Accepted to the European Navigation Converence 2018 (ENC 2018).


# Modulation and Signal-Processing Tradeoffs for Reverse-GPS Wildlife Localization Systems

Andrey Leshchenko and Sivan Toledo
Blavatnik School of Computer Science, Tel-Aviv University

***Abstract*—Reverse-GPS wildlife localization systems are emerging as a key technology for regional high-throughput wildlife tracking. Two such systems have been designed, implemented, and deployed (in six sites on three continents). Both of the existing systems suffer from limitations due to the modulation that is used by transmitters, which are attached to wild animals, and due to the detection and estimation algorithms that they use to detect transmissions and estimate their arrival times. This paper investigates key tradeoffs associated with three different modulation schemes that wildlife tags can plausibly use. The factors that we investigate include the ability to detect weak signals from distant tags, the ability to accurately estimate the time-of-arrival at a given SNR, and the computational cost of these detection and estimation algorithms. Our key contributions are (1) evidence that BPSK modulation is superior in essentially all relevant metrics, except perhaps chip availability, to FSK and OOK; (2) evidence that OOK is a second-best choice and its main drawback is poor performance under interference from other tags; (3) algorithms to inexpensively search the frequency-delay space at moderate and high SNRs. We also report on implementation efforts designed to integrate robust processing of BPSK tags into a wildlife tracking system.**

## I. Introduction

Reverse-GPS wildlife tracking systems are emerging as a key tool to track multiple wild animals at high temporal and spatial resolutions at the regional scale [1], [2], [3]. These systems track animals by attaching to the animals transmitting radio tags. Tags transmit an identifying packet, which we refer to as a ping, at fixed intervals of 1–8s. A system of fixed receivers estimates the arrival times of a ping to the receivers and uses these time-of-arrival (TOA) estimates to estimate the location of the tag. The use of fixed terrestrial receivers and miniature transmitting tags limit the range of such systems to several kilometers to tens of kilometers. These systems have been so far successfully deployed six times. The system of MacCurdy and others has been deployed in the Dutch Wadden Sea [4], [5] and in Mauritania [5]. The system that we have been working on, ATLAS, has been deployed in Israel's Hula Valley [2], [3], in the Dutch Wadden Sea, in England, and in Germany. A third system has been implemented, but does not appear to have been deployed [6]. These systems appear to have largely replaced earlier direction-of-arrival regional wildlife localization systems [7], [8]. TOA emitter localization systems are also used to track wild fish, but using ultrasonic transmitters [9], [10] rather than radio transmitters; the mathematical principles of these systems are similar to those of radio-frequency (RF) systems, but some of the factors that constrain the transmitters are different.

Size and regulation constraints on the transmitting tags used by these systems limit transmit power and ping duration. Therefore, typical signal-to-noise ratios (SNRs) are negative. Receivers detect the pings and estimate their arrival time by correlating the incoming signal (or a signal demodulated from the incoming signal) with an approximate replica of the transmitted packet, which consists of a long Gold code or a long pseudo-random code. Certain constraints on the modulation, detection, and TOA estimation techniques in these systems have limited their performance and robustness. This paper investigate techniques to remedy these weaknesses.

In particular, the system of MacCurdy and others suffered from the use of an incoherent receiver to receive phase-modulated transmissions emitted by tags with poor frequency stability. MacCurdy's receiver used complex baseband correlation to detect and estimate the TOA of incoming transmissions; this technique leads to complete failure to detect the transmissions once their frequency deviates even slightly from the nominal frequency [1], [11]. ATLAS uses frequency modulation and an incoherent receiver. The ATLAS receiver demodulates the frequency-shift-keyed (FSK) signal and correlates the real demodulated signal with a replica of the packet [2]. This makes the receiver relatively immune to transmitted carrier frequency errors, but reduces the sensitivity of the receiver and degrades the accuracy of TOA estimates (relative to complex base-band correlation receivers). The FSK demodulator that ATLAS receivers use also suffers badly from inter-tag interference, thereby limiting the ability to track tags that transmit simultaneously. Krüger's system [6] uses on-off keying (OOK), which we show is less power efficient than phase modulation; this system also uses a carrier-frequency estimation method that is susceptible to interference.

Our aim in this paper is to thoroughly investigate alternative modulation schemes that miniature transmitting tags can plausibly produce and to investigate techniques to detect these transmissions and to estimate their TOAs. We pay particular attention to the computational complexity of detection and TOA estimation, since the power envelope of receivers in such systems is often constrained. We exploit the distinction between the two modes in which a receiver can be relative to a particular tag. In acquisition (or searching) mode, receiver tries to detect pings from a tag with no prior knowledge of either the ping phase or the frequency error in the transmitter. In tracking mode, the receiver knows the ping phase (and hence when to expect the next ping from the tag) and has an estimate of the frequency error in the near past; both are derived from

successful prior detections.

The rest of this paper is organized as follows. Section II explains the fundamental factors that drive the design of reverse-GPS systems and the factors that constrain these designs. Section III describes our model, algorithms to detect transmissions from tags and to estimate their arrival times, and the performance and behavior of these algorithms under both white noise and under interference from other tracking tags. Section IV describe our on-going efforts to implement better tags and better detection and estimation algorithms for ATLAS, and Section V presents our conclusion from this research.

## II. Driving and Constraining Factors

System design for TOA localization systems is driven by the fundamental Cramer-Rao Lower Bound (CRLB). Given a transmitted signal $s$ of duration $T$ seconds and bandwidth $B$ Hertz through a channel that delays it, attenuates it and adds to it Gaussian white noise, the variance of a TOA estimate $\tau$ (an estimate of the delay in the channel) is limited by

$$\text{var}(\tau) \geq \left(\text{SNR} \cdot T \cdot \bar{F}^2\right)^{-1} ,$$

where SNR is the signal-to-noise ratio of the received signal and where $\bar{F}^2$ is the mean squared bandwidth of $s$ [12, pages 53 to 56]. The mean squared bandwidth is defined as

$$\bar{F}^2 = \frac{\int (2\pi F)^2 |S(F)|^2 dF}{\int |S(F)|^2 dF} ,$$

where $S(F)$ is the Fourier transform of $s$.

Achieving or approaching the bound requires knowledge of $s$ in the receiver; incomplete information about $s$ may degrade the estimate. This knowledge can be in the form of a specification of $s$ (e.g., carrier frequency, symbol rate, the symbol sequence, etc) from which the actual signal may deviate to some extent, or from a replica of the actual transmitted signal.

The CRLB indicates what the designer of a TOA localization system can do to achieve high accuracy:

1) Increase the power of the transmitted signal $s$, to achieve high SNR.
2) Increase the duration of $s$.
3) Increase the bandwidth of $s$ (in a way that maximizes $\bar{F}^2$).
4) Minimize uncertainties regarding what is transmitted exactly, especially uncertainties that the receiver cannot resolve.

When the system's own transmissions occupy the channel a significant fraction of the time, increasing the duration of transmissions may reduce the SNR at receivers, especially if transmission times of different tags are not coordinated.

In wildlife localization systems the ability to take the measures listed above is severely constrained, due to the following factors:

1) Increasing transmit power shortens the life-span of battery-operated tags, may require heavier and larger power-supply and power-amplifier components. Furthermore, power is usually constrained by relevant emission regulations.
2) Increasing the duration of transmissions shortens the life-span of batteries and increases inter-tag interference.
3) Increasing the bandwidth of transmissions requires a wideband transmitter, which limits the choice of components for tags. Furthermore, bandwidth is also limited by emission regulation, which may limit either bandwidth or the span of available frequency bands.
4) In wildlife localization systems, information about the transmitted signal $s$ are based on its specification (nominal parameters). As explained below, we cannot expect to have access to a high-quality replica of the signal that was actually transmitted. Reducing uncertainty about the transmitted signal may require modulators that are not available on otherwise attractive transmitter components and frequency control mechanisms that increase size and/or power consumption (e.g., temperature compensated oscillators).

Finally, addressing uncertainties about $s$ in the receiver may require a TOA estimator that is computationally much more expensive than an estimator for a known $s$. This increases the cost of receiver and limits the applicability of the system to off-grid sites (where receivers are powered by solar/wind power), or alternatively reduces the TOA estimate rate within a given receiver power envelope.

The next subsections describe these constraining factors more fully.

### A. Regulatory Constraints

ATLAS tags are designed to comply with ETSI standard EN 300 220 for non-specific short range devices [13]. The tags operate on the 434 MHz band; the standard allows, under operational band H, transmissions of up to 10mW ERP with a duty cycle of 10% between 433.05 and 434.79 MHz; transmitters are not required to ascertain that the channel is not occupied and the transmission may span the entire 1.74 MHz band. The same standard allows for higher power (25mW) and wider bandwidth (3 MHz) at 865–868 MHz, but at a lower duty cycle (1%) and low power density. The standard allows for significantly higher power (500mW) only over a much narrower 0.25 MHz sub-band. Wildlife tags may operate under other regulatory regimes, but these rules are typical for unlicensed equipment that does not participate in sophisticated spectrum-sharing or interference-avoidance protocols.

