Enabling In-Ear Magnetic Sensing: Automatic and User Transparent Magnetometer Calibration

Andrea Ferlini[†], Alessandro Montanari[‡*], Andreas Grammenos[†¶*], Robert Harle[†], and Cecilia Mascolo[†]
[†]University of Cambridge, UK [‡]Nokia Bell Labs, Cambridge, UK [¶]Alan Turing Institute, UK
af679@cl.cam.ac.uk, alessandro.montanari@nokia-bell-labs.com, {ag926, rkh23, cm542}@cl.cam.ac.uk

Abstract*—Earables (in-ear wearables) are a new frontier in wearables. Acting both as leisure devices, providing personal audio, as well as sensing platforms, earables could collect sensor data for the upper part of the body, subject to fewer vibrations and random movement variations than the lower parts of the body, due to inherent damping in the musculoskeletal system. These data may enable application domains such as augmented/virtual reality, medical rehabilitation, and health condition screening. Unfortunately, earables have inherent size, shape, and weight constraints limiting the type and position of the sensors on such platforms. For instance, lacking a magnetometer in all earables reference platforms, earables lack reference points. Thus, it becomes harder to work with absolute orientations. Embedding magnetometers in earables is challenging, as these rely heavily on radio (mostly Bluetooth) communication (RF) and contain magnets for magnetic-driven speakers and docking. We explore the feasibility of adding a built-in magnetometer in an earbud, presenting the first comprehensive study of the magnetic interference impacting the magnetometer when placed in an earable: both that caused by the speaker and by RF (music streaming and voice calls) are considered. We find that appropriate calibration of the magnetometer removes the offsets induced by the magnets, the speaker, and the variable interference due to BT. Further, we present an automatic, user-transparent adaptive calibration that obviates the need for alternative, expensive, and error-prone manual, or robotics, calibration procedures. Our evaluation shows how our calibration approach performs under different conditions, achieving convincing results with errors below 3° for the majority of the experiments.

I. INTRODUCTION

Recent years have seen the rise of wearable technologies, both in the form of specialist devices such as pacemakers and in consumer devices, primarily smartwatches. A growing trend is the use of wireless earbuds that, while designed primarily for personal audio playback, offer a new sensing platform at an important site on the body. So-called *Earables* can be equipped with a variety of sensors and radios making them potentially suitable for a range of applications that go beyond just audio streaming, including indoor navigation, augmented reality, enhanced perception and medical monitoring. Some analysts believe Earables could be as disruptive as smartphones were in the last decade – forecast to hold the largest share (35%) of the wearable market [4]. Today, however, we are in the early stages of understanding their capabilities, both in terms of the sensors that they can offer and their applications.

In this paper we focus on the feasibility and value of adding magnetometers to earables. Today's consumer earables already contain inertial sensors (IMU) like accelerometers and gyroscopes. In other mobile devices, these are paired with magnetometers to provide movement descriptors in a global, absolute frame of reference [36], which would be highly valuable at the head too. Although a reasonable argument could be using the magnetometer in the phone to provide the earbuds' IMU with the references needed to calibrate/re-calibrate them, in practice, this would not work. **Notably, the head does not always move according to the way the body does**; besides, they often do not face the same direction. Hence, relying only on the phone would inevitably provide descriptors that do not necessarily match those when moving the head. Beyond this, adding a magnetometer to an earable would allow for a variety of applications including inertial navigation; magnetic-field-based indoor localization [19]; driver monitoring systems [21]; and medical applications (i.e. intra-body localization [24], [39], speech-language therapy [11], transcranial stimulation [13]). However, the highly constrained earable form factor and practicalities around their calibration have so far prohibited the availability of magnetometers on earables. Common magnetometer calibration techniques are cumbersome and error-prone [15], and regular manual calibration of *both* earbuds (and any personal devices) is unrealistic. Instead, we leverage the heading of the user's phone – typically trustworthy, as we will discuss – to auto-calibrate the magnetometers in the earbuds when the devices are believed to be pointing in the same direction. We ensure the latter constraint by applying the algorithm only when the user is directly interacting with the phone.

The specific contributions of this paper are: **(i)** We explore how magnetometer signals are affected by magnetic disturbances expected in an earable, highlighting the need for good calibration. **(ii)** We propose a novel magnetometer calibration technique that leverages the user's phone sensors (typically well calibrated). The algorithm can run in the background, without user intervention, providing a semi-continuous calibration. **(iii)** We evaluate the performance of the calibration framework, both in terms of accuracy and system performance. Further, we present a proof of concept study with a navigation application. **(iv)** We theoretically and practically assess the computational and energy efficiency of our approach, showing

This work is supported by Nokia Bell Labs through their donation for the Centre of Mobile, Wearable Systems and Augmented Intelligence. AG is supported by the Alan Turing Institute under the EPSRC grant EP/N510129/1 and TU/C/000003.

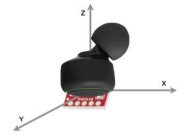
[*]Equal contribution.

our approach is **accurate** yet **computationally inexpensive**, allowing its regular execution on a constrained wearable, consuming $\approx 2.9\%$ extra power over idle.

