Pilots and Other Predictable Elements of the Starlink Ku-Band Downlink

Wenkai Qin*, Mark L. Psiaki†, John R. Bowman†, Todd E. Humphreys*
*Department of Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin
wqin@utexas.edu, todd.humphreys@utexas.edu
†Department of Aerospace and Ocean Engineering, Virginia Tech
mlpsiaki@vt.edu, jrbowman@vt.edu

***Abstract*—We identify and characterize dedicated pilot symbols and other predictable elements embedded within the Starlink Ku-band downlink waveform. Exploitation of these predictable elements enables precise opportunistic positioning, navigation, and timing using compact, low-gain receivers by maximizing the signal processing gain available for signal acquisition and time-of-arrival (TOA) estimation. We develop an acquisition and demodulation framework to decode Starlink frames and disclose the explicit sequences of the *edge pilots*—bands of 4QAM symbols located at both edges of each Starlink channel that apparently repeat identically across all frames, beams, channels, and satellites. We further reveal that the great majority of QPSK-modulated symbols do not carry high-entropy user data but instead follow a regular tessellated structure superimposed on a constant reference template. We demonstrate that exploiting frame-level predictable elements yields a processing gain of approximately 48 dB, thereby enabling low-cost, compact receivers to extract precise TOA measurements even from low-SNR Starlink side beams.**

I. INTRODUCTION

Position, navigation, and timing (PNT) is a crucial utility whose applications span vast sectors of modern life. From agriculture to aviation to financial markets, Global Navigation Satellite Systems (GNSS)—the prevailing technology for PNT—provide the nanosecond-level timing and meter-level positioning necessary for efficient logistics, safe transportation, synchronized telecommunications, and precise financial transactions. Yet legacy GNSS face both a set of critical vulnerabilities and an expanding number of threats. Because legacy civil GNSS transmit weak L-band signals from medium Earth orbit and operate as open-access systems, users risk denial of service and deception by jamming and spoofing attacks. These risks are intensifying as the space domain and electromagnetic spectrum emerge as decisive arenas of 21st-century warfare, with predictable and escalating dangers for civilian users of GNSS-based PNT services [1], [2].

Meanwhile, advances in reusable rocket technology have led to a tenfold reduction in space launch costs, laying the groundwork for a rapidly expanding low Earth orbit (LEO) space economy [3]. Recently launched LEO communications systems (e.g., OneWeb, Starlink, and Amazon Leo) have already attracted millions of subscribers and currently generate billions of dollars in annual revenue. These systems have drawn the interest of PNT researchers seeking a robust, precise, and broadly accessible alternative to legacy GNSS: the consistent availability, high power, and wide bandwidth of LEO communications signals confer inherent resilience to adversarial interference and great potential for PNT precision. But despite the significant potential and low marginal cost of fusing a PNT service with an existing broadband LEO communications service [4], no major constellation operators have announced plans to bring forward a PNT service whose accuracy would match or exceed that of legacy GNSS. Meanwhile, researchers have developed techniques to opportunistically exploit LEO communications signals for PNT, with efforts from several groups already bearing fruit [5]–[13]. While many constellations' signals have been shown to be exploitable [5], [7], [13], Starlink remains the most attractive candidate for opportunistic navigation: with a constellation of over 10,000 satellites, Starlink's dense, geometry-rich coverage reaches nearly the entire globe and offers a promising foundation for multilaterated PNT [14].

Nonetheless, several challenges remain before Starlink-based PNT can be a practically viable alternative to legacy GNSS. One major hurdle concerns the physical footprint of effective Starlink PNT receiver antennas. Supposing Starlink phased array terminal form-factors are indicative of a future PNT-capable terminal, even the smallest option currently offered by SpaceX (Starlink Mini) measures $30\,\mathrm{cm} \times 26\,\mathrm{cm}$, which is too large for many practical applications (e.g., small unmanned aerial vehicles, consumer electronics, and asset trackers). A compact, mass-produced low-noise block (LNB) with an integrated hemispherical feedhorn, such as the unit shown in Fig. 1, offers a more attractive physical profile, with feedhorn outer diameter $6\,\mathrm{cm}$, maximal dimension $12\,\mathrm{cm}$, and mass under $200\,\mathrm{g}$.

