Maximum Likelihood Time of Arrival and Doppler Estimation for Precise Starlink-Based PNT

Wenkai Qin*, Zacharias M. Komodromos†, Samuel C. Morgan*, Todd E. Humphreys*
*Department of Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin
†Department of Electrical and Computer Engineering, The University of Texas at Austin

***Abstract*—We present a maximum likelihood (ML) Doppler and time-of-arrival (TOA) estimation framework for opportunistic tracking of Starlink downlink signals. Extending previous approaches that rely solely on known pilot symbols, we incorporate full-frame ML estimation to harness the data payload, significantly improving Doppler and TOA estimation accuracy. Using live Starlink transmissions, we validate our ML estimator and compare its performance against pilot-based cross-ambiguity function (CAF) and pilot-only ML estimation methods. Results show that the full-frame ML estimator achieves a 10^3 factor improvement in Doppler accuracy over the pilot-only CAF method and 10^2 factor improvement over the pilot-only ML method, reducing post-fit residual RMSE from 1469.20 Hz and 752.43 to 6.34 Hz, respectively. TOA estimation sees a smaller improvement. The findings highlight the value of leveraging the entire OFDM frame for estimation. Additionally, we newly identify two OFDM symbol modulation schemes in use by Starlink.**

I. Introduction

The past decade has seen an explosion of satellite operations. In 2013, only 1,187 active satellites orbited the Earth; as of 2023, this number has ballooned to 9,115 [1]. This rapid growth can be in part attributed to a historic interest in low Earth orbit (LEO) satellites for communications—companies such as SpaceX, OneWeb, and Amazon are pursuing the promise of high-bandwidth, worldwide internet as thousands more vehicles are set to be launched into orbit. Meanwhile, traditional global navigation satellite systems (GNSSs) are facing challenges that threaten its reliability. Due to their low power and wide-area broadcast nature, GNSS signals are inherently susceptible to jamming and interference [2], [3]. Further, GNSSs' constellations have a low vehicle count and are relied upon by many military applications, making them attractive targets in the age of anti-satellite warfare [4], [5]. Recent events have further emphasized the need to protect, augment, and toughen traditional GNSS, as evidenced by the rise of GNSS jamming and spoofing attacks in Europe Mediterranean [6], [7].

LEO communications satellites offer a unique opportunity for GNSS augmentation. Due to their stronger signals, wider bandwidth, and large vehicle count, LEO communications downlinks are inherently resilient to jamming, spoofing, and anti-satellite warfare. As such, researchers have already begun to investigate the feasibility and performance of LEO position, navigation, and timing (PNT) systems [8], [9]. Other researchers exploring the possibility of free-to-use LEO PNT have demonstrated meter-level positioning by opportunistically tracking the Doppler of various communications constellations' transmissions [10]–[15]. Others have It is worth noting that compared to to pseudorange-based PNT techniques, Doppler-based techniques have worse timing accuracy by many orders of magnitude (milliseconds vs. nanoseconds), even under optimistic measurement noise and satellite clock offset rate assumptions [16]–[18]. Recognizing that many PNT applications of practical interest require accurate timing, this paper provides a method for improved observable extraction for both Doppler- and pseudorange-based approaches to PNT.

The opportunistic Starlink PNT receivers currently available in literature extract PNT observables via (1) frequency tracking of leakage tones present between channels, (2) matched filtering against known pilot symbols unveiled by [19] or (3) frequently observed portions of the Starlink frame [20]. The authors of [21] present theoretical root-mean-square error (RMSE) bounds based only on the pilot symbols mentioned by item (2). However, these papers fail to fully achieve the estimate precision available within the Starlink signal, instead exploiting only two of the 302 OFDM symbols transmitted per frame. The authors of [22] demonstrate the importance of harnessing the full OFDM frame in Doppler estimation. When correlating against the template frame presented in [23], the Doppler ambiguity function's first nulls occur at ± 750 Hz, a drastic improvement when compared to the ± 113 kHz achievable when correlating against only two pilot symbols. Nonetheless, the reported performance is only available when a Starlink SV transmits the template frame rather than user data, which necessarily interferes with its communications mission. Quite possibly, the template frame is transmitted less frequently in areas and during times of high Starlink user data traffic. The theoretical framework presented in [24] presents possible maximum-likelihood (ML) and decision-directed estimation framework but falls short of its implementation or empirical validation.

To address these deficiencies in existing opportunistic LEO PNT systems, we apply the theory presented by [24] to implement, demonstrate, and evaluate a Starlink-based, ML Doppler and TOA estimation algorithm. The paper is organized as follows: Section II introduces the signal model. Section III defines the full-frame ML estimation framework. Section IV describes the signal capture system used to generate the results in this paper. Section V provides a detailed explanation of the estimation framework's practical implementation. Section VI presents the key results. Finally, Section VII closes the paper.

Preprint of the 2025 IEEE/ION Position, Location, and Navigation Symposium, Salt Lake City, UT, April 28-May 1, 2025

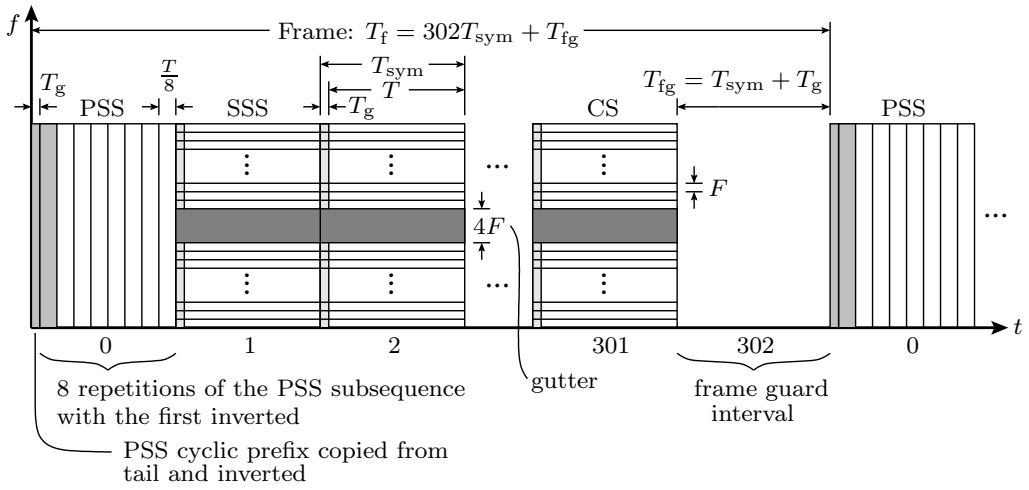


Fig. 1: Frame layout for the Ku-band Starlink downlink along time-frequency dimensions, from [19]. Indices along the horizontal axis enumerate the 303 intervals that constitute a single frame.

II. Signal Model

This section first reiterates key theoretical concepts and measurement models presented by [24], then adapts them for realistic application to Starlink transmissions.

As shown by [19], the Starlink Ku-band downlink utilizes orthogonal frequency-division multiplexing (OFDM) modulation, widely used for high-speed wireless communication systems due to its robustness against multipath interference, efficient spectrum utilization, and adaptability to varying channel conditions. While the full details are presented in [19], we present a brief overview of the Starlink signal structure for the reader's convenience. Fig. 1, replicated from [19], offers a reference for the Starlink frame layout.

Starlink OFDM signals are transmitted over one of eight channels, which altogether span the $10.7 - 12.7$ GHz band allocated for the constellation's downlink transmissions. The ith channel's center carrier frequency $F_{\mathrm{c}i}$ can be derived as $F_{\mathrm{c}i} = 10.7 + \frac{F}{2} + F_\delta(i - 0.5)$ MHz, where F is the OFDM subcarrier spacing and $F_\delta = 250$ MHz is the Starlink channel spacing. A single Starlink OFDM symbol transmitted over any single channel is composed of $N = 1024$ mutually orthogonal data subcarriers spread across its $F_{\mathrm{s}} = 240$ MHz channel bandwidth, resulting in a subcarrier spacing of $F = 234,375$ Hz. Let k represent a single subcarrier's assigned index such that $k \in \mathcal{K} = \{0, 1, \ldots, N-1\}$, and d_k represent the subcarrier offset in frequency from the carrier frequency in units of subcarriers. The subcarrier offset is defined as $d_k = k$ for $k = 0, 1, \ldots \frac{K}{2} - 1$ and $d_k = k - N$ for $k = \frac{K}{2}, \frac{K}{2} + 1, \ldots, K - 1$. Then, one may calculate ith channel's kth subcarrier center frequency as $F_{\mathrm{c}sik} = F_{\mathrm{c}i} + d_k F$.

In OFDM, the kth subcarrier is modulated by a complex-valued frequency-domain coefficient $x_k \in \mathbb{C}$, which encodes one or more bits of information depending on the OFDM modulation scheme (e.g. 1 for BPSK, 2 for QPSK, 3 for 8PSK, etc.). Let $\boldsymbol{x} \triangleq [x_0, x_1, \ldots, x_{N-1}]^{\mathsf{T}}$ represent the payload of an entire OFDM symbol, which may contain either pilot resources, data resources, or nothing (i.e. $\boldsymbol{x} = \mathbf{0}$). The set of OFDM modulation schemes previously known to be used by Starlink are QPSK, 4QAM, and 16QAM [19]. However, this paper shows as of January 2025, $\pi/4$PSK and 32QAM are now used in addition to the aforementioned modulation schemes. In this paper, we differentiate between QPSK and 4QAM OFDM modulation schemes as follows: whereas QPSK refers to the scheme using x_k drawn from the axis-aligned symbol constellation $\mathcal{C}^{\mathrm{QPSK}} \triangleq \{1, j, -1, -j\}$, 4QAM refers to as the scheme using coefficients drawn from the 45°-rotated constellation $\mathcal{C}^{\mathrm{4QAM}} \triangleq \left\{e^{j\pi/4}, e^{3j\pi/4}, e^{5j\pi/4}, e^{7j\pi/4}\right\}$.