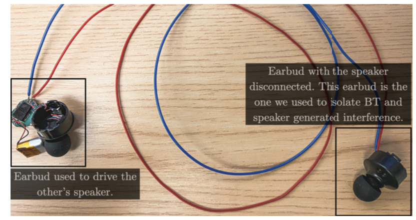
II. Motivation

Magnetometer and Inertial Sensing Historically, the presence of a magnetometer has been key to improve the accuracy of inertial based applications. Inertial sensors drift significantly when integrated over time [31]. For this reason, IMUs are often paired with a magnetometer: while the former measures relative motions (i.e. linear acceleration and rotational velocity), magnetometers sense the Earth Magnetic Field, and are used to find the (absolute) direction of the *Magnetic North* in a global reference frame. This constitutes an absolute anchor to be constantly re-calibrate IMUs. Further, magnetometers are also coupled with IMUs for 3D-motion tracking: without a magnetometer, it becomes extremely hard to have knowledge of the tracked object's heading in a global reference frame [36] (used to correctly initialize the tracking system). Unfortunately, neither IMU calibration, nor 3D-motion tracking (e.g. of the head) are feasible with today's earables, which lack a magnetometer. Fusing the user's smartphone magnetometer data and IMU readings from their earables would naturally result in a wrong estimation, given how user's head and phone often face in different directions. There are many compelling use-cases that can be unlocked from a magnetometer in earables, provided the magnetometer is accurate enough. Concretely, an in-earable magnetometer could enable acoustic AR [18], [34], [42] providing precise navigation thanks to spatial audio based on head orientation. For instance, considering a 4 lanes intersection ($\approx$10 meters wide), an error of 3° on the heading would entail an offset of $10 \times sin(1.5) \approx 0.26m$. Further, by leveraging head rotations to compute the angle of arrival (AoR) of incoming sounds with higher accuracy [17], it could provide improved noise-cancellation. To be effective, AoR estimation errors should be $< 20°$ [26]. This requirement becomes more stringent when using the earable for immersive audio or speaker isolation, especially in multi-source conversations.

Magnetic Interference There are various sources of magnetic interference which would impact a magnetometer in an earable, all being implicitly linked to the user's patterns. Firstly, earbuds usually have one (to drive the speaker) or more magnets (for docking purposes). These, being in close proximity to the magnetometer due to the earable's form factor, can interfere with its readings [22], making inference tasks unreliable at best [15]. Further, earables mostly communicate via Bluetooth (BT). Radio frequencies (RF) communications, like BT, require substantial electrical current which, flowing in the circuitry, generates an electromagnetic field, interfering with the magnetometer. Specifically with earables, RF communications are intense while streaming and just a few beacons otherwise. Practically, BT requires variable current, thus generating variable magnetic fields [22], entailing a hard-to-model electromagnetic interference, highly dependent on the users' patterns. Similarly, when playing music the speaker coil vibrations that generate sound also result in magnetic interference. Like sound, the interference depends on the vibration patterns, and therefore on whether the user is playing music, and what music they are playing. To understand how this practically affects a magnetometer in an earable, we present a thorough analysis of these interfering phenomena (Section III). We look at RF communications (i.e. music streaming form a smartphone to the earbuds), as well as music playback. Music playback was selected because of its popularity and the wide spectrum of tonal patterns of different genres. We analyze the impact of voice calls, too. Interestingly, these patterns, while present, are eclipsed by the interference produced by the RF circuitry (Section III-B).



(a)



(b)

Fig. 1: Basic system setup (Figure 1a) and setup used to isolate the interference generated by BT streaming and speaker (Figure 1b).

III. Magnetometer Calibration and Interference

Proper calibration is key to obtain accurate sensor readings, especially for continuous magnetic sensing: whether we are looking for the heading of an object [37], [41], or tracking magnetic bodies [21], [24], [39], reliable sensor data is crucial.

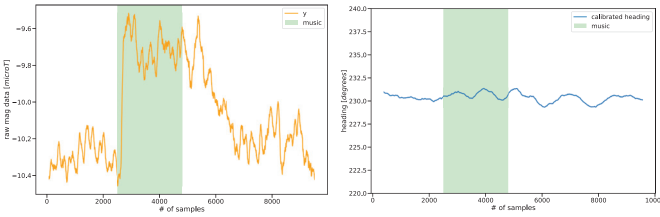
A. Magnetometer Calibration

To calibrate a magnetometer, common approaches seek to estimate the *bias* and *scale factor* for each axes. Magnetometer calibrations can be grouped in static or dynamic [16] and whether they rely on attitude information or not [38]. However, one major obstacle in calibration is interference from other magnetic fields. Concretely, this presents two challenges during calibration, i) interference can rarely be detected thus resulting in incorrect calibration parameters and ii) during use, where interference can result in incorrect bearing estimations.

B. Interference Characterization

Testing device: As an example of earable, we chose eSense [23]. Contrary to other commercial earbuds, eSense permits access to the raw data streamed by the IMU and BLE chips. Since eSense was not equipped with a 3-axis magnetometer, we attached an external one (Freescale MAG3110 [35]) on top of the earbud (Figure 1a) to mimic a realistic position where the sensor might be placed. We sampled the external magnetometer at $80Hz$ using an Arduino Uno. This setup allowed us to collect data while the earbud was performing operations that could interfere with the sensor (i.e. music streaming or phone calls).

Results: We initially looked at the effect of streaming audio to the earbud on the magnetometer. This requires both active speaker movement and significant electrical current associated with the earbud circuitry. Figure 2a shows the effect on



(a) (b)

Fig. 2: Impact of audio streaming on calibrated magnetometers.

the y axis of a pre-calibrated magnetometer before, during (green shaded zone), and after playing music. Playing audio introduced an offset of 10 μT. This is a fraction of the Earth's typical magnetic field, hence the effect on the heading estimate was minimal Figure 2b. Nonetheless, there is a clear audio-induced change. Analogous trend was also evident in the magnetometer signals during voice calls. We can further observe that all of the disturbances were reversed when the audio was stopped, but that this was *not* immediate. Rather, we observed a gradual return to the pre-audio values. From this we can conclude that audio playback in an earbud affects its magnetometer signal. While we observed a relatively small change, it is enough that a one-off calibration procedure during audio playback is to be avoided. Besides, it may also hinder applications that do not use the magnetometer for heading (e.g. magnetic map matching). Furthermore, magnetic fields dissipate quickly with distance and a magnetometer soldered to the earbud board (beyond the scope of this work) might experience greater disruption and heading errors.

