Electromagnetic Signal Injection Attacks on Differential Signaling

Youqian Zhang
youqian.zhang@cs.ox.ac.uk
University of Oxford
Oxford, United Kingdom

Kasper Rasmussen
kasper.rasmussen@cs.ox.ac.uk
University of Oxford
Oxford, United Kingdom

ABSTRACT

Differential signaling is a method of data transmission that uses two complementary electrical signals to encode information. This allows a receiver to reject any noise by looking at the difference between the two signals, assuming the noise affects both signals equally. Many protocols such as USB, Ethernet, and HDMI use differential signaling to achieve a robust communication channel in a noisy environment. This generally works well and has led many to believe that it is infeasible to remotely inject attacking signals into such a differential pair. In this paper, we challenge this assumption and show that an adversary can in fact inject malicious signals from a distance, purely using common-mode injection, i.e., injecting into both wires at the same time.

We explain in detail the principles that an attacker can exploit to achieve a successful injection of an arbitrary bit, and we analyze the success rate of injecting longer arbitrary messages. We demonstrate the attack on a real system and show that the success rate can reach as high as 90%. Finally, we present a case study where we wirelessly inject a message into a Controller Area Network (CAN) bus, which is a differential signaling bus protocol used in many critical applications, including the automotive and aviation sector.

CCS CONCEPTS

• **Security and privacy → Embedded systems security**.

KEYWORDS

electromagnetic interference, signal injection, differential signaling

ACM Reference Format:
Youqian Zhang and Kasper Rasmussen. 2023. Electromagnetic Signal Injection Attacks on Differential Signaling. In *Proceedings of The 18th ACM ASIA Conference on Computer and Communications Security (Conference AsiaCCS'23).* ACM, New York, NY, USA, 12 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Electrical cables are prevalent transmission media enabling communications between devices. Due to their "antenna-like behavior", the cables also pick up environmental electromagnetic waves, compromising signal integrity in the cables [5]. Such a phenomenon motivates the proliferation of studies abusing electromagnetic signals to inject malicious information into wired communications (see details in Section 2). We call such attacks as *electromagnetic signal injection attacks*. Several researchers recommended using *differential signaling* to resist the attacks [12, 28], as it can reject external noise by looking at the difference between two complementary signals, supposing the noise impacts both signals equally. We will show that due to circuits' asymmetry and nonlinearity, differential signaling cannot provide sufficient protection against the electromagnetic signal injection attacks. This allows an attacker, who has no physical access to a victim system, to control a victim system remotely. Note that many popular protocols (e.g., USB, Ethernet, and HDMI) are based on the differential signaling technique, and they serve in countless safety- and security-critical applications (e.g., automobiles, aviation, and robotics); hence, such attacks immediately put these applications at risk.

Specifically, the impact of the injected signals is to cause receiving circuits to detect incorrect bits [43]. Imagine a scenario of a wired USB keyboard. An attacker can use electromagnetic signals to interfere with the USB cable and injects malicious bits (by extension, arbitrary user inputs); as a result, she can deceive the victim system into receiving whatever she wishes. Such attacks on keyboards in critical infrastructures could cause catastrophic consequences, such as wrong prices in stock exchanges, incorrect traffic signals on railways, and overdoses of medicines for patients in hospitals.

Despite various application scenarios or communication protocols, differential signaling receivers work in the same way, making it possible to use a single attack strategy to attack many well-known communication protocols and applications. We start with abstracting and parameterizing a generalized system model from practical differential signaling receivers. Such a system model can help to capture the characteristics of different systems by tuning the parameters. Also, we define a comprehensive attacker model, clarifying an attacker's objectives and capabilities, as well as her limitations in practice. The system- and the attacker models together form a universal tool to describe the attacks (Section 3). Realizing such attacks is complicated: the first challenge is bypassing the interference rejection, and the second challenge is making the receiving circuits detect the intended bit. We systematically and experimentally investigates how to exploit the hardware imperfections to get rid of the two challenges above (Section 4 and Section 6). Further, to evaluate the performance of the attacks, we analyze the success rates of bit injections, discussing critical factors that determine a high success rate (Section 5). At last, we present a case study where an attacker can use the attack principle above to wirelessly inject

Conference AsiaCCS'23, July 10–14, 2023, Melbourne, Australia

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

an arbitrary message into a Controller Area Network (CAN), which is a widely used protocol in automotive and aviation applications (Section 7).

2 BACKGROUND

In this section, we present a brief background on electromagnetic signal injections, explaining why an malicious signal can be injected into circuits and pointing out essentials of an effective injection. As mentioned previously, wires/traces that are for signal transmission between/within circuits can also act as antennas to capture external electromagnetic waves. The capture process is rather complicated, but it has been well studied in the area of "Electromagnetism". Going into details is beyond the scope of this work, but in simple terms, the electromagnetic waves impact the metal conductors by inducing voltage changes in them. The voltage changes are then superimposed with original signals, causing undue waveform changes. Therefore, the antenna-like behavior of the wires allows a remote attacker to inject malicious signals.

Many researchers have thoroughly studied such electromagnetic signal injections and successfully demonstrated them in various systems [12, 13, 18, 23, 27, 28, 31, 36, 37, 44, 45, 47, 50–54, 56]. From the previous studies, in order to achieve effective injections, attack power and frequencies are two basic factors to be considered. First, it requires strong enough attack power. There are many reasons that the attacking signals will be attenuated, such as propagation loss, shielding, and filtering. Consequently, the injected signal may be too weak to cause effective impacts on the victim device. It is essential to emphasize that attenuation provided by shielding and filtering is finite [3, 49]. In our preliminary experiments, we wirelessly injected adversarial signals into cables of branded USB mice, USB keyboards, and routers (we used similar attack setup as shown in Section 7), and such attacks can cause bit errors and stop data transmission. These results imply that deployed attenuation methods (i.e., RF shielding materials and common-mode choke filtering) are not sufficiently enough to block such intentional adversarial signal injection attacks. Second, since the attack power is finite, it is also essential to consider how to maximize the injected power. Usually, this requires that the frequency of the attacking signal must be at the resonant frequency of the target wire. The resonant frequency of the wire can be approximated by its length; nonetheless, a better way to determine the resonant frequency is to have a copy of the receiving circuits and sweep through a range of frequencies [28]. Note that since the lengths of the wires in the victim devices usually range from as short as several millimeters to meters long, the frequencies of the attacking signals are typically in the MHz and GHz frequency bands.

Since the previous studies already demonstrated that it is not hard to inject a malicious signal into a victim system by electromagnetic waves, we do not detail the injection procedures hereafter.

3 SYSTEM MODEL AND ADVERSARY MODEL

We abstract and parameterize a system model of differential signaling from practical circuits. This model allows us to capture the characteristics of different circuits by tuning the parameters. Next, we define an adversary model, which explains an attacker's capabilities and limitations.

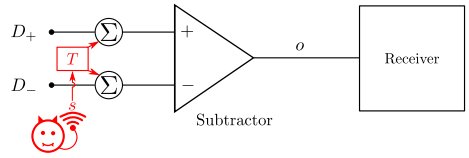
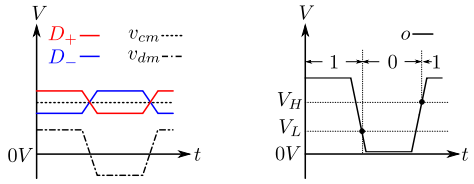


Figure 1: The system model consists of a pair of complementary signals (D_+ and D_-), a subtractor, and a receiver. The difference between the signals is extracted by the subtractor, and next, sent to the receiver. When an attack happens, identical injected signals are superimposed with the complementary signals



(a) Subtractor **(b) Receiver**

Figure 2: (a) The subtractor's two input signals D_+ and D_- can be represented by their differential mode v_{dm} and common mode v_{cm}. (b) The receiver compares o with two thresholds to determine the logic levels.

3.1 System Model

Recall that information is carried by the difference between the differential pair of signals. To obtain the difference, circuits that can "subtract" one signal from the other are used. We define the circuits with such a subtraction function as a "Subtractor", as shown in Figure 1: the subtractor has two inputs, each receiving a signal of the differential pair, and it sends the difference to the following circuits for further processing.

After receiving the subtractor output signal, an essential step is to convert its analog voltage levels into a sequence of bits. This is because the circuits that process information are usually digital (e.g., microprocessors) that only handle logic 1 and logic 0. After obtaining the bits, other functions or tasks such as decoding, error checking, authentication, etc., can be further executed. We define a "Receiver", as presented in Figure 1, to incorporate all functions of such circuits.

3.1.1 Parameterizing Subtractor. We denote the subtractor's two input signals as $D_+(t)$ and $D_-(t)$, and the output signal as $o(t)$. To simplify the notations, we omit time t hereafter. In order to explicitly show the information that is carried by the two input signals, these two input signals can also be rewritten by their differential mode and common mode. The differential mode is defined as the difference between two signals, and we denote it as $v_{dm} = D_+ - D_-$, and it is v_{dm} that represents the transmitted information; the common mode is defined as the average of two signals, and it is denoted as $v_{cm} = \frac{D_+ + D_-}{2}$. Such a relationship between the two input signals and their modes is illustrated in Figure 2a. Note that v_{cm} is a non-zero constant in almost all protocols, and thus, we assume that it is non-zero hereafter unless stated otherwise.