We note that VHF bands in which wildlife telemetry is allowed (e.g., 150 MHz in Europe and 166 MHz in the US) are too narrow for TOA localization systems and that frequencies above 1 GHz suffer from increased path loss that restricts systems to short distances.

To summarize, regulation at sub-1 GHz frequencies typically allow transmit power of 10-25mW and bandwidths of 1.7–3 MHz.

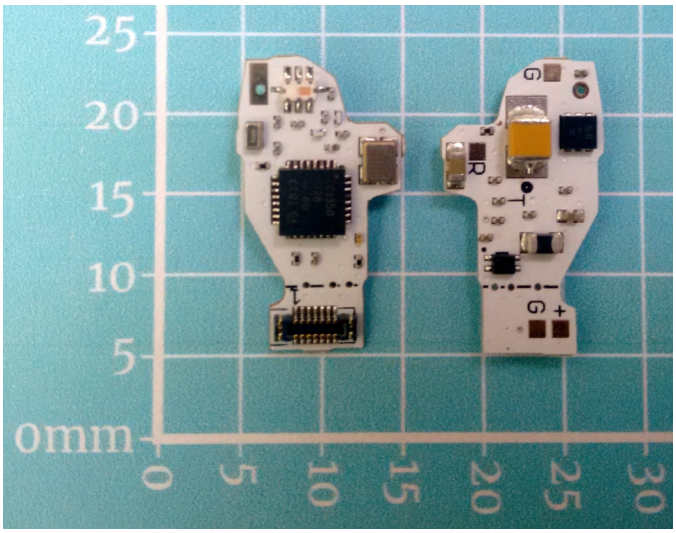


Fig. 1. CC1310-Based ATLAS tags for the 434MHz band. The boards weigh approximately 0.3g, leading to total tag mass of 0.8g and up.

### B. Transmitter Constraints

Size and weight constraints favor the use of single-chip integrated transmitters or transceivers. A chip that also integrates a microcontroller (MCU) is often preferable to a radio-only chip, to further reduce the total size of tags. First generation ATLAS tags used two 4mm-by-4mm chips, a CC1101 integrated transceiver and a separate MSP430 MCU [14]. The second generation of ATLAS tags switched to a single 4mm-by-4mm chip, a CC1310 or CC1350, which combine a radio with an ARM MCU. One CC1310-based tag design is shown in Figure 1. The CC13x0 devices support FSK and GFSK (a band-limited version of FSK) modulation at data rates of up to 4 Mb/s, as well as on-off-keyking (OOK; maximum data rate for OOK is not clearly specified). The (G)FSK modulator in CC13x0 modulates the RF synthesizer directly; the signal is *not* produced digitally at baseband and upconverted.

Chips from other vendors have similar architectures and are subject to similar limitations. For example, the EZRadioPRO family of sub-1 GHz integrated transmitters and transceivers support (G)FSK and OOK modulation at up to 1 Mb/s. We note that in many cases, OOK can be produced by modulating power to a transmitter even if the transmitter does not support OOK or does not support it at the required bitrate [6, Section 5.1.2].

The dominance of (G)FSK in these integrated transceivers, especially at high data rates, is not an accident. These devices have relatively simple incoherent receivers, and in such receivers, phase and/or amplitude modulation schemes do not improve the bit-error rate relative to frequency modulation schemes.

However, binary phase-shift keying (BPSK or PSK) is not difficult to produce directly at RF; essentially, all that is required is an inverter and optionally an amplitude shaper (to limit the bandwidth). Hence, some integrated transmitters and transceivers can produce high data-rate BPSK. One example is the fairly old AX5031 4mm-by-4mm transmitter, which can produce BPSK at up to 2 Mb/s (the transceiver part from the same family, AX5051, only supports 0.6 Mb/s data rates). A newer example is the EFR32 family of RF MCUs, which support BPSK transmissions (only), but which are physically larger than the CC13x0, at 5mm-by-5mm.

The main implication of producing FSK by direct modulation of the synthesizer is that the trajectory of the transmitted signal in phase space is not carefully controlled and therefore not necessarily repeatable. That is, the carrier phase may change arbitrarily between symbols. This essentially rules out the use of complex base-band correlation to detect transmissions and to estimate their TOAs, as we explain below.

Another characteristic of miniature transmitters and transceivers is that their carrier frequency is not very stable; the carrier (as well as the symbol clock) are derived from a crystal oscillator that is usually not temperature compensated and that is subject to dramatic temperature changes when the tag is attached to a wild animal.

### C. Replica Constraints

In TOA wildlife localization system, information about the transmitted signal $s$ is derived from its specification: carrier frequency, symbol rate, modulation parameters, and the symbol sequence. As explained in Section II-B, miniature transmitters may produce signals that deviate significantly from their specification.

Other types of TOA localization systems estimate arrival times using stored replicas of the actual transmitted signal, not from its specification. This eliminates estimation inefficiency caused by discrepancies between the specification and the actual signal. This technique is used in radar systems, in which the transmitter and receiver are co-located; a small fraction of the transmitter's energy is used to create a representation of what was transmitted.

A related technique is used in passive emitter localization systems designed to localize unknown but strong transmissions: the received signals from two receivers are correlated to estimate the time-difference-of-arrival (TDOA) of a signal between the two receivers. Here knowledge of $s$ is embodied by the signal received at another receiver. However, wildlife localization systems cannot use this technique, because the signals emitted by tags are weak and usually received together with strong interferers. In environments with strong interferers, the TDOAs are likely to be those of an interferer, not those of the target.

### D. Receiver Constraints

ATLAS receivers consist of a sampling radios (typically USRP N200) that feeds a computer running Linux or Windows, typically a high-end desktop class computer (quad-core or dual-core Intel i7 processor). Computers in this performance class process RF samples (at about 8 Ms/s) to detect FSK transmissions and to estimate TOAs at roughtly real-time rate. That is, they take about 10ms to process 10ms worth of samples. This performance is achieved by code in C

that uses FFT to correlate and that uses a high-performance multicore FFT library (FFTW [15]). This is a constraining factor, especially when a receiver must acquire the ping phase of many tags. Receivers can use weaker computers, such as Raspberry Pis, but these can only handle a small number of tags.

## III. Detection and TOA Estimation Techniques

This section presents and analyzes techniques to estimate the TOA of packets emitted by wildlife tags.

We denote the transmitted signal by $s(t)$. In practice, $s$ is a real continuous signal, but it is convenient to treat it as a complex (quadrature) signal defined as the Hilbert transform of the transmitted signal. We assume that $s$ is a physical realization of a nominal signal $p$ known to the receiver. The nominal and transmitted signals may differ in two key ways that we explain below. We denote the received signal by $r$, where $r(t+\tau) = \alpha s(t) + \epsilon(t+\tau)$, where $\alpha$ is an attenuation factor and $\epsilon$ represents band-limited noise and interference.

In RF localization systems, the carrier frequency is often much higher than the bandwidth of $p$ so $r$ is down-converted to make sampling and processing feasible or inexpensive. We implicitly assume that $r$ has been down converted to baseband or to a low intermediate frequency and that $p$ is represented in this low frequency.

We refer to the discrete sampled versions of $p$ and $r$ as $p[n]$ and $r[m]$. We assume that the two sampling clocks have the same frequency but that they differ arbitrarily in phase. The assumption that the two frequencies are identical is satisfied by using an accurate (GPS-locked) clock to sample $r$.

The signal $p$ consists of a sinusoidal carrier modulated by $N$ chips (symbols) of a pseudo-random sequence $g[i] = \pm 1$ (possibly a Gold code, but at the lengths that we consider, good pseudo-random codes behaves similarly to Gold codes). We denote the carrier frequency by $f$ (and we denote $\omega = 2\pi f$) and the chip period by $\theta$ (the chip clock frequency is $1/\theta$). We denote by $g(t) = g(\lfloor t/\theta \rfloor)$ a continuous version of $g[i]$.

We consider in this paper three modulation schemes that miniature integrated transmitters and transceivers can produce: FSK, BPSK, and OOK. In BPSK, the nominal signal is

$$p(t) = p_{\text{BPSK}}(t) = Ae^{i(\phi+\omega t)}g(t) \ .$$

In OOK, the nominal signal is

$$p(t) = p_{\text{OOK}}(t) = Ae^{i(\phi+\omega t)}\frac{1}{2}\left(g(t)+1\right) \ .$$

In FSK, we define $\omega(t) = 2\pi(f + g(t)v)$ where $v$ is the deviation; the nominal signal is

$$p(t) = p_{\text{FSK}}(t) = Ae^{i(\phi+\int_0^t \omega(\psi)d\psi)} \ .$$

We assume that $p$ is zero outside $[0, T]$ where $T = n\theta$.

The transmitted signal $s$ may differ from the nominal signal $p$ in two main ways. First, $s$ may be a stretched or compressed version of $p$, $s(t) = p((1+\delta)t)$ for some small $\delta$ with absolute value of several parts per million (ppm). The stretching/compression is a result of an inaccurate reference oscillator from which both the carrier frequency and the symbol clock are derived. Stretching and compression multiply the carrier frequency and the symbol clock by the same factor, but down-conversion changes that, as the next section shows.