To better understand the source of the interference, we examined the frequency spectrum of the magnetometer when playing pure tones. We observed that tones played at (even very) low volume produced noise across all frequencies in the magnetometer. Increasing the volume resulted in additional spikes corresponding to the tone frequency: e.g. a $20Hz$ tone produces a corresponding spike at $20Hz$ (Figure 3). From this we infer that the interference comes partially from the audio playback (speaker/driver circuit produces the spikes) and partly from some non-specific part of the circuitry (giving the general noise). The slow reversal of the audio-induced changes when the audio stops is further evidence that the interference is not solely due to the sound reproduction.

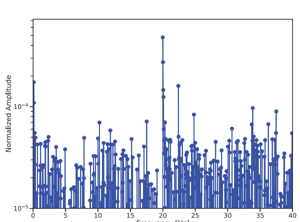


Fig. 3: `FFT` of the magnetometer traces while playing (on high volume) pure tones at $20Hz$.

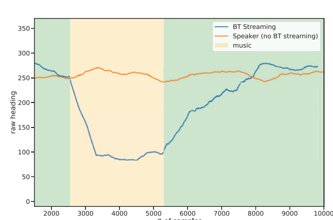


Fig. 4: Impact of BT streaming and speaker on the heading estimated by the raw magnetometer.

We sought to isolate the core circuitry from the speaker. We used two earbuds, A and B. We unsoldered and removed the speaker from A. When streaming audio to A, we observed wideband noise in the magnetometer, persisting beyond the end of the music (Figure 4). In this mode, the main operational circuitry is the BT module and the audio decoder. Since the latter is not used after the music stops, we focused on the BT module. We used a packet sniffer to establish when BT radio was in use, finding that BT packets continued to be sent for a period *after* the audio was manually stopped. This period corresponded directly to the magnetometer recovery phase. We can therefore attribute a substantial part of the interference to the Bluetooth circuitry being active. We then used two wires to connect the speaker terminals on A to the speaker terminals on B. In this way we could assess the impact caused by the speaker alone without interference introduced by the circuitry (Figure 1b). We powered both earbuds on, and then streamed music to A. This caused music to be played on B, where only the core circuitry was active (not BT or other components). The magnetometer on B exhibited a small deviation when music was played, but significantly smaller than that observed when isolating the BT hardware. While the interference induced by the BT is almost identical for different songs, it changes between music players (e.g. Spotify, Apple Music, YouTube) as they adopt different protocols.

Findings Summary: Streaming audio to a modified eSense earbud resulted in a local magnetic field that appeared as interference in the magnetometer readings. The magnitude of the interference was small, but we cannot rule out a larger effect for a magnetometer fully integrated onto the earbud. The interference came primarily from the BT circuitry being active, with a smaller component due to the speaker. Therefore, static magnetometer calibrations in an earbud should not be carried out while the BT radio is active.

IV. A Novel Calibration Algorithm

The previous section established that internal components of an earbud introduce magnetic interference during RF usage, but that it is possible to incorporate a magnetometer such that the heading estimate is minimally affected. However, this is contingent on the magnetometer being either correctly calibrated prior to the interference, or having the calibration dynamically updated, something that is impractical due to the manual calibration procedures conventionally used. In this section we describe a technique to provide calibration that does not need manual intervention and can be updated dynamically.

A. Overview

Performing user-transparent magnetometer calibration is an extremely challenging task [31]. Rather than trying to calibrate earbuds independently, we propose leveraging the user's smartphone to assist in the calibration. We assumes the phone has itself a calibrated magnetometer and can provide reliable global heading. In practice, this assumption is justified since today's phones are able to maintain a calibrated heading by fusing the array of sensors, from IMU to GPS [2]. When people interact with their smartphones, e.g. unlocking it, their head is almost certainly facing the phone. In this case the smartphone and the earbuds are aligned and should report the same bearing, if correctly calibrated (Figure 5). Our approach is to use trusted bearing of the phone to estimate the calibration parameters for the earbuds. Empirically, we find that unlocking an *iPhone* using the *FaceID* usually constitutes the perfect user head-smartphone positional relationship. This work considers unlock interactions, but the technique could be applied whenever phone and head align.

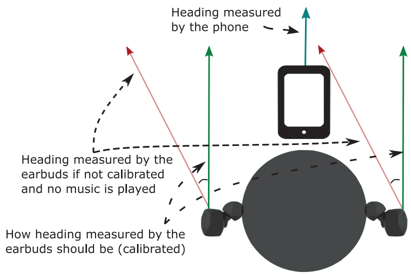


Fig. 5: Intuition behind the proposed calibration technique.

B. Calibration Approximation

We estimate the heading (the angle between the direction of facing and the Magnetic North) by doing: $heading = \operatorname{atan2}(mag_y, mag_x)\frac{180}{\pi}$. This is the standard way of estimating the heading given mag_x, mag_y – the leveled (with respect to the ground) readings of the magnetometer respectively along the x and y axis. The sign, as well as the order, of mag_y and mag_x change depending on the magnetometer orientation. Several commercial earables ensure the bud remains in a still, standard position in the users' ears. Hence, the orientation of each earbud is likely to remain unaltered. We leverage this, and the fact that most of the substantial rotations and changes of heading happen along a plane parallel to the ground, to avoid continuously accounting for tilt compensation. Having defined how we estimate the heading, we can lay the foundations of our approach. Firstly, we apply a standard sensor model for mag_x and mag_y respectively as: $mag_x = S_x(x_{earbud\_raw} - x_{earbud\_offset})$ and $mag_y = S_y(y_{earbud\_raw} - y_{earbud\_offset})$, where x_{earbud_raw} and y_{earbud_raw} are the raw, uncalibrated magnetometer readings in the xy plane, S_x and S_y are scaling factors and x_{earbud_offset} and y_{earbud_offset} are constant offsets. According to our key assumption (Figure 5), we can re-write the way we compute the heading as:

$$h_{phone} = \operatorname{atan}\left(R\frac{(y_{earbud\_raw} - y_{earbud\_offset})}{(x_{earbud\_raw} - x_{earbud\_offset})}\right)\frac{180}{\pi}, \qquad (1)$$

where h_phone is the phone's heading estimate and $R = S_y/S_x$ the scale factors ratio. Our goal is to collect multiple h_{phone} values to solve for the unknowns in this equation. We make the assumption that $R \approx 1$. We expect sensors to be factory calibrated, that should ensure this approximate relationship (more in Section VII). Small perturbations from 1 have minimal effect on the heading ($\arctan(x + \delta x) \approx \arctan(x)$ for small δx). The assumption allows us to reduce the unknowns to two (y_{earbud_offset} and x_{earbud_offset}), requiring as few as $k = 2$ phone interactions to estimate a calibration. In practice, we gather as many phone headings (k) as needed to ensure good calibration quality:

$$h_{phone\_k} = \operatorname{atan}\left(\frac{y_{earbud\_raw\_k} - y_{earbud\_offset}}{x_{earbud\_raw\_k} - x_{earbud\_offset}}\right)\frac{180}{\pi} \qquad (2)$$

By solving the over-determined systems of equations constituted by the k-th phone's headings (Equation (2)), we derive x_{earbud_offset} and y_{earbud_offset}, continuously updated at every interaction. Collecting these *reference* measurements from the phone can occur in the background, without any intervention from the user. Phone readings have to be reliable. In a smartphone the magnetometers are assumed to be calibrated, as both Android and Apple devices fuse the magnetometer readings with GPS (if available) [2]. A new reference is found every time users interact with their phone (i.e. *FaceID* unlock), provided the phone measures an undisturbed magnetic field (i.e. trustworthy heading). This is further borne out by the high average number of daily interaction people have with their mobile phones [20]. While more references are good as they should lead to a better model fit, calibrating with fewer references is valuable. A good fit depends on there being sufficiently distinct h_{phone_k} values, and we do not have control over the users behavioural patterns. Hence, we favour calibrating as soon as possible and refining the calibration with extra measurements later on, without explicit user interaction.

C. Calibration Algorithm

The overall functioning of the our calibration procedure is depicted in Figure 6. At any given time, we monitor for events (e.g. phone-pickup/unlocking) that can potentially provide measurements suitable for calibration. Given the variability of the interference and the consequent volatility characterizing magnetometer calibrations [31], rather than aiming at the perfect calibration routine, instead we strive for the best possible approximation of it. In this way, we can afford to calibrate as soon as possible and continuously monitor the status of the calibration, updating it when necessary. In addition, by using the phone's heading as a reference, there is no need to further process the heading estimated from the magnetometer data in the earbuds by calculating the magnetic declination [3].

Phone pickups and data check: Once a phone-pickup is detected, we perform a data check to ensure the smartphone's data are suitable for the calibration: first, we ensure the magnitude of the phone's readings matches, at least in the order of magnitude, that of the Earth's field at the current location. This is a standard way to check the magnetometer readings trustworthiness. Once secured we are not in a magnetic anomaly, we make sure the phone is in portrait mode. Although this is true for most of the interactions, to calibrate we can not afford using a reference off by 90° (i.e. phone in landscape mode). Lastly, we verify no sharp head movements occurred, which may result in misalignment in the phone's and earbuds' heading, or if the user moved the head but not the phone. This is key to ensure the reference heading is truthful.

Calibration Execution: If the data pass this check, we store them and wait until we have enough references to perform our calibration. At least two equations are needed to linearly solve a system of equations with two unknowns (the offsets). Hence, we need at least two reference headings (Section V). Once satisfied the number entries to perform the calibration, we execute it by applying a least squares fitting (eq. (2)). Before finalising the calibration, we perform a sanity check to ensure there was not any significant interference skewing the fit: we compare, on-the-fly, the instantaneous heading recorded by the phone and the average bearing of the earbuds. The calibration is only committed after the freshly-calibrated earbud magnetometer successfully passes this additional step.

Calibration check and update: Unfortunately, accurately modelling the *life time* of a magnetometer calibration (i.e. how long the calibration will last before the sensor readings will start being off) is extremely difficult. A number of factors can invalid the calibration of a magnetometer, such as the environment, the temperature, and the number of people in a room [29]. We avoid faulty calibration models by regularly checking the validity of the calibration. This is an inexpensive operation we carry out in two ways, depending whether there is a phone unlock event, or not. If so, we compare the bearing of the phone and that reported by the earbuds. If their difference is under a certain threshold, we assume our calibration is still valid, otherwise we drop the existing calibration. The value of the threshold depends on the application and the desired accuracy. Alternatively, if no phone-interaction is detected, we cannot assume the phone's heading is the same of the earbuds'. In this eventuality, we compare the earbud's magnetometer bearing with the earbud's IMU. If there are sharp changes in the magnetometer data, but no rotations or linear accelerations are registered by the IMU (and vice-versa), the existing calibration is likely off. Notably, we look for the magnitude of the motion recorded by the gyroscope and we compare it with the change in bearing reported by the magnetometer. Even if uncalibrated, gyroscopes are fairly precise in measuring relative motion, while they fail in tracking sustained motion.