In practice, the subtractor is essentially a differential amplifier, and it is reasonable to model its output signal as a sum of the amplified differential mode and the amplified common mode [38]. We denote the gains for the differential mode and the common mode as G_{dm} and G_{cm}, respectively; note that the gains are functions of frequencies. The amplified terms are $G_{dm} \cdot v_{dm} + G_{cm} \cdot v_{cm}$.

In addition, there also exist distortion and noise that contaminate the output signal. The distortion originates from nonlinear properties of electronic components (e.g., transistors) that make up the subtractor [38], and we define a function $F(v_{dm}, v_{cm})$ to incorporate the distortion phenomenon. We model the noise as additive Gaussian white noise, denoting it as n. Thus, the subtractor output is expressed as:

$$o = G_{dm} \cdot v_{dm} + G_{cm} \cdot v_{cm} + F(v_{dm}, v_{cm}) + n \tag{1}$$

In Equation 1, the first term $G_{dm} \cdot v_{dm}$ explicitly carries the transmitted information, i.e., v_{dm}. It needs to be emphasized that every subtractor has a finite operating frequency range, within which it is designed to function properly. Inside this operating frequency range, the differential-mode gain remains constant, and as such the subtractor can guarantee a consistent output while handling input signals at different bit rates. The common-mode gain is so small that it makes typical attenuation of 70 dB − 120 dB to the common mode of the inputs [4], thus making $G_{cm} \cdot v_{cm}$ nearly zero. The distortion of the subtractor is also well maintained, and thus the impact of $F(v_{dm}, v_{cm})$ is negligible. Therefore, the subtractor is sufficiently good enough at rejecting the impacts of the common mode within the operating frequency range. However, this no longer holds out of the operating frequency range, and we will detail the reasons in Section 4.

3.1.2 Parameterizing Receiver. The primary function of the receiver is to convert analog voltages into bits as mentioned previously. It is a common way in practical circuits that the logic levels are determined by comparing analog voltage levels with two predetermined thresholds [20]. The reason for using two thresholds instead of a single one is that the difference between two thresholds can prevent the noise from causing wrongly detected logic levels. We denote these two thresholds as V_H and V_L, where $V_H > V_L$. The detection rule is straightforward: as depicted in Figure 2b, when $o \geq V_H$, a logic 1 will be detected; when $o \leq V_L$, a logic 0 will be detected. Specifically, when o is between these two thresholds (e.g., the noise causes the voltages to fluctuate into this region), the detected bit will retain its value. Note that the receiver detects bits periodically.

In addition, it is also essential to cover the circuits before the logic level detection, as an attacker needs to exploit these circuits to achieve a wrongly detected bit. We will detail the attack principles in Section 4.2, but here, we abstract a model and explain the functions of the circuits.

When a signal enters the receiver, it first goes through an electrostatic discharge (ESD) circuit, which is commonly used to protect the input pins of any electronic device from overvoltages. A block diagram of the ESD circuit is presented in Figure 3: it clamps the negative overvoltages to a minimum allowed voltage (e.g., GND), and the positive overvoltages to a maximum allowed voltage (e.g., V_{DD}) [39]. After that, a buffer circuit follows. It is used to get rid

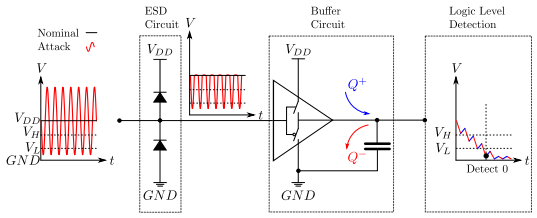


Figure 3: A model of circuits before logic level detection in the receiver. The charging and the discharging of the output parasitic capacitance are asymmetrical, and a net charge $Q^+ - Q^-$ causes the output voltage changes when an attack happens.

of an impedance mismatch between the previous stage and the receiver, and more precisely, it provides isolation and prevents undesired interaction from the previous stage [7]. The buffer circuit is essentially built up with transistors, which work like switches, and its function is abstracted in a way as illustrated in Figure 3: its input signal controls the switches, generating an output signal to reproduce its input signal. In this way, the buffer circuit transfers its input signal to the logic level detection.

3.2 Adversary Model

An attacker's objective is to inject a message with a length of L bits into a victim system. The attacker has no physical access to the victim system, meaning that she cannot modify its circuits, nor can she tap wires to inject attacking signals into it. Because of no physical access, it is not easy to know which bit is transmitted in the wires. However, the attacker can guess the bit, and we will further explain and discuss her guess ability in detail in Section 5. The attacker can wirelessly inject the attacking signals into the victim system by radiating electromagnetic waves, and she can tune her attacking signals, regarding their frequencies, power, etc. The attacker knows the period that the receiver detects a bit. In each bit injection, the duration of the attacking signal is set the same as the period, and hence, the attack can always interfere with the receiver when it detects a bit.

As mentioned previously, an electromagnetic signal injection is a complicated process in practice. We define a transfer function T to explain any changes to the attacking signal $s(t)$ due to the injection process, e.g., frequency selectivity, attenuation, spreading, antenna positions, etc. To simplify the notation, we omit time t, and thus the injected signal is denoted as $T(s)$. Note that there could exist multiple injection places in the victim system. However, only when the injection impacts the signals that carry the information will the attacker be able to manipulate the bits, which are recognized by the receiver. Therefore, it is equivalent to modeling the pair of wires as the injection point as shown in Figure 1. In fact, the differential signaling technique is usually deployed between two ends that are far from each other, and the pair of wires are effective antennas capturing the attacking signals in practice.

4 BIT INJECTION ATTACK

In this section, we will detail how an attack can make the victim system detect an incorrect bit. Note that as mentioned in Section 2, since previous research has thoroughly studied the signal injection

process, we do not further detail it here. However, we will focus on explaining why injected signals can bypass the subtractor, and next, how the receiver responds to the bypassed injected signals.

4.1 Bypassing Subtractor

The injected signals are superimposed with the two input signals of the subtractor, leading to $D_+ + T(s)$ and $D_- + T(s)$. According to the definition of the differential mode and the common mode in Section 3.1.1, the differential mode will not be affected because the injected signals cancel each other out; however, an extra term $T(s)$ that is the average of the identical injected signals is added to the common mode. As explained previously, the common mode of the original input signals has almost no impact on the subtractor output signal. Thus, under the interference of the attack, it is equivalent to approximating the common mode to be $T(s)$, or $v_{cm} = T(s)$. Substitute it into Equation 1, and we can obtain an expression of a malicious subtractor output signal o' as follows:

$$o' = G_{dm} \cdot v_{dm} + G_{cm} \cdot T(s) + F(v_{dm}, T(s)) + n \tag{2}$$

Note that the terms $G_{cm} \cdot T(s) + F(v_{dm}, T(s))$ are malicious changes that are caused by the injected signal $T(s)$, and these terms explicitly represent the bypassed injected signal.

It is essential to point out that the subtractor's common-mode rejection ability is finite, and it is ascribed to two widely accepted reasons. First, the two inputs of the subtractor are not perfectly symmetrical in practice. This results in that common-mode variations in the inputs are converted to differential-mode variations in the output, which is also known as "common-mode to differential-mode conversion" [19, 21, 33, 38, 55]. Second, nonlinearities of the subtractor lead to extra unexpected variations in the output [9, 10]. Especially beyond the operating frequency range of the subtractor, on the one hand, the subtractor's common-mode rejection ability deteriorates dramatically because the common-mode to differential-mode conversion becomes more and more significant at higher frequencies [38], and this indicates that $G_{cm} \cdot T(s)$ becomes larger. On the other hand, the nonlinear phenomenon of the subtractor becomes stronger [9], leading to much more distortion, or larger $F(v_{dm}, T(s))$. The evidence above indicates that if the injected signals are out of the operating frequency range, their impacts on the common mode are much more easily converted into additional malicious voltage changes in the subtractor's output. In this way, the injected signals bypass the subtractor.

It is essential for the attacker to know the waveform of the bypassed injected signal so that she can have a controllable impact on the next stage, i.e., the receiver. However, since the subtractor is not initially designed for use beyond the operating frequency range, it is not easy to precisely predict the waveform. To tackle this challenge, a determined attacker can get a replica of the subtractor and experimentally find the relationship between the output signal and the input signals of the subtractor, and as such, the attacker can estimate the bypassed injected signal. We will demonstrate it in Section 6.2.

4.2 Bit Detected Incorrectly in Receiver

Recalling in Section 3.1.2, we explained that the receiver determines a bit by comparing the subtractor output signal with two thresholds. In order to make the receiver detect an incorrect bit, the bypassed injected signal must make the nominal voltage level (which is represented by $G_{dm} \cdot v_{dm}$) of the subtractor's output signal cross the threshold that determines the opposite bit. Since detecting an incorrect bit is literally flipping a bit, we will use these two synonyms interchangeably hereafter. To make the explanation concise, we assume that the nominal voltage is at V_{DD}, which is above V_H, and it means that the receiver is supposed to receive 1 if no attack presents.