The use of such a compact feedhorn receiver for Starlink-based PNT would confer an additional benefit: its wide beamwidth enables simultaneous reception of multiple satellite transmissions, which allows multilateration, provided that each transmitting satellite can be uniquely associated with its corresponding component in the received signal. Indeed, the systems presented in [9], [15], [16] exclusively employ compact feedhorn receivers to track so-called pilot tones for Starlink-based Doppler positioning. Unfortunately, these pilot tones' strength and availability have significantly decreased since 2023, rendering this approach ineffective [8].

In [11], researchers describe a pseudorange-based Starlink PNT system with an approximately 20-cm feedhorn (yielding

Copyright © 2025 by Wenkai Qin, Mark L. Psiaki,
John R. Bowman, and Todd E. Humphreys

February 2026 preprint of paper submitted for review

15 dBi of gain) coupled to a wideband LNB. However, the resulting position solutions remain relatively coarse, with miss distances on the order of 100 m or more. The system presented in [12] geolocates a compact feedhorn receiver, but it relies on a stationary reference receiver that relays high-bitrate data captured using a 35-dBi, 90-cm parabolic dish antenna to aid signal tracking and disambiguation.

Thus, although the feasibility of Starlink-based PNT has been repeatedly demonstrated, achieving precise PNT with a compact, low-cost antenna and minimal side-channel aiding remains a challenge.

Ultimately, the practicality and performance of a Starlink-based PNT system are subject to constraints on the receiver's antenna gain, which must provide a sufficient signal-to-noise ratio (SNR) to ensure reliable signal acquisition and satisfactory PNT precision. For instance, consider a Ku-band horn antenna with a circular 54 mm diameter aperture. Supposing an operating frequency of 12 GHz and antenna efficiency of $\eta = 0.7$, this antenna's gain is theoretically limited to approximately 15.1 dBi with an 32° half-power beamwidth, as calculated using standard antenna theory [17, Section 2-11], [18, Eq. (12-29a)]. Assuming clear-sky conditions, an LNB noise figure of 1 dB, and published Starlink Ku-band power flux density (PFD) of -145 to -138 dBW per 4 kHz, this antenna yields a maximum pre-correlation SNR in the range of 1 dB to 8 dB [18, Sec. 2.16–2.18], [19].

Our own measurements of Starlink pre-correlation SNR from an assigned beam signal (one directed to the receiver's service cell) captured through a compact feedhorn antenna with an integrated LNB lie several decibels below this range, as shown in Fig. 1. This is because, in practice, SNR is degraded by off-axis reception of the satellite beam, polarization mismatch, receiver implementation losses, atmospheric attenuation, and co-channel interference from other beams and satellites.

Nonetheless, the clearly defined correlation peaks in Fig. 1 show that matched filtering of an assigned beam's signal against a local replica composed of the coherent combination of the primary and secondary synchronization sequences (PSS and SSS—described in [20]) can provide ample processing gain for signal acquisition and tracking, assuming the signal is captured with adequate bandwidth and under low-interference conditions. But this result does not necessarily extend to side beams—those directed to a service cell other than the one in which the receiver resides. The authors of [21] observe that a composite signal's trackable side beams may have SNRs that are 5 to 18 dB below that of the assigned beam—too low to extract useful time-of-arrival (TOA) measurements solely with the PSS+SSS local replica, as will be shown in Section II. The ability to extract sub-meter-level-precise TOA measurements from several side beams would improve geometric dilution of precision (GDOP) and shorten time-to-first-fix.

