were meant to require physical contact for measurements, and should not be reacting to remote electromagnetic transmissions. In other words, this attack is more akin to coupling to a multimeter’s probes to cause wrong voltage readings remotely, rather than causing interference to a WiFi device by transmitting at 2.4 GHz: although both require external stimuli, the mode of injection is different from the intended mode of operation.

Foo Kune et al. also conducted out-of-band attacks against webcams and bluetooth headphones that were up to 1 m away from an 80 mW source [5]. The authors transmitted modulated signals over high-frequency carriers (in the hundreds of MHz) that make use of unintentional antennas on the path between the microphone and the amplifier. Non-linearities then demodulated the input signals and produced intelligible audio output. This output overlapped legitimate conversations via the headset and fooled music identification services as well as automated dial-in services (by emulating key presses via modulated Dual Tone Multiple Frequency (DTMF) signals) [5].

Kasmir and Lopes Estevés similarly targeted smartphone microphones, but with a goal of triggering voice commands (e.g., “OK Google”, “Hey Siri”) by emitting Amplitude-
Waveform $w(t)$

AM Signal $c(t) \cdot (1 + \mu w(t))$

Non-linearities, Filtering, Noise

Carrier $c(t)$

Headset

Attenuation

Fig. 3: Basic operating principle of an electromagnetic out-of-band signal injection attack against a microphone. Amplitude-Modulated (AM) signals are transmitted using an antenna, and are picked up (and attenuated) by headphone cables or PCB traces. Non-linearities in amplifiers coupled with low-pass filters remove the carrier wave and down-convert the target signal. The demodulated signal can be used to break ultrasound protocols [33], fool music services [5], or inject voice commands [38], despite the additional noise introduced.

Modulated (AM) signals [38]. These signals get picked up by the user’s hands-free headset, are then demodulated due to non-linearities, and finally get executed by the software voice-processing service. It is interesting to note that by default, “a long hardware button press is required for launching the service” [38]. However, a Frequency-Modulated (FM) signal at the same frequency can also emulate this headphone button press [38], allowing the attack to be fully carried out remotely.

The attack by Kasmi and Lopes Esteves used front-door coupling, as the “radiation couples to equipment intended to communicate or interact with the external environment” [76]. This is because headphones can be used as FM antennas and can thus not be effectively shielded. It should be noted that the field strength required was in the order of $25 - 30$ V m$^{-1}$, which is close to the limit for human safety, and an order of magnitude higher than the required immunity level ($3$ V m$^{-1}$) [38]. This illustrates that high powers might still be required for reasonable attack distances: in a subsequent work, the authors noted that their attack requires a power of 40 W for a distance of 2 m, and 200 W for a distance of 4 m [40].

Figure 3 shows an example of an electromagnetic out-of-band signal injection attack, which summarizes the attacks against microphone sensors. It shows that the desired attacker waveform $w(t)$ needs to be modulated over a high-frequency carrier $c(t)$, so that the signal can be picked up with relatively low attenuation by the victim device’s wires. For example, amplitude modulation with a modulation depth $0 < \mu \leq 100\%$ can be used to couple to wired headphones. Through non-linearities in the phone’s internal amplifier, as well as low-pass filtering effects, the target waveform is preserved at the output of the digital signal processing (DSP) chip. Although this process introduces noise, the demodulated signal can still be distinguished by online music services [5], imitate voice-initiated commands, which are then executed by the phone [38], or break protocol guarantees [33]. Although the above works targeted microphones, out-of-band electromagnetic signal injections are not sensor- or device-specific. For example, subsequent sections discuss proof-of-concept EM-based injections against temperature sensors [104], [105]. However, most attacks have primarily used amplitude modulation, leading to a question that has yet not been addressed:

Open Question 1: What is the optimal modulation scheme for out-of-band electromagnetic signal injection attacks? In other words, can Frequency Modulation (FM), Phase Modulation (PM), or other schemes be used instead of Amplitude Modulation (AM)?

Another largely-unexplored research area is that of magnetic emissions. Specifically, magnetic attacks had largely been ignored in the literature, until Shoukry et al. demonstrated that it is possible to confuse Anti-Lock Braking Systems (ABS) [6]. This is done by exposing the magnetic-based wheel speed sensors to an in-band attacker-generated magnetic field at close proximity [6]. Doing so can alter speed measurements, potentially veering cars off the road [6].

Although the work by Shoukry et al. is in-band and limited in distance, it inspired subsequent work in out-of-band attacks: Selvaraj et al. recently conducted the first attacks on actuators (rather than sensors) through EM transmissions [7]. Specifically, an unmodulated sawtooth waveform was chosen to cause a “sharp decrease, for a very small amount of time” at the target servo [7]. Because the servo is controlled using Pulse Width Modulation (PWM), a waveform of the same frequency (50 Hz) therefore results in a one-way (clockwise) rotation.

This attack has a few limitations: changing the attacking frequency to 60 Hz causes the servo to “change positions randomly” [7]. Moreover, relatively high powers are required: a 10 V (peak-to-peak) waveform is insufficient, so a 50 W amplifier and a 1-to-6 step-up transformer are necessary. Moreover, one of the servo wires is wrapped around the toroid transferring the EM signal. Although “the same effect was observed when a length of the wire was placed within a solenoid”, “producing an effect at a distance requires the proper selection of a field directivity element” [7].

Open Question 2: How can one precisely control an actuator in both directions, and at a distance?

Selvaraj et al. additionally proposed an analytical model of electromagnetic induction attacks for sensors and actuators, with a focus on the magnetic rather than the electric field [7]. To support their model, they further conducted experiments against General Purpose Input/Output (GPIO) pins of microcontrollers in analog and digital modes. They showed that 1.82 W transmissions of unmodulated signals at frequencies between 0 – 1000 MHz can result in a DC offset, even when the microcontroller is at a distance of up to 1 m from the source. This indicates that an adversary can successfully inject signals over a wide range of frequencies, without having precisely determined the resonance behavior of the system.

Selvaraj et al. [7] were only concerned with the average power received and not time-dependent signals. This is in contrast to work on the demodulating effects of amplifiers...
(Section V), which depends on inter-modulation products and harmonics. Instead, ESD diodes were identified as the culprits for the resulting DC offset, due to clipping non-linearities. However, it is not clear whether the same methodology can induce attacker-desired, time-varying waveforms through modulated (in amplitude or otherwise) transmissions: in the language of Giechaskiel et al. [44], Selvaraj et al. performed an existential injection which disturbs the ADC readings, but not a selective injection of attacker-chosen waveforms, unlike the earlier work of Foo Kune et al. [5].

As a final point of note, researchers have identified that coupling into the wiring interconnects within ICs is possible [74]. However, this disturbance “can be neglected up to several gigahertz”, since ICs are “usually smaller than a few centimeters” [74]. This leads to another research question:

**Open Question 3:** Is it possible to conduct out-of-band signal injection attacks into digital ICs which integrate sensors, ADCs, and microcontrollers?

Although this question has been answered in the affirmative for acoustic attacks (Section VI), it remains open for electromagnetic ones.

V. CONDUCTED SIGNALS

A different class of out-of-band signal injection attacks requires an indirect physical connection between the attacker and the victim, such as a shared power line. Unlike their radiated counterparts of Section IV, these conducted attacks do not require signals to be picked up by unintentional receiving antennas in the path between sensors and microcontrollers. Instead, signals are propagated along conductors primarily on the powering circuit, which can transfer through crosstalk or coupling to paths containing non-linearities. This propagation of electrical disturbances through structures and cables is studied in Transmission-Line Theory [106]–[108], and through the Baum-Liu-Tesche (BLT) equation [109]–[111].

Much like electromagnetic attacks, out-of-band conducted signal injection attacks have also been primarily experimental in nature. Their methodology often follows that of susceptibility literature, which predicts a device’s response to high-frequency radio signals. Systems tested include microcontrollers, ADCs, and other embedded devices which contain I/O and power pins. The goal of such research is to quantify immunity to radiated and conducted EM disturbances, and is typically concerned with the average power received by the embedded system, similar to the work by Selvaraj et al. [7] summarized in Section IV.

To avoid legal and practical considerations related to electromagnetic transmissions, the experimental approach followed is known as Direct Power Injection (DPI), and consists of injecting harmonic disturbances from a few kHz to a couple of GHz and measuring the relationship between forward power and frequency. Multiple works have shown that as the frequency of the input increases, immunity to DPI also increases [58], [61]–[63], [65], [112]. In other words, higher frequencies generally require higher forward power injections for the same level of susceptibility. This was also true of the (remote) injections by Selvaraj et al. [7], discussed above.

A similar methodology can be applied to evaluate the demodulation characteristics of amplifiers and transistors [56], [57], [60], [66], and therefore better predict the fidelity with which attackers can inject target waveforms, both in the conducted and in the radiated settings. This was recently done by Giechaskiel et al. [44] for six ADCs, with a view on how to exploit the demodulating effect for out-of-band signal injection attacks. It was shown that ADCs of three different types from four manufactures, and with different resolutions and sampling frequencies can all demodulate AM waveforms [44]. Generally, it was determined that the fundamental frequency persists along with its harmonics and some high-frequency components, even for carriers which are multiple times the ADCs’ sampling and cutoff frequencies [44].