Second, the rate of transition between symbols or chips in $p$ and $s$ may differ and may not be known precisely. The expressions given above for the modulated signals produce very wideband signals, because of the step changes in $g$. The transmitter may intentionally band-limit $s$ in order to comply with emission regulations, and/or analog components in the transmitter may limit the bandwidth due to their physical limitations. The bandwidth-limiting mechanisms can be modeled mathematically by low-pass filtering $g$. This slows down the frequency transitions in FSK and causes a ramp-down of amplitude towards transition times and ramp-up following transitions in PSK and OOK. This smoothing of $g$ is not necessarily fully specified and tightly controlled at the transmitter.

### A. Simulation and Experimental Parameters

In simulations below, as well as in experiments we carried out with actual transmitters and receivers, we use transmission parameters that correspond to those that are used in the ATLAS system. Specifically, we use sequences of length $N = 8192$, chip rate of 1 Mb/s, sampling rate of 8.333 Ms/s. The carrier frequency of 434 MHz but the signals are down-converted to baseband (zero intermediate frequency) before sampling.

The down-conversion magnifies the effect of reference-clock error on the carrier frequency, but not the effect on the symbol rate. A typical 10ppm clock-rate error changes the 434 MHz carrier by 4.34 KHz; once down-converted to 4 MHz, the carrier-frequency error becomes larger than 1000ppm. The symbol rate has not been translated and its error remains 10ppm. Viewed differently, the 10ppm error displaces the timing of the last symbol by about $8\text{ms} \times 10^{-5} = 0.08\mu\text{s}$ or 8% of one symbol period $\theta$. On the other hand, the 4 MHz down-converted carrier phase rotates during the 8ms packet 32 times too many, implying 64 extraneous phase inversions.

### B. Coherent Detection and TOA Estimation

If $p = s$ (there are little or no frequency and transition-rate errors), the maximizer $\hat{\tau}$ of the absolute value of the correlation of $p$ and $r$ is a maximum likelihood estimator of the delay $\tau$. We approximate the continuous correlation $(p \star r)(\kappa) = \int p(t-\kappa)\bar{r}(t)dt$ where $\bar{r}$ is the complex conjugate of $r$ by interpolating the discrete correlation

$$(p \star r)[k] = \sum_n p[n-k]\bar{r}[n] \ ,$$

which can be computed efficiently in $O((T/\Delta)\log_2(T/\Delta))$ time using the FFT, where $1/\Delta$ is the sampling rate. In this paper we use quadratic interpolation, which interpolates a parabola through $(p \star r)[k_{-1}]$, $(p \star r)[k_0]$, and $(p \star r)[k_{+1}]$, where $k_0$ is the index at which $|(p \star r)[k]|$ is maximized. ATLAS takes the maximum of that parabola as the estimate $\hat{\tau}$ of $\arg\max(p \star r)(\kappa)$ and hence of $\tau$. In this paper we

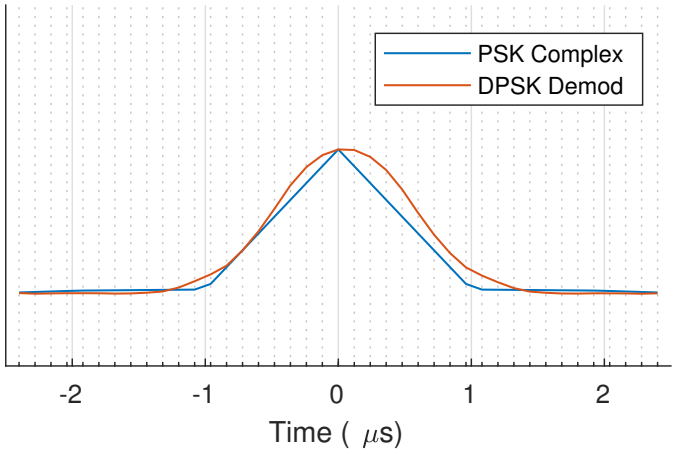


Fig. 2. The shape of the peak of the absolute value of the complex correlation of a PSK signal and of the correlation of a demodulated PSK signal. The vertical gray lines mark sampling times.

use a slightly more sophisticated scheme; we have found that the maximum of the parabola is a biased estimator and that the bias depends on the location of the peak in the inter-sample period. Results reported here correct for this bias by using a piecewise-linear model fitted to simulation results (with a different pseudo-random code). We refer to this TOA estimator as *complex correlation*. We have also experimented with another form of interpolation, which is based on a least-square fitting of an isosceles triangle. This improved the accuracy on unshaped PSK complex correlation, because the shape of the correlation magnitude is indeed a triangle, as can be seen in Figure 2. We have not combined this technique with bias removal and the results in the rest of the paper do not use it.

For OOK, we correlate $r$ with $p_{\text{BPSK}}$. There is clearly no harm in this; in the absence of noise the peak of the correlation should have exactly the same value. However, $p_{\text{BPSK}}$ penalizes the presence of correctly-phased signal when the OOK signal should be zero, whereas $p_{\text{OOK}}$ does not. Therefore, we feel that correlating with $p_{\text{BPSK}}$ produces a more robust estimator.

Smoothing $g$ has little effect on the performance of complex-correlation estimators for PSK and OOK. If transition-rates are not known and/or vary between symbols, complex correlation is still a good TOA estimator for BPSK and OOK, but not for FSK. The problem with FSK is that any variation in the transition rate from $\omega(t)$ at symbol $i$ to $\omega(t)$ at symbol $i+1$ introduces a variation in the relative phase of $s$ during these two symbols (say at the center of the symbol). Hence, magnitude of the complex correlation signal $p \star r$ may fail to show a significat peak at $\tau$.

In BPSK, the phase at the center of each symbol is known from $g$, even if slow transitions cause the phase to be unpredictable near the boundary of symbols. This implies that $|p \star r|$ will have a strong peak at $\tau$. The situation with OOK is even simpler, because the phase of $s$ is always $\phi + \omega t$, even if the amplitude ramps up and down slowly.

However, even in BPSK and OOK, the complex-correlation estimator is very sensitive to stretching and compression of $s$, as shown in the map of the ambiguity function of the BPSK complex correlator in Figure 3 (left). A frequency error of 0.5ppm (only 217Hz for $f = 434$ MHz) is enough to wipe out completely the peak in the complex correlation function. This is a known problem that plagued the system of MacCurdy and others [1]. Sections III-E and III-D propose techniques for estimating the stretching (or compression) of $s$ relative to $p$, which allow complex correlation with a correctly stretched/compressed version of $p$.

Simulation results shown in Figure 4 show that in the absence of frequency discrepancy between $p$ and $s$, complex correlation estimators for PSK and OOK are very effective. At the transmission parameters of ATLAS, they produce standard deviations of less than 1ns at SNRs of 0dB, and they identify the correct correlation peak down to SNRs of less than -30dB. At even lower SNRs the sample variance in the simulation becomes huge because the correlation peak that corresponds to $\tau$ is no longer the maximizer of the correlation. The performance of OOK is about 3dB lower than that of PSK.

### C. Detection and TOA Estimation using Demodulated Signals

Another approach to addressing the difficulties caused by the discrepancies between $p$ and $s$ is to demodulate $r$ and to correlate the real demodulated signal $d$ with $g$. That is, we pass the samples $r[n]$ though a process that aims to recover from it a real estimate $d[n]$ of the real function $g$. We do not aim to recover the symbol clock or the binary symbols. Since the demodulator is local ($d[n]$ depends only on $r[m]$ for $|m-n| \leq 2\theta/\Delta$), $d$ is relatively insensitive to discrepancies between $s$ and $p$. The BPSK demodulator that we use estimates the differential of $g$, not $g$ itself, but the principle is the same.

For FSK, we construct $d$ as follows. We define two matched filters $f_{+1}^{(\theta)}$ and $f_{-1}^{(\theta)}$ for complex sinusoids for frequencies $f+v$ and $f-v$ of length $\theta/\Delta$ (a single symbol) and apply them to $r$, $f_{+1} = f_{+1}^{(\theta)} \star r$. The demodulated signal is

$$d_{\text{FSK}} = \frac{|f_{+1}| - |f_{-1}|}{|f_{+1}| + |f_{-1}|} .$$

This signal is limited to $[-1, +1]$ and reaches these extremes at samples in which the matched filters are aligned with symbols and only if $f+v$ is orthogonal to $f-v$; it they are not orthogonal, $d$ will vary in a smaller range. Once $d$ is computed, we correlate $d$ with $g$, find the maximum of the absolute value of $d \star g$, interpolate to approximate the continuous correlation, and find the maximum of the interpolated real correlation signal.