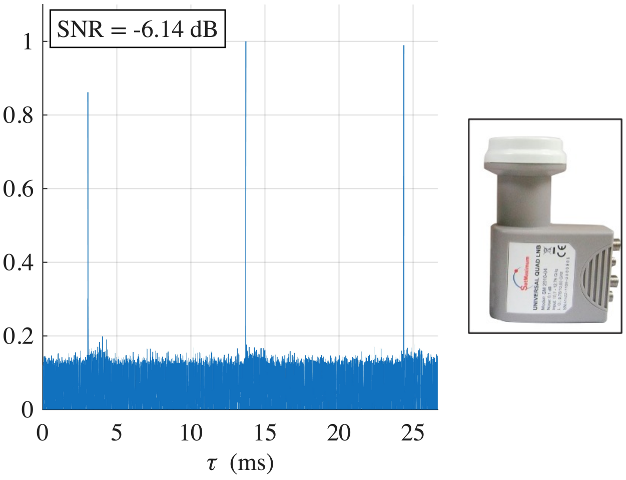

Fig. 1: Normalized matched-filter correlation against a coherent combination of the known primary and secondary synchronization sequences (PSS and SSS) with a signal captured from a compact feedhorn receiver of the type shown. This receiver, intended for consumer-grade satellite television reception in the 10.7–12.75 GHz frequency band, provides a 50–60 dB gain and a noise figure below 1 dB. The SNR value shown indicates the pre-correlation SNR of the received signal associated with the strongest post-correlation peak. For this capture, which spans the full 240 MHz bandwidth of a single Starlink channel, the theoretical processing gain by which the post-correlation SNR exceeds the pre-correlation SNR is 33 dB.

The key to unlocking robust and practical Starlink-based PNT with a compact receiver is to boost the processing gain well beyond that provided by the PSS+SSS combination. Processing gain, a fundamental concept in direct-sequence spread-spectrum (DSSS) systems such as code-division multiple access (CDMA) and GNSS, can be extended to opportunistic exploitation of orthogonal frequency division multiplexing (OFDM) waveforms such as Starlink's, in which context the processing gain scales as the number of known information symbols. With adequate processing gain, the side beam signals required for a fast and precise Starlink PNT solution can achieve sufficient post-correlation SNR to permit precise TOA measurement.

We note that, in principle, an opportunistic PNT user could apply the maximum-likelihood (ML) or decision-directed (DD) methods of [22] to recover TOA measurements despite lack of knowledge of the OFDM information symbols. But these methods break down catastrophically in the low-SNR regime even for low-order modulation (see Fig. 1 in [22]), and so are inapt for Starlink signals received through compact antennas.

While known information symbols are the most straightforward features to exploit *a priori*, knowledge of any invariant property of the Starlink frame likewise increases the processing gain—for instance, the range of modulation types, the temporal arrangement of modulation, or whether individual OFDM symbols are simple in their modulation (e.g., purely QPSK) or composite (e.g., a mixture of QPSK and 16QAM). Accordingly, we aim to identify key predictable features present in the Starlink Ku-band downlink waveform, thereby boosting the achievable processing gain.

Prior work already discovered several predictable features.

Reference [20] revealed the Starlink frame structure, including the PSS and SSS and their explicit values. SpaceX's 2024 patent [23] confirmed this frame structure and additionally disclosed the existence of two bands of pilot symbols near the edges of each Starlink channel's allocated bandwidth. Pilot symbols are complex-valued coefficients that modulate OFDM subcarriers just like other information symbols, but are entirely predictable to the receiver and so can be used for channel estimation. Reference [8] empirically confirmed these pilot bands, hereafter called the edge pilots, and reported additional inter-frame-correlated data not disclosed by [23], but implied that these additional correlated data appear in all frames. This, as we explain herein, does not appear to be true. Moreover, [8] did not provide the actual information symbol values for the edge pilots or any other inter-frame-correlated data, and so does not enable the most potent method for signal detection, namely, matched filtering against a noise-free local signal replica.

The purpose of the current paper is to offer the fullest account yet of predictable structure in the Starlink waveform. Not only do we reveal the exact information symbols of the edge pilots, but we also report a remarkable discovery: after the first few OFDM symbols in each frame, which appear to contain header information, almost all symbols modulated as QPSK are nearly perfectly predictable and thus low-entropy. While this fact is not obvious from a cursory examination of the data, application of certain transformations described in detail herein reveals a highly consistent structure with extended segments of predictable symbols. Fully exploiting these low-entropy signal elements allows matched filtering with sufficient processing gain to support robust PNT with low-cost compact feedhorn receivers.