It is worth noting, however, that the different ADCs do not behave identically. For instance, some ADCs require fine-tuning of the carrier frequency, with 100Hz making a difference as to whether the injected signal is fully demodulated or not [44]. On the other hand, some ADCs are vulnerable across the spectrum, i.e., for all frequencies which do not get severely attenuated to filtering effects [44]. Moreover, as discussed in an extended version of the same paper [116], some ADCs “result in more sawtooth-like output”, and are therefore “more resilient to clean sinusoidal injections”. Finally, the same extended version demonstrates that attacks can also be performed remotely, without following DPI methodology: a 10dBm (10mW) transmission can be demodulated by a receiver amplifier at small distances (5cm).

In their “Trick or Heat” work, Tu et al. [104] also conducted DPI experiments on operation amplifiers, but with a view on how to exploit rectification effects for out-of-band signal injection attacks on temperature sensors. They, too, determined that as the frequency increases, the magnitude of the AC voltage decreases, while “EMI signals at specific frequencies induce a significant DC offset” [104]. Moreover, for a given frequency of injection, power and the induced DC offset are “locally proportional”, though the rate of change “gradually decreases as the power of injected EMI signals grows” [104]. However, power and DC offset are not always positively correlated, even for remote transmissions: for some frequencies, the induced DC offset is negative [104].

Having characterized the behavior of individual amplifiers, Tu et al. turned their attention to different types of thermal sensors, including Negative Temperature Coefficient (NTC) thermistors, shielded and unshielded K-type thermocouples, and Resistance Temperature Detectors (RTDs). With a 35dBm (3.2W) electromagnetic source, Tu et al. succeeded in changing the reported temperature of various devices by at least 0.5°C [104]. The systems attacked included, among others, newborn incubators, soldering irons, and 3D printers, which were placed at distances of up to 6m.

1 Immunity behavior is different for EMI-induced offsets through the ground plane for amplifiers [113] and precision voltage references [114]. See the survey by Ramdani et al. [115] for more information on RF immunity models.
In some experiments, a thick wall was present between the transmitting device and an infant incubator under attack. It was shown that, even in this setup, an adversary can increase the measured skin temperature by 3.4°C or decrease it by 4.5°C [104], again demonstrating the potentially fatal consequences of out-of-band attacks. Most of the attacks by Tu et al. used unmodulated transmissions, and therefore only looked at the relationship between frequency and DC offset, or the relationship between power and DC offset. However, when investigated jointly, amplitude modulation was capable of causing selective injections [44]. In other words, Tu et al. [104] spelled “HI” in the output of the temperature sensor by appropriately modulating the amplitude of the transmission.

In a different strand of research, Lopes Esteves and Kasmi demonstrated how to inject voice commands (“OK Google”) into a smartphone through conducted means [40]. Specifically, the attack exploited the fact that on the device’s circuit board, the phone’s USB charging port is physically close to the audio frontend, where demodulation (envelope detection) can take place due to non-linearities [40]. As a result, back-door coupling occurs, either due to “a re-radiation of the interference from the USB circuitry bypassing the physical isolation by parasitic coupling (crosstalk) or the possible sharing of the \( V_{cc} \) and GND networks on the PCB” [40].

**Open Question 4:** What properties of the power circuit and related layout considerations make systems vulnerable to conducted out-of-band signal injections?

The methodology used was inspired by experiments on the propagation of conducted disturbances and on EM injections into power cables. Specifically, amplitude-modulated signals were injected at various locations of the power network, i.e., at different plug points on the same strip and on extension cords. The phone was left charging either on a computer USB port, or through a wall adapter. Experiments were repeated both with a magnetic injection probe (directly coupling to cables), and a “custom coupler made with capacitors, resistors and a high-frequency transformer” [40]. In all cases, it was determined that the smartphone can demodulate (and execute) commands carried on the 200–250MHz range at distances up to 10m, even with only a 0.5W source. Such conducted attacks therefore significantly lower the power requirements and increase the injection distance compared to the same authors’ remote EM attack on smartphones [38].

True Random Number Generators (TRNGs) which are based on Ring Oscillators (ROs) are also vulnerable to conducted signal injection attacks. ROs are composed of an odd number of logical NOT gates chained together in a ring formation, where the output of the last gate is used as the input to the first gate. The value of the loop between any two stages of the RO oscillates between true and false. The frequency of oscillation is influenced by the delay of the logic gates and the delay between the RO’s stages, which are in turn influenced by small variations in the manufacturing process, as well as voltage, and temperature (PVT) [117].

As a result, by XORing several ring oscillators together, one can exploit the randomness of the phase jitter to create a TRNG [118]. However, due to the frequency dependence on voltage, a suitable signal can lead to frequency locking of the oscillators [119], [120], removing the differences in the randomness of the jitter. Markettos and Moore first conducted the attack in practice in 2009 by directly injecting 24MHz signals into the power supply of two ring oscillators composed of discrete logic chips [35]. Moreover, they succeeded in biasing TRNGs even in secure microcontrollers and smartcards: a sinusoidal wave of 1V peak-to-peak (2.5mW) at 24.04MHz was enough to cause a 5V EMV Chip and Pin smartcard to fail statistical tests of randomness [35].

Time-varying signals are not always necessary: a constant (DC) power supply voltage can also lead to locking of ring oscillators in an under-volted Field-Programmable Gate Array (FPGA). This is because there is a “dependence of the frequency of one oscillator on the current peaks caused by rising and falling edges of the second oscillator” [121]. Figure 4 illustrates what happens when two ring oscillators frequencies lock: during normal operation (Figure 4a), the ring oscillator values “slide past each other, minimising the likelihood of two rings transitioning together” [35]. However, when a frequency- or voltage-based attack causes the ROs to lock (Figure 4b), their relationship becomes predictable, biasing the TRNG.

It should be noted that although under-over-power attacks are usually considered fault attacks, in this case the ring oscillators are still functioning properly, but the entropy of the TRNG is reduced due to less jitter present [39], [122]. An interesting new class of such remote under-voltage attacks on TRNGs has recently surfaced. Because ring oscillators have the potential to increase the delay of FPGA elements by causing voltage drops, they can also cause timing violations, thereby reducing the randomness of TRNGs [123]. Such an attack does not require equipment for physical injections. Instead, the adversary only needs co-located (but logically and physically isolated) circuits on the same FPGA as the target TRNG. This setup reflects multi-tenant cloud designs [123], and presents new challenges for the protection of shared FPGAs against software-only attacks without physical access.

Although the above attacks generally alter the power supply directly, the same outcome can be achieved through EM emanations targeting the wires connecting the various stages of the ring oscillators [36], [37], [49]. This requires micro-probes at very close proximity to the ring oscillators (in the order of 100μm from the FPGA packaging), so as to localize the effects of the injection [36], [49]. However, TRNGs are also vulnerable against EM injections into power supply cables:

![Fig. 4: Example waveforms for two Ring Oscillators (ROs) with frequency locking (a) absent or (b) present.](image-url)
Osuka et al. demonstrated that an injection probe wrapped around the DC power supply cable of a TRNG can also bias the TRNG [42]. Although the design only used two ring oscillators composed of discrete logic chips,\(^2\) injections were successful even when the probe was placed at a distance of 40 cm from the ROs, with a power of only 25.2 dBm (0.33 W).

It is worth highlighting that although the conducted voice command injection attack by Lopes Esteves and Kasmi was performed over shared power lines [40], all existing attacks on TRNGs bypass AC-to-DC rectification and voltage regulation. This leads to the following question for future research:

**Open Question 5:** Is it possible to bias True Random Number Generators through conducted out-of-band signal injection attacks on the primary side of power supplies (mains voltage), as shown in Figure 5?

---

VI. ACOUSTIC EMANATIONS

Research into out-of-band acoustic signal injection attacks has primarily focused on: (a) attacking electro-mechanical devices by causing vibrations at their resonant frequencies; and (b) exploiting microphone non-linearities for inaudible voice commands. In the former category, Micro-Electro-Mechanical Systems (MEMS) gyroscopes and accelerometers have been a popular target for acoustic resonance attacks.

MEMS gyroscopes operate through “vibrating mechanical elements to sense rotation” [124]. In other words, MEMS gyroscopes contain oscillating structures which, when rotated, appear to have a measurable force (called the Coriolis force) exerted on them [125]. These mechanical resonators “generate and maintain a constant linear or angular momentum”, so that “when the gyroscope is subjected to an angular rotation, a sinusoidal Coriolis force at the frequency of drive-mode oscillation is induced in the sense direction” [125]. This force is exerted in a different direction from the moving direction, and, depending on the type of the gyroscope, the angular rotation can be estimated through changes in capacitance, piezoresistive effects, etc. [124], [125]. Out-of-band acoustic signal injections transmit sounds at the resonant frequencies of MEMS sensors, causing them to report incorrect values.

