# Sensor CON-Fusion: Defeating Kalman Filter in Signal Injection Attack

Shoei Nashimoto
Mitsubishi Electric
Kamakura, Kanagawa, Japan
Nashimoto.Shoei@bx.MitsubishiElectric.co.jp

Daisuke Suzuki
Mitsubishi Electric
Kamakura, Kanagawa, Japan
Suzuki.Daisuke@bx.MitsubishiElectric.co.jp

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

Kazuo Sakiyama
The University of Electro-Communications
Choufu, Tokyo, Japan
sakiyama@uec.ac.jp

## ABSTRACT

In recent years, information systems have become increasingly able to interact with the real world by using relatively cheap connected embedded devices. In such systems, sensors are crucial components because systems can observe the real world only through sensors. Recently, there have been emerging threats to sensors, which involve the injection of false information in the physical/analog domain. To counter such attacks, sensor fusion is considered a promising approach because the robustness of a measurement can be improved by combining data from redundant sensors. However, sensor fusion algorithms were not originally designed to consider security, and thus their effectiveness is unclear. For this reason, in this paper, we evaluate in detail the security of sensor fusion. Notably, we consider a sensor fusion scenario that involves measuring inclination, with a combination of an accelerometer, gyroscope, and magnetometer using Kalman filter. Based on a theoretical analysis of the algorithm, two concrete attacks that defeat the sensor fusion are proposed. The feasibility of the proposed attacks is verified by performing experiments in emulated and real environments. We also propose a countermeasure that thwarts the new attacks. Furthermore, we logically prove that the proposed countermeasure detects all possible attacks.

## KEYWORDS

Sensor Fusion, Signal Injection Attack, Kalman Filter, AHRS

**ACM Reference Format:**
Shoei Nashimoto, Daisuke Suzuki, Takeshi Sugawara, and Kazuo Sakiyama. 2018. Sensor CON-Fusion: Defeating Kalman Filter in Signal Injection Attack. In *ASIA CCS '18: 2018 ACM Asia Conference on Computer and Communications Security, June 4–8, 2018, Incheon, Republic of Korea.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3196494.3196506


*ASIA CCS '18, June 4–8, 2018, Incheon, Republic of Korea*

ACM ISBN 978-1-4503-5576-6/18/06.
https://doi.org/10.1145/3196494.3196506

## 1 INTRODUCTION

Over the last decade, smartphones have been one of the main topics in information technology (IT), and it has enabled the connection between the digital and real worlds. The application of IT to embedded systems continues, up to the present. Enabling connectivity to relatively cheap embedded systems is considered a promising approach for the next decade. Consequently, a lot of research and development is being conducted under slogans such as the Internet of things (IoT) and cyber-physical systems.

Sensors provide an interface between the digital and physical worlds, and enable digital system to "see" and "listen" to the real world. Sensors are used extensively in devices such as smartphones, home electronics, and in factory/plant automation. Therefore, the reliability of IoT and cyber-physical systems largely depends on sensors.

There have been several emerging threats that provide false information to sensors in the physical domain. These attacks cannot be thwarted by conventional information security technologies for digital data because the attacks are conducted in the analog domain, i.e., before the data is digitized. Therefore, there is ongoing research pertaining to the security of sensors.

In this paper, a group of attacks that inject false information to sensors is referred to as a *signal injection attack*. Many sensors are susceptible to signal injection attacks, some of which are conducted by spoofing and/or cancelling a genuine signal. This category involves the conventional attacks on radar systems [2, 12, 26], anti-lock braking systems (ABSs) using magnetic speed sensors [21], and the global positioning system (GPS) [5]. Other attacks use signals that exceed implicit or explicit limits of a sensor. Such abnormal signals can cause malfunctions in the signal processing stage after the sensing. Examples of such attacks are (i) the attacks on microphones by injecting ultrasonic sound [1, 28], (ii) the attacks on medical devices by injecting signals that saturates an analog to digital converters (ADCs) [11], and (iii) the laser injection attacks on digital cameras [3]. Another category of attacks uses unsupported signals that becomes measurable after physical or electrical transformations. Electro-magnetic interference (EMI)-based attacks fall within this category [6]. Another notable example in this category includes the attacks on micro electro-mechanical systems (MEMS) sensors using ultrasonic sound [22, 24, 25].

There have been studies that aim to develop countermeasures against signal injection attacks. One promising approach is to make a sensor implicitly secure. An example is the countermeasures for ranging sensors such as light detection and ranging (LiDARs) and radars that randomize transmission signals [2]. Although this is a fundamental treatment, it is not available in every sensor. Notably, many cheap sensors (e.g., gyroscope and temperature sensors) work only passively and thus there is no transmission signal to randomize.

In contrast, there are system-level countermeasures that combine multiple sensors [10, 23], and sensor fusion is considered to be a promising approach [3]. This is because sensor fusion can be used to improve the robustness of measurements by integrating information from several different sources. However, most sensor fusion algorithm designs do not incorporate security features, and a rigorous security evaluation is therefore needed before using them as countermeasures.

In this paper, we focus on sensors for measuring inclination. Son et.al. showed that it is possible to disrupt inclination Measurements that are based on gyroscopes [22]. Meanwhile, there is another inclination measurement used in attitude-heading reference systems (AHRS), which combines readings from the accelerometer, gyroscope, and magnetometer using sensor fusion. The aim is therefore to determine whether the sensor-fusion-based inclination measurement is more secure, and if so, by how much.

Based on the motivating question, in this study, we perform a rigorous security evaluation of an AHRS with sensor fusion. In particular, we focus on a specific sensor fusion algorithm using specific components. It is limited in the sense of the generality, but the sensor fusion algorithm rests on a widely-used algorithm, i.e., Kalman filter. The results show that there are concrete attacks to the sensor fusion. In addition, we comprehensively evaluate the outcome of the attacks by attackers with different capabilities with respect to raw sensors. The feasibility of the attacks is demonstrated by performing experiments in simulated and real environments. Finally, we propose a software-based countermeasure for detecting the proposed attacks.

We summarize our contributions and corresponding sections as follows.

**C1** To the best of our knowledge, this is the first security evaluation of sensor fusion against signal injection attacks.

**C2** Two concrete attacks on a sensor fusion algorithm for measuring inclination based on the indirect Kalman filter (Sect. 3.2).

**C3** Comprehensive evaluation of the attacks under 27 different attacker's access capabilities. Three-level access capabilities (fully controllable, disruptive, and uncontrollable) are considered for each of the accelerometer, gyroscope, and magnetometer (Sect. 3.3).

**C4** Demonstration of the feasibility of the proposed attacks by performing experiments in emulated and real environments (Sect. 4).

**C5** A software-based countermeasure for detecting the proposed attacks (Sect. 5).

## 2 SENSOR FUSION

This section explains the target sensor fusion algorithm for inclination which combines readings from an accelerometer, gyroscope, and magnetometer based on the indirect Kalman filter.

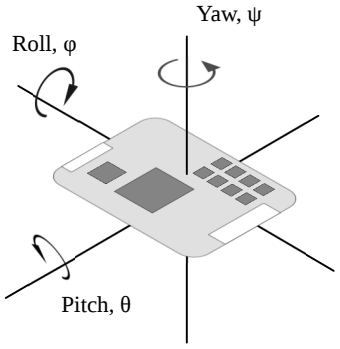


**Figure 1: Inclination and Eular angles**

### 2.1 Sensor Fusion based on Kalman Filter

Sensor fusion is a class of signal processing methods that integrate readings from multiple sensors. There are several sensor fusion algorithms that are employed for different objectives and sensors. One important application is to improve the robustness of measurements obtained unreliable sensors. To do this, a common strategy for the purpose is to measure a target physical quantity redundantly using heterogeneous sensors. In the case of sensor fusion for AHRS, the *inclination* is measured by combining the gyroscope, accelerometer, and magnetometer. Here, the *inclination* refers to the degree by which the ARHS inclines in the global coordination and is commonly expressed by Euler angles i.e., roll, pitch, and yaw as shown in Fig. 1. The inclination measurement is important in many applications, such as motion tracking and attitude control for robots and drones.

Kalman filters are widely used as a generic algorithms for sensor fusion. The purpose of Kalman filters is to estimate a hidden state of a target system from observations with errors. The algorithm needs a system model expressed as a differential equation, which is not always available [16]. Therefore, an alternative approach that is commonly used is the indirect Kalman filter approach, which estimates errors.

Kalman filters are also popular in the application of sensor fusion for inclination [4, 7–9, 13–15, 17, 18, 27, 29]. The purpose of the sensor fusion is to improve the robustness based on redundant measurements. Its principle is briefly explained as follows. By using the three sensors (i.e., accelerometer, gyroscope, and magnetometer), the inclination can be measured in two different ways: (1) to integrate the angular velocity from a gyroscope, and (2) to use gravity and geomagnetism measured by an accelerometer and magnetometer. The first method is good in terms of the responsivity, but it has a problem of accumulated errors because of the integration. The second approach does not have such error accumulation, but it cannot be used unless ARHS is stationary i.e., stable. The concept of sensor fusion involves using the two methods to complement each others.

Kalman filters originated from system control engineering, and are therefore frequently used for state estimation in a control systems. It should be noted that there is a class of attacks called *false*

*data injection*, which injects false information to sensors in a control system. For example, Mo et.al. analyzed a control system using Kalman filters, and showed how to construct sensor inputs that efficiently drives the system into an unstable state [10]. Signal injection attack and false data injection are closely related in that the former focuses more on sensors, while the latter is more system oriented. Sensor fusion is on the boundary of both research fields, and to the best of our knowledge, has not been covered in either field.

## 2.2 Algorithm Details

This section presents technical details of the sensor fusion in AHRS. The description and notations are based on the document by NXP semiconductors describing indirect Kalman filter [18]. NXP semiconductors also provides an open-source software implementation and an off-the-shelf evaluation board, which are used in our experiments in the latter part of the paper. See Appendix B for a detailed derivation.

The sensor fusion algorithm is explained based on the dataflow shown in Fig. 2. Inputs to the algorithm are sensor readings obtained from three sensors. The readings from the accelerometer, magnetometer, and gyroscope at the $k$-th iteration are represented by $\boldsymbol{G}_k$, $\boldsymbol{B}_k$, and $\boldsymbol{Y}_k$, respectively. The purpose of the algorithm is to determine the inclination in the global coordinate system represented by Euler angles, namely $\phi_k$, $\theta_k$, and $\psi_k$. Internally, the inclination can be represented by either a rotation matrix or quaternion. Quaternions and their vector representations are denoted by $q$ and $\boldsymbol{q}$, respectively.