II. SNR-DEPENDENT TOA PRECISION

To assess the conditions under which precise TOA estimation is feasible, we examine the Ziv-Zakai Bound (ZZB), which characterizes estimator performance across a given SNR range by explicitly modeling the transition from small-error behavior in the high-SNR regime, where the Cramér-Rao Bound (CRB) is asymptotically tight, to large-error behavior dominated by noise and ambiguity. In contrast to the CRB, which assumes unbiased estimation in the high-SNR limit, the ZZB captures threshold effects that lead to rapid degradation in TOA precision at low SNR. Reference [24] describes in detail the ZZB formulation and its application to OFDM-based TOA estimation.

Fig. 2 shows TOA precision bounds as functions of pre-correlation SNR for representative Starlink processing strategies. When correlating against the relatively short PSS+SSS local replica using the full-channel bandwidth $B_\mathrm{f} = 240\,\mathrm{MHz}$, the ZZB departs from the CRB for pre-correlation SNR below approximately $-17.4\,\mathrm{dB}$. Narrowing the capture bandwidth to $B_\mathrm{n} = 25\,\mathrm{MHz}$ shifts this divergence point to a substantially higher SNR of approximately $-9.2\,\mathrm{dB}$. Below these thresholds, TOA precision degrades rapidly.

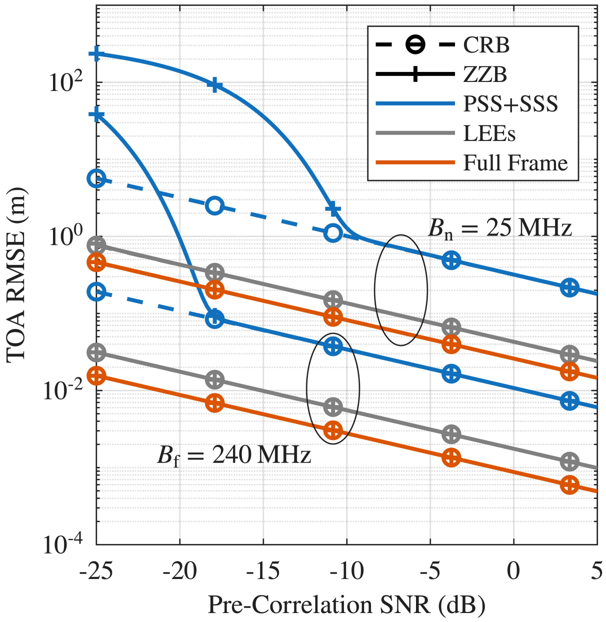

Fig. 2: Lower bounds on single-frame TOA root mean squared error (RMSE) vs. pre-correlation SNR for coherent processing with three representative local replicas: (1) the PSS+SSS combination (blue), (2) low-entropy elements (LEEs) of the frame (gray), and (3) the full Starlink frame (red). Results are shown for two representative capture bandwidths, 25 and $240\,\mathrm{MHz}$. The Ziv-Zakai bounds (ZZBs; solid lines) are tighter than the Cramér-Rao bounds (CRBs; dashed lines) at low SNR because they capture the threshold effects that lead to rapid degradation in TOA precision in that regime. The SNR range shown covers the relevant values for trackable signals from assigned beams (high SNR) to side beams (low SNR). The bounds based on LEEs are discussed further in Section IX.

In contrast, when correlating against a full-frame local replica—one that spans all subcarriers and OFDM symbols of the Starlink frame—the substantially increased number of coherently accumulated samples yields a greater processing gain and suppresses ambiguity sidelobes in the correlation function. As a result, the ZZB remains coincident with the CRB across the entire SNR range shown, even when the capture bandwidth is limited to B_n. These results indicate that full-frame processing maintains CRB-like TOA estimation over a wider SNR range than PSS+SSS-only processing.