For BPSK, we convert $g[n]$ to a differential signal, $h[n] = |g[n] - g[n-1]| - 1$, with value $+1$ when $g[n]$ is different than the previous symbol and $-1$ when they are identical. We define two matched filters of length $2\theta/\Delta$ (two symbols): $p_{\text{same}}^{(2\theta)}$, a continuous-phase complex sinosoid for frequency $f$, and $p_{\text{diff}}^{(2\theta)}$, a complex sinosoid for frequency $f$ with a $180°$ phase shift in the middle. We correlate the matched filters with $r$ to produce filtered signals $p_{\text{same}}$ and $p_{\text{diff}}$ and demodulate

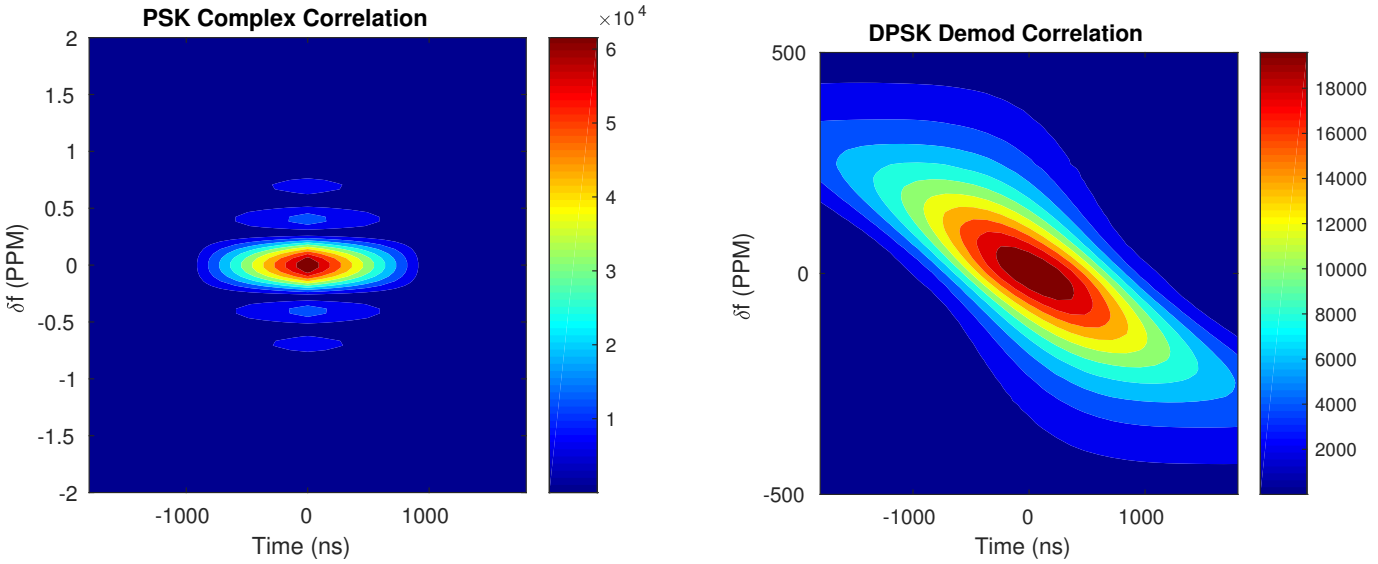


Fig. 3. The ambiguity function (maximum of the absolute value of the $p \star r$ at different frequency errors between $p$ and $s$) for the complex correlator for BPSK (left) and for the real correlator of DBPSK demodulation (right). Note that the scaling of the horizontal axes are identical but that those of the vertical axes are not. The frequency differences on the vertical axes refer to differences in the RF frequency, near 434 MHz.

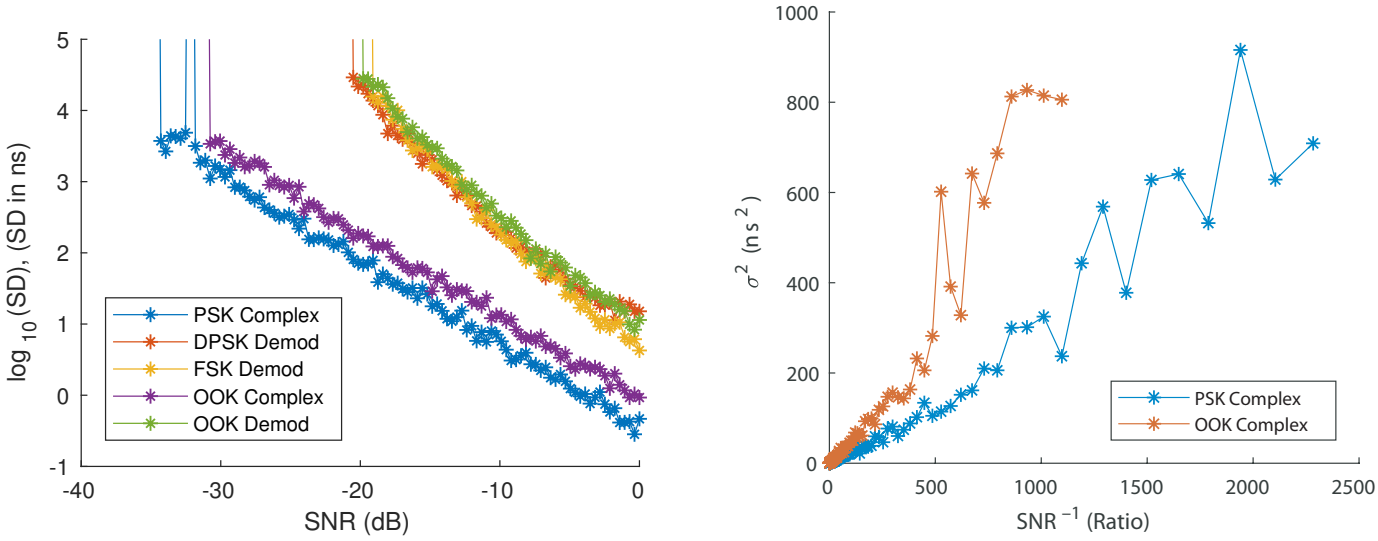


Fig. 4. The accuracy of different delay estimators as a function of the SNR under additive Gaussian white noise. The graph on the left shows the results using logarithmic scales. The graph on the right shows some of these results using the scales that show the linear dependence of var($\hat{\tau}$) on the SNR (the same scales using in the CRLB). In these simulations, $s = p$.

using the expression

$$d_{\text{DPSK}} = \frac{|p_{\text{diff}}| - |p_{\text{same}}|}{|p_{\text{diff}}| + |p_{\text{same}}|} \ .$$

The delay estimate $\hat{\tau}$ is the interpolated maximum of the absolute value of the real correlation between $d_{\text{DPSK}}$ and the differential symbol sequence $h$.

Demodulating OOK is even simpler. The simplest demodulator simply computes the absolute value of samples of $r$. We designed a more robust estimator that works as follows. We first correlate $r$ with a matched filter of length $\theta/\Delta$ (one symbol) at the carrier frequency $f$ and take the absolute value of this filtered signal $f^{(\theta)} \star r$. We then compue the moving average $c$ of this real signal, averaged over windows of 12 symbols. Finally, we subtract the moving average (to remove the DC component) and normalize by $|c|$. That is,

$$d_{\text{OOK}}[i] = \frac{\left(f^{(\theta)} \star r\right)[i] - c[i]}{|c[i]|}$$

where

$$c[i] = \sum_{j=i-6\theta/\Delta}^{i+6\theta/\Delta} \left(f^{(\theta)} \star r\right)[j] \Big/ \left(\frac{12\theta}{\Delta} + 1\right) \ .$$

The ambiguity function $|d_{\text{DPSK}} \star h|$ of the DPSK detector, shown in Figure 3 (right), demonstrates the robustness of

these estimators. The correlation peak is easily detectable even at stretching or compression of the reference clock of the transmitter by 200ppm or more, which far exceeds the errors caused by temperature and aging of miniature low-cost crystal oscillators. We also see that stretching and compression of $s$ causes severe bias in $\hat{\tau}$. The ambiguity function of our FSK detector and OOK detector are not shown but are qualitatively and quantitatively similar.

Also, correlating the output of a demodulator increases $\sigma^2(\hat{\tau}) = \text{var}(\hat{\tau})$ significantly relative to complex correlation estimators. The results in Figure 4 reveal performance degradation of about an order of magnitude in $\sigma(\hat{\tau})$ at high SNRs (near 0) and of more than 2 orders of magnitude at low SNRs. The degradation in location accuracy would be similarly degraded.

### *D. Low-SNR Brute-Force Coherent Acquisition*

To estimate $\hat{\tau}$ using complex correlation under a carrier-frequency error, we need to estimate $\delta$, the carrier-frequency error.