Suppose a single Starlink OFDM symbol is sent over the ith channel with channel coefficients that are constant across frequency and time. The signal experiences LOS time delay τ, phase shift ϕ, and complex additive white Gaussian noise (AWGN) $v_k \sim \mathcal{CN}(0, \sigma^2)$. Assuming negligible intercarrier interference (ICI) due to small Doppler, and intersymbol interference (ISI) due to a sufficiently long cyclic prefix, the baseband received signal y_k is modeled in the frequency domain as

$$y_k = \alpha_k x_k + v_k \tag{1}$$

$$\alpha_k = \sqrt{g}\exp(-j2\pi d_k F\tau + j\phi) \tag{2}$$

where g is the channel gain. The negligible-ICI assumption, inappropriate for the LEO-Earth wireless channel, will be discussed later in this section.

In all wireless OFDM protocols, sequences of OFDM symbols are packaged sequentially in time into groups, commonly referred to as frames. A Starlink downlink frame is composed of $N_{\mathrm{sf}} = 302$ OFDM symbols, each of length $T_{\mathrm{sym}} = 4.4$ μs, plus a frame guard interval T_{fg} for a total frame period of $T_{\mathrm{f}} = 1/750$ s. Let x_{mk} represent the modulation coefficient of the mth OFDM symbol's kth subcarrier, where $m \in \mathcal{M} = \{0, 1, \ldots, N_{\mathrm{sf}} - 1\}$, and let $\boldsymbol{x}_m$ be defined similarly to $\boldsymbol{x}$. All Starlink frames begin with two known OFDM pilot symbols, referred to as the primary synchronization sequence (PSS), which appears at $m = 0$, and the secondary synchronization sequence (SSS), which appears at $m = 1$. Let $\mathcal{M}_{\mathrm{p}} = \{0, 1\}$ represent the subset of $\mathcal{M}$ such that $\boldsymbol{x}_m$ contains known pilot resources, and $\mathcal{M}_{\mathrm{d}} = \{2, 3, \ldots, N_{\mathrm{sf}} - 1\}$ the subset

such that $\boldsymbol{x}_m$ contains unknown data resources. Note that $\mathcal{M} = \mathcal{M}_\text{p} \cup \mathcal{M}_\text{d}$. While the PSS is natively represented in the time domain, the SSS is formatted as a standard QPSK OFDM symbol [19]. The remaining 300 nonzero OFDM symbols $\{\boldsymbol{x}_m \mid 2 \leq m \leq N_\text{sf} - 1\}$ were previously thought to contain unknown data symbols; however, other work has shown that these symbols may be more predictable than previously believed [19], [22], [23]. As such, one could likely construct an *a priori* probability function for the likelihood that any single x_{mk} takes on value $c_\text{sym} \in C$ as

$$P(x_{mk} = c_\text{sym}) \forall c_\text{sym} \in \mathcal{C} \mid \sum_{c_\text{sym} \in \mathcal{C}} P(x_{mk} = c_\text{sym}) = 1 \quad (3)$$

Distillation of (3) for all $m \in \mathcal{M}, k \in \mathcal{K}, c_\text{sym} \in \mathcal{C}$ would likely require mass decoding for a consistency study across hundreds, if not thousands, of Starlink frames.

Now, consider an entire Starlink frame sent over the ith channel that experiences Doppler effects with related Doppler shift frequency F_D. This frequency shift is parameterized as the carrier frequency offset (CFO) parameter $\beta \triangleq -F_\text{D}/F_{ci}$, but note that Doppler effects arising from the considerable relative motion between a Starlink satellite vehicle (SV) and stationary ground receiver are not limited to a Doppler frequency shift—the signal also experiences a time-domain compression/dilation of the baseband signal, which [19] showed was non-negligible for LEO communications constellations transmitting OFDM frames with frame durations on the order of 10^{-3} s and bandwidth on the order of 10^8 Hz. As such, the combined Doppler effects across a Starlink frame with CFO parameter β must be modeled as

$$\tau_m = \tau_0 - mT_\text{sym}\left(\frac{\beta}{1-\beta}\right) \quad (4)$$

$$\approx \tau_0 - mT_\text{sym}\beta \quad (5)$$

$$\phi_m = \phi_0 + mT_\text{sym}\beta F_{ci} \quad (6)$$

where τ_m and ϕ_m are the time delay and phase shift experienced by the mth OFDM symbol, respectively. Accordingly, τ_0 and ϕ_0 are the time delay and phase shift experienced by the OFDM symbol with index $m = 0$, i.e. the PSS. As experimental results show that Starlink channel β-values channel reside on the order of 10^{-6}, the approximation $\frac{\beta}{1-\beta} \approx \beta$ is valid, considering that $|\beta| \ll 1$.

However, such values of β remain too large to assume negligible ICI: assuming a signal transmitted on channel $i = 5$ (with $F_{ci} = 11.825$ GHz) experiences Doppler effects with CFO parameter $\beta = 5 \times 10^{-6}$, the Doppler frequency shift of $F_\text{D} = \beta F_{c5} = 59,125$ Hz is a significant 24.7% of the subcarrier spacing $F = 234,275$ Hz. As such, assume that a receiver first conducts a coarse Doppler acquisition granting *a priori* CFO parameter estimate β_0 such that $\beta = \beta_0 + \delta\beta$, where the error ratio $\delta\beta/\beta_0$ can be as large as approximately 10% in magnitude.

Then, the baseband received signal for the mth OFDM symbol's kth subcarrier modulation y_{mk} can be modeled in the frequency domain as

$$y_{mk} = \alpha_{mk} x_{mk} + v_{mk} \quad (7)$$

$$\alpha_{mk} = \sqrt{g}\exp(-j2\pi d_k F\tau_m + j\phi_m) \quad (8)$$

$$\tau_m = \tau_0 - mT_\text{sym}\delta\beta \quad (9)$$

$$\phi_m = \phi_0 + mT_\text{sym}\delta\beta F_{ci} \quad (10)$$

where x_{mk} is the mth OFDM symbol's kth subcarrier modulation coefficient, AWGN $v_{mk} \sim \mathcal{CN}(0, \sigma^2)$, and the model for τ_m makes a short sequence of approximations under the fact that $\beta_0, \beta_m \ll 1$.

III. Estimation Framework

Following the lead of [24, Sections III, IV], this section derives a generalized ML estimator for the unknown time delay and phase shift experienced by a single, noisy OFDM symbol. It then derives log-likelihood functions specific for respective use in the pilot-only and data-only ML estimators. It then extends the framework described in [24], which does not consider Doppler, by showing how the CFO parameter β can be estimated using a sequence of ML estimates $\hat{\boldsymbol{\theta}}$.

A. Maximum Likelihood Estimation

In general, given a probability distribution function $\Lambda_{\boldsymbol{Y}}(\boldsymbol{\theta}) \triangleq p(\boldsymbol{Y}|\boldsymbol{\theta})$ that describes the likelihood of observing the set of measurements $\boldsymbol{Y} \in \mathbb{R}^{N_y}$ as it relates to unknown parameter vector $\boldsymbol{\theta} \in \mathbb{R}^{N_\theta}$, the ML estimate for $\boldsymbol{\theta}$ is the value of $\boldsymbol{\theta}$ that maximizes $\Lambda_{\boldsymbol{Y}}(\boldsymbol{\theta})$. This is typically written as

$$\hat{\boldsymbol{\theta}}_\text{ML} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \Lambda_{\boldsymbol{Y}}(\boldsymbol{\theta}) \quad (11)$$

Maximization of the log-likelihood function $\log \Lambda_{\boldsymbol{Y}}(\boldsymbol{\theta})$ is often a convenient and equivalent substitute for maximization of the simple likelihood function $\Lambda_{\boldsymbol{Y}}(\boldsymbol{\theta})$ in the framework of ML estimation, as (1) $\Lambda_{\boldsymbol{Y}}(\boldsymbol{\theta})$ often involves products of probability functions and (2) the natural logarithm is a homogeneous function.

$$\hat{\boldsymbol{\theta}}_\text{ML} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \log \Lambda_{\boldsymbol{Y}}(\boldsymbol{\theta}) \quad (12)$$

Suppose a receiver receives and Doppler-precompensates, using an *a priori* value β_0, a noisy OFDM symbol with symbol index $m \in \mathcal{M}$ and frequency-domain representation $\boldsymbol{y}_m \triangleq [y_{m0}, y_{m1}, \ldots, y_{m(N-1)}]^\mathsf{T}$. The signal $\boldsymbol{y}_m$ experiences unknown time delay τ and phase shift ϕ, where the measurement model for y_{mk} is defined by (7-10). In this context, the measurement vector can then be defined as $\boldsymbol{Y} \triangleq \boldsymbol{y}_m$, and the unknown parameter vector as $\boldsymbol{\theta} \triangleq [\tau, \phi]^\mathsf{T}$. Assuming that v_{mk} are independent $\forall\{k_1, k_2 \in \mathcal{K} \mid k_1 \neq k_2\}$, the likelihood function can be written and subsequently decomposed as

$$\Lambda_{\boldsymbol{y}_m}(\boldsymbol{\theta}) = p(\boldsymbol{y}_m|\boldsymbol{\theta}) \quad (13)$$

$$= \prod_{k \in \mathcal{K}} p(y_{mk}|\boldsymbol{\theta}), \quad (14)$$

giving rise to log-likelihood function

$$\log \Lambda_{\boldsymbol{y}_m}(\boldsymbol{\theta}) = \sum_{k \in \mathcal{K}} \log p(y_{mk}|\boldsymbol{\theta}) \quad (15)$$

and ML estimate

$$\hat{\boldsymbol{\theta}}_{\text{ML}} = \underset{\boldsymbol{\theta}}{\text{argmax}} \sum_{k \in \mathcal{K}} \log p(y_{mk}|\boldsymbol{\theta}) \tag{16}$$

With the single-symbol ML estimator now formulated, the log-likelihood function $\log p(y_{mk}|\boldsymbol{\theta})$ needs to be derived under two separate conditions: (1) when $\boldsymbol{x}_m$ carries unknown data resources, and (2) when $\boldsymbol{x}_m$ carries known pilot resources.