One can think of the PSS+SSS-only and full-frame strategies as extreme cases that bookend reasonable coherent processing for TOA measurements from Starlink signals: There is little reason to process less than the coherent combination of the PSS and SSS, which are known and easily constructed [20], whereas coherent processing beyond one full frame is complicated by inter-frame carrier phase discontinuities, which, so far, have resisted modeling. Thus, Fig. 2 covers all reasonable cases between short-duration processing with narrow-bandwidth captures (the PSS+SSS case with B_n) and full-frame processing with full-channel-bandwidth captures (the full-frame case with B_f).

Fig. 2 makes clear that a receiver seeking to suppress the

ZZB-predicted breakdown in TOA precision at low SNR must adopt strategies that provide greater processing gain than that afforded by a PSS+SSS matched filter. Such strategies extend the SNR range over which TOA estimation remains asymptotic and enable the extraction of accurate TOA measurements from weaker side beams that would otherwise be unusable.

III. PROCESSING GAIN

Here we define processing gain, extend it to OFDM, and show how any invariant property of the Starlink frame may be exploited to increase it.

For any signal and processing strategy, define the processing gain as

$$L \triangleq \frac{\mathrm{SNR}_{\mathrm{post}}}{\mathrm{SNR}_{\mathrm{pre}}} \tag{1}$$

where $\mathrm{SNR}_{\mathrm{pre}}$ and $\mathrm{SNR}_{\mathrm{post}}$ denote pre- and post-correlation SNR, respectively. The former is the ratio of the total signal power to the total noise power within the signal bandwidth. The latter is the ratio of the signal-to-noise power within the de-spread bandwidth, or, equivalently, the SNR of the complex product that results from correlation against whatever local replica is used for detection. For DSSS despreading, this definition recovers the familiar relation $L = B_{\mathrm{spread}}/B_{\mathrm{despread}}$, where B_{spread} and B_{despread} are the spread and despread bandwidths, respectively.

Consider a sequence of complex information symbols subject to additive noise

$$r_k = x_k + n_k, \quad k \in \{1, \ldots, N\} \tag{2}$$

where the kth information signal $x_k \in \mathcal{C}$ is drawn uniformly from the modulation constellation $\mathcal{C} = \{c^{(1)}, c^{(2)}, \ldots, c^{(M)}\}$ and the symbols $\{x_k\}_{k=1}^N$ are statistically independent. The constellation is assumed to be normalized and zero-mean such that $\mathbb{E}\left[|x_k|^2\right] = 1$ and $\mathbb{E}\left[x_k\right] = 0$ for all $k \in \{1, \ldots, N\}$. The additive noise samples $\{n_k\}_{k=1}^N$ are complex Gaussian random variables with $n_k \sim \mathcal{CN}(0, \sigma_n^2)$, mutually independent for all $k \in \{1, \ldots, N\}$, and independent of $\{x_k\}_{k=1}^N$.

Define the signal detection statistic as the complex correlation product

$$S \triangleq \sum_{k=1}^{N} r_k^* l_k = \sum_{k=1}^{N} x_k^* l_k + \sum_{k=1}^{N} n_k^* l_k = S_x + S_n \tag{3}$$

where l_k is the kth element of a local replica sequence and $(\cdot)^*$ denotes the complex conjugate. Pre- and post-correlation SNR can then be defined as

$$\mathrm{SNR}_{\mathrm{pre}} \triangleq \frac{\mathbb{E}\left[x_k^* x_k\right]}{\mathbb{E}\left[n_k^* n_k\right]} = \frac{1}{\sigma_n^2}, \quad \mathrm{SNR}_{\mathrm{post}} \triangleq \frac{\mathbb{E}\left[S_x^* S_x\right]}{\mathbb{E}\left[S_n^* S_n\right]}$$

We additionally define the following correlation parameters for which expectations are taken over the joint distribution of x_k and l_k across all $k \in \{1, \ldots, N\}$:

$$\alpha \triangleq \mathbb{E}\left[|x_k|^2 |l_k|^2\right], \quad \rho \triangleq \mathbb{E}\left[|l_k|^2\right], \quad \mu \triangleq \mathbb{E}\left[l_k^* x_k\right] \tag{4}$$