One way to estimate $\delta$ is to correlate $r(t)$ not only with $p(t)$, but also with $p((1+\delta)t)$ for a range of potential $\delta$s. The absolute value of $r(t) \star p((1+\delta)t)$ is a horizontal cross-section through the ambiguity function (see Figure 3). By computing these cross sections every $\delta_0$ difference and covering the range of plausible reference-clock errors, say $\pm$20ppm, we can ensure that at least one cross section is close enough to the peak of the ambiguity function. This is essentially the same technique that is used for Doppler search in GPS receivers. We stress that we use the notation $p((1+\delta)t)$ to denote the RF replica, which we never represent explicitly; we create representations of down-converted replicas, in which the carrier error is multiplied by approximately $f_{RF}/f_{\text{IF}}$ whereas the relative symbol rate error remains $\delta$.

The shape of the ambiguity function determines an upper bound on $\delta_0$. For the parameters of ATLAS transmissions, the ambiguity function for complex PSK correlation shown in Figure 3 suggests setting $\delta_0 \approx 0.1$ppm. To account for frequency errors of $\pm$20ppm, we would need to correlate $r$ with about 400 replicas $p$. This is a significant computational overhead, higher than that of the typical C/A Doppler search in GPS , mostly due to the longer pseudo-random sequences that ATLAS uses. However, at low SNR and if no prior information on $\delta$ is available, this search is necessary.

We propose two techniques to reduce the overhead of the frequency search. One technique tracks $\delta$ through multiple transmissions of the same tag. As long as the temperature of the tag is stable, $\delta$ will be stable, which allows us to use a very limited search. More specifically, given the estimate $\hat{\delta}$ of the previous transmission detected, we search at $\hat{\delta} - \delta_0$, $\hat{\delta}$, and $\hat{\delta} + \delta_0$; if the maximum absolute value of the correlation at $\hat{\delta}$ is higher than at adjacent frequencies, we keep the estimate; otherwise we search for $\hat{\delta}$ again. We expect that in most cases, the three complex correlations will be sufficient to both verify that $\hat{\delta}$ is still a good estimate and to estimate the TOA.

The other technique, which we propose in the next section, exploits the tolerance of correlations with a demodulated signal to frequency errors to dramatically reduce the cost of the search for $\hat{\delta}$.

### *E. High-SNR Acquisition*

A more efficient way to estimate the frequency error $\delta$ is to first estimate the arrival time $\tau$ using a frequency-tolerant delay estimator that correlates a demodulation of $r$. As shown in Figure 3, these TOA estimators are indeed frequency tolerant (the figure shows an estimator for PSK transmissions, but the behaviors of estimators for FSK and OOK are similar).

Once $\tau$ is known, we mix $r(t+\tau)$ with $p(t)$ (pointwise complex multiplication). This produces a complex exponential at the error frequency $\delta$. For example, for PSK we have

$$\begin{aligned} r(t+\tau) \otimes \bar{p}(t) &= (\alpha s(t) + \epsilon(t+\tau)) \otimes \bar{p}(t) \\ &= (\alpha p((1+\delta)t) + \epsilon(t+\tau)) \otimes \bar{p}(t) \\ &= \alpha A e^{i(\phi+\omega(1+\delta)t)} g(t) \otimes A e^{-i(\phi+\omega t)} g(t) \\ &\quad + \epsilon(t+\tau) \otimes A e^{i(\phi+\omega t)} g(t) \\ &= \alpha A^2 e^{-i\omega\delta} + \epsilon(t+\tau) \otimes A e^{i(\phi+\omega t)} g(t) \ . \end{aligned}$$

We now perform an FFT and estimate $\delta$ as $\hat{\delta} = (1/\omega) \arg\max_\delta |\text{FFT}\,(r(t+\tau) \otimes p(t))|$.

Once we have obtained the estimate $\hat{\delta}$, we estimate $\hat{\tau}$ again using complex correlation, to improve the delay estimate.

This method of efficiently searching for the peak of the ambiguity function only works as long as the demodulation-correlator is able to estimate $\tau$. As shown in Figure 4, the demodulation-correlator can identify the delay only down to SNRs of about -20dB. At lower SNRs, the complex correlator can still estimate $\tau$, but we would need to search for $\delta$ using the brute-force method of Section III-D. The method described here only works at relatively high SNRs (down to -20dB or so, for the parameters of the ATLAS system).

### *F. SNR Estimation, Detection Thresholds, and TOA Variance Estimation*

TOA emitter localization systems must be able to determine when they have acquired the signal from the target and can switch to tracking mode, and they must also be able to determine when tracking has failed (in which case they need to switch back to acquisition mode). Another common requirement is to estimate the uncertainty in estimated positions, which depends on the uncertainty in TOA estimates.

The variance of TOA errors for a given signal depends linearly on the noise-to-signal ratio (inverse SNR), as shown by the CRLB. The variance of the errors depends on the SNR of the signal from which we estimate the TOA; this signal is $r$ in the case of complex correlation, and is the output of the demodulator in the case of demodulator correlation. Therefore, we aim to estimate the SNR of these signals, not the SNR at the antenna or at the analog-to-digital converter (ADC).

The method that we use to estimate the SNR was proposed by Weller-Weiser et al. [2] but its properties were not carefully investigated until now. When correlating a demodulated signal

$d[n]$ with $g[m]$ (the case presented in [2]), we normalize $g$ so that $\|g\| = 1$ and use the estimate

$$\begin{aligned} \|\text{signal}\| &= \sum_{i=0}^{N} d\left[i + \frac{\hat{\tau}}{\Delta}\right] g[i] \\ \|\text{noise}\|^2 &= \left(\sum_{i=0}^{N} d\left[i + \frac{\hat{\tau}}{\Delta}\right]^2\right) - \|\text{signal}\|^2 \\ \text{SNR estimate} &= \frac{\|\text{signal}\|^2}{\|\text{noise}\|^2} . \end{aligned}$$

In other words, the total energy (signal plus noise and/or interference) is the Euclidean norm of the demodulated signal during the transmission period. This vector decomposes into two orthogonal parts, the signal, whose length is the dot product between the demodulated signal and the direction of $g$, and the rest, which we take to be noise and interference. When correlating complex signals, we use the same formula but replace $d$ with $r$ and $g$ with $p$ (more precisely, with the frequency-corrected $p$).

The graphs in Figure 5 show that for complex correlation with PSK signals, this SNR estimator is unbiased and that its variance is small for SNRs above about -20dB. At lower SNRs the estimator becomes more noisy, and at some point it flattens out and stops dropping even when the actual SNR keeps dropping. The point at which this failure occurs is the point at which the SNR of the correlation signals drops below 0dB and it is no longer possible to identify the correlation peak associated with the signal. For white noise, this happens at SNRs below about -35dB, consistent with the results in Figure 4. When the transmission is subject to inteference from another tag transmitting a different pseudo-random sequence using the same transmission parameters (frequency, modulation, symbol rate, etc), the SNR estimator flattens out earlier, at about -25dB; at that point, the SNR estimate no longer distinguishes between a very weak signal and a strong inteferer.

The SNR estimator of the demodulated PSK signal is biased with respect to the SNR at the antenna or the ADC, as shown in Figure 5 (right). It also shows similar breakdowns at low SNRs, and it also suffers from an earlier breakdown under interference than under white noise.

The data in Figure 4 shows that the dependence of the TOA variance on the true inverse SNR is approximately linear, as predicted by the CRLB (if we assume that our TOA estimator is efficient). Given the accuracy of the SNR estimator for complex correlation, this indicates that the TOA variance is a linear function of the SNR estimator, and hence can be easily predicted from it using linear regression.

The data presented in Figure 5 also shows that we can use these SNR estimators to reliably decide that the target is present in $r$. At SNR estimates well above the point in which the estimates flattens out, the estimate indicates the presence of $s$ in $r$. Below these flattening out points, it becomes no longer possible to distinguish a strong interferer from a weak target. Figure 6 shows in more details how to set the detection threshold for acquisition and tracking. The figure shows the distribution of the SNR estimator when no target is present in $r$, only white noise (blue) or an interferer (brown). ATLAS stores these histograms, which allows it to find the SNR estimator threshold that corresponds to a particular false-negative probability. We normally use thresholds of $10^{-6}$for acquisition and $10^{-5}$ for tracking.

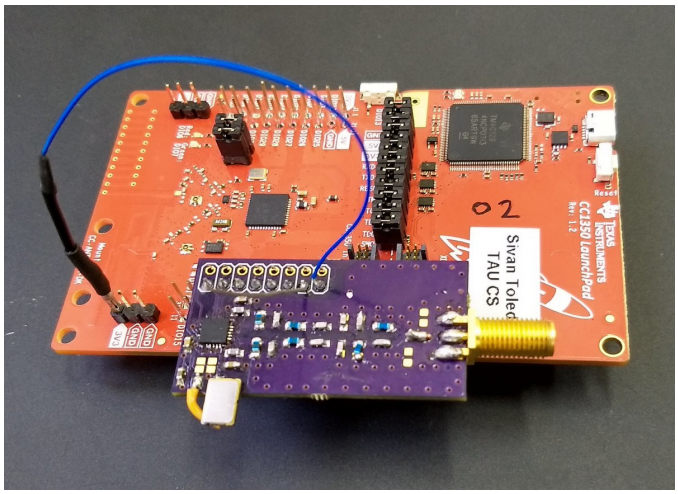


Fig. 9. Our prototype tag based on the AX5031 transmitter, mounted on a custom board (purple) and attached to a CC1350 evaluation board.