In the algorithm, the geomagnetic and gravity vectors play important roles. The geomagnetic vector, $\boldsymbol{m}$, is a unit vector that is parallel to the earth's magnetic field and the gravity vector, $\boldsymbol{g}$, is a unit vector that points to the center of the earth. An important observation is that the inclination is uniquely determined if $\boldsymbol{m}$ and $\boldsymbol{g}$ are given. Conversely, $\boldsymbol{m}$ and $\boldsymbol{g}$ can be deduced if an inclination is given.

Firstly, $\boldsymbol{m}$ and $\boldsymbol{g}$ are calcureted in two different ways. In Fig. 2-(1), the geomagnetic vector $\boldsymbol{m}_k^{6DOF}$ and the gravity vector $\boldsymbol{g}_k^{6DOF}$ are calculated using the 3-axis accelerometer and magnetometer, respectively. Therefore, the degrees of freedom is six (6DOF). In addition, measurement noise matrix $\boldsymbol{Q}_{v,k}$ is calculated. $\boldsymbol{Q}_{v,k}$ mainly consists of two kinds of sensor noises and distrubances as follows.

$$Q_{vB,k} + Q_{d,k} \approx 3(|\boldsymbol{B}_k| - B)^2, \quad (1)$$

$$Q_{vG,k} + Q_{a,k} \approx 3(|\boldsymbol{G}_k| - g)^2, \quad (2)$$

where $B$ and $g$ represent the local geomagnetic field strength and the magnitude of the gravity vector, respectively. In addition, $Q_{vB,k}$ and $Q_{vG,k}$ are noises for magnetmeter and accelerometer, respectively. $Q_{d,k}$ is magnetic disturbance, and $Q_{a,k}$ is acceleration variance.

In Fig. 2-(2), *a priori* estimation of inclination $q_k^-$ is computed by accumulating $\boldsymbol{Y}_k$ to *a posteriori* estimation of inclination $q_{k-1}^+$. Then, based on the *a priori* estimation of inclination $q_k^-$, those of the geomagnetic and gravity vectors are deduced. They are represented as $\boldsymbol{m}_k^-$ and $\boldsymbol{g}_k^-$, respectively.

In Fig. 2-(3), the vectors calculated in two different ways are compared and meaurement errors are evaluated. The error between $\boldsymbol{g}_k^{6DOF}$ and $\boldsymbol{g}_k^-$ is represented by a rotation quaternion $\boldsymbol{q}_{zg\epsilon,k}$,

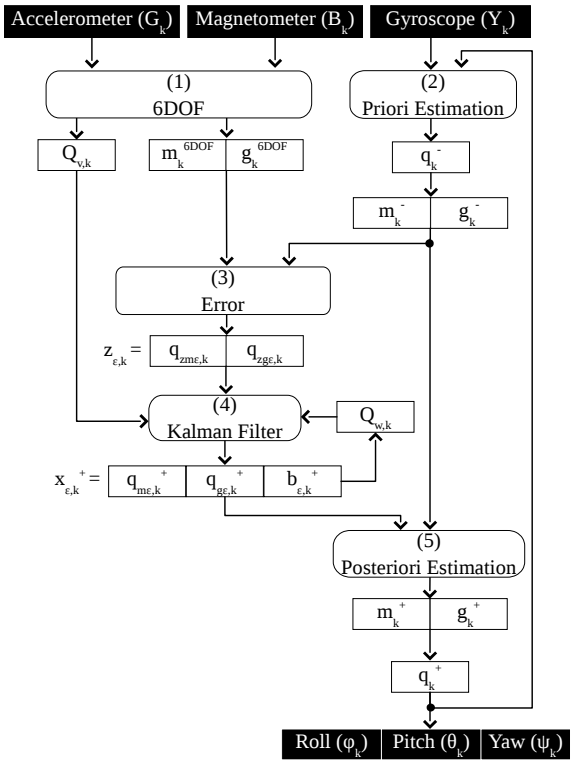


**Figure 2: Dataflow of sensor fusion algorithm**

which rotates $\boldsymbol{g}_k^{6DOF}$ onto $\boldsymbol{g}_k^-$. Similarly, the error $\boldsymbol{q}_{zm\epsilon,k}$ is derived by comparing $\boldsymbol{m}_k^{6DOF}$ and $\boldsymbol{m}_k^-$.

In Fig. 2-(4), a state of the Kalman filter is updated using the Kalman filter gain $\boldsymbol{K}_k$. Instead of directly estimating the process $\boldsymbol{x}_k$ namely inclination, indirect Kalman filter estimates the process error $\boldsymbol{x}_{\epsilon,k}$. The Kalman filter updates its gain $\boldsymbol{K}_k$ using above measurement errors ($\boldsymbol{q}_{zm\epsilon,k}$ and $\boldsymbol{q}_{zg\epsilon,k}$) and two kinds of noises i.e., the measurement noise $\boldsymbol{Q}_{v,k}$ and the process noise $\boldsymbol{Q}_{w,k}$ as follows.

$$\boldsymbol{K}_k = \boldsymbol{Q}_{w,k}\boldsymbol{C}_k^T(\boldsymbol{C}_k\boldsymbol{Q}_{w,k}\boldsymbol{C}_k^T + \boldsymbol{Q}_{v,k})^{-1}, \quad (3)$$

where $\boldsymbol{C}_k$ represents the measurement matrix. In addition, $\boldsymbol{Q}_{w,k}$ mainly consists of previous iteration of *a posteriori* estimate $\boldsymbol{x}_{\epsilon,k-1}^+$. Using the Kalman filter gain, *a posteriori* estimate is calculated as

$$\boldsymbol{x}_{\epsilon,k}^+ = \begin{pmatrix} \boldsymbol{q}_{g\epsilon,k}^+ \\ \boldsymbol{q}_{m\epsilon,k}^+ \\ \boldsymbol{b}_{\epsilon,k}^+ \end{pmatrix} = \boldsymbol{K}_k\boldsymbol{z}_{\epsilon,k} = \boldsymbol{K}_k \begin{pmatrix} \boldsymbol{q}_{zg\epsilon,k} \\ \boldsymbol{q}_{zm\epsilon,k} \end{pmatrix}, \quad (4)$$

where $\boldsymbol{b}_{\epsilon,k}^+$ is an offset vector to correct a drift of the gyroscope.

In Fig. 2-(5), the *a posteriori* estimate of inclination is calculated using *a posteriori* estimate of errors. Firstly, *a posteriori* estimates of the geomagnetic and gravity vectors, namely $\boldsymbol{m}_k^+$ and $\boldsymbol{g}_k^+$, are calculated by correcting the *a priori* estimates with the estimated errors. Then, $\boldsymbol{m}_k^+$ and $\boldsymbol{g}_k^+$ are converted to *a posteriori* estimate of

inclination represented by a rotation matrix $\boldsymbol{R}_k^+$ as follows.

$$\begin{aligned}\boldsymbol{R}_k^+ &= (\boldsymbol{R}_{x,k}^+, \boldsymbol{R}_{y,k}^+, \boldsymbol{R}_{z,k}^+) \\ &= \begin{pmatrix} R_{xx,k}^+ & R_{xy,k}^+ & R_{xz,k}^+ \\ R_{yx,k}^+ & R_{yy,k}^+ & R_{yz,k}^+ \\ R_{zx,k}^+ & R_{zy,k}^+ & R_{zz,k}^+ \end{pmatrix}, \\ &= \left( \boldsymbol{R}_{y,k}^+ \times \boldsymbol{R}_{z,k}^+, \frac{\boldsymbol{g}_k^+ \times \boldsymbol{m}_k^+}{|\boldsymbol{g}_k^+ \times \boldsymbol{m}_k^+|}, \frac{\boldsymbol{g}_k^+}{|\boldsymbol{g}_k^+|} \right). \end{aligned} \tag{5}$$

Finally, the rotation matrix $\boldsymbol{R}_k^+$ is converted to Euler angles, namely roll ($\phi_k$), pitch ($\theta_k$), and yaw ($\psi_k$) as follows.

$$(\phi_k, \theta_k, \psi_k) = \left( \tan^{-1}\left(\frac{R_{yz}^+}{R_{zz}^+}\right), -\sin^{-1}\left(R_{xz}^+\right), \tan^{-1}\left(\frac{R_{xy}^+}{R_{xx}^+}\right) \right). \tag{6}$$

By iterating the above procedures periodically, inclination is successively measured.

## 3 ATTACK ON SENSOR FUSION

In this section, we first present an attacker model. Then, we explain the proposed attacks. Finally, we evaluate the outcome of the proposed attacks considering attackers with different capabilities.

### 3.1 Attacker Model

According to Kerckhoffs' principle, we assume that an attacker has knowledge of the sensor fusion algorithm. In addition, we categorize an attacker's access to a sensor using three levels:

- $\mathcal{C}$ fully Controllable: The attacker can inject arbitrary signal into the target sensor.
- $\mathcal{D}$ Disruptive: The attacker can inject a disruptive (i.e., jamming) signal into a target sensor but precise control is impossible.
- $\mathcal{U}$ Uncontrollable: The attacker does not have any accessibilities to the sensor.

In this paper, the three levels are represented by the symbols $\mathcal{C}$, $\mathcal{D}$, and $\mathcal{U}$, respectively.

Capabilities are determined based on existing attacks on raw sensors. Acoustic injection attacks on accelerometers and gyroscopes were studied in [22, 24, 25] and the results showed that the attacker can control the reading from a specific sensor by injecting modulated sound. We describe such a situation as accelerometer=$\mathcal{C}$ and gyroscope=$\mathcal{C}$. In addition, a sensor can be deceived by plausible signals (e.g., GPS signals for a GPS receiver, or laser for a camera) [3, 5, 11, 21]. This means that magnetometer can be controlled or disrupted by exposure it to an artificial magnetic field. We describe such a situation as magnetometer=$\mathcal{C}$ or $\mathcal{D}$.

### 3.2 Proposed Attacks

Here, we describe the basic idea behind the proposed attack. As mentioned in Sect. 2.2, the sensor fusion algorithm works as a weighted sum between two components wherein the weight is determined by the measurement noise such as $\boldsymbol{Q}_{v,k}$ through $\boldsymbol{K}_k$ (see Eqs. (3) and (4)). An important observation is that one component becomes dominant in extreme cases. When an AHRS is stationary and the measurement noise is extremely low,

$$\boldsymbol{x}_{\epsilon,k}^+ \approx \begin{pmatrix} \boldsymbol{q}_{zg\epsilon,k} \\ \boldsymbol{q}_{zm\epsilon,k} \\ \boldsymbol{b}_{\epsilon,k} \end{pmatrix}, \tag{7}$$

where these vectors represent a hidden state of the process error $\boldsymbol{x}_{\epsilon,k}$. Equation (7)indicates that the Kalman filter ignores the *a priori* gyroscope sensor extrapolation. In contrast, when the AHRS is non-stationary and the measurement noise is extremely high,

$$\boldsymbol{x}_{\epsilon,k}^+ \approx \boldsymbol{0}. \tag{8}$$

In other words, the access capabilities for a subset of raw sensors is sufficient to control the fusion output. For example, the low-noise scenario, the attacker can control the inclination without the need for access to the gyroscope. Moreover, an attacker with either $\mathcal{C}$ or $\mathcal{D}$ access capabilities can intentionally cause a non-stationary situation thereby excluding some uncontrollable sensors from being used to determine the inclination. Hereafter, we refer to attacks on low- and high-noise cases as *stationary* and *non-stationary attacks*, respectively.