Now consider L under each of three conditions for $\{x_k\}_{k=1}^N$: (i) completely unknown, (ii) completely known, or (iii) partially known. These conditions can be modeled through the *a priori* probability mass function (PMF) of x_k over $\mathcal{C}$. In the unknown case, the PMF is uniform over $\mathcal{C}$. In the known case, it collapses to a point mass at the transmitted symbol. In the partially known case, side information or invariant properties of the given waveform may be exploited to restrict the PMF to a subset of the available constellation points or otherwise bias it away from a uniform distribution. In this paper's Supplementary Material we evaluate L under each condition and provide a general result.

For the special case of constant-modulus constellations (e.g., QPSK) and unit-modulus replicas, we have $|x_k| = |l_k| = 1$, from which it follows that $\alpha = \rho = 1$ and

$$L = 1 + (N - 1)|\mu|^2 \tag{5}$$

This result interpolates smoothly between the incoherent limit $L = 1$ for completely unknown data ($|\mu| = 0$), and the fully coherent limit $L = N$ for completely known data ($l_k = x_k$, $|\mu| = 1$).

Note that it makes no difference whether the complex information symbols in (2) represent samples in the time domain, as in DSSS, or the frequency domain, as in OFDM. Since the discrete Fourier transform (DFT) is a linear, unitary transformation that preserves inner products and signal energy (up to a constant normalization), the detection statistic (3) and its properties are preserved through the transformation. Consequently, correlating in the frequency domain across subcarriers yields the same processing gain as correlating in the time domain across samples.

In general, any invariant property of an OFDM waveform can be exploited during the construction of $\{l_k\}_{k=1}^N$ to increase $|\mu|$ and thus L. For example, knowledge that certain OFDM symbols are simple rather than composite in their modulation, or that their information symbols are merely a low-entropy transformation of a known sequence, or knowledge of the guard (cyclic prefix; CP) length N_{g}, can be embedded in the index-dependent PMF for x_k and thus exploited to raise L.

For a generic OFDM waveform, calculation of L is complicated by the CP, which introduces correlation between the information symbols $\{x_k\}_{k=1}^N$, whether viewed in the frequency or time domain. For the particular case of the Starlink Ku-band downlink waveform, calculation is further complicated by time-domain correlation arising due to the repetitive structure of the PSS and to the mid-channel gutter in which four adjacent subcarriers have zero energy. Ignoring the minor effects of such correlation, one can approximate the effective accumulation length N in (5) for full-frame Starlink processing as

$$N \approx N_{\mathrm{sf}}(N_{\mathrm{s}} + N_{\mathrm{g}}) = 318912 \tag{6}$$

where $N_{\mathrm{sf}} = 302$ is the number of nonzero symbols in a frame, $N_{\mathrm{s}} = 1024$ is the number of subcarriers, and $N_{\mathrm{g}} = 32$ is the CP length. Thus, for the case of full-frame correlation with completely known symbols, $L = N \approx 55$ dB.

IV. DATA CORPUS

The signal models and observations about data structures that follow are supported by a large corpus of Starlink Ku-band captures spanning many observation sessions and satellites,

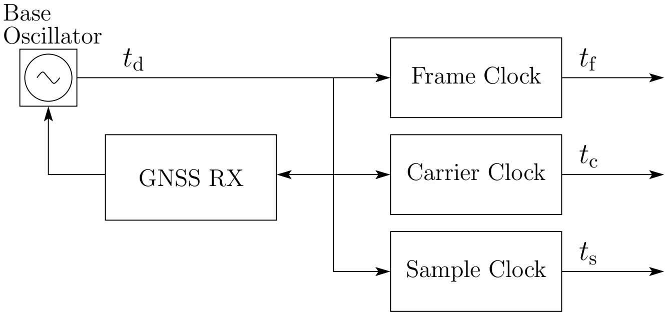

Fig. 3: Cascading clock model for Starlink transmitter showing the relationship between base time t_d, frame clock time t_f, carrier clock time t_c, and sample clock time t_s.

across which the reported properties were observed consistently. The corpus spans all generations of Starlink satellites to date (blocks 1.0, 1.5, and 2.0-mini) and covers captures from October 2024 through June 2025. Details about each capture are given in the Supplementary Material. To encourage follow-on work and ensure reproducibility, the Supplementary Material contains decoded frame data for a high-quality subset of 1009 decoded frames that we call the *exemplar frames*, which are from a single capture of high-SNR data from one Starlink satellite.