On OOK signals, the SNR estimators are much more sensitive to interferers, as shown in Figure 7. Under white noise, the complex OOK correlator is biased and flattens out at high SNRs; these are results of correlating $r$ with $p_{\text{PSK}}$ rather than with $p_{\text{OOK}}$ and are of little significance. The behavior of correlating the output of our OOK demodulator is almost identical to the behavior of demodulated PSK. The SNR estimator of the complex OOK correlator becomes noisy at SNRs about about -15dB in the presence of an interferer. The SNR estimator of the demodulated OOK signal is much noisier, sometimes reporting very low SNRs when the actual SNR is good, even 0dB or above. These weaknesses in the SNR estimator provide more evidence that PSK should be preferred if at all possible. Interference from other tags is also very harmful to FSK packets, as shown in Figure 8.

## IV. Implementation

While the main aim of the research reported in this paper has been to understand the implications of different modulation schemes, which we have largely accomplished in Section III using simulations, our ultimate goal is to implement the most promising techniques in ATLAS. We have not completed this task but we have made some significant steps towards this goal. This section describes these steps.

### A. A Prototype for PSK-Capable Tags

First, we have identified at least one integrated radio that can transmit high-bitrate BPSK packets, the AX5031, and we have built a prototype tag using it. The part can transmit on 434 MHz as well as on 868 and 915 MHz and can transmit

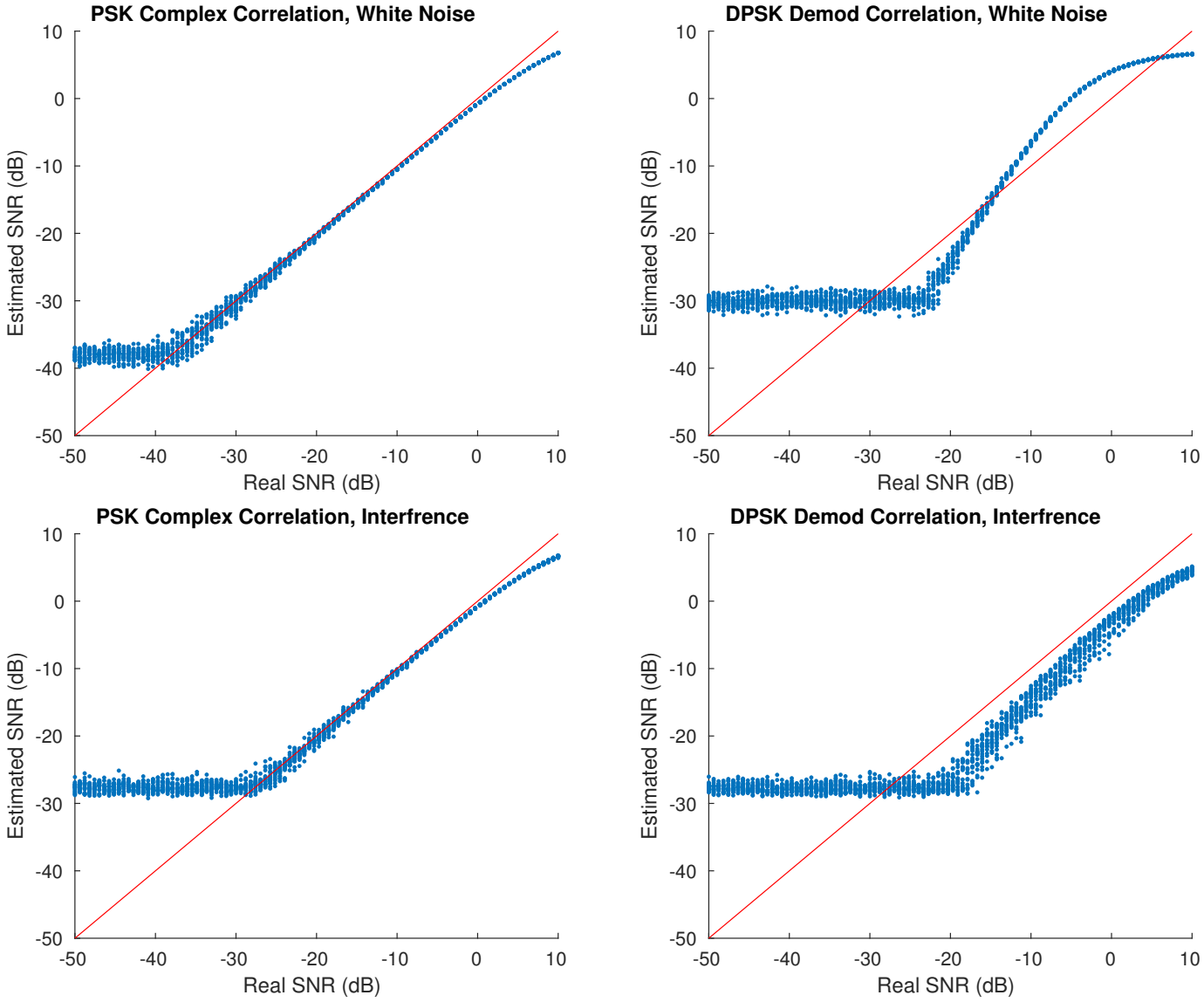


Fig. 5. The relationship between the actual SNR of PSK signals and the estimated SNRs for complex correlation (left) under both white noise (top) and an interferer transmitting a different pseudo-random sequence but with the same transmission parameters (frequency, modulation, etc). The graphs on the right present the same analysis but for correlation of the output of a DPSK demodulator.

PSK packets at up to 2 Mb/s and at up to 15dBm. We designed a custom printed circuit board to evaluate the transmitter and connected it to a commercial evaluation board for the CC1350. We wrote firmware for the CC1350 to periodically transmit BPSK packets from the AX5031 at 1 Mb/s and we verified that our implementation of the algorithms described above (implementations within ATLAS, which are separate from Matlab implementations that we used for the simulations) can acquire and track transmissions from the tag.

We have also used this prototype tag to assess the frequency stability of these miniature integrated transceivers, in order to drive the configuration of brute-force searching.

Figure 10 shows the carrier-frequency trajectory of our prototype tag during normal operation. In this experiment, the tag was stationary on a table at room temperature (around 20C). It was then lifted and then warm air was blown on it. We can see carrier frequency changes of more than 10ppm. The experiment shows that we must expect variations of this magnitude in carrier frequency, but also that when not subject to temperature gradients, the frequency remains stable and can be easily tracked. The carrier frequency was estimated using the ATLAS implementation of the algorithm from Section III-E.

We have also designed a miniature tag that uses a CC1310 as a processor and an AX5031 as a transmitter. The design, shown in Figure 11, is 12 by 23mm. There are two challenges associated with miniaturizing an AX5031-based tag. One is the obvious fact that the tag needs both the AX5031 and a processor. The second is that the AX5031 requires an elaborate matching network consisting of 14 capacitors and inductors. In comparison, the first-generation ATLAS tags

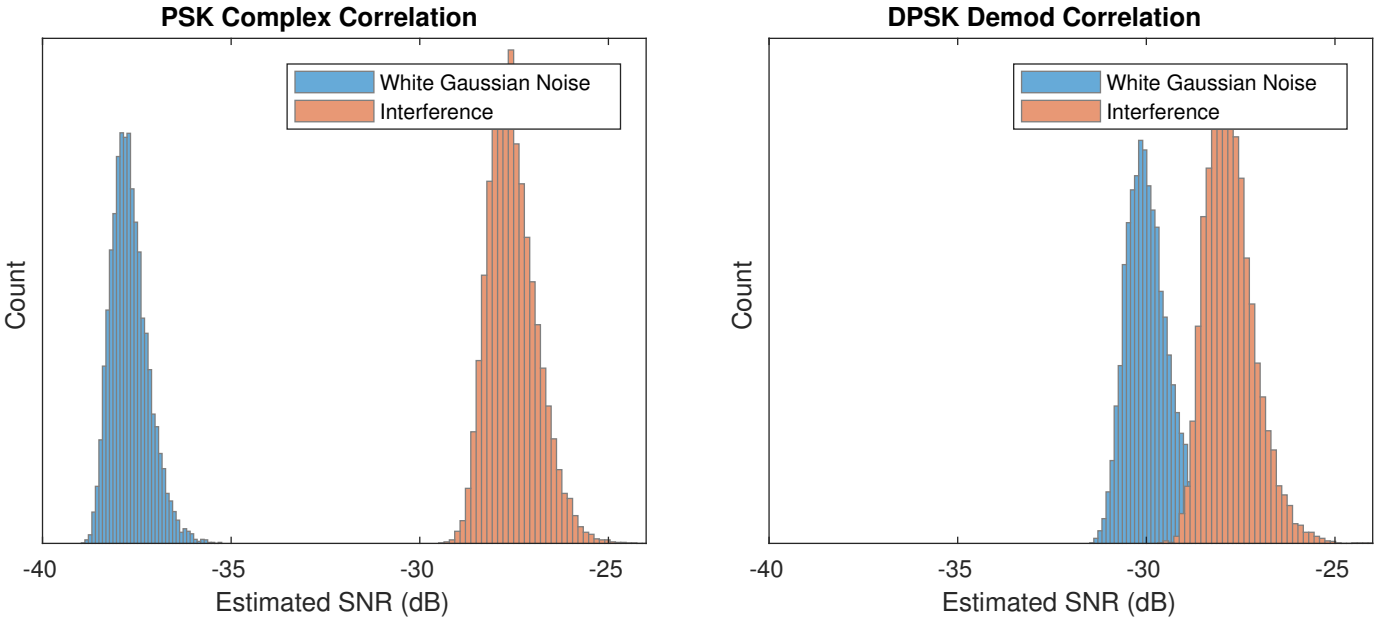


Fig. 6. Histograms of the estimated SNRs of two PSK correlators (complex on the left and demodualted DPSK on the right) when no target is present in the signal, only white noise or an interferer.