B. Log-Likelihood Derivation

First, let τ be normalized by the CFO-compensated OFDM sampling period $T_\text{s} = \frac{1}{(1-\beta_0)F_\text{s}} \approx \frac{1}{F_\text{s}}$, creating $z \triangleq \frac{(1-\beta_0)\tau}{T_\text{s}} \approx \frac{\tau}{T_\text{s}}$ where z is in units of samples.

1) Unknown Data Resources: Consider the case where $\boldsymbol{x}_m$ contains unknown data resources. The likelihood of y_{mk} conditioned on the parameter vector $\boldsymbol{\theta}$ and knowledge of x_{mk} is

$$\begin{aligned} &p(y_{mk}|x_{mk}, \boldsymbol{\theta}) \\ &\quad = \frac{1}{\pi\sigma^2} \exp\left(\frac{-1}{\sigma^2} \left|y_{mk} - \mu_{mk}\nu_k(\boldsymbol{\theta})\right|^2\right), \end{aligned} \tag{17}$$

where $\nu_k(\boldsymbol{\theta}) \triangleq \exp\left(-j2\pi z d_k/N + j\phi\right)$ and $\mu_{mk} \triangleq \sqrt{g}x_{mk}$. Assuming a known prior distribution function for c_sym expressed as $p(x_{mk} = c_\text{sym})$, which is trivially independent from the unknown parameter $\boldsymbol{\theta}$, $p(y_{mk}|\boldsymbol{\theta})$ is decomposed as

$$\begin{aligned} &p(y_{mk}|\boldsymbol{\theta}) \\ &\quad = \sum_{c_\text{sym} \in \mathcal{C}} p(y_{mk}|x_{mk} = c_\text{sym}, \boldsymbol{\theta}) P(x_{mk} = c_\text{sym}) \end{aligned} \tag{18}$$

As determination of $P(x_{mk} = c_\text{sym})$ is not within the scope of this paper, an uniform probability distribution is assumed for the distribution of subcarrier modulations such that $P(c_{\text{sym}i}) = P(c_{\text{sym}j}) = \frac{1}{|\mathcal{C}|}$ for all $c_{\text{sym}i}, c_{\text{sym}j} \in \mathcal{C}$. The likelihood of y_{mk} is then given by

$$p(y_{mk}|\boldsymbol{\theta}) = \frac{1}{|\mathcal{C}|} \sum_{c_\text{sym} \in \mathcal{C}} p(y_{mk}|x_{mk} = c_\text{sym}, \boldsymbol{\theta}), \tag{19}$$

which can be recognized as a Gaussian mixture distribution function. Following the steps taken in [24, Eq. 19], the log-likelihood then becomes

$$\begin{aligned} &\log p(y_{mk}|\boldsymbol{\theta}) \\ &= \log \frac{1}{|\mathcal{C}|} + \log \sum_{c_\text{sym} \in \mathcal{C}} p(y_{mk}|x_{mk} = c_\text{sym}\, \boldsymbol{\theta}) \end{aligned} \tag{20}$$

$$\begin{aligned} &= \log \frac{1}{\pi\sigma^2|\mathcal{C}|} \\ &\quad + \log \sum_{c_\text{sym} \in \mathcal{C}} \exp\left(\frac{-1}{\sigma^2} \left|y_{mk} - \mu_{mk}\nu_k(\boldsymbol{\theta})\right|^2\right) \end{aligned} \tag{21}$$

$$\begin{aligned} &= \log \frac{1}{\pi\sigma^2|\mathcal{C}|} - \frac{1}{\sigma^2}|y_{mk}|^2 \\ &\quad + \log \sum_{c_\text{sym} \in \mathcal{C}} \exp\Bigg(\frac{1}{\sigma^2}\big(2\Re\{y_{mk}^* \mu_{mk} \nu_k(\boldsymbol{\theta})\} \\ &\qquad\qquad - g|c_\text{sym}|^2\big)\Bigg). \end{aligned} \tag{22}$$

2) Known Pilot Resources: Now consider the case where $\boldsymbol{x}_m$ contains pilot resources, i.e. $x_{mk} = p_{mk}$ where p_{mk} is known *a priori*. Then, $\log p(\boldsymbol{y}|\boldsymbol{\theta})$ is evaluated in a straightforward manner:

$$\log p(y_{mk}|\boldsymbol{\theta}) \tag{23}$$

$$= \log\left(\frac{1}{\pi\sigma^2} \exp\left(\frac{-1}{\sigma^2} \left|y_{mk} - \mu_{mk}\nu_k(\boldsymbol{\theta})\right|^2\right)\right) \tag{24}$$

$$\begin{aligned} &= \log \frac{1}{\pi\sigma^2} + \frac{1}{\sigma^2}\big(2\Re\{y_{mk}^* \mu_{mk} \nu_k(\boldsymbol{\theta})\} \\ &\qquad\qquad - |y_{mk}|^2 - g|p_{mk}|^2\big) \end{aligned} \tag{25}$$

where ν_k maintains the previous definition but $\mu_{mk} \triangleq \sqrt{g}p_{mk}$.

C. Least Squares CFO Estimation

While [24] proposes a full-frame ML estimator that grants a unified estimate for τ and ϕ per frame, it fails to estimate Doppler shift F_D. A natural extension of the approach taken in [24] would augment measurement models [24, Eqs. 2, 3] with an additional phase adjustment coefficient of $\exp(j2\pi F_\text{D})$ and carry through the same analysis to arrive at an adjusted ν_k formulation. Otherwise, the ML estimation technique remains the same as discussed. However, consider that ML estimators are often implemented using grid searches when input data is noisy, the likelihood surface is irregular, or both. In the context of a grid search, the above approach adds a factor of N_β to the search space complexity $N_\tau \cdot N_\phi$, where N_β, N_τ, and N_ϕ are the number of elements in the β, τ, and ϕ search spaces. Instead, this paper proposes a two-stage process that first makes a series of estimates for ϕ and τ, then stitches estimates together across an entire frame frame to make a least-squares (LS) estimate of a unitary β for the frame. This process also generates the unknown parameters τ_0 and ϕ_0.

Suppose an entire Starlink frame is pre-compensated with β_0, after which the single-symbol estimator is individually applied to all $\boldsymbol{y}_m \forall m \in \mathcal{M}$ to construct a sequence of estimates for time delay $\hat{\boldsymbol{\tau}} \triangleq [\hat{\tau}_0, \hat{\tau}_1, \ldots, \hat{\tau}_{N-1}]^\mathsf{T}$ and phase shifts $\hat{\boldsymbol{\phi}} \triangleq [\hat{\phi}_0, \hat{\phi}_1, \ldots, \hat{\phi}_{N-1}]^\mathsf{T}$. Assume that the estimates are subject to AWGN error $\boldsymbol{v}_\tau \sim \mathcal{N}(\boldsymbol{0}, \sigma_\tau^2 I)$ and $\boldsymbol{v}_\phi \sim \mathcal{N}(\boldsymbol{0}, \sigma_\phi^2 I)$, where $\boldsymbol{v}_\tau, \boldsymbol{v}_\phi \in \mathbb{R}^{|\mathcal{M}|}$.

Based on the dynamics model (9, 10), one can then form a measurement model for $\hat{\boldsymbol{\tau}}$ and $\hat{\boldsymbol{\phi}}$ based on unknown parametersτ_0, ϕ_0, and $\delta\beta$.

$$\tau_m = \tau_0 - mT_\text{sym}\delta\beta + v_{\tau m} \tag{26}$$

$$\phi_m = \phi_0 + mT_\text{sym}\delta\beta F_{\text{c}i} + v_{\phi m}, \tag{27}$$

where the $(\cdot)_m$ subscript indicates the mth element of $(\cdot)$ and $\hat{(\cdot)}$ notation has been dropped for clarity.

Measurement models (26) and (26) may then be used to form LS estimators for τ_0, ϕ_0 and $\delta\beta$. While it is possible for $\delta\beta$ to be jointly estimated via (26) and (26), practical experimentation showed that solely using $\boldsymbol{\phi}$ for LS $\delta\beta$-estimation generally provided for better results. As such, the LS estimator is formed in a two-step process: First, $\boldsymbol{\phi}$ is used to make LS estimates for ϕ_0 and $\delta\beta$. Second, $\delta\hat{\beta}$ is assumed in estimation of τ_0.