The oscillation of the bypassed injected signal either pushes the voltage level towards or away from V_L, as shown in Figure 3. However, only when the malicious voltage change causes the voltage level to move towards V_L will the receiver wrongly detect a bit. Luckily, the ESD circuit guarantees the direction of the malicious voltage change. This is because the positive part of the bypassed injected signal exceeds the maximum allowed voltage, leading to rectification by the ESD circuit, but the negative part remains, and thus the voltage level moves toward V_L, as shown in the middle of Figure 3.

After being rectified, the malicious voltage change continues propagating through the buffer circuit of the receiver. Note that the frequency of such a malicious signal is further beyond the operating frequency range of the receiver. The fast oscillation of the malicious voltage change will make the switches close and open periodically, thus charging and discharging the parasitic capacitance of its output periodically. However, the charging process and the discharging process are asymmetrical, leading to a quick accumulation of net charge across the output capacitance, and then, it holds still [8, 16]. If a bit is read when the voltage level crosses V_L, 0 is recognized.

In this way, the bypassed injected signal successfully makes the receiver detect 0. In a similar vein, when the nominal voltage level is below V_L (and the receiver expects 0), the bypassed injected signal also makes the receiver detect 1. Researchers pointed out that the impacts of the malicious voltage change are equivalent to adding a constant DC offset to the input signal of the receiver [8, 16]. The magnitude of this equivalent DC offset depends on the frequency and amplitude of the malicious voltage change, as well as the specific circuits that are impacted [16]. This implies that an attacker can successfully flip a bit by properly choosing the attack frequency and power. However, it is also not easy to predict the receiving circuits' responses out of their operating frequency ranges, and hence, it will be difficult to figure out a formula to calculate the effective attack frequency and power. Still, a determined attacker can experimentally obtain such attacking signals by sweeping the frequency and power to find ranges of effective attacking signals. We will demonstrate this in Section 6.3.

5 ANALYSIS OF SUCCESS RATE

When an attacker intends to inject a bit, it is essential to consider which bit is being transmitted in the system: if the transmitted bit is not what the attacker wants, she needs to emit an effective attacking signal to flip the bit; otherwise, she stops the emission, leaving the bit unchanged. In practice, it is not an easy job to know which bit is transmitted. Still, the attacker can make use of her knowledge about the victim system to make a guess of the bit, and her guess will further dictate her actions. In this section, we parameterize the

attacker's guess and the effectiveness of her attacking signals, and next, we analyze the success rate of the injection.

5.1 Parameterization

We denote the bit that is transmitted in the wire as A, and the attacker's guess as G. We use a parameter g to quantify the attacker's knowledge about A, and $g \in [0, 1]$. We define that $g = \frac{1}{2}$ means the attacker knows nothing about the bit, and there is an equivalent chance that the attacker will guess 1 or 0. Furthermore, we define that $g > \frac{1}{2}$ means the attacker knows information that indicates the bit could be 1. A larger g means that the attacker knows more information, and thus, it is more likely to guess 1; when $g = 1$, the attacker is sure that the bit is 1. Conversely, $g < \frac{1}{2}$ means the attacker knows information that indicates the bit could be 0, and a smaller g also implies knowing more information, and thus, it is more likely to guess 0; when $g = 0$, it means the attacker is sure that the bit is 0. We model that G follows a Bernoulli distribution with the parameter g, where G takes 1 with a probability of g and 0 with probability of $1 - g$.

We quantify the performance of an attacking signal by two parameters: u represents the probability of flipping 1 to 0, and v represents the probability of flipping 0 to 1, and $u, v \in [0, 1]$. For a certain victim system, each attacking signal corresponds to a pair of u and v, and all u, v pairs together characterize this specific victim system's responses to attacks. We define *feasible pairs* to incorporate all these pairs. In practice, u and v can be measured experimentally, and we will demonstrate the measurements and the characterization in Section 6.3.

5.2 Success Rate of Bit Injection

Let's begin by considering that the attacker intends to inject a single 1. There are four combinations of A and G, and the attacker's actions and the success rate for each combination are as follows:

- If $A = 1$ and $G = 1$, the attacker makes a correct guess, and since she intends to inject 1, she will not radiate any attacking signal. The success rate is 1, which can be written as $A \cdot G$.
- If $A = 0$ and $G = 1$, the attacker wrongly thinks that the bit is already 1 so that she will not radiate any attacking signal, meaning that she will never flip the bit. Hence, the success rate is 0.
- If $A = 1$ and $G = 0$, the attacker wrongly thinks that the bit is 0 and she will radiate an attacking signal. However, the attacking signal needs to keep the bit unchanged such that it is still 1. Thus, the success rate is $1 - u$, which can be written as $A \cdot (1 - G) \cdot (1 - u)$.
- If $A = 0$ and $G = 0$, the attacker's guess is correct, and the attacker will radiate an attacking signal to flip the bit. The success rate of flipping 0 is v, which can also be written as $(1 - A) \cdot (1 - G) \cdot v$.

We denote the success rate of injecting 1 as P_1, and it can be expressed as a combination of these cases:

$$P_1 = \begin{cases} G + (1 - G) \cdot (1 - u), & \text{if } A = 1 \\ (1 - G) \cdot v, & \text{if } A = 0 \end{cases}$$

Suppose the attacker intends to inject a single 0, we can use a similar way to reach an expression of the success rate of injecting 0, which is denoted as P_0 and expressed as:

$$P_0 = \begin{cases} G \cdot u, & \text{if } A = 1 \\ (1 - G) + G \cdot (1 - v), & \text{if } A = 0 \end{cases}$$

We will focus on the injection of 1 hereafter, as the injection of 0 is a symmetrical process and the explanation is similar.

5.2.1 Impact of g. To investigate the impact of g, we start from the expectation of P_1, which can be easily derived and expressed as:

$$E(P_1) = \begin{cases} u \cdot g + 1 - u, & \text{if } A = 1 \\ -v \cdot g + v, & \text{if } A = 0 \end{cases}$$

Essentially, the larger $E(P_1)$ is, the better. Since we are discussing the impact of g, it is reasonable to assume that u and v are non-zero here; otherwise, g vanishes in $E(P_1)$.

If $A = 1$, $E(P_1)$ increases with g. According to our definition of g, a bigger g means knowing more information about $A = 1$, and thus it is more possible to make a correct guess. The importance of making a correct guess can be easily proved: if $A = 1$, P_1 is maximized when $G = A$. Thus, it can be concluded that if $A = 1$, the larger g is, the more possible that P_1 will be maximized, and the better. Similarly, if $A = 0$, $E(P_1)$ increases while decreasing g, and a smaller g means a higher chance of making a correct guess, and thus more possible to maximize P_1.

Regarding P_0, the analysis is similar and we do not further detail it here. Therefore, to make a correct guess to maximize the success rate, two points need to pay attention to: first, it is crucial that g is in a manner conforming with A, and second, it is always better to know more information about A.

5.2.2 Impact of u and v. As indicated by the equation of P_1, the larger $1 - u$ and v are, the better. However, it needs to be emphasized that in a specific system, u and v are related in a certain way, and an example is shown in Figure 11. From our experiments with different chips in Section 6.3, we observe that there is a trade-off between increasing $1 - u$ and increasing v. Then, here comes the question: Which is the optimal pair?

Determining the optimal pair can be formulated into a multi-objective optimization problem, where $1 - u$ and v are the objectives. The most extensively used method of solving such an optimization problem is called the weighted sum method [6, 32], where the two objectives are combined and converted into one scalar, composite objective function by assigning proper weights to them; note that the sum of the weights equals 1. We select g as the weight for $1 - u$, and $1 - g$ as the weight for v, and the reasons are as follows.

Firstly, if the attacker has no knowledge about A (where $g = \frac{1}{2}$), it is equivalently important to "keep 1 unchanged" and "flip 0". Hence, it requires that the weights are equal, and $g = 1 - g = \frac{1}{2}$ meets the requirement. Secondly, if the attacker knows information indicating that the bit is 1 (where $g > \frac{1}{2}$), "keeping 1 unchanged" is more important, and hence, more weight for $1 - u$ than v. Moreover, when more information is known, the importance of $1 - u$ further increases, and so does the weight. Since $g > 1 - g$ and knowing more information also means that g increases, g can properly quantify the weight of $1 - u$, and accordingly, $1 - g$ quantifies the weight

of v. Thirdly, if the attacker knows information indicating that the bit is 0, we can also deduce that g as the weight for $1-u$ and $1-g$ as the weight for v in a similar way, and we do not further detail the reason. Therefore, searching for the optimal pair of u and v of injecting 1 is solving the following problem:

$$\begin{aligned} \max_{(u,v)} \quad & g \cdot (1-u) + (1-g) \cdot v \\ \text{s.t.} \quad & (u,v) \in \text{feasible pairs} \end{aligned}$$

In a similar vein, concerning injecting 0, the larger u and $1-v$ are, the better. Finding the optimal u and v of injecting 0 is solving the following problem:

$$\begin{aligned} \max_{(u,v)} \quad & g \cdot u + (1-g) \cdot (1-v) \\ \text{s.t.} \quad & (u,v) \in \text{feasible pairs} \end{aligned}$$

In Section 6.3, we will demonstrate how to use the method above to find the optimal pairs and then verify that the optimal pairs will do better than other pairs. Note that since the attacker has no access to the victim system, when she is preparing attacking signals, she needs to conduct experiments on a replica and use the methods above to find the optimal pairs.