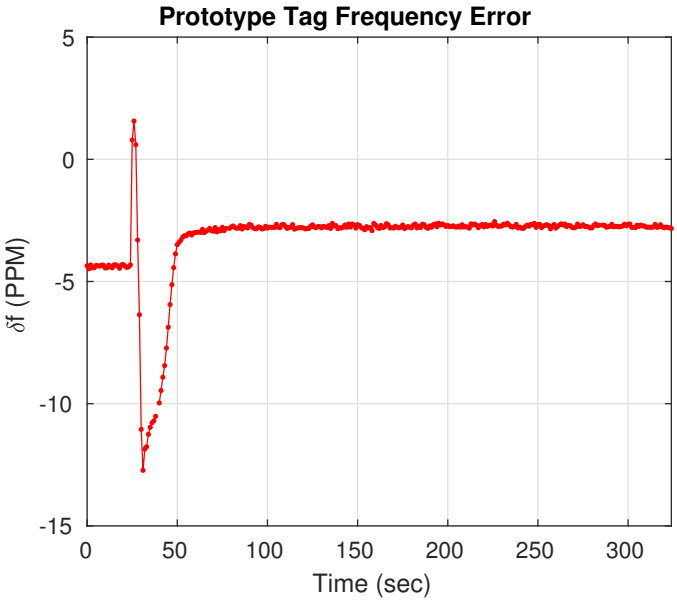


Fig. 10. Carrier-frequency estimation by our algorithm on pings (at a ping rate of 1Hz) produced by our prototype tag. The dramatic spike was created by lifting the tag and then blowing warm air on it by exhaling.

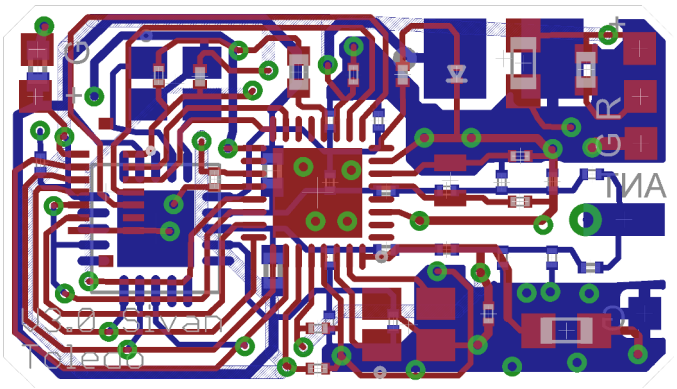

Fig. 11. A board design for a AX5031-based tag, using a CC1310 as a controller. The dimensions are 12 by 23mm. The AX5031 and its periphery is mounted on the bottom of the board (depicted in blue) and the CC1310 is mounted on the top (brown).

used the CC1101 transceiver, for which a single integrated passive device replaces the entire matching network, allowing for more compact designs. Given these constraints, the size of this board is reasonable, and we plan to manufacture a batch of these tags soon.

Another option that we plan to evaluate are new EFR32 RF MCUs; the data sheet of these devices states that they can transmit BPSK, and they are capable of high bitrates. However, BPSK is not yet supported by the development environment for the devices, so we cannot currently test them. These devices come in 5mm-by-5mm packages (larger than the 4mm-by-4mm CC13x0) and integrated passive matching devices are available for them, hopefully allowing for very compact board designs.

### *B. High-Performance Algorithms and Implementation in ATLAS Receivers*

In addition to the Matlab codes that we used for the simulations shown above, we have also implemented most of the algorithms required to handle BPSK packets in ATLAS. In particular, we have implemented the generation of $p_{\text{BPSK}}$, the DPSK demodulator, the complex correlator, and the carrier-frequency estimator from Section III-E. The implementation of the computationally-expensive signal processing algorithms

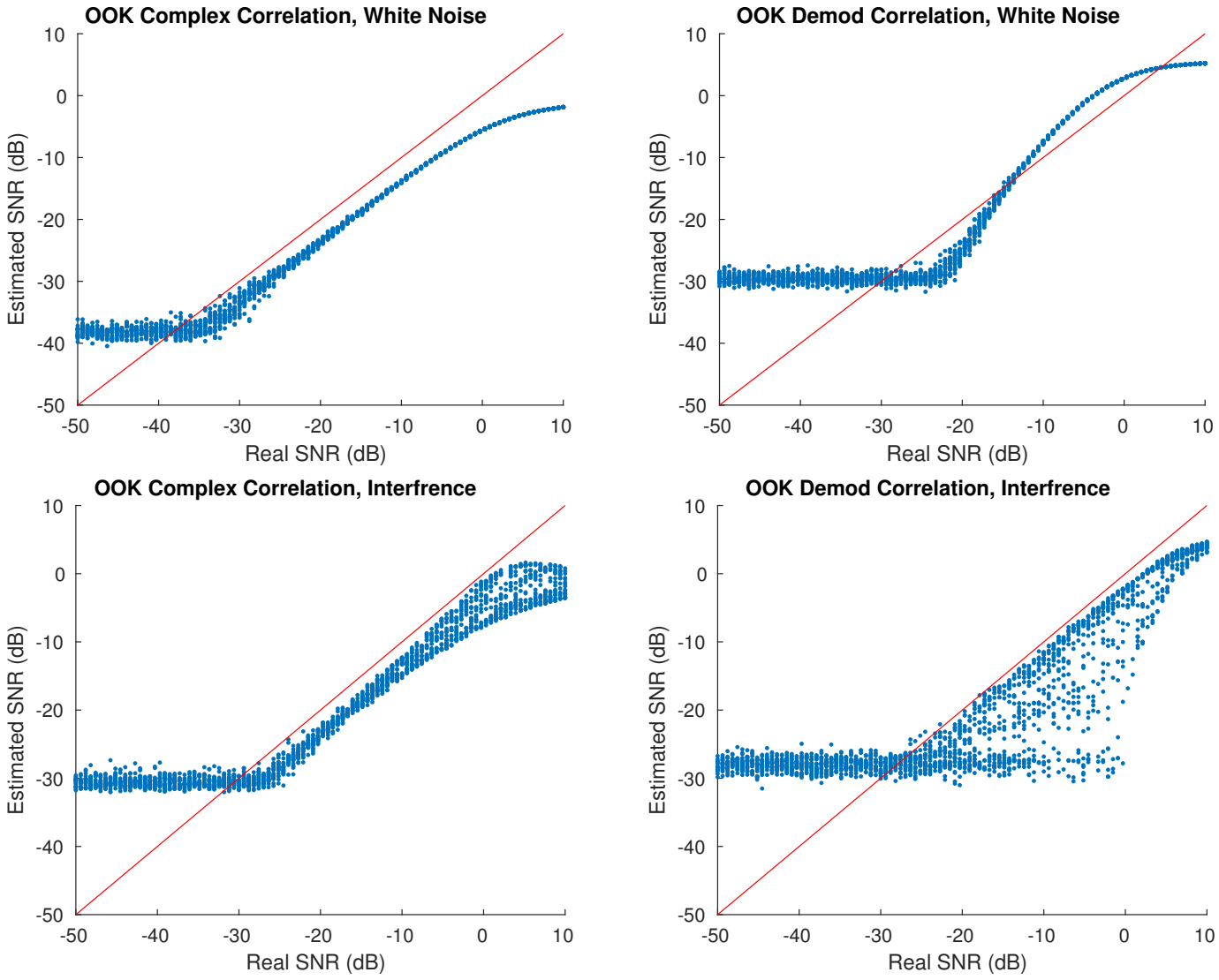


Fig. 7. The relationship between the actual SNR and estimated SNR of OOK signals.

is in C and they utilize a high-performance FFT library [15]. The rest is written in Java. The implementation is integrated with the acquisition and tracking algorithms of ATLAS and it can acquire and track BPSK tags.