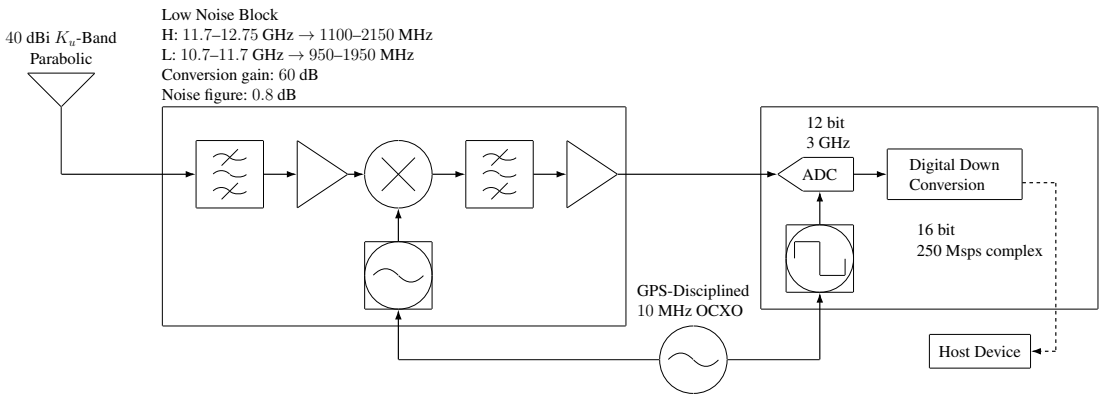


Fig. 2: Block diagram of the Starlink signal capture system.

The LS estimator for ϕ_0 and $\delta\beta$ is

$$\begin{bmatrix}\hat{\phi}_0 \\ \delta\hat{\beta}\end{bmatrix} = (H^{\mathsf{T}}H)^{-1}H^{\mathsf{T}}\boldsymbol{\phi} \tag{28}$$

$$H = \begin{bmatrix}1 & 0 \\ 1 & T_{\text{sym}}F_{\text{c}i} \\ \vdots & \vdots \\ 1 & (N_{\text{sf}}-1)T_{\text{sym}}F_{ci}\end{bmatrix}, \tag{29}$$

and the LS estimator for τ_0, taking $\delta\hat{\beta}$ as *a priori* from (28), is The LS estimator for ϕ_0 and $\delta\beta$ is

$$\hat{\tau}_0 = (H^{\mathsf{T}}H)^{-1}H^{\mathsf{T}}(\boldsymbol{\tau} - \boldsymbol{b}) \tag{30}$$

$$H = \mathbf{1}_{|\mathcal{M}|} \tag{31}$$

$$\boldsymbol{b} = \delta\hat{\beta}\left[0, 1, \ldots, N_{\text{sf}}-1\right]^{\mathsf{T}}, \tag{32}$$

where $\mathbf{1}_{|\mathcal{M}|}$ is the 1-vector of size $|\mathcal{M}| \times 1$.

IV. Signal Capture

In this paper, live Starlink transmissions are used to implement a tangible proof-of-concept demonstration of the ML estimation framework proposed by [24]. This section outlines the signal capture setup used to receive and record the signals used for software testing and validation.

Fig. 2 shows a block diagram of the signal capture system. The Starlink capture chain begins with an offset parabolic dish antenna mounted on an azimuth-elevation actuated mount that uses publicly accessible Starlink ephemerides in the form of two-line elements to steer the antenna boresight towards Starlink SVs passing overhead our signal capture site in Austin, Texas.

To avoid possible estimation performance issues stemming from low SNRs, the results presented in this paper use only dominant signals from assigned beams possessing signal-to-noise ratios (SNRs) of 16 dB or higher, using the parlance of [21]. However, future work could show the performance of the proposed estimator under varying SNR conditions to verify the theoretical and simulated estimation bounds proposed in [24]. The dish antenna's approximately 3° beamwidth helps ensure that signals captured mostly originate from a single SV. Any inadvertently-captured side beams presented as secondary signals with typical SNRs -10 dB or lower than those of assigned beams, greatly simplifying the process of isolating a composite signal's dominant signal.

The antenna focuses captured signals onto a feedhorn connected to a low-noise block (LNB) with a conversion gain of 60 dB and a noise figure of 0.8 dB. The LNB is dual-band, downconverting either 10.7–11.7 GHz (the lower band) to 950–1950 MHz, or 11.7–12.75 GHz (the upper band) to 1100–2150 MHz. The antenna's nominal gain is 40 dBi at 12.5 GHz, but there are losses of at least 4-5 dB due to lack of a circular-to-linear polarizer and feedhorn misalignment. A downstream radio frequency signal analyzer (RFSA) conducts further bandpass filtering, 16-bit complex sampling, and digital downconversion with a total bandwidth of $F_{\text{sr}} = 250$ Msps, although only ∼200 Msps are usable due to the RFSA's internal front-end filtering. Note that the same GPS-disciplined 10 MHz oven-controlled crystal oscillator was used to simultaneously drive the Ku-band to L-band downmixing performed by the LNB and the wideband sampling performed by the RFSA.

V. Implementation

This section opens with a description of two key challenges to be addressed, then details all major portions of the ML estimation framework's software implementation.

A. Key Challenges

Before delving into the specifics of software implementation, several key challenges in applying the theoretical concepts presented in Section III to live signals are highlighted. The approaches taken to address these issues are mentioned before further elaboration in Sections V-B through V-F.

1) Unknown Modulation Scheme: The framework presented by Section III generally assumes *a priori* knowledge of any $\boldsymbol{x}_m$'s corresponding modulation scheme $\mathcal{C}_m$. For opportunistic receiver of Starlink signals, both these assumptions are false. Starlink is known to begin all frames with $\mathcal{C}_m = \mathcal{C}^{\text{QPSK}}$ for $m = \{1, 2, 3, 4, 5\}$ [2], but its behavior $\forall m \geq 6$ is known to be unpredictable. To provide the data-only ML estimator with the required $\mathcal{C}_m$, a modulation detector is implemented, as detailed in Section V-D.

2) Phase Estimate Ambiguity: Close inspection of the data-only likelihoood function formulation (19) reveals that there could exist phase ambiguities of some magnitude θ_{a} in $\hat{\phi}$, depending on the shape of $\mathcal{C}_m$. Note that calculate of the log-likelihood function involves a sum of $p(y_{mk}|x_{mk} = c_{\text{sym}}, \boldsymbol{\theta})$ across all $c_{\text{sym}} \in \mathcal{C}$, necessarily losing information regarding which exact value of c_{sym} was the greatest contributor to $p(y_{mk}|\boldsymbol{\theta})$. Say, for example, that $\mathcal{C}_m = \mathcal{C}^{\text{BPSK}} = \{-1, 1\}$. Then,

$$\begin{aligned} p(y_{mk}|\tau, \phi) &= \frac{1}{2}\big[p(y_{mk}|x_{mk} = -1, \tau, \phi) \\ &\quad + p(y_{mk}|x_{mk} = 1, \tau, \phi)\big] \qquad (33) \\ &= \frac{1}{2}\big[p(y_{mk}|x_{mk} = 1, \tau, \phi + \pi) \\ &\quad + p(y_{mk}|x_{mk} = -1, \tau, \phi + \pi)\big] \qquad (34) \\ p(y_{mk}|\tau, \phi) &= p(y_{mk}|\tau, \phi + \pi), \qquad (35) \end{aligned}$$

thus granting a $\theta_{\text{a}} = \pi$ phase estimation ambiguity in $\hat{\phi}$. One can think of this as rotating the constellation diagram of $\boldsymbol{y}_m$ to best fit the constellation shape of $\mathcal{C}_m$, without further care for the particular orientation of $\boldsymbol{y}_m$ within the confines of $\mathcal{C}_m$'s constellation shape.

For the QPSK, 4QAM, 16QAM, and 32QAM modulation schemes, this phenomenon results in a $\pi/2$ phase estimate ambiguity. For the 8PSK scheme, a $\pi/4$ phase estimate ambiguity arises. The presence of θ_{a} limits the entire estimator's pull-in range for β estimation, as $\beta F_{\text{c}i} T_{\text{sym}} \ll \theta_{\text{a}}$ is needed to prevent inter-symbol cycle slips. Hard decoding decisions may be used to resolve the issue of phase ambiguity.

B. Signal Preprocessing

Starlink signals are captured through the system described in Section IV. Before the estimation framework described in Section III can be applied, some signal preprocessing is required. First, the captured signal is resampled at the Starlink information symbol rate $F_{\text{s}} = 240$ MHz such that all downstream processing may proceed with minimal ambiguity.

As stated in Section II, application of the single-symbol ML estimators outlined in Section III requires Doppler-precompensation of y_{mk} to minimize ICI. As such, a local replica p_k consisting of the coherent time-domain concatenation of the PSS and SSS is used to conduct preliminary acquisition using a standard cross-ambiguity function (CAF) [25], granting coarse initial Doppler and TOA estimates β_{0n} and t_{*0n} for the nth frame originating from the dominant signal of a short, seconds-long capture. This process is outlined in Algorithm 1.

The ML estimation framework operates on a single Starlink frame at once. To isolate a single frame's signal, signal data corresponding to time range $[t_{*0}, t_{*0} + T_{\text{f}}]$ is extracted from a seconds-long signal recording. To pre-compensate, this single-frame signal is then resampled from F_{s} to $(1 - \beta_0)F_{\text{s}}$, and shifted by $F_{\text{D}} = \beta F_{\text{c}i}$, where i is found by matching the recording's center frequency F_{cr} to known Starlink channel bands as published in [19]. This grants a Doppler-compensated time domain signal corresponding to a single Starlink frame. Note that while some CFO error still exists such that $|\delta\beta| > 0$, this level of Doppler precompensation sufficiently mitigates the effects of ICI.