5.2.3 Measuring Susceptibility. Although the attacker cannot access the victim system, a system designer of the victim system can do so. Thus, she can measure and obtain the optimal pairs, and then, use them to estimate the success rate. Note that the success rate is also a metric that sufficiently quantifies the susceptibility of the victim system. A higher/lower success rate means that the victim system is more/less susceptible to an injection. Thus, the system designers can use this analysis to quantitatively evaluate the security of their systems, and are able to change components or data modulation scheme to reduce adversarial success.

5.3 Success Rate of Message Injection

Recall that the attacker's objective is injecting L bits into the victim system, and the success rate of injecting a message will decrease exponentially with the message length. However, it needs to be pointed out that in Section 4.2, we explained that an attack can cause voltage changes to accumulate quickly and then holds still. Therefore, suppose the attacker will perform identical attacks (i.e., the same attacking signal, the same intended injected bit) on a sequence of transmitted bits that are consecutive and identical, once the first bit injection is successful, the success rates of the following bit injections will increase. This is because the first successful bit injection attack gets rid of many uncertainties with respect to the guess, the effectiveness of the attacking signal, timing, etc. We can approximate the success rate after the first successful bit injection to be 1 until the end of the consecutive injections. Such an approximation may overestimate the success rate of a message injection, as unpredictable responses in the victim system may still lead to a failure of a bit injection. An example of using such a method to estimate the success rate of a message injection will be shown in Section 7. It needs to be emphasized that system designers would rather overestimate the success rate than underestimate it because when they deploy measures to improve the security the nominal protection will make the victim system less susceptible in practice.

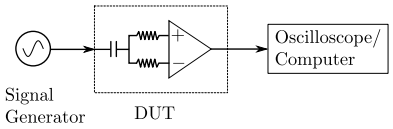


Figure 4: A testbed consists of a signal generator, a device under test (DUT), and an oscilloscope/computer.

6 EXPERIMENTS

In this section, we will demonstrate that injected signals can easily bypass the subtractor. Then, we will show that the bypassed injected signals cause incorrectly detected bits in the receiver.

6.1 Testbed

To test different chips of the subtractor and the receiver, we build a testbed before our experiments. The testbed's functions are generating attacking signals and measuring responses of a device under test (DUT).

A setup of the testbed is shown in Figure 4. A signal generator produces an attacking signal and injects it into the DUT through a wire. Such a signal injection setup is also known as Direct Power Injection (DPI) methodology [17]. Moreover, we use an oscilloscope to capture and measure the waveforms of the injected signals and the DUT's output signals, and a computer is used to process and analyze the measurements.

It is essential to point out that DPI is quite common in electromagnetic interference (EMI) research, and it allows us to precisely control the injected frequency and power so that we can measure the responses of the chips reliably. After figuring out how the chips respond to the attacks, Section 7 presents a *wireless* attack on a real-world CAN bus system to highlight that our attack works in practice.

6.2 Subtractor

We choose five different off-the-shelf subtractor chips, which are TJA1050, MCP2551, SN65HVD230, MX485, and SN751768P. They support CAN bus or RS485/422, and they are widely used in many critical applications such as automotive, medical equipment, and industrial devices. The subtractor chip is configured in a way as shown in the DUT block in Figure 4: two same resistors are added to the input of the subtractor, and these two transistors are equivalent to the terminated resistors in practice that are required by the differential signaling standards. Note that the voltage difference between the two inputs is internally configured to keep unchanged.

Regarding the injected signal, it is coupled to the midpoint between the two resistors by a capacitor. Doing so is equivalent to injecting a common-mode interference into the subtractor. The injected signal is sinusoidal, and its frequency is swept from 10 kHz to 100 MHz, and its peak-to-peak voltage of the injected signal is set to be 1 V, 2 V, and 4 V. Note that other waveforms such as square and sawtooth are potentially effective, but due to the limit of our signal generator, we cannot produce high-frequency and powerful signals with these special waveforms, so we stick to sinusoidal signals in our experiments.

6.2.1 Impacts of Injected Frequency and Power. As explained in Section 3.1.1, when no attack happens, the subtractor's output signal

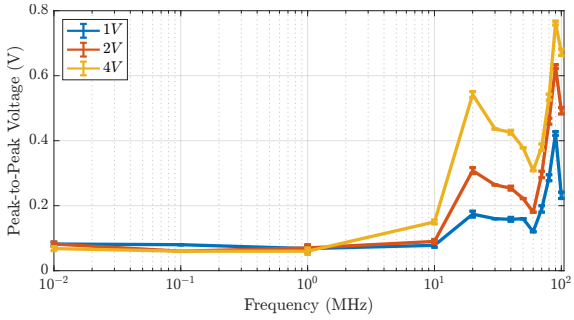


Figure 5: Test a subtractor chip TJA1050 with injected signals that have frequencies ranging from 10 kHz to 100 MHz and peak-to-peak voltages from 1 V to 4 V. The y-axis represents the output of the subtractor.

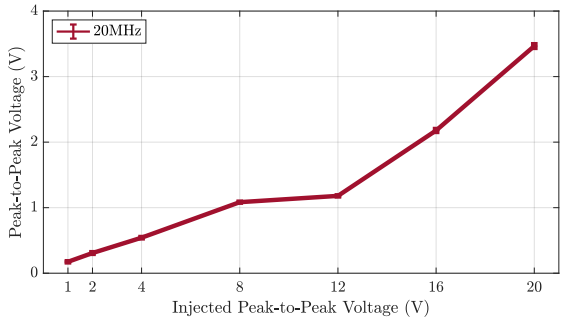


Figure 6: The strength of the bypassed injected signal increases while the injected power increases.

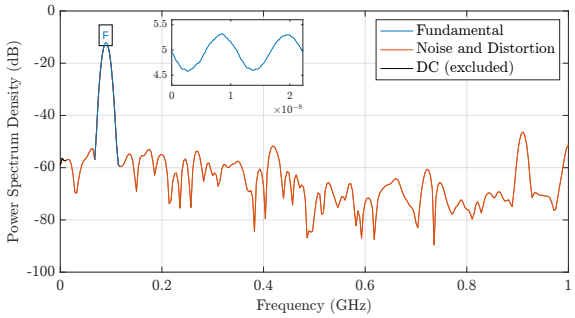


Figure 7: Set the injected frequency to be 90 MHz and the injected power to be 4 V. The frequency domain of the bypassed injected signal is presented. A time domain screenshot of the bypassed signal is in the floating window, where the x-axis is time (s) and the y-axis is voltage (V).

remains consistent with the differential mode of its input signals. In our configuration above, since the voltage difference between the two inputs is constant, the subtractor's output signal is also constant. Note that the noise essentially exists, but it is too small to significantly disturb the output signal. When an injected signal applies, the subtractor's output signal will start oscillating, and such an oscillation represents the bypassed injected signal. Note that the bypassed injected signal is explained by $G_{cm} \cdot T(s) + F(v_{dm}, T(s))$ in Equation 2. Therefore, we can use the peak-to-peak voltage of the subtractor's output signal to quantify the strength of the bypassed injected signal.

Taking a subtractor chip TJA1050 as an example, when there is no attack, the peak-to-peak voltage of the output signal is 0.06 V, which reflects the noise level. When an attacking signal is injected into the chip, the averaged peak-to-peak voltage and its standard deviation are shown in Figure 5. Between 10 kHz and 1 MHz, the output is as close as the noise level. This is because the common-mode interference is well handled within the operating frequency range. However, when the frequency is increased above 10 MHz, the peak-to-peak voltage has an increasing trend along with the frequency. These results explicitly show that the subtractor's common-mode rejection ability deteriorates out of the operating frequency range. In addition, two local maximums appear at 20 MHz and 90 MHz, as shown in Figure 5. This means that the injected signals at these two frequencies bypass this subtractor chip more efficiently than other frequencies. From the perspective of an attacker, she can take advantage of properly choosing the injected frequency to achieve a bypass using less attack power.

While increasing the injected power, the peak-to-peak voltage of the output also increases, implying a stronger bypassed injected signal. However, as shown in Figure 5, with the injected power of 4 V, the highest peak-to-peak voltage of the bypassed injected signal is still below 1 V. To achieve a stronger bypassed injected signal in the subtractor output, we use an RF power amplifier to increase the injected signal up to 20 V at 20 MHz, which is an efficient frequency that the subtractor lets the injected signal bypass. The output of the subtractor is shown in Figure 6. The results indicate that with increasing the injected power, the strength of the bypassed injected signal also increases. Also, it can be observed that the strength of the bypassed injected signal is roughly proportional to the injected power, and this allows the attacker to estimate the strength of the bypassed injected signal.