## V. Conclusions

Our conclusions from this study are fairly clear. BPSK modulation has emerged as a clear winner for reverse-GPS wildlife tracking systems, as long as the system employs a ping detection algorithm that is tolerant to carrier-frequency errors in tags. When implemented using complex correlation (which require that the carrier-frequency errors are estimated and adjusted for), BPSK is highly immune to both white noise and to interference from other tags and offers sensitivity and TOA estimation accuracy that are superior to those of FSK and OOK, the other modulation schemes that miniature UHF transmitters and transceivers can emit. Furthermore, we have shown how to estimate the frequency errors of tags using a computationally inexpensive algorithm as long as SNRs are not very low. Finally, because the performance of our demodulated DPSK correlator is similar to that of our demodulated FSK correlator, we can say that BPSK has no disadvantages relative to FSK in this application, other than perhaps the wider availability of miniature FSK transmitters.

OOK is theoretically similar to BPSK but less efficient in terms of use of power and much more susceptible to interference from other tags. The inefficiency is both due to the fact that half the transmitted power is a carrier that is useless for tag identification and for TOA estimation, and due to the fact that the synthesizer consumes power even when the transmitter is in the off state; in miniature transmitters, most of the power is consumed by the synthesizer, not by the RF power amplifier. However, there are more integrated transmitters and transceivers that can transmit OOK than devices that can transmit BPSK; this is one reason that we explored this modulation.

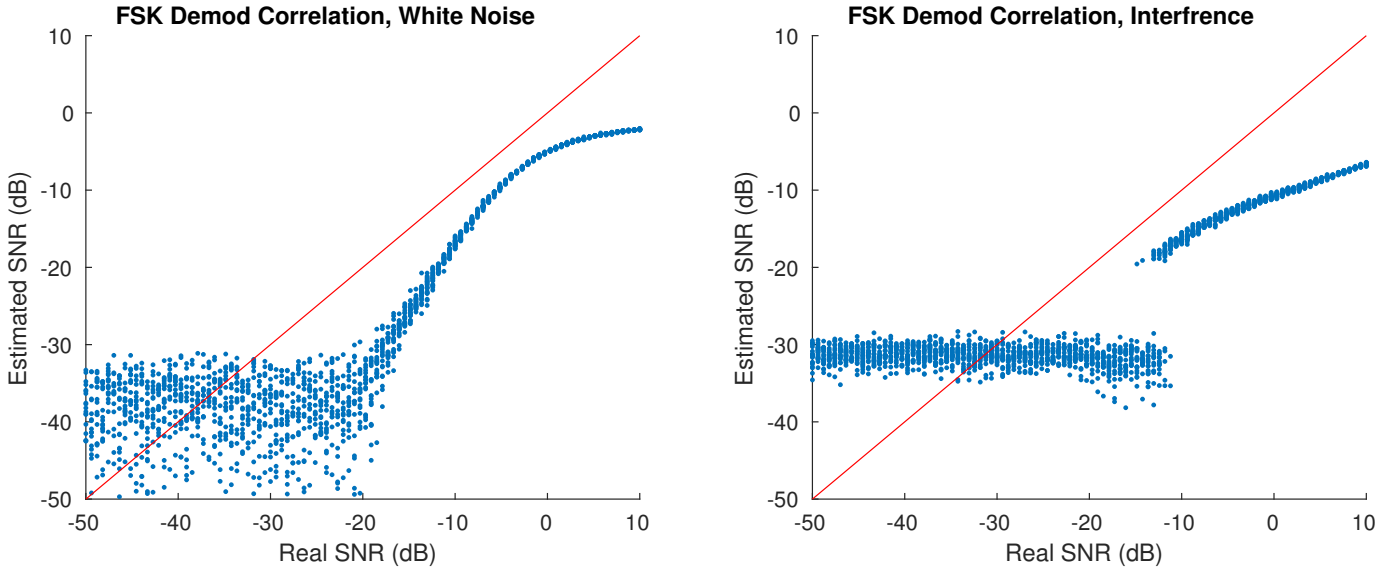


Fig. 8. The relationship between the actual SNR and estimated SNR of FSK signals. We do not show the performance of the complex correlator since it is unreliable for FSK.

There appears to be no good reason to use FSK, the modulation scheme that ATLAS currently use, other than the fact that virtually every miniature integrated transmitter and transceiver can produce FSK.

*Acknowledgments:* This research was supported in part by the Miverva Center for Movement Ecology and by grants 965/15 and 863/15 from the Israel Science Foundation (funded by the Israel Academy of Sciences and Humanities).

## REFERENCES

[1] R. MacCurdy, R. Gabrielson, E. Spaulding, A. Purgue, K. Cortopassi, and K. Fristrup, "Automatic animal tracking using matched filters and time difference of arrival," *Journal of Communications*, vol. 4, no. 7, pp. 487–495, 2009.

[2] A. Weller-Weiser, Y. Orchan, R. Nathan, M. C. A. J. Weiss, and S. Toledo, "Characterizing the accuracy of a self-synchronized reverse-GPS wildlife localization system," in *Proceeings of the 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, Vienna, Austria, Apr. 2016, pp. 1–12.

[3] S. Toledo, O. Kishon, Y. Orchan, A. Shohat, and R. Nathan, "Lessons and experiences from the design, implementation, and deployment of a wildlife tracking system," in *Proceeings of the IEEE International Conference on Software Science, Technology and Engineering (SWSTE)*, Beer Sheva, Israel, Jun. 2016, pp. 51–60.

[4] A. I. Bijleveld, R. B. MacCurdy, Y.-C. Chan, E. Penning, R. M. Gabrielson, J. Cluderay, E. L. Spaulding, A. Dekinga, S. Holthuijsen, J. ten Horn, M. Brugge, J. A. van Gils, D. W. Winkler, and T. Piersma, "Understanding spatial distributions: negative density-dependence in prey causes predators to trade-off prey quantity with quality," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 283, 2016.

[5] T. Piersma, R. B. MacCurdy, R. M. Gabrielson, J. Cluderay, A. Dekinga, E. L. Spaulding, T. Oudman, J. Onrust, J. A. van Gils, D. W. Winkler, and A. I. Bijleveld, "Fijnmazige positiebepaling van individuen in groepen: de principes en drie toepassingen van toa-tracking," *Limosa*, vol. 87, pp. 156–167, 2014.

[6] S. W. Krüger, "An inexpensive hyperbolic positioning system for tracking wildlife using off-the-shelf hardware," Master's thesis, North-West University, Potchefstroom, South Africa, 2017.

[7] R. Kays, S. Tilak, M. Crofoot, T. Fountain, D. Obando, A. Ortega, F. Kuemmeth, J. Mandel, G. Swenson, T. Lambert, B. Hirsch, and M. Wikelski, "Tracking animal location and activity with an automated radio telemetry system in a tropical rainforest," *The Computer Journal*, vol. 54, pp. 1931–1948, 2011.

[8] Š. Řeřucha, T. Bartonička, P. Jedlička, M. Čížek, O. Hlouša, R. Lučan, and I. Horáček, "The BAARA (Biological AutomAted RAdiotracking) system: A new approach in ecological field studies," *PLoS ONE*, vol. 10, p. e0116785, 2015.

[9] Z. Deng, M. A. Weiland, T. Fu, T. A. Seim, B. L. Lamarche, E. Y. Choi, T. J. Carlson, and M. B. Eppard, "A cabled acoustic telemetry system for detecting and tracking juvenile salmon, part 2: Three-dimensional tracking and passage outcomes," *Sensors*, vol. 11, pp. 5661–5676, 2011.

[10] M. A. Weiland, Z. Deng, T. A. Seim, B. L. Lamarche, E. Y. Choi, T. Fu, T. J. Carlson, A. I. Thronas, and M. B. Eppard, "A cabled acoustic telemetry system for detecting and tracking juvenile salmon, part 1: Engineering design and instrumentation," *Sensors*, vol. 11, pp. 5645–5660, 2011.

[11] R. B. MacCurdy, R. M. Gabrielson, and K. A. Cortopassi, "Automated wildlife tracking," in *Handbook of Position Location: Theory, Practice, and Advances*, S. A. R. Zekavat and R. M. Buehrer, Eds. Wiley, 2012, pp. 1129–1167.

[12] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory*. Prentice Hall, 1993.

[13] European Telecommunications Standards Institute (ETSI), "Short range devices (SRD) operating in the frequency range 25 MHz to 1000 MHz; part 2: Harmonised standard covering the essential requirements of article 3.2 of directive 2014/53/EU for non specific radio equipment," EN 300 220-2 V3.1.1, 2017.

[14] S. Toledo, O. Kishon, Y. Orchan, Y. Bartan, N. Sapir, Y. Vortman, and R. Nathan, "Lightweight low-cost wildlife tracking tags using integrated tranceivers," in *Proceeings of the 6th Annual European Embedded Design in Education and Research Conference (EDERC)*, Milano, Italy, Sep. 2014, pp. 287–291.

[15] M. Frigo and S. G. Johnson, "The design and implementation of FFTW3," *Proceedings of the IEEE*, vol. 93, no. 2, pp. 216–231, 2005, special issue on "Program Generation, Optimization, and Platform Adaptation".