Finally, the standard OFDM steps of (1) stripping the cyclic prefix and (2) taking a fast Fourier transform of the captured data are performed, granting $\boldsymbol{y}_m$.

Algorithm 1: Coarse Doppler-TOA Acquisition

Input: Captured signal $\boldsymbol{y}$, recording sampling rate F_{sr}, recording center frequency F_{cr}, Doppler search space F_{Dsearch}
Output: Estimated frame start index k_0, estimated CFO parameter β_0

```
Generate local synchronization sequence p_k;
Define time vector t_{p_k} based on sampling rate;
Initialize correlation tracking arrays;
foreach Doppler shift F_D in F_Dsearch do
    Apply frequency shift: p_k ← p_k · exp(j2πF_D t_{p_k});
    Compute cross-correlation of y with p_k;
    Extract lag values and correlation peaks;
    Store maximum correlation and corresponding lag index;
end
Find Doppler shift F_D that maximizes correlation;
Compute frame start index k_0 by scaling lag index;
Compute F_ci from F_cr; Compute β_0 = −F_D/F_ci;
return k_0, β_0;
```

C. Pilot Symbol Estimator

The pilot symbol estimator is used to generate $\hat{\boldsymbol{\theta}}_m = [\hat{z}_m, \hat{\phi}_m]^{\mathsf{T}}$ for the first two symbols of each Starlink frame, $\boldsymbol{y}_0$ and $\boldsymbol{y}_1$. First, the channel gain g is estimated as the median per-subcarrier power and the PSS and SSS are scaled to obtain $\mu_{mk} = \sqrt{g} p_{mk}$. Then, the log-likelihood $\log \Lambda_{\boldsymbol{y}_m}(\boldsymbol{\theta})$ is computed according to (22) for each member of a two-dimensional z-ϕ search space. Note that for the pilot symbol estimator, the search space defined as between ± 2 samples in the z-axis, and $\pm 2\pi$ rad in the ϕ-axis. An example empirical log-likelihood for the pilot symbol estimator is shown in the top plot of Fig. 3. Finally, $\hat{z}_{\text{ML}}$ and $\hat{\phi}_{\text{ML}}$ are taken as the pair corresponding to the maximum value of $\log \Lambda_{\boldsymbol{y}_m}(\boldsymbol{\theta})$ and stored in $\hat{\boldsymbol{z}}$ and $\hat{\boldsymbol{\theta}}$. The processes described in Sections V-C through are outlined in Algorithm 1.

D. Modulation Scheme Detector

As previously discussed, $\mathcal{C}_m$ must be detected per-symbol before data ML estimation can proceed. First, $\boldsymbol{y}_m$ is rendered in constellation and adjusted using the pilot-estimated subsample delay $\hat{z}_1$ as

$$\boldsymbol{y}_m \leftarrow \boldsymbol{y}_m \exp\left(-j2\pi \frac{\hat{z}_1 d_k}{N}\right), \qquad (36)$$

which helps to de-scatter the constellation representation. This effect can be observed in the transition between the left and middle plots of Figs. 4 and 5. This step is necessary for reliable differentiation between high-BPS modulation schemes, particularly 16QAM and 32QAM, for which even a subsample delay

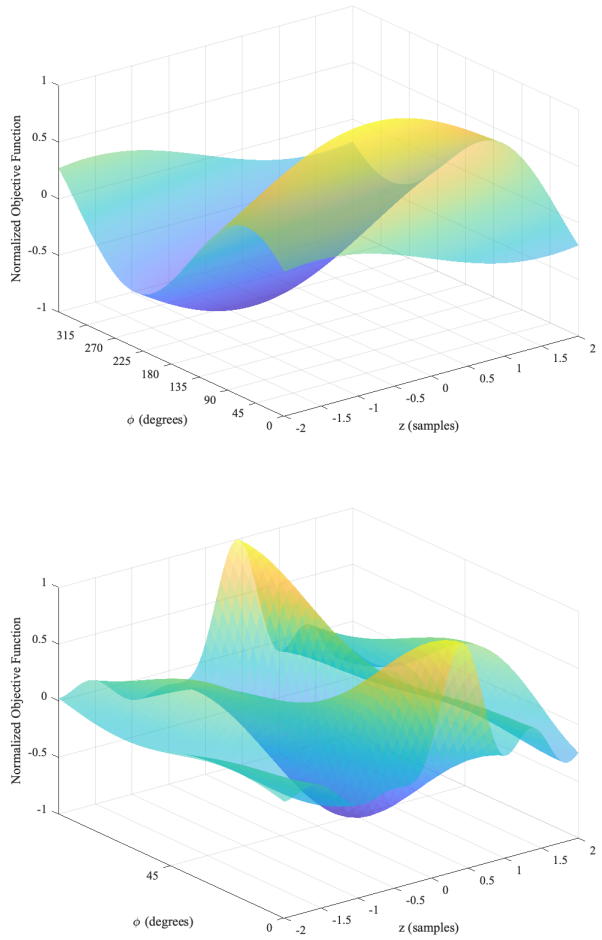


Fig. 3: Empirical log-likelihood functions for a single pilot symbol (top) and single data symbol (bottom). Note that the data symbol's log-likelihood function is evaluated only over $0 \leq \phi < 90°$, whereas the pilot symbol is evaluated over $0 \leq \phi < 360°$. Nonetheless, in both cases, only a single peak is observed.

as small as $z = 0.1$ could render the constellation diagram too noisy for precise modulation determination. For each possible BPS value $n_{\text{bps}} \in \mathcal{N}_{\text{bps}} \triangleq \{2, 3, 4, 5\}$, k-means clustering is conducted with $k = 2^{n_{\text{bps}}}$ with 20 replicates, and the clustering performance is rated using the median silhouette score (MSS), which measures the overall agreement of $\boldsymbol{y}_m$ with its assigned cluster map [26]. MSS is calculated as

$$s[k] = \frac{b[k] - a[k]}{\max(a[k], b[k])} \tag{37}$$

$$\text{MSS} = \text{med}(s) \tag{38}$$

where:

- $s[k]$ is the silhouette score for point y_{mk}.
- $a[k]$ is the average distance from y_{mk} to all other points in its own cluster.
- $b[k]$ is the minimum average distance from y_{mk} to points in different clusters, minimized over clusters.

The $\mathcal{C}_m$ was chosen based on the n_{bps} with maximum MSS according to the following table:

Bits Per Subcarrier	Modulation Scheme
2	QPSK/4QAM
3	8PSK
4	16QAM
5	32QAM

TABLE I: Designation of modulation Scheme ccording to optimal BPS.

When $n_{\text{bps}} = 2$, $\boldsymbol{x}_m$ could adopt either a 4QAM or QPSK modulation scheme since both schemes use the same BPS. To resolve this, the algorithm initially assumes the modulation remains unchanged from the last occurrence of either 4QAM or QPSK, i.e. ML estimation is conducted under the assumption $\mathcal{C}_m = \mathcal{C}_M \in \{\mathcal{C}^{\text{QPSK}}, \mathcal{C}^{\text{4QAM}}\}$, where $\mathcal{C}_M$ is the most recent occurrence of QPSK or 4QAM modulation.

If this assumption is incorrect, the estimated phase of the current symbol $\hat{\phi}_m$ should exhibit a characteristic phase shift of approximately $\pi/4$ rad relative to $\hat{\phi}_{m-1}$. To verify whether a modulation change has occurred, the angle difference $\Delta\theta_{\pi/2}(\hat{\phi}_m, \hat{\phi}_{m-1})$ is measured after data symbol estimation, where $\Delta\theta_{\pi/2}$ is the angle difference function under $\pi/2$ wrapping as defined below.

$$\Delta\theta_{\pi/2}(\theta_1, \theta_2) = \text{mod}\left(\theta_2 - \theta_1 + \frac{\pi}{4}, \frac{\pi}{2}\right) - \frac{\pi}{4} \tag{39}$$

If $\Delta\theta_{\pi/2}(\hat{\phi}_m, \hat{\phi}_{m-1})$ is greater than a predefined tolerance θ_{tol}, $\hat{\phi}_m$ and $\mathcal{C}_m$ are retroactively adjusted to fit the new hypothesis.

E. Data Symbol Estimator

The data symbol estimator takes essentially the same steps as the pilot symbol estimator with two exceptions: (1) An additional software loop sums over all possible $c_{\text{sym}} \in \mathcal{C}_m$ to calculate $\log \Lambda_{\boldsymbol{y}_m}(\boldsymbol{\theta})$, formulated according to (25). This is practically implemented as a matrix summation for speed. (2) The search space is defined as between $\pm\pi/2$ rad in the ϕ-axis, because $p(y_{mk}|\tau, \phi) = p(y_{mk}|\tau, \phi + n\pi/2)$ for 4QAM, QPSK, 16QAM, 32QAM as previously discussed. An example empirical log-likelihood for the data symbol estimator is shown in the bottom plot of Fig. 3. Together, the modulation scheme detector and data symbol estimator are run per-symbol for all data symbols, granting $\hat{\boldsymbol{\theta}}_m = [\hat{z}_m, \hat{\phi}_m]^{\mathsf{T}} \forall m \in \{2, 3, \ldots, N_{\text{sf}} - 1\}$.