Note that such a bypassing phenomenon does not only occur in the TJA1050 chip but also in many other subtractor chips.

6.2.2 Impacts of Noise and Distortion. As indicated by Equation 2, the distortion $F(v_{dm}, T(s))$ plus the noise n will make the bypassed injected signal differs from the injected signal regarding waveforms. It is essential to know how much the bypassed injected signal is distorted because the bypassed injected signal will act on the receiver straight away and its waveform determines how the receiver responds.

To measure the impacts, we use a signal to noise-and-distortion (*SINAD*) ratio as a metric, which is calculated by the following equation:

$$SINAD = 10 \times \log_{10} \frac{P_s}{P_{n+d}} \tag{3}$$

where P_s is the power of the fundamental frequency of the signal, and P_{n+d} is the power of noise and distortion. The *SINAD* ratio is a widely used measure that quantifies the quality of a signal that is particularly degraded by the noise and distortion [25, 57]. The higher the *SINAD* ratio of a signal is, the better the signal quality is, and hence, less distortion in the signal. In Figure 7, a frequency domain of a bypassed injected signal is presented to show the difference between the fundamental frequency and the noise plus the distortion, and the *SINAD* is around 27 dB. In this figure, the

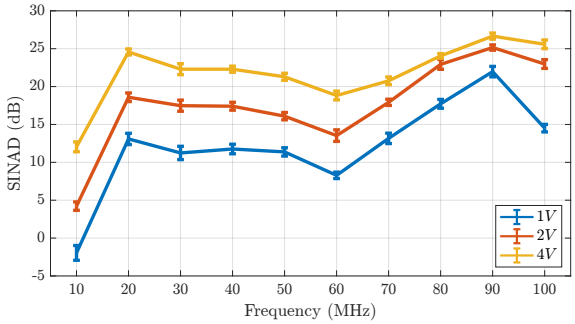


Figure 8: The signal to noise-and-distortion (*SINAD*) ratio quantifies the signal quality of the bypassed injected signal. The larger the ratio, the less the signal is distorted.

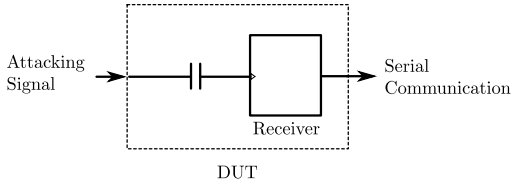


Figure 9: DPI is used to pump the attacking signal into the receiver, and the receiver sends measurements to a computer by serial communication.

distortion exists in the bypassed injected signal as harmonics, but they are too small to distort the bypassed injected signal significantly, which can also be observed from the time domain of the signal (please refer to a floating window at the top-left corner in the figure).

We also take the TJA1050 chip as an example, and in Figure 8, we show its *SINAD* ratio when different injected signals apply. The ratio is low when the injected frequency and the injected power are small (e.g., 10 MHz and 1 V), and this is because only a tiny amount of injected signal can bypass the subtractor, as explained previously. While either increasing the injected power or the frequency, the ratio has an increasing trend. In addition, the *SINAD* ratio is at least 10 dB for most of the measurements. Such a result implies that this chip demonstrates weak distortion and noise, which do not need to be worried too much while modeling its output signal. However, it does not mean that every subtractor chip has such weak distortion and noise, and an attacker still needs to handle them carefully.

6.3 Receiver

In various systems, microcontrollers are usually the devices that realize the receiver functions. We select three different microcontroller chips to test, which are nRF52833, ATMEGA328P, and ATSAM3X8E. We use the testbed that is shown in Figure 4 to study how the injected signal impacts bits that are recognized by the receiver chips. However, there are small modifications in the DUT block. First, as shown in Figure 9, the input is changed to single-ended. Second, because these chips support serial communication with the computer, the DUT directly sends recognized bits to the computer through a serial communication line.

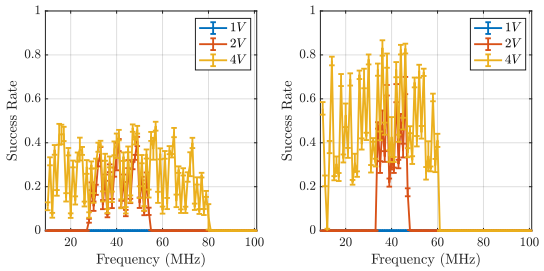


Figure 10: Success rates of bit injections in nRF52833. (a) Flip bits from 1 to 0. (b) Flip bits from 0 to 1.

We set the injected frequency from 10 MHz to 100 MHz with a step of 1 MHz. Note that since the subtractor chips have demonstrated that they can well remove the common-mode interference below 10 MHz, we do not further test that frequency range. The peak-to-peak voltage of the injected signal is set to be 1 V, 2 V, and 4 V. For each combination of the injected power and the injected frequency, 10 measurements are recorded; in each measurement, 256 bits are collected by the chip, and we calculate the percentage of successfully flipped bits as the success rate; then, the mean and the standard deviation of the success rates are calculated and presented.

Taking an nRF52833 chip that works at $V_{DD} = 3\,\text{V}$ as an example, it has $V_H = 2.1\,\text{V}$ and $V_L = 0.9\,\text{V}$ according to its datasheet [35]. Recalling in Section 3.1.2, V_H and V_L are two thresholds that are used to determine logic levels. To flip 1, the voltage change needs to be at least $3\,\text{V} - 0.9\,\text{V} = 2.1\,\text{V}$; conversely, to flip 0, it needs to be at least $2.1\,\text{V} - 0\,\text{V} = 2.1\,\text{V}$. The experimental results of flipping 1 are shown in Figure 10a, and the results of flipping 0 are shown in Figure 10b.

When the injected signal is 1 V, no bit flip is observed. This is because the injected signal is too weak to cause the voltage change beyond the threshold. When the injected signal is increased above 2 V, bit flips happen. Although the injected signal of 2 V is still weaker than the required threshold of 2.1 V, recall that as explained in Section 4.2 the voltage change can accumulate quickly and lead to a voltage change over the threshold ultimately, and consequently, the bit flips happen. When the injected power is increased to 4 V, the success rate becomes higher. Also, the frequency range where bit flip happens widens when the injected signal becomes much stronger. The results also imply that this chip is more susceptible in a frequency range that is centered at 40 MHz, and it is relatively easier to cause bit injections in this frequency range with less attack power.

In the other two chips, it is also observed that the success rates of bit injections are related to both the power and the frequency of the injected signal. The results show that the higher the power is, the higher the success rate is, and the wider the frequency range in which bit flips happen. Note that in Figure 10, the success rates show a periodic pattern in terms of the injected frequency: a peak appears every 2 MHz. Such a repeated pattern has nothing to do with the testing circuits outside the chip because the periodic pattern is not observed in other chips. It is speculated that some deterministic properties of that chip lead to this periodic pattern, but it is also

$$\max_{u,v} \quad g \cdot (1-u) + (1-g) \cdot v$$

g	Optimal (u, v)
$0.00 - 0.88$	$(0.09, 0.83)$
$0.89 - 0.90$	$(0.07, 0.69)$
$0.91 - 1.00$	send nothing

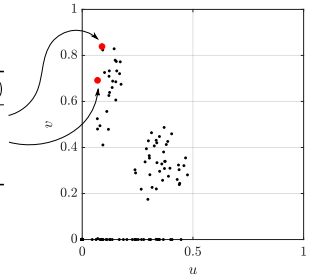
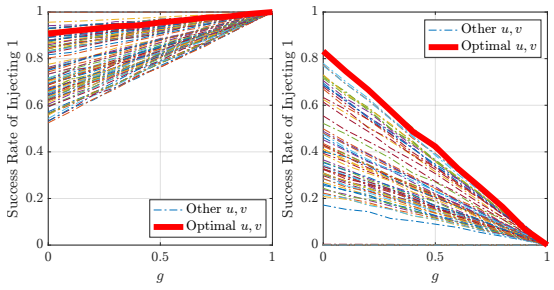


Figure 11: The pairs of u and v characterize the chip's responses to the attacks. Regarding injecting 1, the optimal pair with respect to g can be decided by solving the optimization problem.



(a) **(b)**

Figure 12: (a) If $A = 1$, the success rate of injecting 1 with using different pairs of u, v. (b) If $A = 0$, the success rate of injecting 1 with using different pairs of u, v.

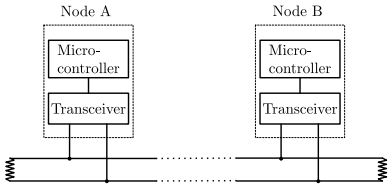


Figure 13: Nodes are connected to the same bus in CAN.

trivial to figure out the properties because knowing them does not help attack other chips.

6.3.1 Characterizing Receiver's Response. Recalling in Section 5, we introduce parameters u and v, and they can be used to characterize a victim device's responses to attacks. It is not difficult to find that the success rate of flipping 1 is u (see Figure 10a), and the success rate of flipping 0 is v (see Figure 10b). Thus, we can obtain the feasible pairs of u and v, and we plot them in Figure 11, which visualizes the chip's (nRF52833) responses to the attacks.