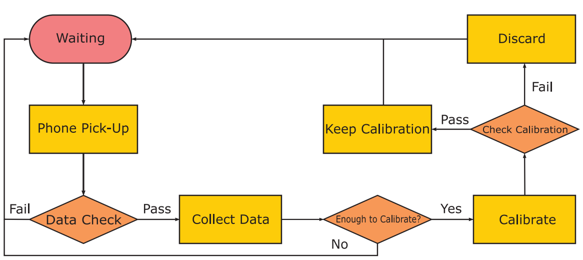
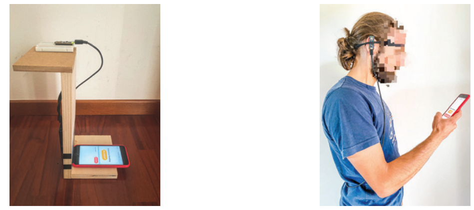


Fig. 6: Auto-Calibration procedure.

V. Calibration Evaluation

In this section we provide a detailed evaluation of the proposed calibration procedure both in controlled conditions and with an in-the-wild with a case study. We conclude by presenting some theoretical considerations on the proposed algorithm complexity, supporting this analysis with power consumption experimental results of our calibration routine.



(a) (b)

Fig. 7: Setup used to benchmark the proposed calibration technique (7a) and volunteer wearing the Arduino as if they were earbuds (7b). This is the setup used for our in-the-wild use test.

A. Micro Benchmarks

We start our evaluation with a list of micro benchmarks. Figure 7a reports the setup we used to benchmark our calibration. Once assessed how the calibration removes the interference caused by both BT and music playback using the hardware described in Section III, we focus on benchmarking the calibration technique by using the magnetometer embedded in an Arduino Nano 33BLE [1]. For reproducibility, we build a stand (Figure 7a) to simulate the positional relationship between the magnetometer and the smartphone (iPhone 8Plus).

Static Scenario: We begin assessing how different factors may affect the calibration accuracy. We look at the impact of spacing between references (i.e. $spacing(i, i+1) = |h_{phone_i} - h_{phone_{i+1}}|$), the number of references, and the number of data points fed into the calibration algorithm (i.e. time spent sampling magnetometer data at every h_{phone_i}). Contrarily to what we originally thought, we do not observe any significant gain with larger spacing of the reference headings. Interestingly, we noticed that we do not require more than $2-3$ reference headings to achieve errors on average smaller than 2°, without getting into the territory of diminishing returns – i.e. paying more in terms of energy consumption Section V-C) without benefiting a substantial accuracy boost. Lastly, we assessed how very a few data points (sampled over 0.5s) are already sufficient for our calibration to reach its highest accuracy.

Dynamic Case: Consider a person rotating their head: their bearing would change of an angle equal to that of head rotation. A calibration must remain valid for every direction faced during the movement. We make sure magnetometer and ground truth are always in the same positional relationship evaluating that using the stand in Figure 7a. Concretely, we look at whether the calibrated heading diverge from the ground truth whenever there is motion. For all the experiments, we use only two references, collected over windows of $1/20$s.

Practically, we begin by looking at small movements of a few degrees (both positive and negative) and outline our findings in Figure 8a. The uncalibrated error is significant, averaging 30° over the duration of our experiment (mean rolling error in Figure 8b), being at best is 10° off ground truth and at worst 45°. Besides, the magnitude of the motions seems not to reflect those of the real movements. Conversely, we can observe that the calibrated heading trace is very stable and close to the ground truth heading averaging an error smaller than 5° at any given time. We repeated that for larger movements (Figure 9a), expecting large movements to induce greater errors in the heading. The uncalibrated heading trace is at one point in excess of over 90°, providing unusable data for most applications (Figure 9b); while the calibrated heading trace closely follows the ground truth, with an overall mean error of just a few degrees, always *smaller* than 5°.

B. In-The-Wild Example Case-Study: Navigation

We believe, in the case of navigation, earables could be better suited than smartphones. The rationale behind that lays in the ability of earbuds to track head movements: a desirable feature at complex intersections. Further, earables can provide extra robustness by recording 2 independent measurements of the same heading. Notice that an end-to-end earable-based navigation system is out of the scope of this work, instead, we use this toy example to evaluate, in-the-wild, our calibration.

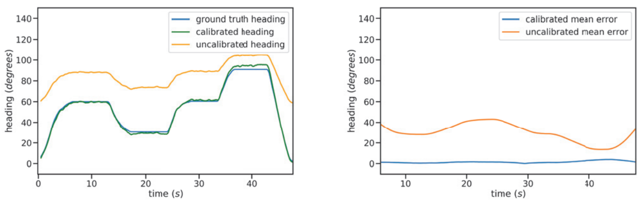
Assessing the goodness of our calibration in-the-wild, we deal with both rotational movements and human motion (i.e. linear acceleration). Concretely, a user (ethics approval grated by the departmental ethics board) *wore* two Arduino (with a build-in magnetometer) as earbuds (Figure 7b), while holding the phone in their hands.The volunteer was told to walk as desired in two distinct locations the first one being indoor (in a house) while the latter outdoor (over a block). We had no control over the potential source of interference in the environment. This experiment showcases how the proposed calibration is capable of enabling earable-based in-the-wild navigation, without constraining nor bounding the user, both indoor and outdoor. Figure 10 reports our results when estimating the heading with calibrated and uncalibrated magnetometer traces and their average errors. Through our calibration, we are able to achieve accuracy up to few degrees (below 3° for most of the time) for the whole duration of the experiment. Notably, considering the 4 lanes intersection example, our $\approx 3°$ error on the heading would lead to a $\approx \pm 0.26m$ error. We believe $\pm 0.26m$ is an acceptable tolerance for pedestrian navigation.