Basic idea of the attacks is to deceive sensor(s) being dominant when measurement noise is low and high, respectively. More specifically, a stationary attack controls the accelerometer and/or magnetometer to make an AHRS misunderstood that it is stationary. Similarly, a non-stationary attack controls the gyroscope with disrupting the accelerometer and/or magnetometer to make the AHRS misunderstood that it is non-stationary.

Although the focus of this paper is sensor fusion on ARHS, the same principle applies to other sensor fusion algorithms. In other words, the above-mentioned characteristic is a common weakness of sensor fusion strategy, which uses multiple sensors complementarily.

#### 3.2.1 *Stationary Case.*

The measurement noise $\boldsymbol{Q}_{v,k}$ becomes low when the conditions in Eqs. (1) and (2) are approximately equal to zero. In this case, from Eq. (7), the *a posteriori* estimates are corrected without the need for a gyroscope. Thus, the estimates, namely $\boldsymbol{m}_k^+$ and $\boldsymbol{g}_k^+$, correspond to the readings from the megnetometer $\boldsymbol{B}_k$ and accelerometer $\boldsymbol{G}_k$, respectively. From Eqs. (5) and (6), the relationship between sensor readings and Euler angles becomes as follows.

$$\phi_k \approx \tan^{-1}\left(\frac{G_{y,k}}{G_{z,k}}\right), \tag{9}$$

$$\theta_k \approx -\sin^{-1}\left(\frac{G_{x,k}}{|\boldsymbol{G}_k|}\right), \tag{10}$$

$$\psi_k \approx \tan^{-1}\left(\frac{(-B_y G_z + B_z G_y)|\boldsymbol{G}_k|}{G_z(B_x G_z - B_z G_x) - G_y(B_y G_x - B_x G_y)}\right). \tag{11}$$

In this stationary case, to perform a $C$-attack on sensor fusion, both accelerometer=$C$ and magnetometer=$C$ are required. Meanwhile, the gyroscope is omitted, and thus can be $\mathcal{U}$. Given a target inclination (roll, pitch, and yaw), the attacker can determine the corresponding inputs to the accelerometer and magnetometer by using Eqs. (9)–(11). Even when only either accelerometer=$C$ or magnetometer=$C$, the attacker can still partially control the inclination. If accelerometer=$C$, an attacker can get (roll, pitch, yaw) = ($C$,

$C$, $\mathcal{D}$), ($\mathcal{D}$, $C$, $C$), or ($C$, $\mathcal{D}$, $C$). Alternatively, if magnetometer=$C$, then the attacker can get yaw=$C$.

*Example.* We want to set all the Euler angles to 60 [°]. Eqs. (1) and (2) give $|\boldsymbol{B}_k| = B$ [μT] and $|\boldsymbol{G}_k| = 1$ [g]. Substituting $\phi_k = \theta_k = \psi_k = 60$ into Eqs. (9)–(11) gives

$$\boldsymbol{G}_k = (-0.8660, 0.4330, 0.2500) \text{ [g]}, \quad (12)$$

$$\boldsymbol{B}_k = (10.00, -2.33, 38.45) \text{ [μT]}. \quad (13)$$

Note that the following assumption is used: $B = 40$ [μT] and $B_{x,k} = 10$ [μT].

#### 3.2.2 Non-Stationary Case.

When the AHRS is non-stationary (e.g., external magnetic sources such as magnets, or a fast and sudden movement), the right-hand-side values of Eqs. (1) and (2) become much greater than zero, i.e., the measurement noise $\boldsymbol{Q}_{v,k}$ is high. In this case, no correction is needed for the *a posteriori* estimate of errors as indicated by Eq. (8). As a result, the *a posteriori* estimates of two kinds of vectors namely $\boldsymbol{m}_k^+$ and $\boldsymbol{g}_k^+$ correspond to the *a priori* estimates namely $\boldsymbol{m}_k^-$ and $\boldsymbol{g}_k^-$. Because the *a posteriori* and *a priori* estimates are exactly same, we get

$$q_k^+ \approx q_k^- = q_{k-1}^+ \otimes \Delta q((\boldsymbol{Y}_k - \boldsymbol{b}_{k-1}^+)\delta t), \quad (14)$$

where $\otimes$ represents quaternion product. Eq. (14) means $q_k^+$ can be derived by rotating its previous iteration $q_{k-1}^+$ by a rotation quaternion derived from $\boldsymbol{Y}_k$ and $\boldsymbol{b}_{k-1}^+$.

In this non-stationary case, the use of gyroscope=$C$ is sufficient to control the sensor fusion output. In order to satisfy the prerequisite, both the accelerometer=$\mathcal{D}$ and magnetometer =$\mathcal{D}$ are needed. Eq. (14) shows that the inclination is fully controlled by the gyroscope. Therefore, the attacker can easily obtain the gyroscope reading $\boldsymbol{Y}_k$ that sets the inclination to a target value.

The attack is still possible when only either the accelerometer or magnetometer is disruptive. When the accelerometer=$\mathcal{D}$, an attacker can get (roll, pitch, yaw) = ($C$, $C$, $\mathcal{D}$), ($\mathcal{D}$, $C$, $C$), or ($C$, $\mathcal{D}$, $C$). Alternatively when magnetometer=$\mathcal{D}$, then the attacker gets yaw=$C$. That is explained as follows. According to Eqs. (5) and (6), the roll and pitch are determined using solely the gravity vector, which is derived from the gyroscope and accelerometer. Conversely, the yaw is determined using both of the gravity and geomagnetic vectors, and the geomagnetic vector is derived from the gyroscope and magnetometer. Therefore, when the accelerometer has a low measurement noise, it affects all Euler angles. Similarly, when the magnetometer has a low measurement noise, it affects the yaw only.

In addition, when gyroscope=$\mathcal{D}$ and accelerometer=$\mathcal{D}$, the roll, pitch, and yaw are $\mathcal{D}$. Similarly, when gyroscope=$\mathcal{D}$ and magnetometer= $\mathcal{D}$, then only the yaw is $\mathcal{D}$.

*Example.* Similarly to the example in Sect. 3.2.1, we want to set all the Euler angles to 60 [°]. A non-stationary attack controls how the inclination is rotated along each axis of the gyroscope. Note that the axes of inclination (i.e., global coordinate system) differs from those of the gyroscope (i.e., sensor coordination). Thereby, for simplicity, each axis should be determined individually. Firstly, a non-stationary condition is given from Eqs. (1) and (2) as

$$|\boldsymbol{B}_k| \gg B \text{ [μT]}, \ |\boldsymbol{G}_k| \gg 1 \text{ [g]}. \quad (15)$$

Then, from the relation between the quaternion and a rotation matrix, Eq. (6) can be expressed as

$$\phi_k = \tan^{-1}\left(\frac{2(q_{0,k}^+ q_{1,k}^+ + q_{2,k}^+ q_{3,k}^+)}{1 - 2((q_{1,k}^+)^2 + (q_{2,k}^+)^2)}\right), \quad (16)$$

$$\theta_k = \sin^{-1}\left(2(q_{0,k}^+ q_{2,k}^+ - q_{3,k}^+ q_{1,k}^+)\right), \quad (17)$$

$$\psi_k = \tan^{-1}\left(\frac{2(q_{0,k}^+ q_{3,k}^+ + q_{1,k}^+ q_{2,k}^+)}{1 - 2((q_{2,k}^+)^2 + (q_{3,k}^+)^2)}\right), \quad (18)$$

where $q_k^+ = \{q_{0,k}^+, q_{1,k}^+, q_{2,k}^+, q_{3,k}^+\}$. Thus, each component of the quaternion can be determined to correspond to the initial angles $(\phi_k, \theta_k, \psi_k) = (0, 0, 0)$ [°] and the target angles $(60, 60, 60)$ [°], respectively. Moreover, the relation between the initial and target quaternions can be expressed as Eq. (14). For example, the following steps accomplish the purpose: (i) controlling $Y_{z,k}$ to make the yaw 8 [°], (ii) controlling $Y_{x,k}$ to make the roll 26 [°], (iii) controlling $Y_{y,k}$ to make the pitch 60 [°].

If $Y_{z,k}$, $Y_{x,k}$, and $Y_{y,k}$ can be controlled by 10 [°/s], the control is accomplished as follows. By applying $\boldsymbol{Y}_k = (0, 0, 10)$ for 0.4 [s], the Euler angles are changed from $(\phi_k, \theta_k, \psi_k) = (0, 0, 0)$ to $(\phi_k, \theta_k, \psi_k) = (0, 0, 4.0)$. Then, $\boldsymbol{Y}_k = (10, 0, 0)$ is applied for 2.6 [s], and we get $(\phi_k, \theta_k, \psi_k) = (26.0, 0, 4.0)$. Finally, we apply $\boldsymbol{Y}_k = (0, 10, 0)$ for 7.6 [s], and the resulting Euler angles are $(\phi_k, \theta_k, \psi_k) = (60.53, 59.77, 60.81)$.

### 3.3 Consequences of the Attacks for Different Attacker's Access Capabilities

In Sect. 3.1, the access capability of an attacker is classified into three levels namely $C$, $\mathcal{D}$, and $\mathcal{U}$. Because there are three access levels for each sensor comprising AHRS, there are $3^3 = 27$ cases in total. Table 1 summarizes the comprehensive evaluation of the 27 cases. Each row corresponds to the attacker's ability to access raw sensors. The results for both stationary and non-stationary attacks are shown. Using these results, we can predict the effect of the proposed attack for given degrees of access to raw sensors. As mentioned in Sects. 3.2.1 and 3.2.2, in some cases, an attacker has options regarding which angle to control. The table expresses the circumstances with multiple rows.

## 4 EXPERIMENT

The feasibility of the proposed attack was verified by performing two experiments. The first experiment is conducted using simulations, where a signal injection attack is modelled by digitally overwriting sensor readings. This shows the effectiveness of the proposed attacks under an ideal condition. The second experiment was conducted in a real environment in which acoustic and magnetic signal injection attacks were carried out. This reveals an obstacle in the real world, e.g., inadequate control over sensors.