Note that $u = 0$ and $v = 1$ are an ideal pair, which represents an attacking signal that forces any bit to 1. The closer a pair is to it, the easier the injection of 1 will be. Similarly, $u = 1$ and $v = 0$ is the other ideal pair, which represents an attacking signal that forces any bit to 0. As shown in Figure 11, the feasible pairs' distribution is skewed to $u = 0$ and $v = 1$, meaning that it is much easier to inject 1 than 0 into this chip. Since injecting 1 and injecting 0 are symmetrical processes and the analysis will be similar, we focus on injecting 1 hereafter.

Recalling in Section 5.2.2, we formulate the method of determining the optimal pair of u and v. To determine the optimal pair regarding injecting 1, we assume that g is always correct, and we present the results in Figure 11. When g is below 0.88, the optimal pair is $u = 0.09$ and $v = 0.83$. Such an attacking signal can successfully flip 0 with a probability of 0.83 and keep 1 unchanged with a probability of $1 - 0.09 = 0.91$. With further increasing g, as explained in Section 5.2.2, the attacker is becoming more and more sure that the bit is 1, and hence, u decreases to 0.07. When g is greater than 0.9, the solution indicates that the attacker will send nothing. Next, we simulate attacks and show how the optimal pair outperforms.

6.3.2 Simulation and Success Rate. First, we simulate attacks with the optimal pair. The transmitted bit A is set to either 1 or 0, and g ranges from 0 to 1. We average the simulated success rates of each g and present the results in Figure 12. In Figure 12a, when $A = 1$, the success rate increases with g; in Figure 12b, when $A = 0$, the success rate decreases with g. The simulation results match with our model of $E(P_1)$ in Section 5.2.1, and in addition, the importance of having a g that is in a manner conforming with A is also explicitly shown.

Next, we repeat the simulation with other pairs of u and v, and compare them with the optimal pair, and the results are shown in Figure 12. In Figure 12a, when $A = 1$, some pairs outperform the optimal pairs, but these pairs are those that have small u and small v: they are good at keeping 1 unchanged, but they cannot flip 0 effectively. Therefore, as shown in Figure 12b, when $A = 0$, the optimal pair outstrips others.

To decide whether the optimal pair outperforms any other pair significantly, we can conduct multiple t-tests. Since the success rate has a linear relationship with g as shown in both Figure 12a and Figure 12b, we use the averaged success rate at $g = \frac{1}{2}$ as a metric to represent the attack performance. Note that the simulation is repeated 100 times for each pair, thus 100 samples for each pair. Next, we run t-tests to test against the alternative hypothesis that the optimal pair has a higher averaged success rate, namely, outperforms the other pair. The significance level is set to 0.05, which is conventionally accepted as the threshold. These tests show that they reject the null hypothesis, except the pair of $u = 0.092$ and $v = 0.82$. It is not surprising because it is the pair that is close to the optimal pair of $u = 0.09$ and $v = 0.83$, as shown in Figure 11.

7 MESSAGE INJECTION INTO CAN

A Controller Area Network (CAN) is a protocol that is devised to allow many devices to communicate with each other on a two-wire bus, and it is now deployed in many different applications from medical instruments to automotive. The CAN is a broadcast type of bus, and any device, also known as a node, can freely send/receive data. This feature makes it possible for an attacker to broadcast whatever she wants on a CAN bus. In this section, we first briefly introduce the basics of the CAN, and then we demonstrate how to inject an arbitrary message into the CAN.

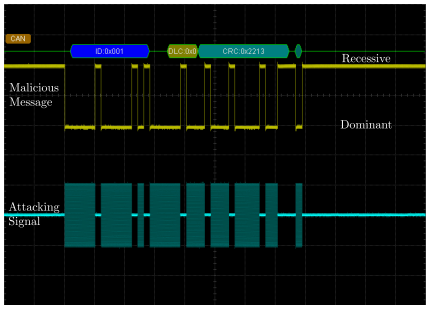


Figure 14: A malicious message and its corresponding attacking signal.

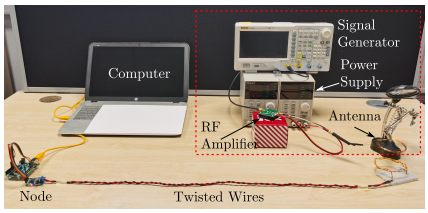


Figure 15: A practical setup of message injection attack. The devices in the red rectangle form an attacker's setup

7.1 CAN Basics

A basic structure of the CAN is shown in Figure 13. In a node, a transceiver is an interface between the wires and the microcontroller, and its function is to convert the differential signals into a single signal that the microcontroller can use while receiving data, or the other way around while transmitting data. The microcontroller handles signals on a software level, including identifying the type of the data, error checks, bus arbitration, etc. On the physical level, when the voltage levels of the differential signals are the same, a recessive state (1) is defined; otherwise, a dominant state (0). When no message is broadcast, the bus remains at the recessive state.

7.2 Message Injection

It is not difficult to find that such a CAN system matches our system model: the two wires in the CAN system correspond to the two input wires in our system model; the transceiver is the subtractor; the microcontroller is the receiver. Thus, it is possible for an attacker to use the bit injection attack to inject arbitrary messages.

We assume that the attacker has $g = 1$, i.e., she knows the line is always at a recessive state (see Section 8.1). With an arbitrary message that the attacker wants to inject, the first step is to convert it into a sequence of bits according to the rules of the CAN protocol. Based on the bits, an attacking signal is generated. For example, the attacker wants to inject a malicious message that is shown in Figure 14, which contains an identity (ID) field with a value of 0x001, a data length (DLC) field with a value of 0x0, and a cyclical redundancy check (CRC) field with a value of 0x2213. Note that this malicious message is just an example, and the attacker can craft any valid message as she wishes. Since the line is always at a recessive state, the attacker only needs to radiate electromagnetic interference when dominant bits need to be injected. Therefore, the attacker can craft an attacking signal as shown in Figure 14. When such an attacking signal is injected into the wires, it first bypasses the transceiver, and further force the microcontroller to receive dominant bits. The microcontroller will check the received message; if no error is detected, it will ultimately recognize the malicious message.

We use commercially off-the-shelf electronic devices to build a CAN bus system. As shown in Figure 15, a node is connected to one end of two twisted wires. In the node, a TJA1050 is used as the transceiver, and an ATMEGA328P that is integrated with an MCP2515 CAN controller is used as the microcontroller. This node is programmed to always listen to the wires. Moreover, the node is connected to a computer through serial communications so that the received message can be recorded and shown on the computer. As for the attacking signal, a signal generator is connected to an RF power amplifier, and the amplified signal is radiated by a coil antenna. In order to inject the attacking signal into the wires effectively and efficiently, the coil antenna is put at around 5 cm above the wires. Note that this is limited by both local RF equipment regulations and the gain of RF amplifier, but a determined attacker will not be regulated by laws, and she can also increase her attack power by extra cost, thus conducting the attack at a farther distance. Please see more discussion in Section 8.3. The frequency of the electromagnetic waves is set to be 22 MHz and the amplitude to be 20 V, which has the highest u that is around 0.74 according to preliminary experiments before the attack. Then, the message injection attack is conducted, and consequently, 3 malicious messages will be successfully recognized every 1000 attacks in 2 seconds, and the success rate is 0.003.

Such a success rate matches our expectations. Since there are 29 bits to be injected in this message injection, if we regard each bit injection as independent, the expected success rate will be $0.74^{29} \approx 0.0002$. However, as explained in Section 5.3, once the first bit injection of several consecutive injections is successful, the success rate for the following injections will be higher, and we approximate the success rate to 1. As shown in Figure 14, to inject this message, 9 groups of consecutive flips from 1 to 0 are needed, and hence, the expected success rate is around $0.74^{9} \approx 0.06$. A practical result should lie between 0.0002 and 0.06, and thus, it is reasonable to obtain a success rate of 0.003 in practice. Note that this success rate is for injecting a complete well-formed CAN message. It might not be high, but it still allows an attacker to inject a full message every 2 seconds. As discussed in Section 8.3, we did the experiments as a fairly low power level so as to not radiate too much power on unlicensed frequencies. With that in mind, it is a pretty good result.

It is essential to emphasize that the setup uses commercial receivers and hardware that are similar to those in automobiles, and the experimental results sufficiently reflect the feasibility of the signal injection attacks on CAN bus. Moreover, provided that it would be hard to guarantee the safety of other drivers if we did the experiment at high power outside, we chose to let the laboratory experiments be enough.

8 DISCUSSION

In this section, we discuss methods of gaining knowledge of transmitted bits, synchronization, power restrictions, and future countermeasures.

8.1 Gaining Knowledge

Knowing transmitted bits indeed gives an attacker advantage in achieving high success rates of injections. There are multiple methods of obtaining information about the bits. For example, a recent work [41] showed that whatever the state of a CAN bus is, causing two bit errors that are separated by a fixed number of clock cycles can force the bus into an idle state. In addition to actively interfering with the victim system, the attacker can also use existing information about the victim system to figure out what is transmitted. For example, a preamble of a packet is usually predetermined and published in the protocol, and it is relatively easy to know the transmitted bits in the preamble. Despite the fact that the payload or checksum may be hard to guess, the attacker could also use a magnetic field probe to (wirelessly) listen for electromagnetic leakage from the wires so as to obtain the bits by analyzing and processing the leakage [12].