C. Computational and Power Consumption Considerations

A calibration has to be accurate, yet computationally inexpensive; however, all of the established calibration techniques require to perform a regression to fit a model based on the observations received. Especially, online schemes do the fitting every time the calibration is performed; in our case, as dictated from Figure 6, we check if a calibration is needed at every suitable phone pickup. Normally, such techniques use a form of least-squares fitting having a formal complexity of $\mathcal{O}(c^2 n)$, where c the number of features and n the number of vectors in $\mathbb{R}^c$ used when performing the fitting. However, in our case, we require very a few vectors $n < 10$, all in $\mathbb{R}^2$, representing the readings of the magnetometer on the xy plane. Hence, the amortised complexity of our procedure ends up being even more affordable in practice, even if we have to run the calibration procedure often. We evaluated this claim with a power consumption experiment with a Raspberry Pi Zero.Set side by side to the overhead of radio communications over idle ($\approx 1.7\%$ TX and $\approx 2.4\%$ RX), our scheme imposes a comparable low overhead, only consuming an additional $\approx 2.9\%$ over idle, with a total of $456.38mW$ (idle: $442.58mW$).

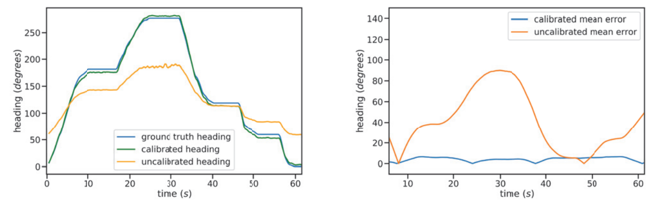
VI. Related Work

Sensing with Earables The applicability of earables for health condition screening has been investigated many [25] looked at their potential in monitoring energy expenditure and heart rate. [14] introduced the idea of an ear-worn multi-modal platform to sense brain, cardiac, and respiratory functions. More recently, [9] proposed a PPG-equipped ear-piece to monitor blood pressure. Ear-worn devices have also been researched to monitor sleep stages [28] or detect eating activities [5]–[7].
Magnetometer Calibration Although in-ear magnetometers are mostly affected by a static interfering component induced by permanent magnets in the buds' case, they are also impacted by a dynamic component caused by RF communications and audio playback. Isolating the magnetometer with special materials, preventing magnetic disturbance, is not a viable option: without considering the cost, it would likely isolate the magnetometer from the magnetic field we wanted to measure in a first instance [27]. Similarly, filtering approaches [40] do not work for perturbations generated by RF circuitry [27]. Besides, increasing the air gap between sources of interference and magnetometer [12], [27] is not feasible: assuming these can be modeled as magnetic dipoles, the strength of the magnetic field they generate decreases with r^3, where r is the radius of a sphere with the magnetic dipole as center [22]. Yet, this does not consider the design constraints and miniaturization trends of earables, and does not account for the external magnetic disturbances [16], [38]. Likewise, the same argument holds for factory calibrations [16]. Sensor calibration is a well studied topic with a rich body of literature. However, only a few works specifically tackle calibration for mobile devices [15], and, to the best of our knowledge, none investigates calibration strategies specifically for earables – affected by variable interference. The aim of a magnetometer calibration is to find some parameters to only measure the Earth's magnetic field [37]. Thus, the sensor can be used, for example, as heading source. A magnetometer calibration can be either static or dynamic [16]; attitude dependent or not [38]. Static calibrations are often done with the aid of specialized equipment (e.g. a proton magnetometer [30], a robot arm [33]), or by manual direction placement [10]. Historically, the most common approach is what known as *compass swinging* [8]. This only works for 2D-magnetometers, and requires the user to be instructed to rotate the compass in specific orientations [41]. Dynamic calibrations, more practical for mobile devices, are usually based on additional information from the system (e.g. IMU or GPS). Many are iterative and carried out at run-time, often trading accuracy for adaptation. Examples are ellipsoid fitting [32], Kalman-filter-based iterative algorithms [16], and stochastic optimization approaches [38], with the most famous being the *figure 8* calibration: the user has to move the magnetometer along an 8-shaped trajectory, to collect enough data to run an ellipsoid fitting algorithm. How-



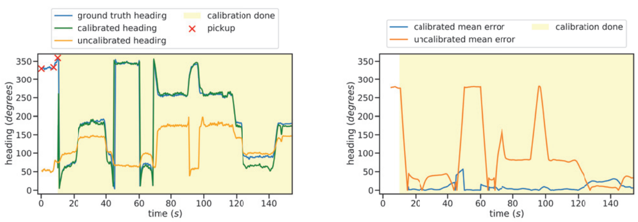
(a) (b)

Fig. 8: Heading estimation (8a) and mean errors (8b) for small angles.



(a) (b)

Fig. 9: Heading estimation (9a) and mean errors (9b) for large angles.



(a) (b)

Fig. 10: In-the-wild heading estimation (10a) and mean errors (10b).

ever, these result cumbersome for the user, and often error-prone [15]. Hence, we present a completely user-transparent, adaptive, magnetometer calibration, which specifically targets earables, without requiring any specialized equipment, other that the user's phone, natural companion of every earbuds.

VII. Final Remarks

This work proposes a method to calibrate in-ear magnetometers in a user-transparent and efficient way and shows the feasibility of having magnetometers in earables. We conclude by discussing the limitations of our calibration technique.

Presence of a Phone: Our technique requires the earbuds user to carry a phone. Phones are ubiquitous and carried almost everywhere. Moreover today's earables are not stand-alone, requiring a companion device they are connected to – usually a phone. We leave as a future work devising a user-transparent magnetometer calibration for a future stand-alone earbud.

Tilt Compensation: Magnetometers usually provide a 2 degrees of freedom orientation parallel to the ground [36]. If the sensor is not leveled, tilt compensations is needed. Common strategies usually exploit gravity measurements from an accelerometer to compute the sensor tilt and map it back in the correct reference frame [36]. While earables experience less significant rotations than other devices by virtue of their attachment to the head, tilt compensation is still required. In our work, although not explicitly stated, we always make sure the x and y we are using are leveled with respect to the xy plane. Further, in our system, head dips of $\alpha°$ can be modeled as rotations about the y axis, resulting in $(mag_x \cos(\alpha), mag_y)$. $\cos(\alpha) \approx 1$ for small values of α, we assume little head dips (like those when normally interacting with a phone) only marginally affect our system.