### 4.1 Setup

We used an NXP Semiconductor FRDM-K64F-AGM01 sensor fusion evaluation board as our evaluation platform [20]. It consists of a mother board (FRDM-K64F) that has an ARM Cortex-M4 microcontroller (MCU) and a daughter board (FRDM-STBC-AGM01) that has sensors. The daughter board has an FXAS21002C 3-axis gyroscope

**Table 1: Classification of Effects of Attacks.**

| Input | | | Output | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Stationary | | | Non-stationary | | |
| Accelerometor | Gyroscope | Magnetometer | Roll | Pitch | Yaw | Roll | Pitch | Yaw |
| $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ |
| $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{D}$ | $\mathcal{C}$<br>$\mathcal{D}$<br>$\mathcal{C}$ | $\mathcal{C}$<br>$\mathcal{C}$<br>$\mathcal{D}$ | $\mathcal{D}$<br>$\mathcal{C}$<br>$\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ |
| $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{C}$<br>$\mathcal{D}$<br>$\mathcal{C}$ | $\mathcal{C}$<br>$\mathcal{C}$<br>$\mathcal{D}$ | $\mathcal{D}$<br>$\mathcal{C}$<br>$\mathcal{C}$ | $\mathcal{C}$<br>$\mathcal{D}$<br>$\mathcal{C}$ | $\mathcal{C}$<br>$\mathcal{C}$<br>$\mathcal{D}$ | $\mathcal{D}$<br>$\mathcal{C}$<br>$\mathcal{C}$ |
| $\mathcal{C}$ | $\mathcal{D}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{D}$ | $\mathcal{D}$ | $\mathcal{D}$ |
| $\mathcal{C}$ | $\mathcal{D}$ | $\mathcal{D}$ | $\mathcal{C}$<br>$\mathcal{D}$<br>$\mathcal{C}$ | $\mathcal{C}$<br>$\mathcal{C}$<br>$\mathcal{D}$ | $\mathcal{D}$<br>$\mathcal{C}$<br>$\mathcal{C}$ | $\mathcal{D}$ | $\mathcal{D}$ | $\mathcal{D}$ |
| $\mathcal{C}$ | $\mathcal{D}$ | $\mathcal{U}$ | | | | | | |
| $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ |
| $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{D}$ | $\mathcal{C}$<br>$\mathcal{D}$<br>$\mathcal{C}$ | $\mathcal{C}$<br>$\mathcal{C}$<br>$\mathcal{D}$ | $\mathcal{D}$<br>$\mathcal{C}$<br>$\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ |
| $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | | | | | | |
| $\mathcal{D}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ |
| $\mathcal{D}$ | $\mathcal{C}$ | $\mathcal{D}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{C}$ |
| $\mathcal{D}$ | $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{C}$<br>$\mathcal{D}$<br>$\mathcal{C}$ | $\mathcal{C}$<br>$\mathcal{C}$<br>$\mathcal{D}$ | $\mathcal{D}$<br>$\mathcal{C}$<br>$\mathcal{C}$ |
| $\mathcal{D}$ | $\mathcal{D}$ | $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{C}$ | $\mathcal{D}$ | $\mathcal{D}$ | $\mathcal{D}$ |
| $\mathcal{D}$ | $\mathcal{D}$ | $\mathcal{D}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{D}$ | $\mathcal{D}$ | $\mathcal{D}$ |
| $\mathcal{D}$ | $\mathcal{D}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{D}$ | $\mathcal{D}$ | $\mathcal{D}$ |
| $\mathcal{D}$ | $\mathcal{U}$ | $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ |
| $\mathcal{D}$ | $\mathcal{U}$ | $\mathcal{D}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ |
| $\mathcal{D}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ |
| $\mathcal{U}$ | $\mathcal{C}$ | $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{C}$ |
| $\mathcal{U}$ | $\mathcal{C}$ | $\mathcal{D}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{C}$ |
| $\mathcal{U}$ | $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ |
| $\mathcal{U}$ | $\mathcal{D}$ | $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{D}$ |
| $\mathcal{U}$ | $\mathcal{D}$ | $\mathcal{D}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{D}$ |
| $\mathcal{U}$ | $\mathcal{D}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ |
| $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{C}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ |
| $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{D}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ |
| $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ | $\mathcal{U}$ |

$\mathcal{C}$: fully Controlable, $\mathcal{D}$: Disruptive, $\mathcal{U}$: Uncontrollable

and an FXOS8700CQ comprising a three-axis accelerometer and three-axis magnetometer.

The sensor fusion algorithm was implemented as software in the MCU on the mother board. The software for the sensor fusion is a part of the IoT Sensing Software Development Kit (ISSDK) distributed by NXP Semiconductors along with the evaluation board [19].

#### 4.1.1 *Setup for Emulating Signal Injection.*

The emulation environment was developed in order to model an ideal attacker who has complete control over sensors. To do this, the firmware on the MCU is modified to enable the overwriting of sensor readings. More specifically, we can set desired offset values for each sensor axis by sending a special packet to the MCU. Once the offset values are configured, the offset values are automatically added to raw sensor readings. Then, the modified sensor readings are fed to the sensor fusion library. Note that the modification is limited to the sensor input and the sensor fusion library is untouched. Using the modified firmware, we can evaluate the effects on sensor fusion when readings from sensors are controlled by an attacker.

#### 4.1.2 *Setup for Acoustic and Magnetic Signal Injections.*

To evaluate the feasibility of an actual attack, an experimental setup was developed. As mentioned in Sect. 3.3, the accelerometer and gyroscope may be susceptible to acoustic waves, while the a magnetometer can be attacked by an artificial magnetic fields. The experimental setup is composed of both acoustic and magnetic signal generators. A block diagram and overview of the setup are shown in Figs. 3 and 4, respectively.

In order to generate an acoustic signal, the waveform generated at the function generator is amplified and fed to a speaker. Piezo speakers are used in order to avoid magnetic disturbance. Parameters for the sinusoidal wave, such as the frequency and voltage,

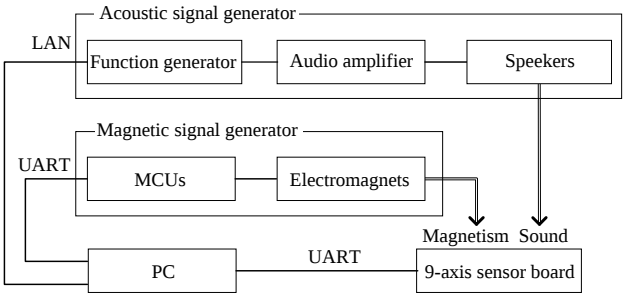


**Figure 3: Experimental environment (block diagram)**

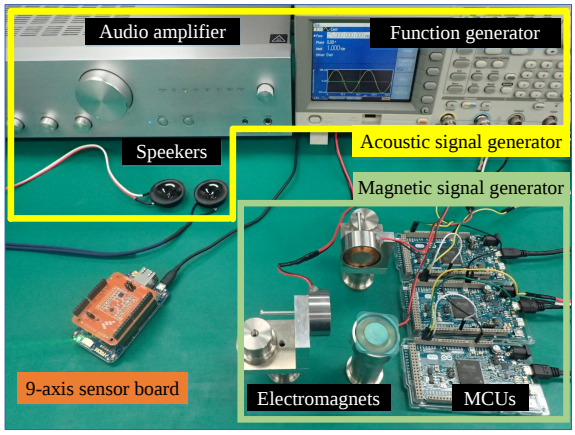


**Figure 4: Experimental environment (overview)**

are controlled by a PC connected to the function generator. The magnetic signal is generated as follows. Three electromagnets are connected to the MCU boards. The strength and direction of the magnetic field are controlled by changing the voltage supplied to the electromagnets using the MCU boards. The PC controls the magnetic field using MCUs. The parts that generate the acoustic and magnetic signals can be controlled independently and used simultaneously.

See Appendix A for preliminary experiments that were performed to determine the sensitivity of sensors on the FRDM-K64F-AGM01 against acoustic and magnetic signal injections.

## 4.2 Experiment using Emulated Signal Injections

In this section, the feasibility of the stationary and non-stationary attacks is verified in the emulation environment.

### 4.2.1 *Stationary Attack.*

The example shown in Sect. 3.2.1 was conducted in the emulation environment. This is a stationary attack by an attacker who has full control over the accelerometer and magnetometer.

Plots that were obtained as a result of the attack are shown in Fig. 5. Fig. 5-(a) is the inclination that resulted from sensor fusion. Meanwhile, Figs. 5-(b) to -(d) are raw sensor obtained from the gyroscope, accelerometer, and magnetometer, respectively.

In the experiment, readings from the accelerometer and magnetometer are changed to the target values given by Eqs. (12) and (13). To emphasize the behavior of sensor fusion, each axis is changed one-by-one with a slight delay. Firstly, the accelerometer's axes, namely $G_y$, $G_x$, and $G_z$, are changed in the order as shown in Fig. 5-(c). Then, the magnetometer's axes, namely $B_y$, $B_x$, and $B_z$, are changed in the order as shown in Fig. 5-(d). The inclination before the attack is (roll, pitch, yaw) = (-0.4, 0.4, -2.0) [°]. After all of the axes are changed, the inclination is changed to (62.3, 58.1, 60.0) [°] which is very close to the desired value, i.e., (60.0, 60.0, 60.0) [°]. It should be noted that the output from the gyroscope is untouched during the experiment as shown in Fig. 5-(b). The result demonstrates that the stationary attack is feasible.

### 4.2.2 *Non-Stationary Attack.*

The example shown in Sect. 3.2.2 was conducted in the emulation environment. This is the non-stationary attack by an attacker with the following capabilities: gyroscope=$C$, accelerometer=$\mathcal{D}$, and magnetometer=$\mathcal{D}$.

Figure 6-(a) depicts the time variation of the inclination, while Figs. 6-(b) to -(d) are raw sensor outputs. Firstly, specific axes in the accelerometer and magnetometer (i.e., $G_z$ and $B_z$) are increased as shown in Figs. 6-(c) and -(d). The sensor fusion algorithm recognizes this as a non-stationary situation because norms of readings from the accelerometer and magnetometer exceed $1g$ and $1B$, respectively.

Then, the gyroscope is controlled. As mentioned in Sect. 3.2.2 , each axis, i.e., $Y_z$, $Y_x$, and $Y_y$, are changed one-by-one as shown in Fig. 6-(b). The inclination before the attack is (roll, pitch, yaw) = (-0.8, -0.1, -1.1) [°], and this is changed to (57.1, 60.3, 58.8) [°] after the attack, as shown in Fig. 6-(a). The resulting inclination is very close to the target value (60.0, 60.0, 60.0) [°]. This result confirms the feasibility of the non-stationary attack.