8.2 Synchronization

Synchronizing attacking signals with legitimate signals in cables is challenging because it requires the attacker to know the time information. Luckily, our attack does not need synchronization. This is because our attack principle, i.e., causing rapid voltage accumulation and holding the voltage level still, allows an attacker to make the attacking signal's duration the same as the bits' duration (as mentioned in Section 3.2) so as to affect the bits without deliberate synchronization. Note that in the previous work regarding injecting bits into a single-ended signaling system, perfect synchronization is essential due to their attack principle, i.e., to detect an incorrect bit, a receiving circuit must detect the bit when a periodic voltage change (caused by the attack) reaches a specified level [12].

8.3 Restricted Attack Power and Distance

A well-known trade-off between distance and power (free space path loss) indicates that a close attacker with low power is equivalent to a faraway attacker with high power. Since we are limited legally in how much power we can emit on frequencies within which we do not have a license, we have chosen only to stay close and go as high as necessary to show that our method works. Indeed, more power will extend attack distance; however, we chose to focus on the novelty of injecting into differential pairs rather than characterize the distance/power relationship.

8.4 Future Countermeasures

In previous literature, abundant detection/mitigation methods against electromagnetic signal injection attacks have been proposed for various systems: for example, adding extra detection circuits to spot attacks by monitoring abnormal electromagnetic activities [1, 2, 11, 50], encoding critical signals secretly and detecting attacks if the signal integrity corrupts [26, 42, 48, 56], or, exploiting measurements from sensors plus tricky algorithms to detect/mitigate attacks [15, 23, 24, 34, 51]. Similar ideas theoretically apply to differential signaling, but millions of devices are not yet protected from the electromagnetic signal injection attacks. Future work is encouraged to fill this critical gap. Note that deploying countermeasures requires extra hardware/software, which will challenge many practical constraints of systems, such as budget caps, size/weight requirements, and computational resources. It is essential for system designers to weigh up between countermeasures' performance and the constraints/costs.

9 RELATED WORK

Using electromagnetic signal injection attacks to manipulate analog signals, e.g., sensor measurements, have been studied extensively [14, 18, 22, 23, 26, 28, 37, 46, 50, 51, 56]. However, regarding digital signals, early work mainly focused on studying how the electromagnetic interference impacts communication qualities [29, 30, 40], and it is more recently that a few studies started to use electromagnetic signal injection attacks to achieve arbitrary manipulation in single-ended wires. Selvaraj et al. [44, 45] conducted attacks onto pulse signals so as to control the actions of actuators, e.g., servos, which are widely used to operate robotic arms or aileron of drones. Similar attacks can maliciously control the speeds of DC motors, which can be found in insulin pumps and smart locks [56]. Dayanıklı et al. [12, 13] demonstrated that they could manipulate the pulse signals that control switches in AC-DC Converters, which play a critical role in the power delivery system of electric vehicles. Köhler et al. [27] showed the feasibility of using electromagnetic signal injection attacks to interrupt necessary control communication between an electric vehicle and its charger, causing charging sessions to abort.

10 CONCLUSION

Although differential signaling was proposed for making communication cables more immune to external interference, in this work, we show that electromagnetic signal injection attacks can inject arbitrary information into a differential signaling system. Because of the input asymmetry and nonlinearities of the subtractor, the rejection ability of the differential signaling technique is not sufficiently good for high-frequency signals to prevent attackers from injecting adversarial signals. Moreover, in the receiver, the ESD circuit's rectification plus the buffer circuit's net charge accumulation results in high-frequency signals ultimately being incorrectly detected as either 0's or 1's depending on the frequency. Our experiments have demonstrated the attack principles, and how to properly choose the frequency and the power of the attacking signal, in order to achieve successful injection. We analyze the success rate of injection of more complicated bitstrings, taking into account any knowledge that the attacker might have about the existing data transmissions in the cable. We show how this knowledge and the choice of attacking signals will affect the success rate. This analysis can also be used defensively by system designers who want to evaluate the security of their own systems. Finally, we demonstrate arbitrary message injection into a CAN bus, allowing an attacker to dictate the actions of the victim system.

REFERENCES

[1] Christian Adami, Christian Braun, Peter Clemens, M Joester, S Ruge, M Suhrke, HU Schmidt, and HJ Taenzer. 2014. HPM Detector System with Frequency Identification. In *2014 International Symposium on Electromagnetic Compatibility (EMC Europe)*. IEEE, 140–145.

[2] Christian Adami, Christian Braun, Peter Clemens, Michael Suhrke, HU Schmidt, and Achim Taenzer. 2011. HPM Detection System for Mobile and Stationary Use. In *EMC Europe 2011 York*. IEEE, 1–6.

[3] Ahmed Alamin. 2022. Common Mode Chokes Basics and Applications. https://abracon.com/uploads/resources/Common-Mode-Chokes-Basics-and-Applications.pdf

[4] Analog Devices. 2009. *Op Amp Common-Mode Rejection Ratio (CMRR).* Technical Report. Analog Devices.

[5] Constantine A Balanis. 2015. *Antenna Theory: Analysis and Design.* John Wiley & Sons, 41 – 44.

[6] Jurgen Branke, Jürgen Branke, Kalyanmoy Deb, Kaisa Miettinen, and Roman Slowiński. 2008. *Multiobjective Optimization: Interactive and Evolutionary Approaches.* Vol. 5252. Springer Science & Business Media, 11 – 13.

[7] Bruce Carter and Thomas R Brown. 2001. *Handbook of Operational Amplifier Applications.* Texas Instruments Dallas, TX, 15–16.

[8] P Crovetti and F Fiori. 2013. IC Digital Input Highly Immune to EMI. In *2013 International Conference on Electromagnetics in Advanced Applications (ICEAA).* IEEE, 1500–1503.

[9] Paolo S Crovetti. 2011. Finite Common-mode Rejection in Fully Differential Nonlinear Circuits. *IEEE Transactions on Circuits and Systems II: Express Briefs* 58, 8 (2011), 507–511.

[10] Paolo Stefano Crovetti and Franco Fiori. 2006. Finite Common-mode Rejection in Fully Differential Operational Amplifiers. *Electronics Letters* 42, 11 (2006), 615–617.

[11] JF Dawson, ID Flintoft, P Kortoci, Linda Dawson, AC Marvin, MP Robinson, Mirjana Stojilovic, Marcos Rubinstein, Benjamin Menssen, Heyno Garbe, et al. 2014. A Cost-efficient System for Detecting An Intentional Electromagnetic Interference (IEMI) attack. In *2014 International Symposium on Electromagnetic Compatibility.* IEEE, 1252–1256.

[12] Gökçen Y Dayanıklı. 2021. *Electromagnetic Interference Attacks on Cyber-Physical Systems: Theory, Demonstration, and Defense.* Ph. D. Dissertation. Virginia Tech.

[13] Gökçen Y Dayanıklı, Rees Hatch, Ryan M Gerdes, Hongjie Wang, and Regan Zane. 2020. Electromagnetic Sensor and Actuator Attacks on Power Converters for Electric Vehicles. In *IEEE Workshop on the Internet of Safe Things.* IEEE.

[14] J Lopes Esteves and C Kasmi. 2018. Remote and Silent Voice Command Injection on a Smartphone through Conducted IEMI: Threats of Smart IEMI for Information Security. *Wireless Security Lab, French Network and Information Security Agency (ANSSI), Tech. Rep* (2018).

[15] Kai Fang, Tingting Wang, Xiaochen Yuan, Chunyu Miao, Yuanyuan Pan, and Jianqing Li. 2022. Detection of Weak Electromagnetic Interference Attacks Based on Fingerprint in IIoT Systems. *Future Generation Computer Systems* 126 (2022), 295–304.

[16] Franco Fiori. 2011. Susceptibility of CMOS Voltage Comparators to Radio Frequency Interference. *IEEE transactions on electromagnetic compatibility* 54, 2 (2011), 434–442.

[17] Ilias Giechaskiel and Kasper Rasmussen. 2019. Taxonomy and Challenges of Out-of-band Signal Injection Attacks and Defenses. *IEEE Communications Surveys & Tutorials* 22, 1 (2019), 645–670.

[18] Ilias Giechaskiel, Youqian Zhang, and Kasper Rasmussen. 2019. A Framework for Evaluating Security in the Presence of Signal Injection Attacks. In *European Symposium on Research in Computer Security.* Springer, 512–532.

[19] Gianluca Giustolisi, Giuseppe Palmisano, and Gaetano Palumbo. 2000. CMRR Frequency Response of CMOS Operational Transconductance Amplifiers. *IEEE Transactions on instrumentation and Measurement* 49, 1 (2000), 137–143.

[20] David Harris and Sarah L Harris. 2010. *Digital Design and Computer Architecture.* Morgan Kaufmann, 2–52.