F. Doppler-TOA Extraction

Remembering that $\hat{\phi}$ exhibits an $n\pi/2$ ambiguity, $\hat{\boldsymbol{\theta}} = [\hat{\boldsymbol{z}}, \hat{\boldsymbol{\phi}}]^{\mathsf{T}}$ are first preprocessed to eliminate outliers and unwrap the phase estimate before use for Doppler-TOA extraction. Outliers are identified via a RANSAC one-dimensional polynomial fit on $\boldsymbol{z}$. Use of the phase-wrapped $\hat{\boldsymbol{\phi}}$ (blue circles in Fig. 6) for outlier identification is unfavorable, as the data follows a sawtooth waveform with sharp jumps at $\hat{\phi} = \pi/2$. Outliers often arose due to insufficient SNR, leading to misidentification of the modulation scheme and/or improper realignment of $\boldsymbol{y}_m$'s constellation with that of $\mathcal{C}_m$. Examples

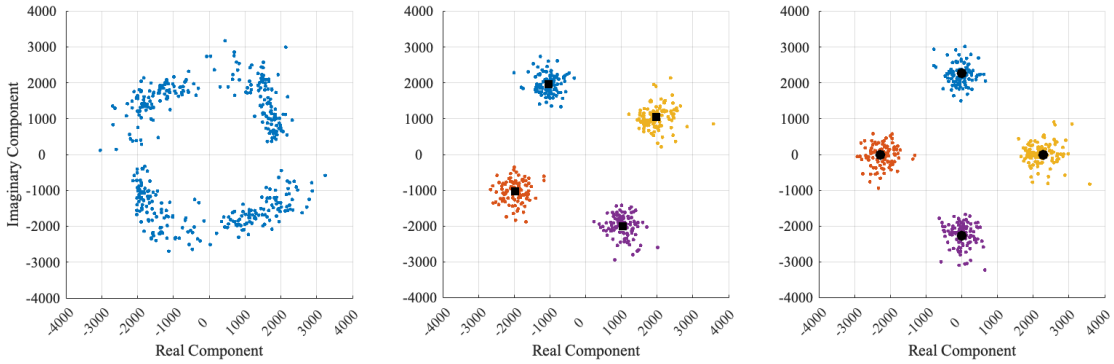


Fig. 4: Tripanel constellation representation of single OFDM symbol $\boldsymbol{y}_m$ with $\mathcal{C}_m = \mathcal{C}^{\text{QPSK}}$ after standard OFDM processing (left), pre-compensation with $\hat{z}_1$ from SSS ML estimate (center), and after correction with $\hat{z}_m$ and $\hat{\phi}_m$ (right). Colors indicate k-means clusters. In the middle panel, the black squares show the centroid of each k-means cluster. In the right panel, the black circles show the position of $\mathcal{C}_m$.

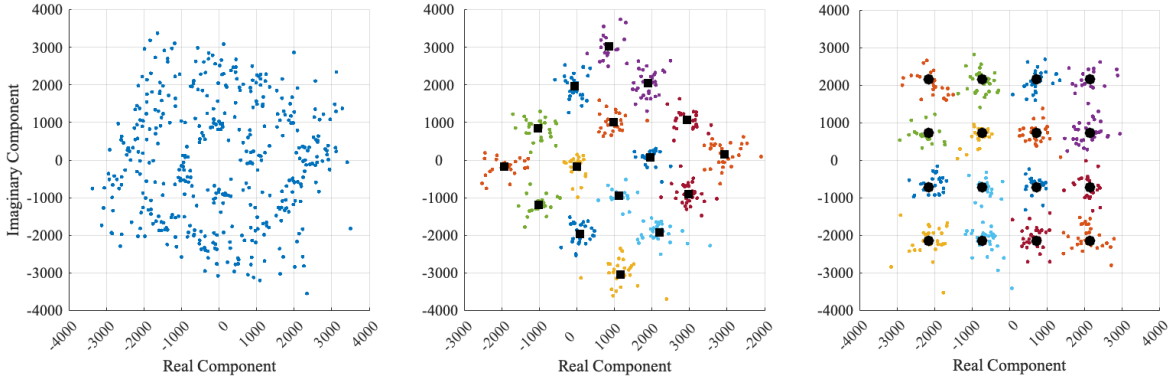


Fig. 5: As Fig. 4 but for an OFDM symbol where $\mathcal{C}_m = \mathcal{C}^{\text{16QAM}}$.

of excluded data can be seen in Fig. 6 as red x-marks. The remaining data $\hat{\boldsymbol{\phi}}$ is then unwrapped with a $\pi/2$ wrapping point.

Using $\hat{\boldsymbol{\theta}}_m = [\hat{z}_m, \hat{\phi}_m]^{\mathsf{T}}$, the theory presented by Section III-C is equivalently implemented as LS one-dimensional polynomial fitting. The fit trends can be seen in Fig. 6 as yellow lines. These steps are outlined in Algorithm 3. This process generates the time delay correction $\hat{\tau}_0$ and CFO parameter correction $\delta\hat{\beta}$, which may be used to refine the initial estimates given by coarse acquisition.

$$\hat{t}_* = t_{*0} - \hat{\tau}_0 \tag{40}$$

$$\hat{\beta} = \beta_0 + \delta\hat{\beta} \tag{41}$$

The CFO parameter estimate $\hat{\beta}$ from the nth frame is used to conduct signal preprocessing for the $(n+1)$st frame, preventing uncontrolled drift of β away from β_0.

VI. Results

A. New Starlink OFDM Modulation Schemes

Two OFDM modulation schemes were observed that have not yet been published, to the author's knowledge. 32QAM was regularly observed in captures ranging from September 2024 to January 2025, as shown in Fig. 10.

Further, what initially appeared to be 8PSK modulation in a capture taken on January 13, 2025 turned out to be subcarrier-sliced $\pi/4$, as can be appreciated from Fig. 11.

B. Doppler-TOA Tracking Performance

Using 1.7 seconds of STARLINK-1274's downlink signal data as captured on January 13, 2025, the Doppler-TOA tracking performance of the implemented ML estimator is compared against two other candidates: (1) the standard CAF-based approach using PSS+SSS, and (2) an ML estimation approach using PSS+SSS. The latter was calculated by excluding $\hat{z}_m$ and $\hat{\phi}_m$ $\forall m \in \mathcal{M}_{\text{d}}$, resulting in $\hat{\boldsymbol{z}}, \hat{\boldsymbol{\phi}} \in \mathbb{R}^{2\times 1}$.

Algorithm 2: ML Estimation of Delay and Phase

Input: Captured signal $\boldsymbol{y}$, valid indices k_{valid}
Output: Delay estimate $\hat{\boldsymbol{z}}$, Phase estimate $\hat{\boldsymbol{\theta}}$
Perform OFDM processing on $\boldsymbol{y}$;
Shorten subcarriers to valid indices k_{valid};
Pilot symbol ML estimation:
Estimate channel gain g;
Set search spaces Z and Φ;
foreach *symbol* $\boldsymbol{y}_m \mid m \in \{0, 1\}$ **do**
  foreach $z \in Z$ **do**
    foreach $\phi \in \Phi$ **do**
      Compute $\log \Lambda_{\text{pilot}}(\boldsymbol{y}_m, z, \phi)$;
    end
  end
end
Find and save maximum likelihood estimates $\hat{z}$ and $\hat{\phi}$ to $\hat{\boldsymbol{z}}[m]$ and $\hat{\boldsymbol{\theta}}[m]$;
Data symbol ML estimation:
Set search spaces Z and Φ;
foreach *symbol* $\boldsymbol{y}_m \mid m \in \{2, 3, \ldots, N_{sf} - 1\}$ **do**
  Determine modulation scheme $\mathcal{C}_m$; Estimate channel gain g;
  foreach $z \in Z$ **do**
    foreach $\phi \in \Phi$ **do**
      foreach $c_{sym} \in \mathcal{C}_m$ **do**
        Increment $\log \Lambda_{\text{data}}(\boldsymbol{y}_m, z, \phi)$;
      end
    end
  end
  Find and save maximum likelihood estimates $\hat{z}$ and $\hat{\phi}$ to $\hat{\boldsymbol{z}}[m]$ and $\hat{\boldsymbol{\theta}}[m]$;
  Handle 90-degree phase ambiguity adjustments;
end

Then, to estimate β for each frame, the LS approach is not applied, as a one-dimensional line fit is only minimally identifiable with just two measurements. These tracking results are shown in Figs. 8 and 9. To evaluate the estimate variation, a second-order polynomial was fit to the trace of each estimate parameterized by the frame slot number, which identifies a frame's position within the whole FAI and is labeled as N_{ag} in Fig. 7. The post-fit residual RMSEs were evaluated and are presented Table II.

While the improvement in F_D estimation is substantial, the variance in the t_* estimate shows little reduction. Notably, the t_* estimate measured manifests timing issues present upon transmission by the Starlink SV, which are known to manifest abrupt adjustments and significant jitter, even under nominal behavior in the parlance of [21]. Further analysis using a two-receiver capture system, both with known locations, and a time difference of arrival approach could possibly better reveal the precision improvements afforded by the ML technique presented.

While these RMSE values are not yet close to the theoretical Cramer-Rao bounds reported in [21], there is still practical improvement over the current standard CAF-based methods.