## 4.3 Experiment with Acoustic and Magnetic Signal Injections

In this section, the feasibility of the proposed attack is verified in a real environment. In our setup, the ability to access each sensor is as follows: accelerometer=$\mathcal{D}$, gyroscope=$C$ and magnetometer=$C$ (see Appendix A for details). Therefore, the attacker is expected to achieve roll=$C$, pitch=$C$, yaw=$C$ with the non-stationary attack as summarized in Table 1. Because the outcome of the stationary attack is expected to be limited (i.e., roll=$\mathcal{U}$, pitch=$\mathcal{U}$, Yaw=$C$), only the non-stationary attack is examined. Note that the ability to access raw sensors is specific to our target evaluation board. Some accelerometers are reported to be controllable by signal injection attacks [24].

The inclination before the attack is (roll, pitch, yaw) = (2.0, 5.3, -2.6). Similarly to the previous experiment, the purpose of the attack is to set the inclination to (roll, pitch, yaw) = (60.0, 60.0 60.0) [°]. The results are shown in Fig. 7. Figs. 7-(a) depicts the time variations of the inclination. Figs. 7-(b) to -(d) show sensor readings that were obtained from the gyroscope, accelerometer, and magnetometer. Firstly, a non-stationary scenario is created by injecting disruptive signals to the accelerometer and magnetometer as shown in Figs. 7-(c) and -(d).

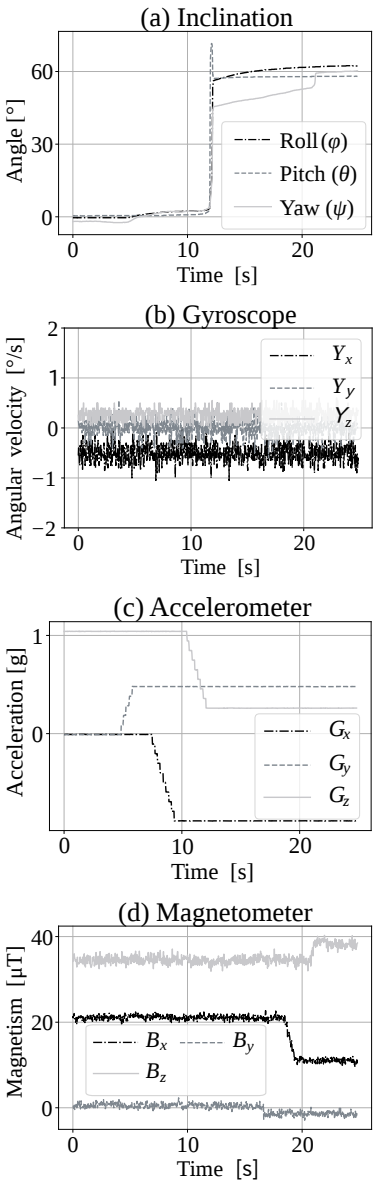


**Figure 5: Stationary attack in emulated environment**

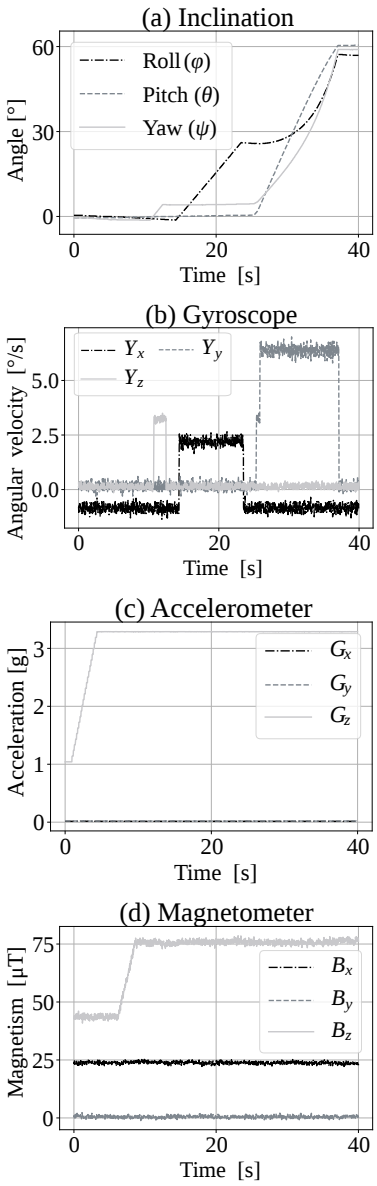


**Figure 6: Non-stationary attack in emulated environment**

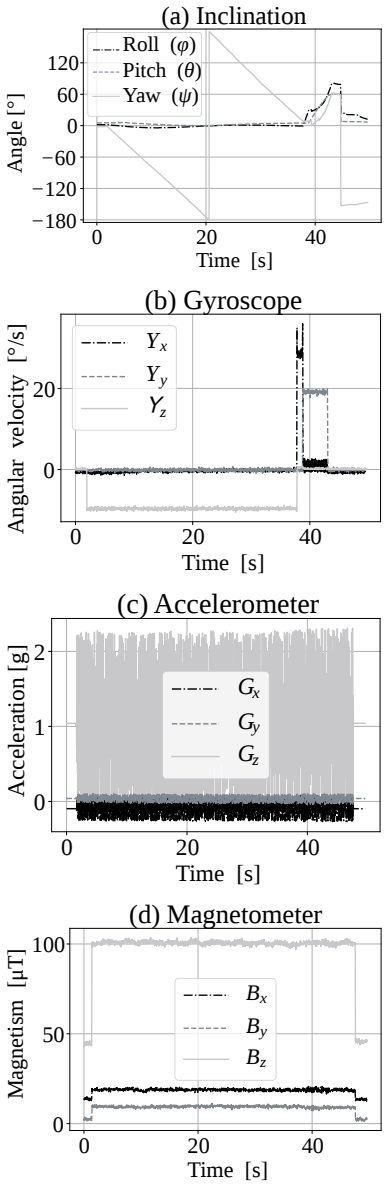


**Figure 7: Non-stationary attack in real environment**

After the non-stationary scenario is created, the gyroscope is precisely controlled. For simplicity, the $Y_x$, $Y_y$, and $Y_z$ axes are controlled individually in series. Fig. 7-(a) shows that the roll, pitch, and yaw approach to the target values. Finally, the inclination reached to (78.5, 60.1, 60) [°] at around Time=42 [s] as shown in Fig. 7-(a). The inclination is successfully controlled toward the target value; however, there is an error in the roll angle. In addition, the resulting inclination does not last long. After about 1 s, the inclination suddenly changes.

The non-ideal behavior which is not shown the real environment is caused because the power of the speaker is insufficient to prevent the accelerometer from being disrupted. This causes the resulting yaw angle to become 78.5 [°] instead of 60.0 [°]. In addition, because there is a disruption when using a sinusoidal wave, the effect of the disruption is periodic, and thus there is a period in which the disruption is weak. During this period, the sensor fusion algorithm recognizes that the environment has become less noisy, and starts to correct the inclination using the accelerometer and magnetometer, as shown at around Time=42 [s] when there is a sudden change in inclination. This results indicate that the attacker can conduct the attack even when there is only a partial disruption. As a result, the

experimental results indicate the feasibility of the non-stationary attack in a real environment.

## 5 COUNTERMEASURE

In this section, we propose the use of a software-based countermeasure against the proposed attacks. The basic idea behind the countermeasure is to detect the attacks by monitoring the inconsistency between readings from sensors. More specifically, gravity and geomagnetic vectors measured in two different ways inevitably become inconsistent as a result of the proposed attack. Therefore, the errors between the two measurements, i.e., the magnetic error $q_{zm\epsilon,k}$ and gravity error $q_{zg\epsilon,k}$ can be used for the detection.

The effectiveness of the proposed countermeasure is verified using experimental data. The magnetic error $q_{zm\epsilon,k}$ and gravity error $q_{zg\epsilon,k}$ in normal and abnormal conditions are shown in Figs. 8 and 9, respectively. Notably, the abnormal case is captured from a real environment, while the non-stationary attack in Sect. 4.3 is being performed. The results show that errors are observed, while the attack is about 10 times larger than those of the normal case. Therefore, the attack is clearly detectable.

The reason for the inconsistency in the detection is that due to the injection of a signal injection to a suitable subset of raw sensors. Therefore, the proposed countermeasure detects the most cases in Table 1. An exception is when all of the raw sensors are fully controllable, i.e., accelerometer=$C$, magnetometer=$C$ , and gyroscope=$C$. In that case, the attacker can potentially control the sensor fusion output without being detected. However, the proposed countermeasure is still valuable because sensors are not always controllable. The target sensor evaluated in this study paper is a good example in that the accelerometer is not controllable, as shown in Appendix A.

## 6 RELATED WORK

This section explains conventional studies investigated acoustic injection attacks [22, 24, 25].

### 6.1 Acoustic Injection Attacks

In rapid increase of demands for security of sensors, Son et.al. unveiled a threat of acoustic injection attacks to MEMS sensors [22]. They demonstrated intentional sound noise disrupted MEMS gyroscopes on drones, and thus drones lost control. Similarly, Trippel et.al. showed the attacks could also disrupt MEMS accelerometers [24]. In addition, they utilized amplitude modulated (AM) and phase modulated (PM) acoustic waves to obtain full control of an accelerotmer. With AM or PM, proof-of-concept attacks were carried out to demonstrate a smartphone-controlled model car and a fitness tracker could be controlled. Based on these studies, Wang et.al. showed MEMS gyroscopes also could be controlled using modulated acoustic waves [25]. Moreover, they demostrated various kinds of smart devices e.g., virtual reality headsets, drones and self balancing robots could be disrupted.

Acoustic injection attacks focus on a structure of MEMS sensors. MEMS gyroscopes and MEMS accelerometers consist of a sensing mass and springs, and such spring-mass structures have resonant frequency [22, 24, 25]. The attacks inject acoustic waves which frequency matches the resonant frequency of a target sensor. As a result, the sensor vibrates and unexpected measurements are digitized through an amplifier, a low pass filter and an ADC. When the amplifier is insecure, the measurements saturate, and thus digitized values become constant and biased. When the low pass filter is insecure, unnecessary frequencies are passed, and thus digitized measurements sinusoidally fluctuate [24].

### 6.2 Countermeasures

To mitigate acoustic injection attacks, two kinds of countermeasure are considered. One is hardware-based, the other is software-based. Hardware-based countermeasures have been investigated: 1) surrounding a MEMS sensor with sound isolation materials [22, 24, 25], 2) tuning the resonant frequency [22], 3) comparing outputs from multiple sensors [22, 25], 4) designing a MEMS structure resilient to acoustic waves [24, 25], 5) detecting the resonating sound with a microphone [25]. Each countermeasure has limitations such as increase of cost for additional components or decrease sensitivity of the sensor.

Trippel et.al. proposed software-based countermeasures using software controlled ADCs. One is randomized sampling, the other is out-of-phase sampling. Each sampling method changes the sampling period and therefore it prevents sinusoidal analog measurements from being translated to biased digitized values. This approach has an limitation that it is ineffective for saturated output from an insecure amplifier [24].