Early research into the properties of MEMS gyroscopes had shown that high-power acoustic noise at or near the resonant frequency can degrade the performance of the sensor [126]–[128]. However, the security effects of intentional sound transmissions were not explored until the 2015 “Rocking Drones” paper by Son et al. [10]. Initially, the effect was a simple denial-of-service (DoS) attack on drones. It was caused by the transmission of single-tone sound waves at the resonant frequency of drones’ gyroscopes, so there was no control over their movements. The distance was also short, at 10 cm using a speaker producing a Sound Pressure Level (SPL) of up to 113 dB at the target frequencies.

Proof-of-concept control was then first demonstrated in a Black Hat conference presentation against gyroscopes in virtual reality (VR) headsets and self-balancing vehicles [34]. Tu et al.’s “Injected and Delivered” paper later became the first academic work to control gyroscopes in a much more fine-grained fashion [43]. This research allowed control for long periods of time (up to the minute range), and at long distances (up to 7.8 m with a maximum SPL of 135 dB) [43].

Tu et al. noticed that single-tone frequencies (like the ones used by Son et al. [10]) result in an oscillating discrete (digitized) output, which destabilizes equipment. In other words, a simple transmission at the resonant frequency is a type of DoS attack because the angular velocity (as measured by the gyroscope) fluctuates between positive and negative values (Figure 6a). However, it is possible to remove these negative components by decreasing the transmission amplitude during the corresponding measurements. This is called a Side-Swing attack, which “proportionally [amplifies] the induced output in the target direction and attenuate[s] the output in the opposite direction” [43] (Figure 6b).

Instead of attenuating signals during half of the transmission period, one can also control “the induced output by manipulating the phase of the digital signal with repetitive phase pacing” in a Switching attack [43] (Figure 6c). As this is accomplished in practice by changing the tonal frequency instead of attenuating the amplitude, a Switching attack contributes twice as much to the overall change in direction as a Side-Swing attack. This is shown by looking at the accumulating heading angle in Figure 7. It should be noted that by accounting for drifts in the sampling rate of the ADC (which are amplified during adversarial injections [43]), both attacks can control the gyroscopic output for longer periods of time.

In response to the rising interest in acoustic vulnerabilities, Khazaaleh et al. [129] created a mathematical model to explain the resonance response of gyroscopes. They showed that “the misalignment between the sensing and driving axes of the gyroscope is the main culprit behind the vulnerability of the gyroscope to ultrasonic attacks” [129]. More precisely, because “the sensing direction is not exactly orthogonal to the driving direction, some of the energy gets coupled to the sensing direction” [129]. This causes a false reading, which is typically corrected “by employing a demodulator in the readout circuit” [129]. When the transmission frequency is slightly different from the sensing frequency, the gyroscope generates “measurable output”, whose frequency equals “the difference between the driving frequency and the frequency of the acoustic signal” [129]. As shown experimentally, this model also explains why it is better to transmit near the reso-

\(^2\) Much like the original work by Markettos and Moore [35]. However, Bayon et al. [36], [49] targeted a more realistic TRNG composed of 50 ROs.
Fig. 6: Different acoustic injection approaches against gyroscopes. For (a) a Denial-of-Service (DoS) attack, a single-tone transmission at the gyroscope’s resonant frequency suffices. This results in oscillating digital measurements of angular velocity, and can destabilize equipment [10], [43]. To remove the negative measurement components, one can either (b) decrease the transmission amplitude in a Side-Swing attack [43], or (c) change the frequency of transmission for a Switching attack [43]. The cumulative effects of these approaches are shown in Figure 7.

![Graph showing the effects of different attacks on gyroscope measurements](image)

Fig. 7: Effects of the three attacks of Figure 6. In a Denial-of-Service (DoS) attack, the accumulating heading angle fluctuates, while it continues increasing for both attacks proposed by Tu et al. [43]. In Switching attacks, the angle increases at twice the rate of Side-Swing attacks.

It also suggests that low-pass filters or differential measurements through additional proof masses are ineffective countermeasures against out-of-band acoustic signal injection attacks [129]. Although Tu et al. succeeded in controlling gyroscopes in phones, scooters, stabilizers, screwdrivers, and VR headsets among others, only DoS attacks were successful against accelerometers [43]. MEMS accelerometers consist of spring-mass systems, so acceleration results in a deflection of the seismic mass. This deflection “is detected by means of capacitive elements, the capacitances of which change with deflection, or by piezoresistive elements that detect strain induced by the motion of the seismic mass through a change in resistor values” [130]. Acoustic vibrations at the resonant frequencies of the spring-mass systems can also displace the suspended mass, making them vulnerable to out-of-band attacks. According to Trippel et al., insecure amplifiers and low-pass filters (LPFs) prior to the accelerometer ADCs can demodulate both Amplitude-Modulated (AM) and Phase-Modulated (PM) attacker injections [13]. These insecurities are the results of clipping non-linearities and permissive filtering respectively, and allow for both biasing and control attacks.

Trippel et al. used accelerometers to spell words, naming their work “ WALNUT” for the output of the spoofed sensor measurements. They were also able to control off-the-shelf devices, such as remote-controlled (RC) cars, and Fitbit fitness tracking wristbands [13]. Although spoofing step counts might seem innocuous, companies often offer financial rewards for health-related activity [13], so cheating devices (which do not yet exploit out-of-band effects) are already being sold [131]. Most attacks by Trippel et al. [13] were performed at distances of 10cm, with a speaker producing an SPL of 110dB. The duration of control over the output of the MEMS sensors was often limited to 1–2s (and up to 30s) due to sampling rate drifts. Moreover, it was shown that the three axes do not behave identically to acoustic injections. For example, there are some MEMS devices for which only the x-axis responds to acoustic transmissions, while others are vulnerable in all three axes, but at different resonant frequencies.

Although Trippel et al. [13] attacked a single sensor in one direction at a time, Nashimoto et al.’s “Sensor CON-Fusion” investigated whether sensor fusion using a Kalman Filter can improve the robustness of measurements [41]. It was shown that “while sensor fusion introduces a certain degree of attack resilience, it remains susceptible” to combined acoustic and electromagnetic injections [41]. Specifically, Nashimoto succeeded in simultaneously controlling the roll, pitch, and yaw (the three angular axes in aircraft nomenclature) by fusing the outputs of an accelerometer, a gyroscope, and magnetometer. However, in non-simulated environments, “there is an error in the roll angle”, “the resulting inclination does not last long” [41]. Although fusion is further explored in the context of defense mechanisms (Section IX), the above discussion leads to the following research question:

**Open Question 6:** Is it possible to use acoustic injections to precisely control MEMS gyroscope and accelerometer measurements in all three directions simultaneously and/or for longer periods of time?
In a different strand of research, ten years after a video demonstrating that shouting in a data center causes unusually high disk I/O latency [132], Shahrad et al. showed that acoustic transmissions can cause vibrations in Hard-Disk Drives (HDDs) [53]. These vibrations result in read and write errors at distances up to 70 cm with a sound level of 102.6 dBA [53]. As a result, they can make systems unresponsive, even leading to Blue Screen of Death (BSOD) errors [53].

Although Shahrad et al. primarily focused on the effect of the angle of transmission [53], research conducted in parallel more precisely pinpointed the root cause of the issue using Finite Element Analysis [12]. Specifically, it was shown that (audible) acoustic waves “can displace a read/write head or disk platter outside of operational bounds, inducing throughput loss”, even though the displacement is of only a few nanometers [12]. In addition, in their “Blue Note” paper, Bolton et al. also used ultrasonic transmissions to attack the shock sensors which are meant to protect HDDs during sudden drops by “parking the read/write head” [12]: modern HDDs contain “piezo shock sensors or MEMS capacitive accelerometers” to “detect sudden disturbances” [12], and can also be attacked through ultrasound transmissions at their resonant frequencies. Through these malicious acoustic attacks, SPL levels of up to 130 dB cause HDDs at distances of 10 cm to become unresponsive, thus disabling laptops and video recorders [12].

Not all acoustic attacks transmit at resonant frequencies. By contrast, other research targets microphone non-linearities to cause inaudible sound to be recorded. Microphones operate by converting the mechanical deformation of a membrane (caused by the air pressure of a sound wave) into a capacitive change, which produces an Alternating Current (AC) signal [11]. This process also has non-linearities which produce second-order components [11], including harmonics and intermodulation products, as discussed in Section III.

Early work on acoustic attacks primarily focused on adversarial control of machine learning in speech recognition systems [135], [136]. Such research did not take advantage of non-linearities, and was in-band, as the transmissions were audible (although indecipherable by humans). However, later investigations revolved around ensuring that the transmitted frequencies are beyond the human-audible range (20 kHz). This was first accomplished by Zhang et al. in their “DolphinAttack” paper, which exploited non-linearities in microphone sensors [11], as shown in Figure 8.

The work by Zhang et al. showed how to transform (modulated) ultrasound transmissions into valid commands which were executed by speech recognition systems such as Apple Siri, Google Now, Microsoft Cortana, and Amazon Alexa [11]. The same authors later expanded on their attack by testing different setups, and increasing the attack distance from 1.7 m to 19.8 m [134]. This was done by replacing the 125 dB source with an ultrasonic transmitter array and amplifier outputting 1.5 W to increase sound pressure. Above this transmission power, the attack becomes audible due to non-linearities in the transmission medium (air) and the source speakers [11].