[21] Richard Jaeger. 1976. Common-mode Rejection Limitations of Differential Amplifiers. *IEEE Journal of Solid-State Circuits* 11, 3 (1976), 411–417.

[22] Yan Jiang, Xiaoyu Ji, Kai Wang, Chen Yan, Richard Mitev, Ahmad-Reza Sadeghi, and Wenyuan Xu. 2022. WIGHT: Wired Ghost Touch Attack on Capacitive Touchscreens. In *2022 IEEE Symposium on Security and Privacy (SP).* IEEE Computer Society, 1537–1537.

[23] Chaouki Kasmi and Jose Lopes Esteves. 2015. IEMI Threats for Information Security: Remote Command Injection on Modern Smartphones. *IEEE Transactions on Electromagnetic Compatibility* 57, 6 (2015), 1752–1755.

[24] Chaouki Kasmi and Jose Lopes-Esteves. 2015. Automated Analysis of the Effects Induced by Radio-frequency Pulses on Embedded Systems for EMC Functional Safety. In *2015 1st URSI Atlantic Radio Science Conference (URSI AT-RASC).* IEEE, 1–1.

[25] Walt Kester. 2009. Understand SINAD, ENOB, SNR, THD, THD+ N, and SFDR So You Don't Get Lost in the Noise Floor. *MT-003 Tutorial* (2009).

[26] Sebastian Köhler, Richard Baker, and Ivan Martinovic. 2022. Signal Injection Attacks against CCD Image Sensors. In *Proceedings of the 2022 ACM on Asia Conference on Computer and Communications Security.* 294–308.

[27] Sebastian Köhler, Richard Baker, Martin Strohmeier, and Ivan Martinovic. 2022. BROKENWIRE: Wireless Disruption of CCS Electric Vehicle Charging. *arXiv preprint arXiv:2202.02104* (2022).

[28] Denis Foo Kune, John Backes, Shane S Clark, Daniel Kramer, Matthew Reynolds, Kevin Fu, Yongdae Kim, and Wenyuan Xu. 2013. Ghost Talk: Mitigating EMI Signal Injection Attacks against Analog Sensors. In *IEEE Symposium on Security and Privacy.* IEEE, 145–159.

[29] Ke-Jie Li, Yan-Zhao Xie, Fan Zhang, and Yu-Hao Chen. 2019. Statistical Inference of Serial Communication Errors Caused by Repetitive Electromagnetic Disturbances. *IEEE Transactions on Electromagnetic Compatibility* 62, 4 (2019), 1160–1168.

[30] Michał Maćkowski. 2009. The Influence of Electromagnetic Disturbances on Data Transmission in USB Standard. In *International Conference on Computer Networks.* Springer, 95–102.

[31] A Theodore Markettos and Simon W Moore. 2009. The Frequency Injection Attack on Ring-oscillator-based True Random Number Generators. In *International Workshop on Cryptographic Hardware and Embedded Systems.* Springer, 317–331.

[32] R Timothy Marler and Jasbir S Arora. 2010. The Weighted Sum Method for Multi-objective Optimization: New Insights. *Structural and multidisciplinary optimization* 41, 6 (2010), 853–862.

[33] G Meyer-Brotz and A Kley. 1966. The Common-mode Rejection of Transistor Differential Amplifiers. *IEEE Transactions on Circuit Theory* 13, 2 (1966), 171–175.

[34] Devaprakash Muniraj and Mazen Farhood. 2019. Detection and Mitigation of Actuator Attacks on Small Unmanned Aircraft Systems. *Control Engineering Practice* 83 (2019), 188–202.

[35] Nordic Semiconductor. 2019. nRF52833 Objective Product Specification. https://docs.rs-online.com/f99e/A700000006639409.pdf

[36] Saki Osuka, Daisuke Fujimoto, Yu-ichi Hayashi, Naofumi Homma, Arthur Beckers, Josep Balasch, Benedikt Gierlichs, and Ingrid Verbauwhede. 2018. EM Information Security Threats against RO-based TRNGs: The Frequency Injection Attack Based on IEMI and EM Information Leakage. *IEEE Transactions on Electromagnetic Compatibility* 61, 4 (2018), 1122–1128.

[37] Kasper Bonne Rasmussen, Claude Castelluccia, Thomas S Heydt-Benjamin, and Srdjan Capkun. 2009. Proximity-based Access Control for Implantable Medical Devices. In *Proceedings of the 16th ACM conference on Computer and communications security.* 410–419.

[38] Behzad Razavi. 2002. *Design of Analog CMOS Integrated Circuits.* Tata McGraw-Hill Education, 100 – 133.

[39] Jean-Michel Redouté and Michiel Steyaert. 2009. *EMC of Analog Integrated Circuits.* Springer Science & Business Media, 72 – 82.

[40] Fei Ren, Y Rosa Zheng, Maciej Zawodniok, and Jagannathan Sarangapani. 2007. Effects of Electromagnetic Interference on Control Area Network Performance. In *2007 IEEE Region 5 Technical Conference.* IEEE, 199–204.

[41] Matthew Rogers and Kasper Rasmussen. 2022. Silently Disabling ECUs and Enabling Blind Attacks on the CAN Bus. In *Embedded Security in Cars (escar).*

[42] Henri Ruotsalainen, Albert Treytl, and Thilo Sauter. 2021. Watermarking Based Sensor Attack Detection in Home Automation Systems. In *2021 26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA).* IEEE, 1–8.

[43] Frank Sabath. 2010. Classification of Electromagnetic Effects at System Level. In *Ultra-Wideband, Short Pulse Electromagnetics 9.* Springer, 325–333.

[44] Jayaprakash Selvaraj. 2018. *Intentional Electromagnetic Interference Attack on Sensors and Actuators.* Ph. D. Dissertation. Iowa State University.

[45] Jayaprakash Selvaraj, Gökçen Y Dayanıklı, Neelam Prabhu Gaunkar, David Ware, Ryan M Gerdes, Mani Mina, et al. 2018. Electromagnetic Induction Attacks Against Embedded Systems. In *Proceedings of the 2018 on Asia Conference on Computer and Communications Security.* ACM, 499–510.

[46] Haoqi Shan, Boyi Zhang, Zihao Zhan, Dean Sullivan, Shuo Wang, and Yier Jin. 2022. Invisible Finger: Practical Electromagnetic Interference Attack on Touchscreen-based Electronic Devices. In *2022 IEEE Symposium on Security and Privacy (SP).* IEEE Computer Society, Los Alamitos, CA, USA, 1548–1548. https://doi.org/10.1109/SP46214.2022.00119

[47] Hocheol Shin, Yunmok Son, Youngseok Park, Yujin Kwon, and Yongdae Kim. 2016. Sampling Race: Bypassing Timing-based Analog Active Sensor Spoofing Detection on Analog-digital Systems. In *10th {USENIX} Workshop on Offensive Technologies ({WOOT} 16).*

[48] Yasser Shoukry, Paul Martin, Yair Yona, Suhas Diggavi, and Mani Srivastava. 2015. PyCRA: Physical Challenge-response Authentication for Active Sensors under Spoofing Attacks. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security.* 1004–1015.

[49] Frederick M Tesche, Michel Ianoz, and Torbjörn Karlsson. 1996. *EMC Analysis Methods and Computational Models.* John Wiley & Sons.

[50] Yazhou Tu, Sara Rampazzi, Bin Hao, Angel Rodriguez, Kevin Fu, and Xiali Hei. 2019. Trick or heat? Manipulating Critical Temperature-based Control Systems Using Rectification Attacks. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security.* 2301–2315.

[51] Kai Wang, Richard Mitev, Chen Yan, Xiaoyu Ji, Ahmad-Reza Sadeghi, and Wenyuan Xu. 2022. GhostTouch: Targeted Attacks on Touchscreens without Physical Touch. In *31st USENIX Security Symposium (USENIX Security 22).* USENIX Association, Boston, MA. https://www.usenix.org/conference/usenixsecurity22/presentation/wang-kai

[52] Yuanda Wang, Hanqing Guo, and Qiben Yan. 2022. GhostTalk: Interactive Attack on Smartphone Voice System Through Power Line. In *Network and Distributed Systems Security (NDSS) Symposium.*

[53] David A Ware. 2017. *Effects of Intentional Electromagnetic Interference on Analog to Digital Converter Measurements of Sensor Outputs and General Purpose Input Output Pins.* Ph. D. Dissertation. Utah State University.

[54] Chen Yan, Hocheol Shin, Connor Bolton, Wenyuan Xu, Yongdae Kim, and Kevin Fu. 2020. SoK: A Minimalist Approach to Formalizing Analog Sensor Security. In *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, 233–248.

[55] MING-GUANG YI. 1980. Common-mode Rejection Ratio of Differential Amplifiers. *IEEE Journal of Solid-State Circuits* 15, 2 (1980), 214–221.

[56] Youqian Zhang and Kasper Rasmussen. 2022. Detection of Electromagnetic Signal Injection Attacks on Actuator Systems. In *25th International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2022)*. ACM. https://doi.org/10.1145/3545948.3545949

[57] Hank Zumbahlen. 2008. *Linear Circuit Design Handbook*. Elsevier-Newnes, 448–449.