## 7 CONCLUSIONS AND FUTURE WORK

In this study, we performed a rigorous security evaluation of sensor fusion for inclination, and we combined an accelerometer, gyroscope, and magnetometer based on the indirect Kalman filter. Though our contributions are limited in the sense that we evaluated a specific sensor fusion algorithm with a specific sensor board, an AHRS using (indirect) Kalman filter is popular and widely used. So we conclude that while sensor fusion introduces a certain degree of attack resilience, it remains susceptible to attacks. We proposed two attacks, and systematically analyzed their consequences in cases where there were different levels of access to raw sensors (see Sect. 3.3). The proposed attacks were verified by performing experiments in simulated and real environments. We also proposed attack-detection techniques to counter the attacks.

There are several problems that will be the focus of further study.

*Stationary Attack in Real Environment:* In this paper, the noiseless attack was evaluated only in the emulated environment because the target accelerometer is not controllable with acoustic or magnetic signal injections. Therefore, its evaluation in a real environment is necessary in future research. For this purpose, there is the need to examine accelerometers that are controllable with acoustic signal injection.

*Attack on Other Sensor Fusion Algorithms:* The sensor fusion for inclination based on the indirect Kalman filter is just one example of a sensor fusion algorithms. The security of other sensor-fusion algorithms should be evaluated if underlying raw sensors are susceptible to signal injection attacks.

*Feasibility of the Attacks in Real-World Applications:* In this paper, we focused on the attack resilience of the sensor-fusion algorithm,

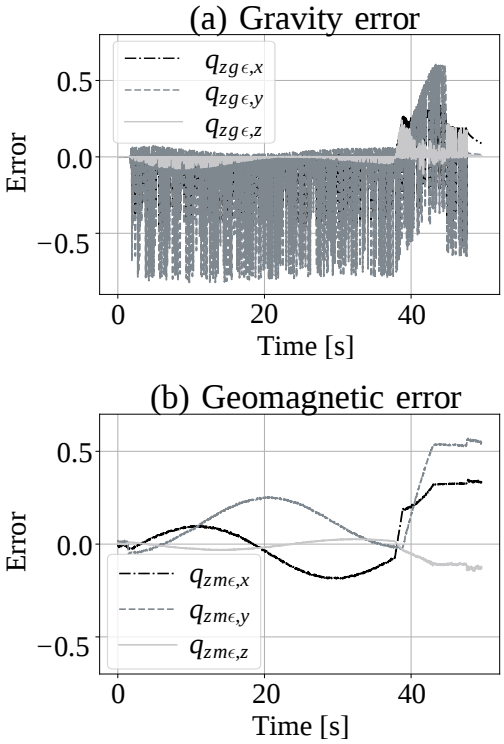


**Figure 8: Error in noisy attack**

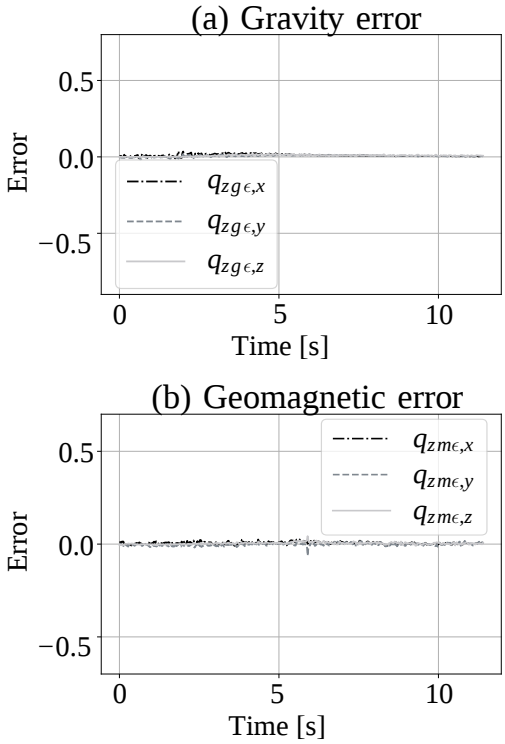


**Figure 9: Error in normal usage**

and thus conducted experiments using an evaluation board in a laboratory environment. For real applications such as drones or robots, there are additional control and signal processing layers on top of the sensor fusion for AHRS. Consequences of the proposed attack in such applications are non-trivial, and thus require further research.

*Black Box Attack:* In this paper, we assumed that the sensor fusion algorithm is known to the attacker following the common principle in the security research. However, some sensor modules have integrated MCUs that implement sensor fusion. In such cases, the algorithm and/or implementation are not available to the attacker. The applicability of the proposed attack to such black-box implementations is an interesting question. In addition, the construction of a similar attack just by observing the behavior of a black-box implementation may be a challenging study for future work.

*Sensitivity of Magnetometer to Acoustic Signals:* In the evaluation, we observed that the magnetometer is sensitive to acoustic signal injection, similar to the gyroscope and accelerometer, as shown in Appendix A. The behavior itself is irrelevant to the proposed attack; however, to the best of our knowledge, is an interesting phenomenon that has not been reported in previous works. The mechanism behind the phenomenon and its applicability to other magnetometers are possibilities avenues for future research.

## ACKNOWLEDGEMENTS

This paper is based on results obtained from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

## REFERENCES

[1] Nicholas Carlini, Pratyush Mishra, Tavish Vaidya, Yuankai Zhang, Micah Sherr, Clay Shields, David Wagner, and Wenchao Zhou. 2016. Hidden Voice Commands.. In *USENIX Security Symposium*. 513–530.

[2] Ruchir Chauhan. 2014. *A platform for false data injection in frequency modulated continuous wave radar.* Utah State University.

[3] Drew Davidson, Hao Wu, Robert Jellinek, Vikas Singh, and Thomas Ristenpart. 2016. Controlling UAVs with Sensor Input Spoofing Attacks.. In *WOOT*.

[4] Peter D Hanlon and Peter S Maybeck. 2000. Multiple-model adaptive estimation using a residual correlation Kalman filter bank. *IEEE Trans. Aerospace Electron. Systems* 36, 2 (2000), 393–406.

[5] Andrew J Kerns, Daniel P Shepard, Jahshan A Bhatti, and Todd E Humphreys. 2014. Unmanned aircraft capture and control via GPS spoofing. *Journal of Field Robotics* 31, 4 (2014), 617–636.

[6] 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 *Security and Privacy (SP), 2013 IEEE Symposium on*. IEEE, 145–159.

[7] Hyung-Jik Lee and Seul Jung. 2009. Gyro sensor drift compensation by Kalman filter to control a mobile inverted pendulum robot system. In *Industrial Technology, 2009. ICIT 2009. IEEE International Conference on*. IEEE, 1–6.

[8] Wei Li and Jinling Wang. 2013. Effective adaptive Kalman filter for MEMS-IMU/magnetometers integrated attitude and heading reference systems. *The Journal of Navigation* 66, 1 (2013), 99–113.

[9] Faraz M Mirzaei and Stergios I Roumeliotis. 2008. A Kalman filter-based algorithm for IMU-camera calibration: Observability analysis and performance evaluation. *IEEE transactions on robotics* 24, 5 (2008), 1143–1156.

[10] Yilin Mo and Bruno Sinopoli. 2010. False Data Injection Attacks in Control Systems. In *First Workshop on Secure Control Systems*.

[11] Young-Seok Park, Yunmok Son, Hocheol Shin, Dohyun Kim, and Yongdae Kim. 2016. This Ain't Your Dose: Sensor Spoofing Attack on Medical Infusion Pump.. In *WOOT*.

[12] Aanjhan Ranganathan, Boris Danev, Aurélien Francillon, and Srdjan Capkun. 2012. Physical-layer attacks on chirp-based ranging systems. In *Proceedings of*

*the fifth ACM conference on Security and Privacy in Wireless and Mobile Networks*. ACM, 15–26.

[13] Daniel Roetenberg, Henk Luinge, and Peter Veltink. 2003. Inertial and magnetic sensing of human movement near ferromagnetic materials. In *Mixed and Augmented Reality, 2003. Proceedings. The Second IEEE and ACM International Symposium on*. IEEE, 268–269.

[14] Daniel Roetenberg, Henk J Luinge, Chris TM Baten, and Peter H Veltink. 2005. Compensation of magnetic disturbances improves inertial and magnetic sensing of human body segment orientation. *IEEE Transactions on neural systems and rehabilitation engineering* 13, 3 (2005), 395–405.

[15] Daniel Roetenberg, Per J Slycke, and Peter H Veltink. 2007. Ambulatory position and orientation tracking fusing magnetic and inertial sensing. *IEEE Transactions on Biomedical Engineering* 54, 5 (2007), 883–890.

[16] Stergios I Roumeliotis, Gaurav S Sukhatme, and George A Bekey. 1999. Circumventing dynamic modeling: Evaluation of the error-state kalman filter applied to mobile robot localization. In *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on*, Vol. 2. IEEE, 1656–1663.

[17] Angelo M Sabatini. 2006. Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing. *IEEE Transactions on Biomedical Engineering* 53, 7 (2006), 1346–1356.

[18] NXP Semiconductors. 2016. *AN5023 - Sensor Fusion Kalman Filters*.

[19] NXP Semiconductors. 2016. IoT Sensing Software Development Kit: Embedded Software Framework. (2016). Retrieved December 15, 2017 from https://www.nxp.com/support/developer-resources/software-development-tools/sensor-developer-resources/sensor-toolbox-sensor-development-ecosystem/iot-sensing-software-development-kit-embedded-software-framework:IOT-SENSING-SDK

[20] NXP Semiconductors. 2016. Sensor Toolbox Development Platform for FXAS21002C and FXOS8700C 9-Axis Solution. (2016). Retrieved December 15, 2017 from http://www.nxp.com/products/developer-resources/hardware-development-tools/freedom-development-boards/sensor-toolbox-development-platform-for-fxas21002c-and-fxos8700c-9-axis-solution:FRDM-STBC-AGM01

[21] Yasser Shoukry, Paul Martin, Paulo Tabuada, and Mani Srivastava. 2013. Non-invasive spoofing attacks for anti-lock braking systems. In *International Workshop on Cryptographic Hardware and Embedded Systems*. Springer, 55–72.

[22] Yunmok Son, Hocheol Shin, Dongkwan Kim, Young-Seok Park, Juhwan Noh, Kibum Choi, Jungwoo Choi, Yongdae Kim, et al. 2015. Rocking Drones with Intentional Sound Noise on Gyroscopic Sensors.. In *USENIX Security Symposium*. 881–896.

[23] Ciza Thomas (Ed.). 2011. *Sensor Fusion - Foundation and Applications*. InTech.

[24] Timothy Trippel, Ofir Weisse, Wenyuan Xu, Peter Honeyman, and Kevin Fu. 2017. WALNUT: Waging doubt on the integrity of mems accelerometers with acoustic injection attacks. In *Security and Privacy (EuroS&P), 2017 IEEE European Symposium on*. IEEE, 3–18.