Earlier, Roy et al. had also noted that non-linearities in speakers make it harder for an adversary to increase the attack distance: “increasing the transmit power at the speaker triggers non-linearities at the speaker’s own diaphragm and amplifier, resulting in an audible [output]” [55]. Instead, multiple speakers in the form of an ultrasonic speaker array can be used to attack voice recognition systems including Amazon Alexa and Google Now at a distance of up to 7.6 m using a 6 W source [55]. The attack works by partitioning the audio spectrum across the various speakers in a way that “reduces the audible leakage from any given speaker” while minimizing the total leakage power [55]. This prevents any of the non-linearities (and the transmitted signal itself) from being audible. It should be noted that if multiple non-cooperating ultrasonic sources are emitting simultaneously, intermodulation distortions can create audible by-products [68], allowing for the detection of potential attacks.

Parallel to the 2017 DolphinAttack paper [11], similar research was in progress at Princeton [133]: Song and Mittal also succeeded in injecting inaudible voice commands to an Amazon Echo and an Android phone. Although they accomplished relatively long distances (3.5 m with an input power of 23.7 W), their work remains in poster format. Moreover, prior to their inaudible voice commands work [55], Roy et al. also used inaudible ultrasound transmissions to record audible sounds. Specifically, in their “BackDoor” work, they proposed a high-bandwidth covert channel that operates at up to 1.5 m using a 2 W source [54]. Although covert channels are not discussed in this survey, the work by Roy et al. is included due to the methodology which naturally led to their later work.

More concretely, instead of using amplitude modulation over a single frequency like Zhang et al. [11], Roy et al. simultaneously played two ultrasound tones whose shadows create audible sounds (only sensed by microphones) due to non-linearities [54]. They showed that amplitude modulation could not be used due to non-linearities in the ultrasound transmitters themselves, which would result in audible signals. Instead, further pre-computation was required to remove the “ringing effect”, where “the transmitted sound becomes slightly audible even with FM modulation” [54]. These results point towards the next open question [134]:

**Open Question 7:** How can non-linear acoustics in the transmission medium (air and speakers) be avoided to further extend the range of inaudible attacks?
VII. OPTICAL AND THERMAL MANIPULATIONS

Although electromagnetic, conducted, and acoustic attacks form the majority of out-of-band signal injection attacks, there has been some research on optical attacks, as well as temperature attacks which bias RO-based TRNGs. In the former category, attacks exploit permissive filtering and poor shielding in interfaces which only expect ambient environmental conditions. Most papers so far have targeted sensors in an in-band fashion: out-of-scope research includes attacks on sonars, radars, and LiDARs [9], [14], [31], [32], as well as visible-light attacks on cameras in unmanned aerial vehicles (UAVs) [52] and cars [9], [14], [32].

Researchers have also hypothesized that excessive light injections would blind car cameras and confuse auto-controls in automated vehicles [137]. Indeed, limited success against cameras has been achieved using Ultraviolet (UV) and Infrared (IR) lasers up to 2 m away [14]. However, attacks were only possible in dark environments, and the results were not reproducible with invisible lasers against other makes and models of CMOS cameras [9].

In a different strand of research, Park et al. showed that some medical infusion pumps are not well-protected against adversarial optical injections [8]. Specifically, in order to measure the flow rate of the medicine being administered, pumps are fitted with drip measurement sensors. These sensors consist of an IR emitter and receiver facing each other. When a drop passes through the sensor, the IR receiver temporarily senses less light due to diffusion, allowing for the rate to be measured. However, because the sensor is not well-enclosed, an adversary can shine an IR laser into the sensor, causing these drops to be undetected, thereby saturating the sensor. By then un-blinding the sensor, the attackers can also trick the firmware into detecting fake drops, and bypass alarms.

Adversaries can therefore selectively both over- and under-infuse a patient for an extended period of time, and for a variety of normal flow volumes. However, this can only be done with coarse-grained control over the real flow rate. Most of the experiments by Park et al. were conducted at a distance of 10 cm. Success was nevertheless reported up to 12 m away using a 30 mW IR laser pointer [8]. These results show that optical attacks can reach meaningful distances, but are, of course, limited by line-of-sight considerations.

It should be noted that the attack by Park et al. should be considered to be out-of-band, as the pump was not meant to receive external stimuli, in contrast to, for example, LiDARs, which are supposed to interact with external objects. In other words, a LiDAR depends on its surroundings to reflect its transmitted pulses and therefore infer the distance to the interfering objects. On the other hand, the drip sensor and (part of) the intravenous (IV) tube could be enclosed and shielded from the environment. This naturally leads to the defense mechanisms proposed by Park et al. [8], which are discussed in Section IX, and which beg the following question:

Open Question 8: Are all out-of-band optical attacks a matter of poor filtering and shielding?

Under somewhat unrealistic assumptions, the answer to the above question might be “no”. In their typical mode of operation, Light Emitting Diodes (LEDs) convert current to light. However, they can also function in the reverse by producing current when illuminated [139], [140]. In preliminary experiments, Loughry recently showed that this behavior can be exploited for an optical covert channel [138]. The setup is certainly unconventional: both LED pins are connected to microcontroller GPIO pins, which are configured as inputs with (internal) pull-up and pull-down resistors, as shown in Figure 9. According to Loughry, seven out of ten LEDs tested responded to laser and light of different wavelengths, with some LEDs producing measurable current at both ends (anode and cathode) [138]. Whether out-of-band signal injections or fault attacks can exploit this effect is yet to be seen [138].

The final class of attacks exploits the dependence of ring oscillators (ROs) on temperature to reduce the entropy of the TRNG. It is only mentioned here for completeness, as distance requirements would dictate physical access to the device under attack. Early work in the area showed that statistical randomness tests of RO-based TRNGs would fail for certain FPGA temperatures [141]. More detailed experimentation conducted a few years later using different heat-transfer methods (resistor heater, Peltier cooler, and liquid nitrogen) then showed that “the hotter the temperature, the larger the bias” [142]. Martín et al. [39] also investigated the effect of temperature across multiple TRNG designs based on Self-Timed Rings (STRs), which do not exhibit the frequency locking effects discussed earlier [143]. It was shown that the effects of temperature increases on the randomness of STR-based TRNGs were not significant, due to a combination of a decrease in frequency and an increase in jitter due to thermal noise [39].

Since devices were operating within their specifications, these biasing effects could be considered out-of-band attacks which operate at limited distances. This is in contrast to, for example, Martín et al.’s work which investigates the entropy of TRNGs in response to ionizing radiation [144]. However, remote conducted attacks on TRNGs are possible using local voltage drops [123]. Whether it is possible to reproduce these effects using heating circuits [145] remains an open question:

Open Question 9: Can software-only thermal effects bias RO-based TRNGs in multi-tenant FPGAs?

Overall, the efficacy of out-of-band optical thermal injections has been limited so far, due to the limited nature of vulnerable interfaces and proximity considerations respectively.
Fig. 10: Evolutionary and thematic taxonomy of out-of-band signal injection attacks, categorized by topic and methodology. Cross-influences between out-of-band signal injections and some key related works are also shown. Highly influential papers are discussed more extensively in Section VIII.
VIII. TAXONOMY OF ATTACKS

This section presents a taxonomy of out-of-band signal injection attacks, tracing their evolution through time and topic, and identifying commonalities in their methodology and source of vulnerability. We first highlight thematic and evolutionary cross-influences between the various works studied in this survey (Figure 10), and then categorize the key hardware imperfections that make them possible (Table II). We start with a citation graph of out-of-band attacks and related work in Figure 10, where an edge $X \rightarrow Y$ indicates that $X$ is cited by $Y$. To reduce clutter, if a paper $X$ is cited by both $Y$ and $Z$, and $Y$ is cited by $Z$, then no arrow from $X$ to $Y$ is drawn.

Figure 10 reveals that cross-influences are not limited to attacks. Instead, there is a general trend of earlier research observing the effects of non-adversarial interference, with later work actively exploiting the same phenomenon for signal injection attacks. For example, multiple works had identified the effects of electromagnetic interference on medical devices [82], [83], [89], [90] (with more papers discussed in Section IV), but Foo Kune et al. [5] were the first to recognize the effect as a security concern rather than a safety and reliability one. Similarly, Dean et al. commented on the effect of acoustic noise on gyroscopes [127], [128], but Son et al. [10] later used the same effect to destabilize drones.

Observation 1: Out-of-band signal injection attacks identify the effect of noise on systems and find novel ways to amplify it through hardware imperfections.

The graph further shows that out-of-band signal injection attack topics and methodology has both been inspired by and has itself inspired research exploiting more traditional avenues of attack. For example, investigations into acoustic signal injection attacks [11], [54], [55] have been influenced by research exploiting machine learning algorithms which respond to commands which are audible but indecipherable by humans [135], [136]. In addition, out-of-band acoustic attacks on MEMS sensors [10], [13] have both inspired and been influenced by in-band optical attacks on unmanned aerial vehicles (UAVs) [52]. Moreover, after the effects of acoustic noise on the security of gyroscopes were first identified [10], subsequent work improved the level of control over the output [43], and attacked additional types of MEMS sensors [13], and electro-mechanical devices such as HDDs [12].