Scaling Factors: Our approach does not estimate the sensor scale factors. Instead, we assume the ratio of scale factors would be 1. This allowed us to reduce the number of distinct headings to estimate the calibration, favouring usability and low complexity. Our justifications for the assumption are: (i) scale factors are used to address soft iron distortions which, although having a non-negligible effect when looking at the intensity of the Earth's magnetic field, are substantially less significant than hard iron distortion when looking at the heading; (ii) we expect earbud manufacturers to perform a factory calibration of their sensors, which would build in compensation for any soft or hard iron biases internal to the earbuds themselves. Calibration parameters do of course change with environmental factors, necessitating in-field calibrations. However, the changes are likely perturbations around the factory calibration. Therefore the *ratio* of observed scale factors – on which the heading computation depends – would be expected to be approximately 1. Notably, this was the case experimentally for all of the magnetometers we tested. Perturbations to this ratio have a limited effect on the estimated heading, so our assumption has minimal effect on the error, whilst reducing the complexity of our calibration procedure; (iii) assuming $R \approx 1$, we do not estimate the true scale factors. While this have minimal effect on the heading accuracy, it does mean we do not obtain a reliable estimate of the overall magnetic field magnitude, implying we cannot use the field magnitude to discard readings when in a magnetic anomaly. Nonetheless, we are able to accurately compute the earbuds heading, key enabler to many applications, starting from IMU calibration. In the future we hope to extend our model to scale factors, too. In the meantime, we note that anomalies may be detected through large rotation discrepancies between the gyroscope and the magnetometer.

Interactions with the Phone: As a proof of concept, in this work, we rely on *FaceID* unlocks to ensure the positional relationship phone-earbuds is what we require for our calibration to work properly (Figure 5). However, unlocks are not the only events leading to this specific positional relationship. By reliably detecting more of such occurrences (even during a single interaction) would be possible to further increase the granularity of our measurements, which might be suitable for some application which need more frequent references.

Acknowledgments

We would like to thank the anonymous reviewers and our shepherd Dr. Archan Misra for his valuable comments and guidance.

References

[1] Arduino nano 33 ble imu. https://www.st.com/resource/en/datasheet/lsm9ds1.pdf. Accessed: 2020-05-07.

[2] Getting Heading and Course Information apple developer. https://developer.apple.com/documentation/corelocation/getting_heading_and_course_information. Accessed: 2020-05-08.

[3] Magnetic declination. https://en.wikipedia.org/wiki/Magnetic_declination. Accessed: 2020-06-12.

[4] Statista forecast unit shipments of wearable devices worldwide from 2017 to 2019 and in 2022 (in million units), by category. https://www.statista.com/statistics/385658/electronic-wearable-fitness-devices-worldwide-shipments/. Accessed: 2020-05-05.

[5] O. Amft, M. Stäger, P. Lukowicz, and G. Tröster. Analysis of chewing sounds for dietary monitoring. In *International Conference on Ubiquitous Computing*, pages 56–72. Springer, 2005.

[6] A. Bedri, R. Li, M. Haynes, R. P. Kosaraju, I. Grover, T. Prioleau, M. Y. Beh, M. Goel, T. Starner, and G. Abowd. Earbit: using wearable sensors to detect eating episodes in unconstrained environments. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, 1(3):1–20, 2017.

[7] S. Bi, T. Wang, N. Tobias, J. Nordrum, S. Wang, G. Halvorsen, S. Sen, R. Peterson, K. Odame, K. Caine, et al. Auracle: Detecting eating episodes with an ear-mounted sensor. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(3):1–27, 2018.

[8] N. Bowditch. *The American practical navigator: an epitome of navigation/vol. 1*. 2018.

[9] N. Bui, N. Pham, J. J. Barnitz, Z. Zou, P. Nguyen, H. Truong, T. Kim, N. Farrow, A. Nguyen, J. Xiao, R. Deterding, T. Dinh, and T. Vu. Ebp: A wearable system for frequent and comfortable blood pressure monitoring from user's ear. In *The 25th Annual International Conference on Mobile Computing and Networking*, MobiCom '19, New York, NY, USA, 2019. Association for Computing Machinery.

[10] F. Camps, S. Harasse, and A. Monin. Numerical calibration for 3-axis accelerometers and magnetometers. In *2009 IEEE International Conference on Electro/Information Technology*, pages 217–221. IEEE, 2009.

[11] C. Cheng, X. Huo, and M. Ghovanloo. Towards a magnetic localization system for 3-d tracking of tongue movements in speech-language therapy. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 563–566. IEEE, 2009.

[12] A. Ferlini, A. Montanari, C. Mascolo, and R. Harle. Head motion tracking through in-ear wearables. In *Proceedings of the 1st International Workshop on Earable Computing*, EarComp'19, page 8–13, New York, NY, USA, 2019. Association for Computing Machinery.

[13] M. S. George, S. H. Lisanby, and H. A. Sackeim. Transcranial magnetic stimulation: applications in neuropsychiatry. *Archives of General Psychiatry*, 56(4):300–311, 1999.

[14] V. Goverdovsky, W. von Rosenberg, T. Nakamura, D. Looney, D. J. Sharp, C. Papavassiliou, M. J. Morrell, and D. P. Mandic. Hearables: Multimodal physiological in-ear sensing. *Scientific reports*, 7(1):1–10, 2017.

[15] A. Grammenos, C. Mascolo, and J. Crowcroft. You are sensing, but are you biased? a user unaided sensor calibration approach for mobile sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(1):1–26, 2018.

[16] K. Han, H. Han, Z. Wang, and F. Xu. Extended kalman filter-based gyroscope-aided magnetometer calibration for consumer electronic devices. *IEEE Sensors Journal*, 17(1):63–71, 2016.