[25] Zhengbo Wang, Kang Wang, Bo Yang, Shangyuan Li, and Aimin Pan. 2017. Sonic Gun to Smart Devices - Your Devices Lose Control Under Ultrasound/Sound. In *BlackHat USA*.

[26] Chen Yan, Wenyuan Xu, and Jianhao Liu. 2016. Can you trust autonomous vehicles: Contactless attacks against sensors of self-driving vehicle. *DEF CON* 24 (2016).

[27] Xiaoping Yun and Eric R Bachmann. 2006. Design, implementation, and experimental results of a quaternion-based Kalman filter for human body motion tracking. *IEEE transactions on Robotics* 22, 6 (2006), 1216–1227.

[28] Guoming Zhang, Chen Yan, Xiaoyu Ji, Taimin Zhang, Tianchen Zhang, and Wenyuan Xu. 2017. Dolphinatack: Inaudible voice commands. *arXiv preprint arXiv:1708.09537* (2017).

[29] Rui Zhang and Leonhard M Reindl. 2011. Pedestrian motion based inertial sensor fusion by a modified complementary separate-bias Kalman filter. In *Sensors Applications Symposium (SAS), 2011 IEEE*. IEEE, 209–213.

# A PRELIMINARY EXPERIMENT

This section describes a preliminary experiment for determining the sensitivity of sensors on FRDM-K64F-AGM01 against acoustic and magnetic signal injections. The sensors on the board are FXAS21002C and FXOS8700CQ, which have not been evaluated in previous works [22, 24].

## A.1 Acoustic Injection

Using the setup described in Sect. 4.1, the preliminary experiment is conducted as follows. The target board is placed in a stationary chamber. Then, a sinusoidal wave with a certain frequency is injected for 1 s while the sensor reading is recorded. The procedure is repeated by varying the frequency from 2 kHz to 40 kHz in increments of 10 Hz.

Figure 10 illustrates the frequency responses of the inclination and each sensor. In each subfigure, the horizontal axis represents the frequency of the sinusoidal wave, and the vertical axis represents the sensor reading. There are spikes at relatively higher frequencies in the gyroscope, as shown in Figs. 10-(d) to -(f). In contrast, the accelerometer shows spikes at relatively lower frequencies, as shown in Figs. 10-(g) to -(i). These results reflect the nature of the sensors. The gyroscope is usually designed to have a high reactivity compared to the accelerometer, and it thus has higher resonant frequencies.

The sensitive frequencies observed in Fig. 10 were further investigated in the time domain. Then, it is determined that desired offset values can be added to the axes of the gyroscope independently. This means that the accessibility of the gyroscope is $C$. In contrast, only sinusoidal waves can be injected to the accelerometer output. Furthermore, a disruption to only the $G_z$-axis of the accelerometer is sufficient to mount the non-stationary attack. Therefore, the accessibility of the accelerometer is $\mathcal{D}$, which is when $G_z$ is disrupted.

It is interesting to note that some spikes are observed in the magnetometer. The behavior itself is irrelevant to the proposed attack, but it is a phenomenon that has not been reported to date. Some MEMS-based magnetometers are sensitive to vibrations induced by acoustic triggers; however, the internal design of the FXOS8700CQ is not clear.

## A.2 Magnetic Injection

The sensitivity of FXOS8700CQ against magnetic injection attacks was evaluated. Using the setup described in Sect. 4.1, the experiment was conducted as follows. Three electromagnets were placed in the X, Y, and Z axes that surround the magnetometer. The voltage supply to the electromagnets, namely $V_x$, $V_y$, and $V_z$, are changed from 2.5 [V] to -2.5 [V], while recording the sensor output.

Figure 11 shows the results of the experiment. Figs. 11-(d) to -(f), -(g) to -(i), and -(j) to (l) show readings obtained from the gyroscope, accelerometer, and magnetometer, respectively. For reference, Figs. 11-(m) to -(o) show traces of $V_x$, $V_y$, and $V_z$ applied for the electromagnets.

The results show that all of the axes of the magnetometer are affected. Moreover, the measured magnetic fields, $B_x$, $B_y$, and $B_z$ are clearly proportional to $V_x$, $V_y$, and $V_z$, respectively, as shown in Figs. 11-(j) to -(o). Although there is crosstalk between the axes, the corresponding axis (i.e. $B_i$ for $V_i$) is mainly affected. In addition, as expected, the gyroscope and accelerometer are immune to the magnetic field. As a result, we can conclude that the accessibility of the magnetometer is $C$.

# B DETAILED DERIVATION

This section provides an additional derivation for the sensor fusion algorithm in Fig. 2.

In Fig. 2-(1), the geomagnetic vector $\boldsymbol{m}_k^{6DOF}$ is given by

$$\boldsymbol{m}_k^{6DOF} = \frac{\boldsymbol{B}_k}{|\boldsymbol{B}_k|}. \tag{19}$$

The gravity vector $\boldsymbol{g}_k^{6DOF}$ is given by

$$\boldsymbol{g}_k^{6DOF} = \frac{\boldsymbol{G}_k}{|\boldsymbol{G}_k|}. \tag{20}$$

In addition, the measurement noise matrix $\boldsymbol{Q}_{v,k}$ is defined as the covariance of the measurement noise vector, $\boldsymbol{v}_k$, as follows.

$$\begin{aligned}\boldsymbol{Q}_{v,k} &= E[\boldsymbol{v}_k \boldsymbol{v}_k^T] = E\left[\begin{pmatrix}\boldsymbol{v}_{qzg,k}\\ \boldsymbol{v}_{qzm,k}\end{pmatrix}\begin{pmatrix}\boldsymbol{v}_{qzg,k}\\ \boldsymbol{v}_{qzm,k}\end{pmatrix}^T\right]\\ &= \begin{pmatrix} E\left[\boldsymbol{v}_{qzg,k}(\boldsymbol{v}_{qzg,k})^T\right] & E\left[\boldsymbol{v}_{qzg,k}(\boldsymbol{v}_{qzm,k})^T\right]\\ E\left[\boldsymbol{v}_{qzm,k}(\boldsymbol{v}_{qzg,k})^T\right] & E\left[\boldsymbol{v}_{qzm,k}(\boldsymbol{v}_{qzm,k})^T\right]\end{pmatrix}. \end{aligned}\tag{21}$$

$\boldsymbol{v}_{qzm,k}$ represents an error in the estimates of the geomagnetic vector calculated in two ways. Similarly, $\boldsymbol{v}_{qzg,k}$ represents that of the gravity vector. Each component of $\boldsymbol{Q}_{v,k}$ is calculated as follows.

$$E\left[\boldsymbol{v}_{qzg,k}(\boldsymbol{v}_{qzg,k})^T\right] = \frac{1}{12}\{(Q_{vG,k} + Q_{a,k}) + \alpha^2(Q_{vY} + Q_{wb})\}\boldsymbol{I}_3, \tag{22}$$

$$E\left[\boldsymbol{v}_{qzm,k}(\boldsymbol{v}_{qzm,k})^T\right] = \frac{1}{12}\{(Q_{vB,k} + Q_{d,k}) + \alpha^2(Q_{vY} + Q_{wb})\}\boldsymbol{I}_3, \tag{23}$$

$$E\left[\boldsymbol{v}_{qzg,k}(\boldsymbol{v}_{qzm,k})^T\right] = E\left[\boldsymbol{v}_{qzm,k}(\boldsymbol{v}_{qzg,k})^T\right] = \boldsymbol{0}, \tag{24}$$

$$\alpha = \frac{\pi\delta t}{180}, \tag{25}$$

where $\boldsymbol{I}_3$ is a $3 \times 3$ identity matrix, and $\delta t$ is a sampling interval of the Kalman filter. Eqs. (21)-(25) show that $Q_{v,k}$ is represented by two kinds of components: one kind is sensor noises $Q_{vG,k}$, $Q_{vB,k}$, and $Q_{vY}$ for the accelerometer, magnetometer, and gyroscope, respectively. Another kind is the use of components that disturb accurate measurements: acceleration $Q_{a,k}$ for the accelerometer, magnetic disturbance $Q_{d,k}$ for magnetometer, and *zero rate offset* noise $Q_{wb}$ for the gyroscope. Here, the *zero rate offset* is drifts occurring in the stationary condition. Noises in the accelerometer and magnetometer can be represented as Eqs. (1) and (2). $Q_{vY}$ and $Q_{wb}$ are constants, but they differ for each gyroscope.

In Fig. 2-(2), the true angular velocity $\boldsymbol{\omega}_k^-$ is estimated by correcting $\boldsymbol{Y}_k$ based on an offset vector $\boldsymbol{b}_k^-$ as follows.

$$\omega_k^- = \boldsymbol{Y}_k - \boldsymbol{b}_k^- \tag{26}$$

Then, *a priori* estimation of inclination $q_k^-$ is computed by accumulating $\boldsymbol{\omega}_k^-$ to *a posteriori* estimation of inclination $q_{k-1}^+$.

$$q_k^- = q_{k-1}^+ \otimes \Delta q(\boldsymbol{\omega}_k^- \delta t), \tag{27}$$