Observation 2: After an exploitable hardware imperfection has been identified, determining its root cause opens up new avenues of attacks across different domains, and with alternative methodologies.

Figure 10 further highlights a few key works which in the opinion of the authors have played a central role in the development of the field. The first set of such papers [5], [10], [11], [35], [36] was chosen due to the high number of citations they have received overall ($\geq 95$) and from other out-of-band attack research ($\geq 5$).\footnote{Citation counts are current as of 7 Nov. 2019 according to Google Scholar.} Specifically, the works by Markettos and Moore [35] and by Bayon et al. [36] were chosen because they successfully biased TRNGs in the conducted and EM settings respectively, going beyond earlier theoretical work on oscillator locking [119], [120]. Their work led to the development of a new branch of attacks, which has so far developed rather independently of other out-of-band signal injections, as shown in Figure 10.

Foo Kune et al.’s work on adversarial electromagnetic interference [5] also features prominently in Figure 10, having been cited by almost all out-of-band signal injection attacks that were published after it (with the exception of TRNG research). Foo Kune et al. were the first to successfully exploit non-linearities and unintentional antennas in remote electromagnetic injection attacks, which were non-adversarial in prior work (e.g., [82], [83]) or only mentioned in passing [33].

With over 160 citations since 2017, the research by Zhang et al. [11] has been very influential in the realm of out-of-band acoustic attacks against microphones. Unlike earlier work on covert communication [54], and indecipherable-yet-audible commands [135], [136], Zhang et al.’s “DolphinAttack” exploited microphone non-linearities for inaudible injections.

Research on acoustic attacks is perhaps more mature against MEMS sensors, in great part due to early work by Son et al., who first showed how to disrupt gyroscopes [10]. Moreover, Trippel et al.’s research has also significantly furthered the state-of-the-art in acoustic injections by controlling the output of accelerometers for short periods of time [13]. As a result, the work of Tu et al., which showed how to extend the duration of control [43], is in the second set of works highlighted in Figure 10. This set contains recent studies whose novelty and potential has not yet received mainstream attention.\footnote{Trippel et al.’s 2017 work [13] lies between the two categories, already having over 75 citations, 12 of which are from other attack papers.} For example, the techniques proposed by Tu et al. to overcome ADC sampling rate drifts should be applicable to other methods of injection, and against different types of targets. We place the work by Bolton et al. [12] in the same category, as it managed to bridge research on HDD attacks (e.g., [53]) with attacks on MEMS sensors, and contained significant insights into why resonance attacks against hard drives work.

The final work highlighted in Figure 10 is the in-band attack of Shoukry et al. against an Anti-lock Braking System (ABS) [6]. It has been included not just for its high citation count ($> 135$, of which 10 are out-of-band attacks), but because it is the first EM paper to focus on the magnetic field rather than the electric field. As a result, it serves as inspiration to recent out-of-band magnetic attacks [7], [41], which we hope will be explored more in the future (Section XI).

We also summarize the various out-of-band signal injection attacks along with factors which contribute to them in Table II. The table further notes the maximum power used and distance achieved for an attack, including the level of control over the resulting signal. Effects range from theoretical attacks which are only partially realized to practical attacks which disrupt, bias, or completely control the output. As Table II indicates, information on the attack setup was often hard to find, sometimes completely missing, and often had to be identified by
TABLE II: Sources of vulnerability, methods, and effects for out-of-band signal injection attacks along with maximum power and distance. For each potential source of vulnerability, ✓ signifies that the paper claims the vulnerability contributes to the attack, while – that it does not. The attack methods used are acoustic, conducted, electromagnetic, and optical. The effects achieved include disruption, bias, and control of the output, while denotes a theoretical or partially-realized attack. Resonance (†) includes frequency locking for ring oscillators, and unintentional wire antennas for EM attacks.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>Ref.</th>
<th>Target</th>
<th>Power</th>
<th>Distance</th>
<th>Method</th>
<th>Effect</th>
<th>Resonance (†)</th>
<th>Non-Linearity</th>
<th>Improper Filtering</th>
<th>Poor Shielding</th>
<th>Insecure Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rasmussen et al.</td>
<td>2009</td>
<td>[33]</td>
<td>Medical</td>
<td>(unspecified)</td>
<td>(unspecified)</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Petit et al.</td>
<td>2015</td>
<td>[14]</td>
<td>Camera</td>
<td>(unspecified)</td>
<td>(unspecified)</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Cao et al.</td>
<td>2016</td>
<td>[122]</td>
<td>TRNG</td>
<td>(under-volt)</td>
<td>(direct)</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Wang et al.</td>
<td>2017</td>
<td>[34]</td>
<td>MEMS</td>
<td>(unspecified)</td>
<td>(unspecified)</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Nashimoto et al.</td>
<td>2018</td>
<td>[41]</td>
<td>MEMS</td>
<td>(unspecified)</td>
<td>(unspecified)</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Mahmoud and Stoilović</td>
<td>2019</td>
<td>[123]</td>
<td>TRNG</td>
<td>(140K ROs)</td>
<td>(internal)</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Markettos and Moore</td>
<td>2009</td>
<td>[35]</td>
<td>TRNG</td>
<td>0.002 W</td>
<td>(direct)</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Bayon et al.</td>
<td>2012</td>
<td>[36]</td>
<td>TRNG</td>
<td>0.003 W</td>
<td>100μm</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Buchovecka et al.</td>
<td>2013</td>
<td>[37]</td>
<td>TRNG</td>
<td>≤ 0.563 W</td>
<td>“bear”</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Foo Kune et al.</td>
<td>2013</td>
<td>[5]</td>
<td>Med. &amp; Mic.</td>
<td>10.000 W</td>
<td>1.67 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Kasimi and Lopes Esteves</td>
<td>2015</td>
<td>[38]</td>
<td>Microphone</td>
<td>200,000 W</td>
<td>4.00 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Bayon et al.</td>
<td>2016</td>
<td>[49]</td>
<td>TRNG</td>
<td>0.003 W</td>
<td>100μm</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Park et al.</td>
<td>2016</td>
<td>[8]</td>
<td>Medical</td>
<td>0.030 W</td>
<td>12.00 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Roy et al.</td>
<td>2017</td>
<td>[54]</td>
<td>Microphone</td>
<td>2.000 W</td>
<td>1.50 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Song and Mittal</td>
<td>2017</td>
<td>[133]</td>
<td>Microphone</td>
<td>23.700 W</td>
<td>3.54 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Lopes Esteves and Kasmi</td>
<td>2018</td>
<td>[40]</td>
<td>Microphone</td>
<td>0.500 W</td>
<td>10.00 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Osuka et al.</td>
<td>2018</td>
<td>[42]</td>
<td>TRNG</td>
<td>0.331 W</td>
<td>0.40 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Roy et al.</td>
<td>2018</td>
<td>[55]</td>
<td>Microphone</td>
<td>6.000 W</td>
<td>7.62 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Selvaraj et al.</td>
<td>2018</td>
<td>[7]</td>
<td>GPIO</td>
<td>1.820 W</td>
<td>1.00 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Giechaskiel et al.</td>
<td>2019</td>
<td>[44]</td>
<td>ADC</td>
<td>0.010 W</td>
<td>0.05 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Tu et al.</td>
<td>2019</td>
<td>[104]</td>
<td>Temp. &amp; Amp.</td>
<td>3.162 W</td>
<td>6.00 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Yan et al.</td>
<td>2019</td>
<td>[134]</td>
<td>Microphone</td>
<td>1.500 W</td>
<td>19.80 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Son et al.</td>
<td>2015</td>
<td>[10]</td>
<td>MEMS</td>
<td>113dB</td>
<td>0.10 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Trippel et al.</td>
<td>2017</td>
<td>[13]</td>
<td>MEMS</td>
<td>110dB</td>
<td>0.10 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Bolton et al.</td>
<td>2018</td>
<td>[12]</td>
<td>HDD</td>
<td>130dB</td>
<td>0.10 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Shahrad et al.</td>
<td>2018</td>
<td>[53]</td>
<td>HDD</td>
<td>103dB</td>
<td>0.70 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Tu et al.</td>
<td>2018</td>
<td>[43]</td>
<td>MEMS</td>
<td>135dB</td>
<td>7.80 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
<tr>
<td>Khazaaleh et al.</td>
<td>2019</td>
<td>[129]</td>
<td>MEMS</td>
<td>94dB</td>
<td>0.11 m</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
<td>✖️</td>
</tr>
</tbody>
</table>

looking up the datasheets of the signal generators, antennas, and amplifiers used. This lack of experimental details is further discussed in the context of future research in Section XI.

We identify five key aspects of vulnerability that attacks exploit: resonance; non-linearity; improper filtering; poor shielding; and insecure algorithms. All attacks which do not target microphones and optical sensors depend on resonance of some sort: this can be acoustic resonance of mechanical structures, electromagnetic resonant frequencies of unintentional antennas, or the existence of locking frequencies for ring oscillators. Other attacks depend on non-linearities of amplifiers, microphones, and speakers to demodulate high-frequency signals. This is because resonant frequencies are often much higher than those of desired injection waveforms.