V. SIGNAL MODEL

To facilitate understanding of the OFDM signal processing techniques presented in the sequel, this section presents a joint time- and frequency-domain signal model for the Starlink Ku-Band downlink. The model assumes the basic frame structure detailed in [20]. In brief, Starlink Ku-band downlink data are packaged into frames transmitted at a maximum rate of 750 Hz. Each frame consists of $N_\mathrm{sf} = 302$ OFDM symbol slots of nominal length $T_\mathrm{sym} = 4.4\,\mu\mathrm{s}$ followed by an unoccupied frame guard interval $T_\mathrm{fg} = T_\mathrm{sym} + T_\mathrm{g}$, where $T_\mathrm{g} = 2/15$ µs is the CP interval. A sequence of frames from a given Starlink satellite vehicle (SV) to a given service cell is conveyed by a downlink beam whose SV-to-cell assignment remains fixed over a 15-second fixed assignment interval (FAI) [21].

A. Clock Models

Unlike GNSS and other DSSS-based waveforms, Starlink's symbols, frames, and underlying carrier are not driven directly by the same clock—at least not for block 1.0, 1.5, and 2.0 SVs [21]. Accordingly, we adopt the cascading clock model shown schematically in Fig. 3. The GNSS-disciplined base time t_d drives the downstream frame, carrier, and sample clocks. These in turn drive signal generation onboard Starlink SVs. The frame clock sets the start time of frames within an FAI. The carrier clock drives generation of the carrier signal used to mix the baseband OFDM waveform to radio frequency (RF). The sample clock drives the digital-to-analog converter used to reconstruct the digital waveform as an analog baseband waveform.

Time according to these three clocks is related to true time t, or time according to an ideal clock, by

$$t = t_\mathrm{f}(t) - \delta t_\mathrm{f}(t) \tag{7}$$

$$t = t_\mathrm{c}(t) - \delta t_\mathrm{c}(t) \tag{8}$$

$$t = t_\mathrm{s}(t) - \delta t_\mathrm{s}(t) \tag{9}$$

where $\delta t_\mathrm{f}(t)$, $\delta t_\mathrm{c}(t)$, and $\delta t_\mathrm{s}(t)$ are the frame, carrier, and sample clock offsets, respectively. Given that this paper addresses the signal processing of individual frames rather than the timing behavior of frame sequences as in [21], the frame clock is not discussed in further detail.

The time derivatives of $\delta t_\mathrm{c}(t)$ and $\delta t_\mathrm{s}(t)$ are referred to as the carrier and sample clock drifts. For the temperature-controlled crystal oscillators (TCXOs) apparently used in Starlink SVs [21], these may be modeled as constant over a frame interval $T_\mathrm{f} = 1/750$ s. This gives rise to a linear clock offset model that is valid for $t_0 \leq t \leq t_0 + T_\mathrm{f}$:

$$\delta t_\mathrm{s}(t) = \delta t_{\mathrm{s}0} + \dot{\delta t}_\mathrm{s}(t - t_0) \tag{10}$$

$$\delta t_\mathrm{c}(t) = \delta t_{\mathrm{c}0} + \dot{\delta t}_\mathrm{c}(t - t_0) \tag{11}$$

Here, $\delta t_{\mathrm{c}0}$ and $\delta t_{\mathrm{s}0}$ are the carrier and sample clock offsets at time t_0, while $\dot{\delta t}_\mathrm{c}$ and $\dot{\delta t}_\mathrm{s}$ are the respective clock drifts, modeled as piecewise constant from frame to frame. Substituting (11) into (8) and solving for $t_\mathrm{c}(t)$ yields

$$\begin{aligned} t_\mathrm{c}(t) &= t + \delta t_{\mathrm{c}0} + \dot{\delta t}_\mathrm{c}(t - t_0) \\ &= (1 + \dot{\delta t}_\mathrm{c})t + (\delta t_{\mathrm{c}0} - \dot{\delta t}_\mathrm{c} t_0) \end{aligned} \tag{12}$$