[17] C. N. Hansen. *Understanding active noise cancellation*. CRC Press, 2002.

[18] A. Härmä, J. Jakka, M. Tikander, M. Karjalainen, T. Lokki, J. Hiipakka, and G. Lorho. Augmented reality audio for mobile and wearable appliances. *Journal of the Audio Engineering Society*, 52(6):618–639, 2004.

[19] J. Haverinen and A. Kemppainen. Global indoor self-localization based on the ambient magnetic field. *Robotics and Autonomous Systems*, 57(10):1028–1035, 2009.

[20] D. Hintze, P. Hintze, R. D. Findling, and R. Mayrhofer. A large-scale, long-term analysis of mobile device usage characteristics. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(2):1–21, 2017.

[21] H. Huang, H. Chen, and S. Lin. Magtrack: Enabling safe driving monitoring with wearable magnetics. In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, pages 326–339, 2019.

[22] D. S. Jones. *The theory of electromagnetism*. Elsevier, 2013.

[23] F. Kawsar, C. Min, A. Mathur, and A. Montanari. Earables for personal-scale behavior analytics. *IEEE Pervasive Computing*, 17(3):83–89, 2018.

[24] I. S. Khalil, A. Adel, D. Mahdy, M. M. Micheal, M. Mansour, N. Hamdi, and S. Misra. Magnetic localization and control of helical robots for clearing superficial blood clots. *APL bioengineering*, 3(2):026104, 2019.

[25] S. F. LeBoeuf, M. E. Aumer, W. E. Kraus, J. L. Johnson, and B. Duscha. Earbud-based sensor for the assessment of energy expenditure, heart rate, and vo2max. *Medicine and science in sports and exercise*, 46(5):1046, 2014.

[26] S. Liebich, J.-G. Richter, J. Fabry, C. Durand, J. Fels, and P. Jax. Direction-of-arrival dependency of active noise cancellation headphones. In *ASME 2018 Noise Control and Acoustics Division Session presented at INTERNOISE 2018*. American Society of Mechanical Engineers Digital Collection, 2018.

[27] J. Metge, R. Mégret, A. Giremus, Y. Berthoumieu, and T. Décamps. Calibration of an inertial-magnetic measurement unit without external equipment, in the presence of dynamic magnetic disturbances. *Measurement Science and Technology*, 25(12):125106, 2014.

[28] A. Nguyen, R. Alqurashi, Z. Raghebi, F. Banaei-kashani, A. C. Halbower, and T. Vu. A lightweight and inexpensive in-ear sensing system for automatic whole-night sleep stage monitoring. In *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM*, SenSys '16, page 230–244, New York, NY, USA, 2016. Association for Computing Machinery.

[29] T. Ozyagcilar. Calibrating an ecompass in the presence of hard and soft-iron interference. *Freescale Semiconductor Ltd*, pages 1–17, 2012.

[30] H. Pang, S. Luo, M. Pan, Q. Zhang, and R. Xie. Calibration of three-axis magnetometer diversionary error based on equipment and lms adaptive algorithm. In *Sixth International Symposium on Precision Engineering Measurements and Instrumentation*, volume 7544, page 75445P. International Society for Optics and Photonics, 2010.

[31] S. Poddar, V. Kumar, and A. Kumar. A comprehensive overview of inertial sensor calibration techniques. *Journal of Dynamic Systems, Measurement, and Control*, 139(1):011006, 2017.

[32] T. Pylvänäinen. Automatic and adaptive calibration of 3d field sensors. *Applied Mathematical Modelling*, 32(4):575–587, 2008.

[33] E. L. Renk, M. Rizzo, W. Collins, F. Lee, and D. S. Bernstein. Calibrating a triaxial accelerometer-magnetometer-using robotic actuation for sensor reorientation during data collection. *IEEE Control Systems Magazine*, 25(6):86–95, 2005.

[34] F. Ribeiro, D. Florencio, P. A. Chou, and Z. Zhang. Auditory augmented reality: Object sonification for the visually impaired. In *2012 IEEE 14th international workshop on multimedia signal processing (MMSP)*, pages 319–324. IEEE, 2012.

[35] F. Semiconductor. Xtrinsic mag3110 three-axis, digital magnetometer. *MAG3110 Datasheet*, 2013.

[36] S. Shen, M. Gowda, and R. Roy Choudhury. Closing the gaps in inertial motion tracking. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*, pages 429–444, 2018.

[37] R. Sodhi, J. Prunty, G. Hsu, and B. Oh. Automatic calibration of a three-axis magnetic compass, Nov. 18 2008. US Patent 7,451,549.

[38] M. Tahir, A. Moazzam, and K. Ali. A stochastic optimization approach to magnetometer calibration with gradient estimates using simultaneous perturbations. *IEEE Transactions on Instrumentation and Measurement*, 68(10):4152–4161, 2018.

[39] I. Umay, B. Fidan, and B. Barshan. Localization and tracking of implantable biomedical sensors. *Sensors*, 17(3):583, 2017.

[40] R. Versteeg, M. McKay, M. Anderson, R. Johnson, B. Selfridge, and J. Bennett. Feasibility study for an autonomous uav-magnetometer system. Technical report, IDAHO NATIONAL LAB IDAHO FALLS, 2007.

[41] A. Wahdan, J. Georgy, W. F. Abdelfatah, and A. Noureldin. Magnetometer calibration for portable navigation devices in vehicles using a fast and autonomous technique. *IEEE Transactions on Intelligent Transportation Systems*, 15(5):2347–2352, 2014.

[42] Z. Yang, Y.-L. Wei, S. Shen, and R. R. Choudhury. Ear-ar: indoor acoustic augmented reality on earphones. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, pages 1–14, 2020.