In addition, many works identify improper filtering, particularly prior to ADCs and amplifiers, as well as poor shielding as factors for out-of-band attacks. Finally, in some cases, insecure sampling and processing algorithms exacerbate the problem by making it easier for an adversary to trick the system under attack into performing a dangerous action. These sources of vulnerability form the basis for many of the proposed countermeasures, which we discuss in detail in Section IX.

IX. ANALYSIS OF COUNTERMEASURES

Although the literature on out-of-band attacks is quite broad, research on defenses has been more sparse. Section IX-A first summarizes the state-of-the-art in countermeasures specific to out-of-band injections. Section IX-B then expands our discussion by introducing general protective and preventive approaches which remain applicable in this context.

A. Out-of-Band Defense Mechanisms

The works that have investigated countermeasures against out-of-band signal injection attacks have noted that a combi-
TABLE III: Summary of evaluated (✓), proposed (+), and absent (−) countermeasures against acoustic, conducted, electromagnetic, and optical out-of-band signal injection attacks.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>Ref.</th>
<th>Target</th>
<th>Robust Hardware</th>
<th>Better Sampling</th>
<th>Sensor Fusion</th>
<th>Improved Filtering</th>
<th>More Shielding</th>
<th>Anomaly Detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Markettos and Moore</td>
<td>2009</td>
<td>[35]</td>
<td>TRNG</td>
<td>✓</td>
<td>−</td>
<td>−</td>
<td>✓</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Rasmussen et al.</td>
<td>2009</td>
<td>[33]</td>
<td>Medical</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>✓</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Bayon et al.</td>
<td>2012</td>
<td>[36]</td>
<td>TRNG</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Buchovecka et al.</td>
<td>2013</td>
<td>[37]</td>
<td>TRNG</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Kasmi and Lopes Esteves</td>
<td>2015</td>
<td>[38]</td>
<td>Microphone</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>✓</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Petit et al.</td>
<td>2015</td>
<td>[14]</td>
<td>Camera</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>✓</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Shoukry et al.</td>
<td>2015</td>
<td>[69]</td>
<td>Active Sensors</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>✓</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Son et al.</td>
<td>2015</td>
<td>[10]</td>
<td>MEMS</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>✓</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Bayon et al.</td>
<td>2016</td>
<td>[49]</td>
<td>TRNG</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Cao et al.</td>
<td>2016</td>
<td>[122]</td>
<td>TRNG</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Shin et al.</td>
<td>2016</td>
<td>[48]</td>
<td>Medical</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Roy et al.</td>
<td>2017</td>
<td>[54]</td>
<td>Microphone</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Song and Mittal</td>
<td>2017</td>
<td>[133]</td>
<td>Microphone</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Wang et al.</td>
<td>2017</td>
<td>[34]</td>
<td>MEMS</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Bolton et al.</td>
<td>2018</td>
<td>[12]</td>
<td>HDD</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Lopes Esteves and Kasmi</td>
<td>2018</td>
<td>[40]</td>
<td>Microphone</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Nashimoto et al.</td>
<td>2018</td>
<td>[41]</td>
<td>MEMS</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Osuka et al.</td>
<td>2018</td>
<td>[42]</td>
<td>TRNG</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Shahrad et al.</td>
<td>2018</td>
<td>[53]</td>
<td>HDD</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Tu et al.</td>
<td>2018</td>
<td>[43]</td>
<td>MEMS</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Giechaskiel et al.</td>
<td>2019</td>
<td>[44]</td>
<td>ADC</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Khazaaleh et al.</td>
<td>2019</td>
<td>[129]</td>
<td>MEMS</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Mahmoud and StoiLović</td>
<td>2019</td>
<td>[123]</td>
<td>TRNG</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Muniraj and Farhood</td>
<td>2019</td>
<td>[147]</td>
<td>Servo</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Yen et al.</td>
<td>2019</td>
<td>[134]</td>
<td>Microphone</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Tharayil et al.</td>
<td>2019</td>
<td>[148]</td>
<td>MEMS</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Zhang and Rasmussen</td>
<td>2020</td>
<td>[105]</td>
<td>Generic Sensors</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

Observation 3: The effectiveness of proposed countermeasures remains mostly theoretical, as practical implementations are often limited in scope, with superficial discussion of monetary and computational costs.

Robust Hardware: In response to resonance and non-linearity vulnerabilities, various works have proposed preventive improvements in the hardware itself to make it more robust and less susceptible to attacks. One of these improvements against electromagnetic attacks reduces asymmetries in differential inputs to a system. By doing so, attacker transmissions are received almost identically by the two unintentional receiving antennas, and are severely attenuated. For example, Markettos and Moore recommend reducing the asymmetries in ring oscillators through “carefully balanced transistors”, or the use of differential ones, which are “less affected by supply and substrate noise” [35]. Similarly, Foo Kune et al. found that using differential rather than single-ended comparators attenuated signals by up to 30dB [5]. Although signals could still be injected, the power requirements to do so increased significantly, thereby raising the bar for attackers.

Another approach is to change the sensors themselves, rather than attempt to improve the physical layout of a circuit. For example, both Shahrad et al. [53] and Bolton et al. [12] note that replacing Hard-Disk Drives (HDDs) with Solid-State Drives (SSDs) thwarts acoustic resonance attacks due to a lack of moving parts. In a similar vein, Zhang et al. noted that the iPhone 6 Plus resisted their inaudible voice commands, since it is “designed to suppress any acoustic signals whose
Zhang and Rasmussen [105]. Oversampling by a factor of 2\(n\) and selectively turning sensors on and off allows detection of out-of-band electromagnetic attacks: without knowing the secret sequence (1001 here for \(n = 4\)), the adversary will cause inconsistent (A) or unexpected (B) non-zero samples.

![Fig. 11: High-level overview of the defense mechanism by Zhang and Rasmussen [105].](image)

Finally, better frontends with fewer non-linearities are less sensitive to EMI noise [38], [44] and sonic injections [11], [13], [34]. They can therefore make it harder for adversaries to inject their desired signals into the system. Such general designs are discussed in greater detail in Section IX-B.

**Better Sampling:** Many papers have proposed improvements in the sampling technique to make it harder for an adversary to predict how a high-frequency signal will be converted to a low-frequency one. In 2015, Shoukry et al. proposed an alternative method of sampling active sensors called “PyCRA” (for Physical Challenge-Response Authentication) to detect signal injection attacks [69]. Active sensors “perform some action to evoke and measure a physical response from some measurable entity”, and include, for instance, magnetic encoders measuring angular velocity. Shoukry et al.’s proposal revolves around physical challenges to prove the absence of adversarial transmissions. Specifically, when the actuator is off (silenced), there should be no measured quantity unless an attack is taking place. By only shutting down the actuator for a small period of time, PyCRA can detect attackers without compromising the quality of sensor measurements and actuation results. Adversaries cannot stop transmissions in time due to physical and computational delay limits, allowing PyCRA to identify them [69]. It should be noted that Shin et al. have suggested that PyCRA would require high computational overhead in practice, both in terms of the minimum sampling rate needed to hit those physical limits, and for the detectors themselves [48]. Moreover, PyCRA requires active sensors, so it primarily protects against in-band attacks.

However, a similar proposal by Zhang and Rasmussen recently showed how to protect both powered and non-powered passive sensors [105]. The key idea is to use a secret bitstream to selectively turn off the sensor and observe whether the measured signal has been altered by electromagnetic injections, as shown in Figure 11. More concretely, for each sensor measurement, \(2n\) ADC samples are taken, corresponding to an \(n\)-bit secret sequence. Each secret bit is Manchester-encoded, so that a 0-bit is represented as the pair \((1, 0)\), corresponding to turning the sensors off for one sample, and then turning them on (“biasing”) for another sample. A 1-bit is similarly encoded as the pair \((1, 0)\), first turning on the sensor, and then turning it off during the second sample.

When the sensor is turned off, all samples should be close to zero, within some noise- and device-dependent tolerance. Moreover, for fast-enough sampling frequencies and slow-enough sensor signals, the \(n\) samples when the sensor is on should be close to each other. As a result, to inject a single measurement successfully, an attacker needs to correctly predict the \(n\) secret bits, which only happens with probability \(2^{-n}\) for a randomly chosen bit sequence. By using a switch, non-powered passive sensors can also be adapted to use this approach. Moreover, spikes in the frequency domain allow Zhang and Rasmussen to detect attacker transmissions even for non-constant sensor signals [105]. Overall, by oversampling for each sensor measurement, noise can be distinguished from adversarial signals with probabilistic guarantees.

An alternative approach to prevent attackers from injecting their desired waveforms into a system is to add randomness to the sampling process, especially for ADCs which are only vulnerable for limited carrier frequencies [44]. The effect is essentially one of “having an inaccurate ADC” [13], allowing a moving average to filter out injected periodic signals. This is similar to sampling with a “dynamic sample rate”, defeating the side-swing and switching attacks of Tu et al. [43], which were explained in Section VI. Out-of-phase sampling has also been proposed as a band-stop filter to reject frequencies near an accelerometer’s resonant frequency, thereby removing attacker-injected DC offsets [13].