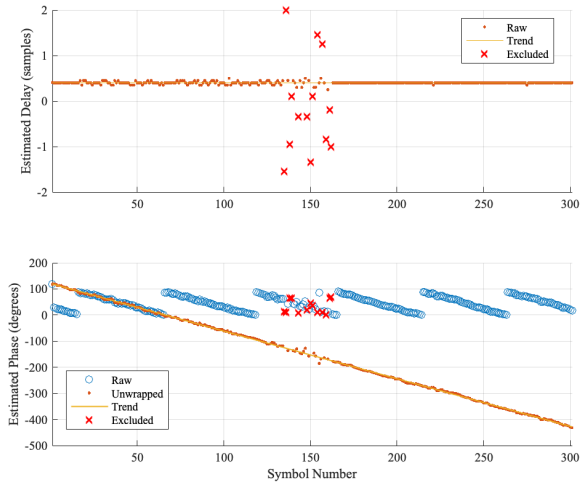


Fig. 6: Example full-frame results for $\hat{\boldsymbol{z}}$ (top), $\hat{\boldsymbol{\phi}}$ (bottom). Note the excluded estimates in red, which all originate from 32QAM modulation schemes, as well as the the sawtooth shape of the wrapped $\hat{\boldsymbol{\phi}}$. The unwrapped $\hat{\boldsymbol{\phi}}$ is seeded with value $\hat{\phi}_1$, the phase shift estimated using the PSS-based pilot symbol estimator, which does not exhibit the $\pi/2$ phase ambiguity characteristic of the data symbol estimator.

Algorithm 3: Extract Doppler-Delay Adjustments

Input: Delay estimate $\hat{\boldsymbol{z}}$, Phase estimate $\hat{\boldsymbol{\theta}}$
Output: Initial sample offset z_0, initial phase offset ϕ_0, CFO parameter adjustment $\delta\beta$
Remove outliers:
Perform RANSAC on $\hat{\boldsymbol{z}}$ to get m_{inlier};
Invert m_{inlier} to get m_{excluded};
Remove excluded estimates from $\boldsymbol{\phi}$, $\boldsymbol{z}$;
LS estimation:
Perform polyfit($\hat{\boldsymbol{z}}$) to get fit coefficients p_z ;
$\boldsymbol{\phi}$ = unwrap90($\boldsymbol{\phi}$) ;
Perform polyfit($\hat{\boldsymbol{\phi}}$) to get fit coefficients p_ϕ ;
Compute F_{ci} from F_{cr}; Compute $\delta\beta = -F_{\text{D}}/F_{\text{ci}}$;
Set $z_0 = p_z(2)$;
Set $\phi_0 = p_\phi(2)$;

The marked improvement of the full-frame ML estimator over the pilot-only ML estimator shows the value of harnessing the full frame to estimate Doppler and TOA.

VII. Conclusion

This paper presents a maximum likelihood Doppler and time-of-arrival estimation framework for opportunistic tracking of Starlink signals, extending previous work by incorporating both pilot and data symbols into the estimation process. Through theoretical derivation, practical implementation, and empirical validation using live Starlink signals, we demonstrate a significant improvement in Doppler estimation accuracy compared to standard pilot-only approaches. Our results show that full-frame ML estimation substantially reduces

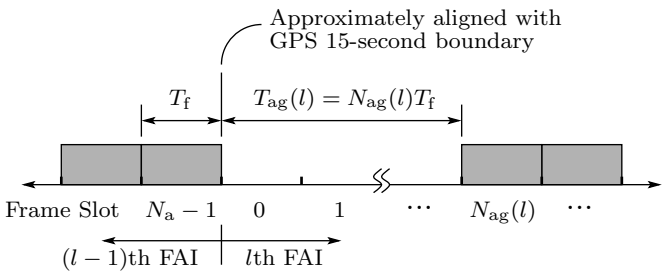


Fig. 7: Frame sequence timing diagram showing the transition from the $(l-1)$th FAI to the lth FAI.

Method	F_D RMSE (Hz)	t_* RMSE (ns)
Pilot-Based CAF	1469.20	2.090
Pilot-Only ML	752.43	1.629
Full-Frame ML	6.34	1.626

TABLE II: Post-fit residual RMSEs by estimation method.

Doppler estimation RMSE, highlighting the advantage of leveraging data payloads in addition to known pilot symbols. However, while TOA estimation benefits from the proposed approach, its improvement remains modest, suggesting the need for further refinements such as super-resolution decision-directed techniques. Future work will focus on refining the estimation process under varying SNR conditions and investigating advanced data decoding strategies to further enhance positioning accuracy.

Acknowledgments

Research was supported by the Department of Defense through the National Defense Science and Engineering Graduate (NDSEG) Fellowship Program, the U.S. Department of Transportation under Grant 69A3552348327 for the CARMEN+ University Transportation Center, and by affiliates of the 6G@UT center within the Wireless Networking and Communications Group at The University of Texas at Austin.

References

[1] "Number of active satellites from 1957 to 2022," https://www.statista.com/statistics/897719/number-of-active-satellites-by-year/, 2023, accessed: 2023-10-20.

[2] T. E. Humphreys, "Interference," in *Springer Handbook of Global Navigation Satellite Systems*. Springer International Publishing, 2017, pp. 469–503.

[3] R. Mitch, R. Dougherty, M. Psiaki, S. Powell, B. O'Hanlon, J. Bhatti, and T. E. Humphreys, "Know your enemy: Signal characteristics of civil GPS jammers," *GPS World*, Jan. 2012.

[4] E. C. Dolman, "New frontiers, old realities," *Strategic Studies Quarterly*, vol. 6, no. 1, pp. 78–96, 2012.

[5] W. J. Broad and D. E. Sanger, "China tests anti-satellite weapon, unnerving U.S." https://www.nytimes.com/2007/01/18/world/asia/18cnd-china.html, January 2007, accessed: 2023-11-1.

[6] S. Gebrekidan, K. R. Lai, P. Robles, and J. White, "Why GPS is under attack," https://www.nytimes.com/interactive/2024/07/02/world/gps-threats.html, July 2024, accessed: 2024-10-23.

[7] R. Huntley, "GNSS jamming and spoofing are a daily occurrence," Oct 2024. [Online]. Available: https://www.eetimes.eu/gnss-jamming-and-spoofing-are-a-daily-occurrence/

[8] P. A. Iannucci and T. E. Humphreys, "Fused low-Earth-orbit GNSS," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 1–1, 2022.

[9] L. Ries, M. C. Limon, F.-C. Grec, M. Anghileri, R. Prieto-Cerdeira, F. Abel, J. Miguez, J. V. Perello-Gisbert, S. D'Addio, R. Ioannidis, A. Ostillio, M. Rapisarda, R. Sarnadas, and P. Testani, "LEO-PNT for augmenting europe's space-based PNT capabilities," in *2023 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, 2023, pp. 329–337.

[10] Z. M. Kassas, N. Khairallah, and S. Kozhaya, "Ad astra: Simultaneous tracking and navigation with megaconstellation LEO satellites," *IEEE Aerospace and Electronic Systems Magazine*, 2024.

[11] M. Neinavaie and Z. M. Kassas, "Cognitive sensing and navigation with unknown OFDM signals with application to terrestrial 5G and Starlink LEO satellites," *IEEE Journal on Selected Areas in Communications*, 2023.

[12] N. Jardak and R. Adam, "Practical use of Starlink downlink tones for positioning," *Sensors*, vol. 23, no. 6, p. 3234, 2023.

[13] C. Yang and A. Soloviev, "Starlink Doppler and Doppler rate estimation via coherent combining of multiple tones for opportunistic positioning," in *2023 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. IEEE, 2023, pp. 1143–1153.

[14] S. Kozhaya and Z. M. Kassas, "On the fundamental tracking performance and design considerations of radio navigation," *IEEE Journal on Selected Areas in Communications*, 2024.

[15] H. Sallouha, S. Saleh, S. De Bast, Z. Cui, S. Pollin, and H. Wymeersch, "On the ground and in the sky: A tutorial on radio localization in ground-air-space networks," *IEEE Communications Surveys and Tutorials*, vol. 27, no. 1, pp. 218–258, 2025.

[16] M. L. Psiaki, "Navigation using carrier Doppler shift from a LEO constellation: TRANSIT on steroids," *NAVIGATION*, vol. 68, no. 3, pp. 621–641, 2021.

[17] B. McLemore and M. L. Psiaki, "Navigation using Doppler shift from LEO constellations and INS data," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 5, pp. 4295–4314, 2022.

[18] A. Baron, P. Gurfil, and H. Rotstein, "Implementation and accuracy of Doppler navigation with LEO satellites," *NAVIGATION: Journal of the Institute of Navigation*, vol. 71, no. 2, 2024.

[19] T. E. Humphreys, P. A. Iannucci, Z. M. Komodromos, and A. M. Graff, "Signal structure of the Starlink Ku-band downlink," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 1–16, 2023.

[20] S. Kozhaya, S. Joe, and Z. M. Kassas, "Unveiling Starlink for PNT," *NAVIGATION*, vol. 72, no. 1, 2025.

[21] W. Qin, A. M. Graff, Z. L. Clements, Z. M. Komodromos, and T. E. Humphreys, "Timing properties of the Starlink Ku-band downlink," *IEEE Transactions on Aerospace and Electronic Systems*, 2025, submitted for review.

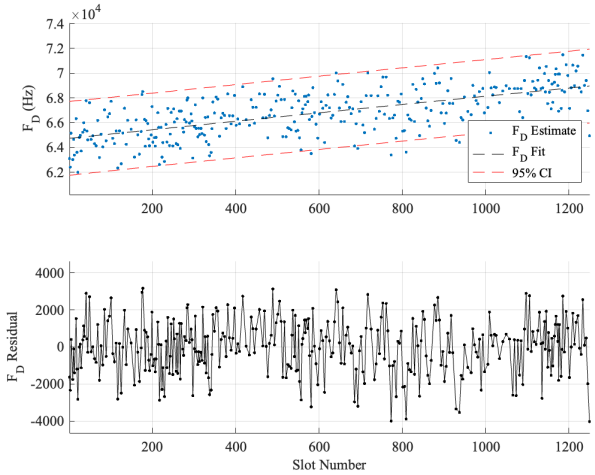
[22] S. Kozhaya, J. Saroufim, and Z. M. Kassas, "Starlink for PNT: A trick or a treat?" in *Proceedings of the 36th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2023)*, 2024.