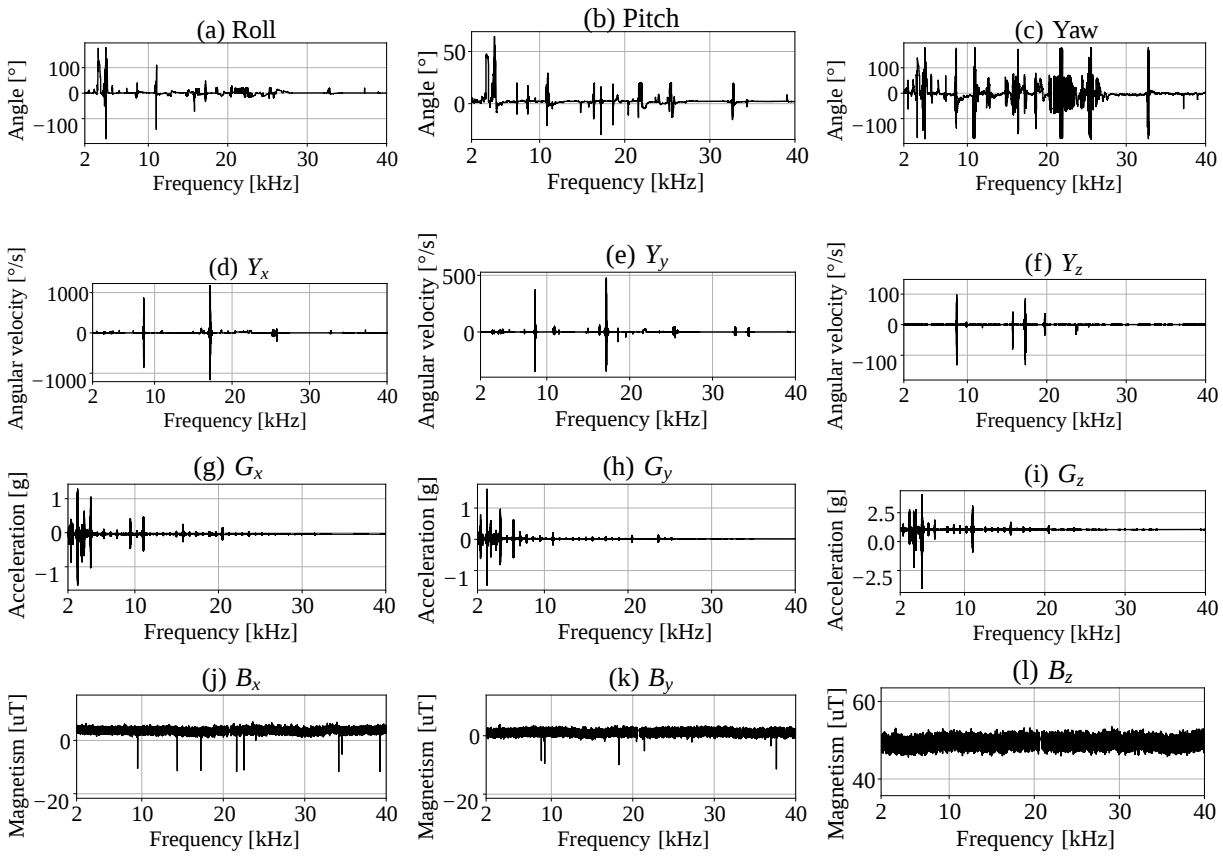


**Figure 10: Result of acoustic effect on AHRS**

where $\Delta q(\boldsymbol{\omega}_k \delta t)$ is a quaternion given by

$$\begin{aligned} \Delta q(\boldsymbol{\omega}_k \delta t) &= \{q_0, \boldsymbol{q}\} \\ &= \{q_0, q_1, q_2, q_3\} \\ &= \{\sqrt{1-q_1^2-q_2^2-q_3^2}, \frac{\boldsymbol{\omega}_k}{|\boldsymbol{\omega}_k|} \sin\left(\frac{|\boldsymbol{\omega}_k|\delta t}{2}\right)\}, \end{aligned} \tag{28}$$

and $\otimes$ represents the quaternion product.

$$\begin{aligned} q_0 &= q_{a,0}q_{b,0} - q_{a,1} * q_{b,1} - q_{a,2}q_{b,2} - q_{a,3}q_{b,3} \\ q_1 &= q_{a,0}q_{b,1} + q_{a,1} * q_{b,0} + q_{a,2}q_{b,3} - q_{a,3}q_{b,2} \\ q_2 &= q_{a,0}q_{b,2} - q_{a,1} * q_{b,3} + q_{a,2}q_{b,0} + q_{a,3}q_{b,1} \\ q_3 &= q_{a,0}q_{b,3} + q_{a,1} * q_{b,2} - q_{a,2}q_{b,1} + q_{a,3}q_{b,0} \end{aligned} \tag{29}$$

In the case of Eq. (27), $q_a$ is $q_k^+$ and $q_b$ is $\Delta q(\boldsymbol{\omega}_k \delta t)$. Finally, the *a posteriori* estimates $\boldsymbol{m}_k^-$ and $\boldsymbol{g}_k^-$ are deduced from $q_k^-$.

In Fig. 2-(3), the error $\boldsymbol{q}_{zm\epsilon,k}$ is derived by comparing $\boldsymbol{m}_k^{6DOF}$ and $\boldsymbol{m}_k^-$ as mentioned in Sect. 2.2. Similarly, the error $\boldsymbol{q}_{zg\epsilon,k}$ is derived from $\boldsymbol{g}_k^{6DOF}$ and $\boldsymbol{g}_k^-$.

In Fig. 2-(4), the Kalman gain $\boldsymbol{K}_k$ is updated, then *a posteriori* estimates of the process errors $\boldsymbol{x}_{\epsilon,k}^+$ are derived as mentioned in Sect. 2.2.

In Fig. 2-(5), *a posteriori* estimates of the geomagnetic and gravity vectors, namely $\boldsymbol{g}_k^+$ and $\boldsymbol{m}_k^+$, which were calculated by correcting the *a priori* estimates with the estimated errors as follows.

$$\boldsymbol{m}_k^+ = q_{m\epsilon,k}^+ \boldsymbol{m}_k^- (q_{m\epsilon,k}^+)^*, \tag{30}$$

$$\boldsymbol{g}_k^+ = q_{g\epsilon,k}^+ \boldsymbol{g}_k^- (q_{g\epsilon,k}^+)^*, \tag{31}$$

where the symbol $^*$ represents the conjugate. Here, those of the gyroscope (i.e., $\boldsymbol{m}_k^+$ and $\boldsymbol{g}_k^+$) are corrected by $q_{m\epsilon,k}^+$ and $q_{g\epsilon,k}^+$, which respectively reflect measurements by the magnetometer and accelerometer. The degree of correction is determined by the Kalman filter gain $\boldsymbol{K}_k$. Then, $\boldsymbol{m}_k^+$ and $\boldsymbol{g}_k^+$ are converted to *a posteriori* estimates of the inclination represented by a rotation matrix $\boldsymbol{R}_k^+$ in Eq. 5. Assume the rotation matrix is represented by NED (north, east, and down) coordinate system. The z-axis of $\boldsymbol{R}_k^+$, namely $\boldsymbol{R}_{z,k}^+$ is in the downward direction, and thus it is in parallel to the gravity vector $\boldsymbol{g}_k^+$. Therefore, it is given by

$$\boldsymbol{R}_{z,k}^+ = \frac{\boldsymbol{g}_k^+}{|\boldsymbol{g}_k^+|}. \tag{32}$$

The y-axis $\boldsymbol{R}_{y,k}^+$ is in the eastward direction, and it is orthogonal to both $\boldsymbol{m}_k^+$ and $\boldsymbol{g}_k^+$, and is given by

$$\boldsymbol{R}_{y,k}^+ = \frac{\boldsymbol{g}_k^+ \times \boldsymbol{m}_k^+}{|\boldsymbol{g}_k^+ \times \boldsymbol{m}_k^+|}. \tag{33}$$

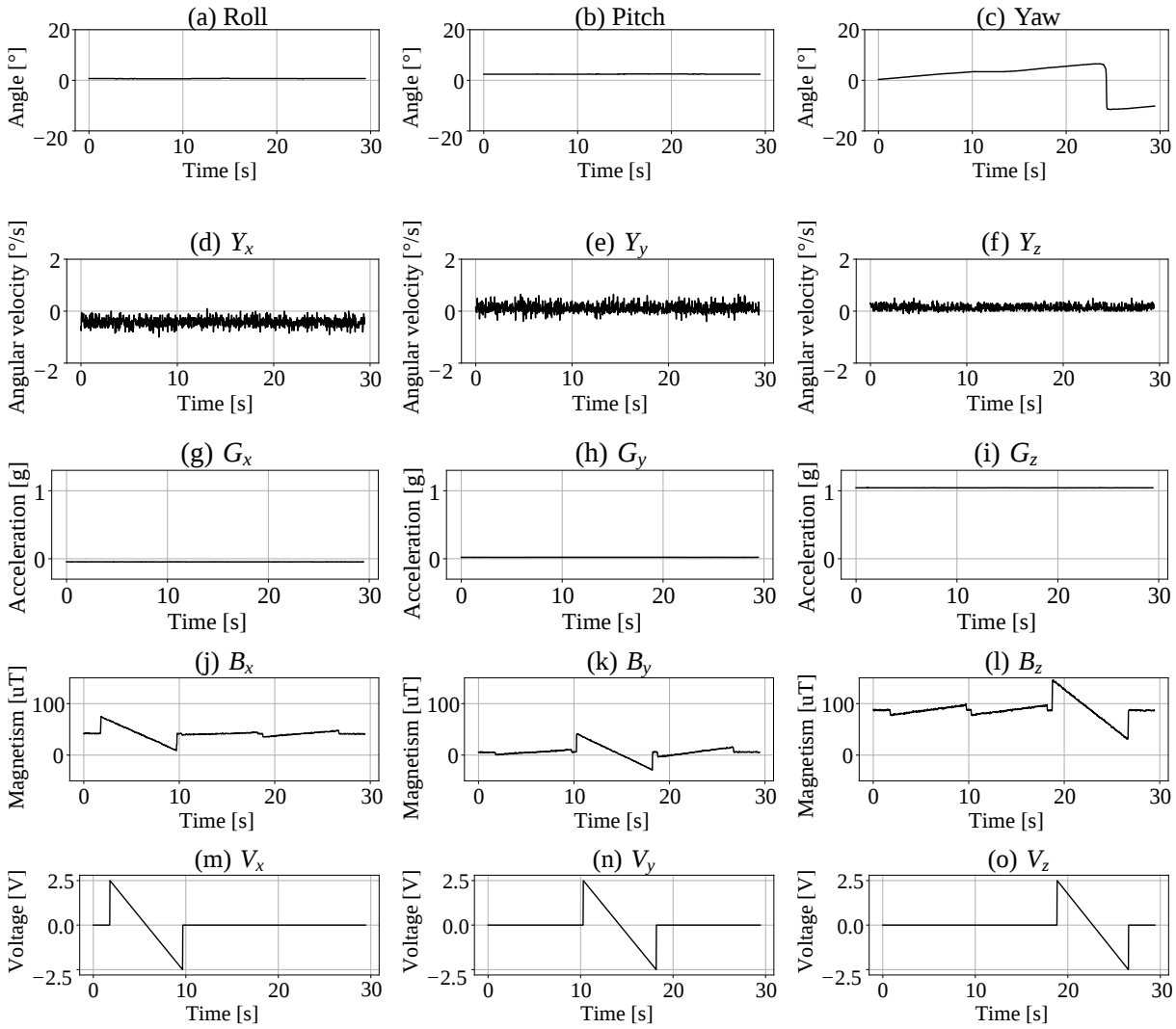


**Figure 11: Result of magnetic effect on AHRS**

This is because the geomagnetic vector $\boldsymbol{m}_k^+$ points northward and downward. Finally, the remaining x-axis $\boldsymbol{R}_{x,k}^+$ which is orthogonal to $\boldsymbol{R}_{y,k}^+$ and $\boldsymbol{R}_{z,k}^+$ is uniquely determined as

$$\boldsymbol{R}_{x,k}^+ = \boldsymbol{R}_{y,k}^+ \times \boldsymbol{R}_{z,k}^+. \tag{34}$$

As a result, the rotation matrix $\boldsymbol{R}_k^+$ is converted to Euler angles as stated in Eq. 6.