It should be noted that protecting against signal injection attacks into actuators has not been studied as extensively in literature. However, Muniraj and Farhood recently proposed a detection method based on a watermarking scheme that slightly alters the actuation parameters [147], similar to the proposal of Zhang and Rasmussen [105]. Under attack, the measured response of the system does not match the effect of the watermark, allowing detection. The same paper also suggested pseudo-randomly changing the pulse frequency of the Pulse Width Modulation (PWM) signals, making an actuator attack harder to accomplish. Overall, these methods only alter the shape of the waveform that an adversary can inject rather than the root cause of the vulnerability itself. They are therefore not sufficient countermeasures by themselves.

**Sensor Fusion:** A few works have suggested that using sensors of different types (fusion) or multiple sensors of the same type (duplication) with different vulnerable frequency ranges will make injections harder [12], [14], [34], [43], [48]. This is because an adversary would need to mount multiple simultaneous attacks, which potentially interfere destructively. However, in “Sensor CON-Fusion”, Nashimoto et al. [41] showed that a fusion algorithm based on Kalman filters could be circumvented (Section VI). As a result, better techniques are needed to protect against adversarial injections, instead of simply faulty readings [149], [150]. Tharayil et al. proposed an improved such fusion algorithm, which takes into account the mathematical relations between the underlying physical quantities. This allows them to link measurements by a gyroscope and a magnetometer in a way which can detect adversarial
injections without any hardware modifications [148].

Other researchers have proposed that additional sensors be used to measure and counteract attacker signals. These additional components should not be identical to the vulnerable sensors: as Khazaaleh et al. show, many MEMS gyroscopes integrate a second, “identical proof mass to perform differential measurements” and “eliminate unwanted vibrations” [129]. However, they still remain vulnerable to ultrasonic attacks.

Instead, “an additional gyroscope [...] that responds only to the resonant frequency” may be able to remove the resonance effect from the main gyroscope [10]. Similarly, microphones might be able to detect (and potentially cancel) resonant frequencies to protect MEMS gyroscopes [34] and HDDs [12]. This approach would be hard in practice, at least for hard drives: the area to be protected would need to cover “the read/write head [completely] as it moves across the disk”, and the sound wave to be generated would be potentially large, raising many issues about its implementation [12].

Filtering: Most papers studied highlight the need for better filters to reduce the vulnerable frequency range against conducted [35], [40], [44], [104], electromagnetic [5], [7], [42], [44], [104], [105], acoustic [13], [43], and optical attacks [14]. However, only Foo Kune et al. [5] have performed systematic experiments studying the effect of filtering on the efficacy of out-of-band signal injection attacks. To start with, Foo Kune et al. noted that adding a low-pass filter in their experiments against Bluetooth headsets allowed audio signals to pass, et al. noted that “while a light sensor can function in a permissive choice in the design of voice interfaces can help mitigate signal injection attacks. For example, safe defaults when the sensor output is deemed as untrustworthy [5], [8] can reduce the effects of successful attacks on health-and safety-critical actions taken by systems. Similarly, less permissive choices in the design of voice interfaces can prevent non-targeted attacks from succeeding. As an example, adding voice authentication and custom keywords can prevent
command injections into smartphones [38], [40], [44].

**Observation 4:** Until more resilient components replace vulnerable ones, defense-in-depth is necessary to protect against signal injection attacks. This can be accomplished through better filtering and shielding to prevent attacks, and through better sampling, fusion, and anomaly detection algorithms to identify them.

### B. Other Defensive Approaches

As out-of-band signal injection attacks are closely connected with different areas of research (Section X), there is extensive overlap in the proposed countermeasures. For example, before Foo Kune et al. [5] proposed an adversarial EMI detector, Wan et al. [153], [154] introduced a similar design to “increase the immunity of a microcontroller-based system in a complex electromagnetic environment”. Moreover, to protect against LiDAR attacks, Shin et al. proposed sensor fusion and redundancy, fail-safe defaults, better shielding (by reducing receiving angle), and randomized pinging directions and waveforms [31]. Similarly, Davidson et al. proposed sensor fusion and an improved optical flow algorithm to protect against optical in-band sensor spoofing [52]. Moreover, Blue et al. detected (audible) command injections by identifying a frequency band which is produced by electronic speakers, but is absent in human speech [152]. In fact, detecting unique features of the sensed property is a common defense mechanism for general sensor manipulation attacks, such as those against Smart Grid power plants [155], [156], or unmanned aircraft systems [147]. However, such approaches require a theoretical system model, and assume an adversary who cannot inject data obeying this model.

**Observation 5:** Defense mechanisms for in-band attacks, excessive environmental noise, and faulty sensors are often directly applicable to out-of-band signal injection attacks and vice versa.

In a different strand of research, Redouté and Richelli have proposed some guidelines for improving immunity against EM interference attacks [157], [158]. These recommendations could be applied in the context of general out-of-band attacks:

1. **Filter induced signals before the non-linear device.** This suggestion is not limited to amplifiers, but can be used in other setups, including power transistors [159]. It has been shown to result in an up to a 12.5× reduction in EMI-induced offsets [157], [160], [161], but may require bulky passive components, adding noise to the circuit.
2. **Linearize the stage generating the DC shift,** for example, by using amplifiers with a wider common mode input range, resulting in better linear behavior [162].
3. **Prevent the accumulation of DC shift,** for instance by addressing the slew rate asymmetry and parasitic capacitances [163]–[165].
4. **Compensate and remove the induced offset,** for example, using cross-connected differential pairs [164].

As discussing all possible EMI-resistant amplifier designs is out-of-scope, the interested reader should refer to various comparative works [165]–[167] as a starting point. Similarly, one should refer to advances in gyroscope technologies [124], [168], [169] which do not use MEMS constructions, or reduce sensitivity to random vibrations: as Khazaaleh et al. noted [129], removing the “misalignment between the sensing and driving axes” will make systems more secure against out-of-band acoustic attacks.

**Observation 6:** More accurate and sensitive hardware that is robust to environmental influences is a natural defense mechanism against out-of-band attacks.

### X. ADDITIONAL RELATED WORK

As this study contains the first survey of out-of-band signal injection attacks, this section shows the close connections with side-channel leakage and electromagnetic interference. For example, using insights into the resonant frequencies of gyroscopes, Farshteindiker et al. showed that unprivileged websites could act as covert channel receivers, even at very low sampling frequencies of 20Hz [170]. Block et al. improved the design by not requiring external equipment for the attack, instead relying on the smartphone’s speaker and accelerometer [171]. Matyunin et al. then used the same effect for cross-device tracking using ultrasonic transmissions at or near the resonant frequencies of gyroscopes [172]. Moreover, Michalevsky et al. showed that MEMS gyroscope measurements are sensitive to acoustic signals in their vicinity [173]. As a result, they can be used to distinguish between different speakers, and, in part, the content of the speech [173] due to conducted vibrations of the loudspeakers used [174], [175].

In other words, the same source of vulnerability which can be used to destabilize [10] and control [13], [34], [43] gyroscopes and accelerometers can be used for covert channel communication [170], [171], tracking [172], and speaker identification [173], [174]. Similarly, instead of using microphone non-linearities for command injections [11], [55], [134], Shen et al. [176] and Chen et al. [177] leveraged them to protect users’ privacy by jamming nearby recording devices.

The countermeasures proposed in the works above mirror those of Section IX-A, and include anti-aliasing filters, shielding, and sensor fusion. Moreover, suggestions to increase noise in side- and regular-channel emissions parallel out-of-band defense mechanisms based on reducing the sampling accuracy. For example, decreasing “the fidelity of the input audio” can prevent against inaudible voice injection attacks [136]. Similarly, fonts which minimize emissions at high frequencies [178]–[180] exploit the human eye sensitivity to “low spatial frequencies” [178]. As EM emanations of video display units (“TEMPEST”) [181]–[183] mostly convey “the high-frequency part of the video signal” [178], images are transformed in a way that is almost transparent to human viewers, but prevents the reconstruction from side-channel listeners. Researchers have likewise shown that adding certain patterns to video frames [184] or the flashing of LED lights [185]...
can reduce the fidelity of reconstructed images from camera recordings, while not influencing regular viewers as much. Finally, in some respects, anomaly detection resembles statistical and machine learning approaches to detect covert channels, and can therefore draw inspiration from seemingly unrelated disciplines, such as timing and storage network covert channels [186].

Observation 7: The sources of vulnerabilities for out-of-band signal injection attacks are often the same as those for hardware-based covert- and side-channel attacks. This allows the same techniques to be reused for attacks and defenses across disciplines.

Other research has indicated that devices which are typically used as actuators can actually effectivley function as sensors. LEDs can function as photodiodes [138], while speakers [187] and HDDs [188] can both be converted into microphones. Although all three attacks have so-far required the assistance of malware, further research is required to identify the implications for out-of-band signal injection attacks.

Observation 8: Reuse of off-the-shelf equipment in unconventional setups expands the surface for signal injection attacks exploiting hardware imperfections.