[23] Z. M. Komodromos, W. Qin, and T. E. Humphreys, "Weak signal acquisition and tracking of the Starlink Ku-Band downlink to enable global PNT," in *Proceedings of the ION Joint Navigation Conference (JNC)*, 2024.

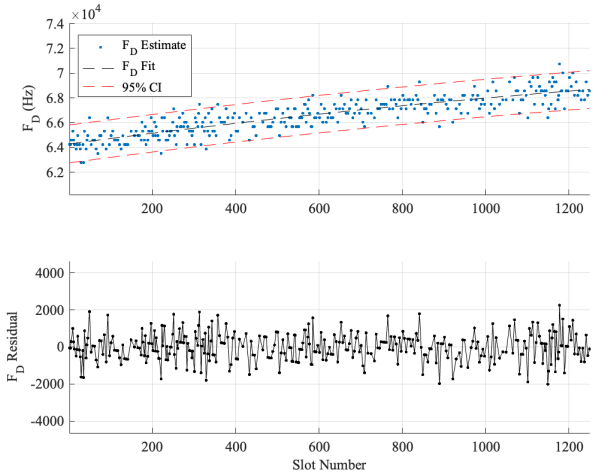
[24] A. M. Graff and T. E. Humphreys, "OFDM-based positioning with unknown data payloads: Bounds and applications to LEO PNT," *IEEE Transactions on Wireless Communications*, 2024, submitted for review.

[25] A. Rihaczek, *Principles of high-resolution radar*. McGraw-Hill, 1969.

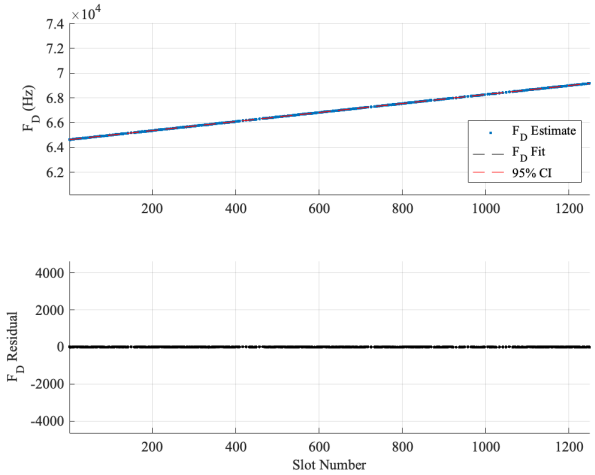
[26] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987.



(a) Pilot-based CAF acquisition. Equivalent to coarse acquisition using the coherent concatenated PSS + SSS local replica.



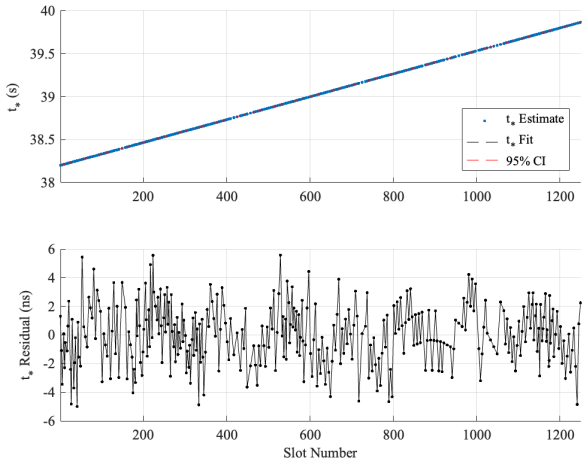
(b) Pilot-only ML estimation. In this case, $\hat{\boldsymbol{z}}, \hat{\boldsymbol{\phi}} \in \mathbb{R}^{2\times 1}$, the minimum number of measurements to estimate β.



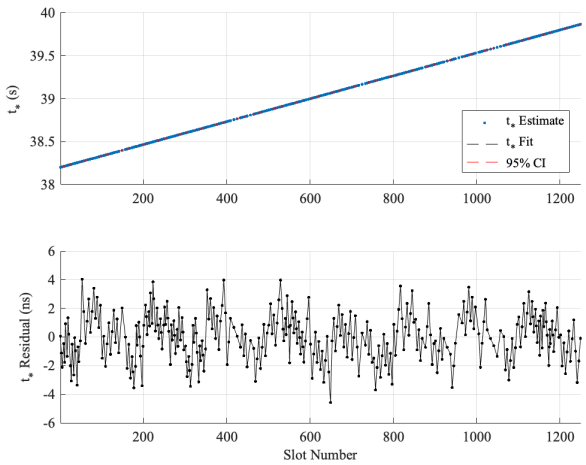
(c) Full-frame ML estimation. The full technique as proposed in Section III

.

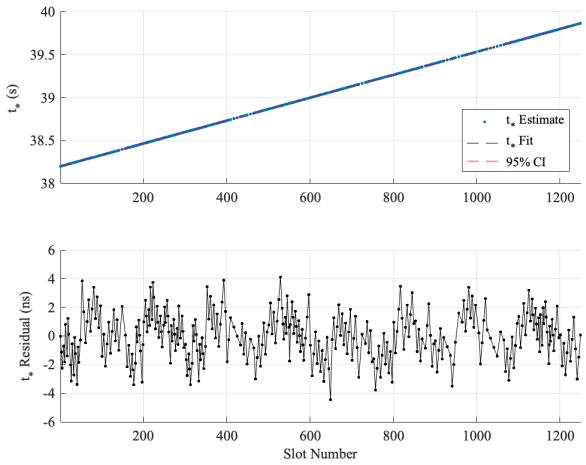
Fig. 8: β tracking performance using various estimation techniques. All axes are matched across plots.



(a) See caption of Fig. 8a.



(b) See caption of Fig. 8b.



(c) See caption of Fig. 8c.

Fig. 9: As Fig. 9 but for t_* tracking performance.

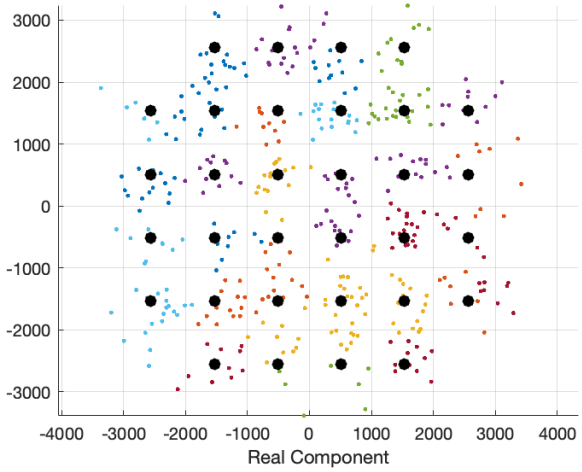


Fig. 10: Example of 32QAM modulation, observed in a decoded signal originally captured in January 2025.

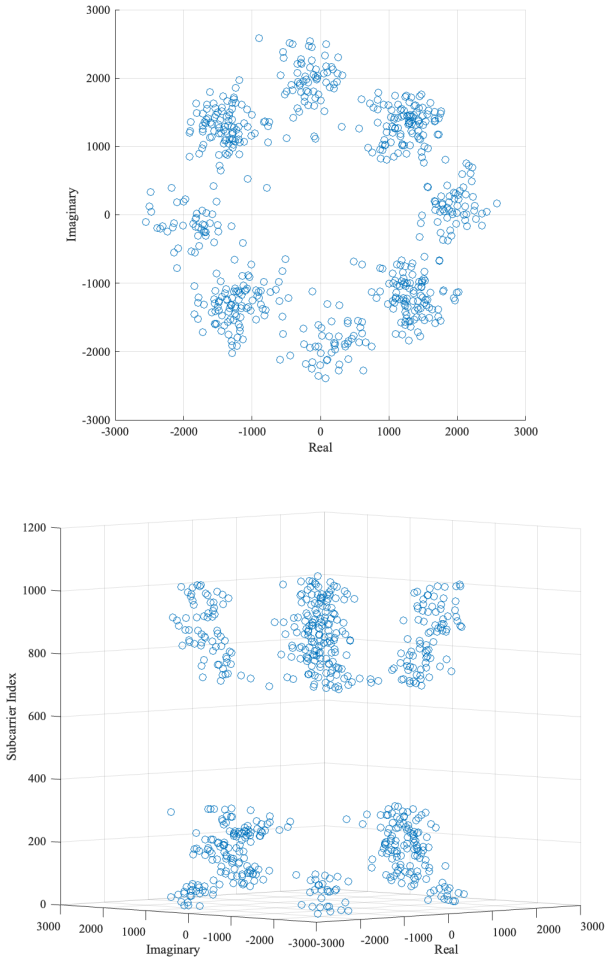


Fig. 11: Side and top views of subcarrier-sliced $\pi/4$PSK. The top view (top) is the typical constellation representation in the complex plane, where the symbol seems to take on 8PSK modulation. However, closer inspection of the side view (bottom) reveals that the modulation clearly switches between 4QAM and QPSK modulations twice, at two discrete subcarrier indices.