The lines between electromagnetic interference and out-of-band signal injection attacks are also blurred. This is in part because the self-classification of attacks depends primarily on the research community with which an author is aligned, rather than the end result of the injection. For example, the voice injection command attacks of Kasmi and Lopes Esteves [38], [40] are categorized by their authors as Intentional Electromagnetic Interference (IEMI) attacks, despite the relatively low power used, and the lack of upsets or destruction of equipment. Similarly, Osuka et al. considered their work to be in the IEMI realm [42], even though they biased the randomness of a TRNG. This fact also partially explains why research on out-of-band attacks against TRNGs has largely ignored attacks against other targets and vice versa.

In general, this mismatch of expectations often results in unexplored avenues of research, as can be seen, for example, in the IEMI attacks on UAVs of Lopes Esteves et al. [189]: although there is a strong inverse correlation between the battery temperature reading and the strength of the electric field, the authors do not further investigate how to precisely control the sensor output. Moreover, as was explained in Section IX-B, research into electromagnetic interference can provide insights into how to build more resilient hardware, even when the hardware is only tested against “unintentional parasitic signals and does not take a malicious behavior of an attacker into account” [40].

Observation 9: The proposed terminology based on the outcome rather than the method of injection can help systematize attack and defense approaches, and reveal previously unexplored connections.

XI. Future Directions

Despite the amount of research conducted on out-of-band signal injection attacks, there is no common methodology to evaluate how susceptible systems are to them. This is in contrast to related disciplines, such as side-channel analysis [75], direct power injection and near-field scan immunity [59], fault injection attacks [4], and IEMI attacks [190]. Indeed, although many papers sweep through frequencies to find the resonant ones [5], [13], some do not adopt this terminology [10], [12], [53], and do not specify how wide the frequency steps should be. This proves to be particularly problematic, as some attack windows “are as narrow as a few Hertz” [53].

Recently, Tu et al. [43] provided a more detailed methodology for acoustic injection attacks, which starts with a profiling stage. During this phase, single-tone sounds are transmitted, and are swept at an interval of 10Hz. The devices targeted remain stationary during the profiling stage. Further increments of 1Hz or smaller can be used near the resonant frequencies to estimate the sampling frequency of the ADC, and account for its drift. The next stage involves synchronizing to a frequency which is close to a multiple of the ADC sampling rate. This step is followed by manipulating the attack parameters, and adjusting them in response to drifts. Although this approach provides some common ground for evaluation, several questions remain unanswered, especially when assessing countermeasures to claim that a system is secure. These questions include: what the frequency range itself should be; what the step should be for wide ranges; what modulation method to use and with what parameters; and whether there are other factors that would need to be examined during experimentation. For instance, for electromagnetic attacks, the incident angle of the EM field and the distance of attacks can have a profound impact on their success, especially as they relate to generalizing from the near- to the far-field.

Observation 10: A precise experimental procedure which specifies sweep, modulation, and other parameters is needed for out-of-band signal injection attacks.

The question of the maximum feasible attack distance has mostly been of theoretical interest, with practical attacks often limited to a few centimeters. Even though EM attacks should in theory have a longer range than acoustic and optical attacks, the converse appears to be true in the experiments conducted by the works studied in this survey (Table II). There is also a worrying trend of assuming that more power and more expensive equipment easily translates to a long-range attack. For example, Tu et al. [43] claim that with more speakers, gyroscopes can be attacked from an 8× longer distance, but as Roy et al. [55] showed, doing so is not a trivial engineering concern, if the inaudibility of injections is to be maintained. Similarly, although Foo Kune et al. [5] claim that a 20dB gain directional antenna and a 1W source can attack equipment at distances of up to 50m, these estimates seem optimistic: according to Lopes Esteves and Kasmi [40], a 200W source is required for a distance of 4m for remote command injections [38]. What is more, high-powered EM sources have
the potential to cause faults in other equipment and be harmful to human life. As a result, determining how to inject precise signals from a distance is particularly challenging. These problems become even more pronounced for magnetic attacks on actuators, which have been limited so far.

**Observation 11:** Dedicated facilities and test equipment for long-range experimentation are needed.

Note that as many of the systems targeted are safety- and mission-critical, we can expect that, in the future, some devices may be required to undergo a certification process. Indeed, a CERT alert warning of MEMS susceptibility to ultrasonic resonance [21] highlights that out-of-band signal injection vulnerabilities are a concern for governments and corporations alike. Regulations will thus pave the way for an expanding industry around facilities and test equipment for EMC immunity against adversarial injections.

As the currently published work often leaves experimental details under-specified, reproducibility becomes a significant challenge: for instance, as discussed in Section VIII, details on the power used were often not readily available, but required searching through datasheets. Moreover, the duration of attacks was also often not specified. Trippel et al. [13] reported that some of their attacks against accelerometers only work for a couple of seconds before the attack fails. Sampling rate drifts thus necessitated manual tuning, or more sophisticated attacking techniques, such as those proposed by Tu et al. [43]. However, without details of the function generator specifications, it would be hard to know whether some of the issues are caused by poor clock accuracy of the generator. This problem is bound to become even more pronounced when using Software-Defined Radio (SDR) and other low-end commodity hardware for attack weaponization.

Minor variations in the construction of devices can also have significant effects on the sensors’ behavior, and will potentially impact the reproducibility of attacks. For instance, Dey et al. [191] showed that otherwise identical accelerometers can be tracked due to slightly different performance characteristics. Giechaskiel et al. [44] recently introduced security definitions to address the lack of directly comparable metrics describing the outcome of injection attacks. However, the overall absence of experimental details, coupled with monetary costs and legal requirements associated with using the electromagnetic spectrum, make security research into out-of-band signal injection attacks a challenging space for new researchers to enter.

**Observation 12:** Reproducibility through common metrics which allow for direct comparison of the effects of injection and standardized experimental setups are necessary to advance the state-of-the-art.

Besides the defense mechanisms of Section IX, to protect future devices for attack, new security-sensitive products must take a fundamentally different approach to trusting the outputs of sensors. In the words of Fu and Xu, there is a need to “shift from component-centric security to system-centric tolerance of untrustworthy components”, perhaps taking note of advances in fault-tolerant literature [67]. Fu and Xu also recommend that sensor outputs be “continuously checkable by software for adversarial influence”, such as through internal debugging information that is hidden from accessible APIs [67]. They further highlight the need for interdisciplinary teams and education [67]. Indeed, until new hardware is deployed, many cross-disciplinary solutions will be required to prevent, detect, and mitigate attacks. As out-of-band signal injection attacks become more powerful, collaboration will be necessary to address the multifaceted research influences of the field.

**Observation 13:** Interdisciplinary research quantifying the effectiveness of countermeasures is needed to inform future hardware and software design choices.

## XII. Conclusion

Our ever-increasing reliance on sensors and actuators highlights the need for a comprehensive look into electromagnetic, conducted, acoustic, and optical out-of-band signal injection attacks. These attacks cause a mismatch between a physical property being measured by a sensor or acted upon by an actuator and its digitized version. Out-of-band signal injection attacks can be used to control or disrupt drones, hard drives, and medical devices, among others, with potentially fatal consequences on human life. In light of the importance of such attacks, this paper took the first step towards unifying the diverse and expanding research through a taxonomy of attacks, defenses, and terminology. Our work revealed interdisciplinary influences between seemingly disparate topics, and also made several observations that can inform future research in the area. Overall, better experimental and reporting procedures are necessary for direct comparisons of the effects of attack and defense mechanisms.

## References


[150] A. Richelli, G. Delaini, M. Grassi, and J.-M. Redouté, “Susceptibility of operational amplifiers to conducted EMI injected through the ground


Ilias Giechaskiel recently submitted his PhD thesis at the Department of Computer Science at the University of Oxford. He was a Clarendon and Cyber Security Scholar at Kellogg College, and was also funded by the EPSRC and the Oxford CDT in Cyber Security. His dissertation, “Leaky Hardware: Modeling and Exploiting Imperfections in Embedded Devices” was supervised by Prof. Kasper B. Rasmussen. Ilias holds a Master’s in Advanced Computer Science (with distinction) from the University of Cambridge, and a Bachelor’s in Mathematics (summa cum laude) from Princeton University. During his PhD, Ilias was also a Visiting Assistant in Research at Yale University, where he worked under Prof. Jakub Szefer on FPGA covert channels. His interests in hardware security extend beyond academia: Ilias has participated in numerous security Capture-the-Flag (CTF) competitions, and has interned at Bloomberg, Microsoft, Dropbox, Microsoft Research, and Jump Trading.

Kasper B. Rasmussen is an Associate Professor in the Computer Science Department at the University of Oxford. He joined the department in 2013 and was awarded a University Research Fellowship from the Royal Society of London in 2015. Prior to being at Oxford, Kasper spent two years as a post-doc at University of California, Irvine. Kasper completed his Ph.D. with Prof. Srdjan Capkun in the Department of Computer Science at ETH Zurich, where he worked on security issues related to secure time synchronization and localization, with a particular focus on distance bounding. His thesis won the “ETH Medal” for outstanding dissertation from the Swiss Federal Institute of Technology, and he was additionally awarded the Swiss National Science Foundation (SNSF) Fellowship for prospective researchers.