Similarly, $t_\mathrm{s}(t)$ becomes

$$t_\mathrm{s}(t) = (1 + \dot{\delta t}_\mathrm{s})t + (\delta t_{\mathrm{s}0} - \dot{\delta t}_\mathrm{s} t_0) \tag{13}$$

B. Transmitted Signal Model

As mentioned, each Starlink frame contains $N_\mathrm{sf} = 302$ occupied OFDM symbol intervals of length T_sym. We refer to the time-domain signal that applies over the ith symbol interval of the mth frame as $x_{mi}(t)$, where $i \in \mathcal{I} \triangleq \{0, 1, \ldots, N_\mathrm{sf} - 1\}$. The first symbol interval is occupied by $x_{m0}(t)$, a known time-domain sequence referred to as the PSS; the second is occupied by $x_{m1}(t)$, a known time-domain OFDM symbol referred to as the SSS. The remaining $\{x_{mi}(t) \mid i \in [2, N_\mathrm{sf} - 1]\}$ are time-domain OFDM symbols containing a mix of header information, pilots, user data, and—as will be shown—low-entropy codes.

Each $x_{mi}(t)$ for $i \in [1, N_\mathrm{sf} - 1]$ consists of $N_\mathrm{s} = 1024$ mutually orthogonal data subcarriers spread across the channel bandwidth $F_\mathrm{s} = 240\,\mathrm{MHz}$, resulting in a carrier spacing of $F = F_\mathrm{s}/N_\mathrm{s}$. Let $k \in \mathcal{K} \triangleq \{0, 1, \ldots, N_\mathrm{s} - 1\}$ represent a single subcarrier's assigned index. We adopt the typical OFDM subcarrier index-to-frequency offset mapping $d[k]$, defined as

$$d[k] = \begin{cases} k, & k \in [0, \frac{N_\mathrm{s}}{2} - 1] \\ k - N_\mathrm{s}, & k \in [\frac{N_\mathrm{s}}{2}, N_\mathrm{s} - 1] \end{cases} \tag{14}$$

Each subcarrier is modulated by a complex-valued information symbol $X_{mik} \in \mathcal{C}_{mi}$, which encodes one or more bits of information depending on the modulation scheme corresponding to the constellation $\mathcal{C}_{mi}$ (e.g., BPSK, 4QAM, etc., or a mixture of these). Four mid-channel subcarriers, those with indices in $\mathcal{K}_{\mathrm{g}} = \{0, 1, 1022, 1023\}$, form a gutter in which $X_{mik} = 0$ for $k \in \mathcal{K}_{\mathrm{g}}$. Let $\mathcal{I}_1 \triangleq \mathcal{I} \setminus 0$ be the symbol index set excluding the PSS, which is not a standard OFDM symbol. Then for $i \in \mathcal{I}_1$ the baseband time-domain signal over $0 \leq t < T_{\mathrm{sym}}$ can be written as

$$x_{mi}(t) = \frac{1}{\sqrt{N_{\mathrm{s}}}} \sum_{k \in \mathcal{K}} X_{mik} \exp\left(j2\pi d[k] F (t - T_{\mathrm{g}})\right) \tag{15}$$

where $T_{\mathrm{g}} = N_{\mathrm{g}}/F_{\mathrm{s}}$ is the CP interval, $T_{\mathrm{sym}} = T + T_{\mathrm{g}}$ is the full symbol interval, and $T = N_{\mathrm{s}}/F_{\mathrm{s}}$ is the useful (non-cyclic) symbol interval. Note that despite the introduction of $d[k]$ in (15), our definition of X_{mik} is identical to that of [20] for sampling at $t = 0, 1/F_{\mathrm{s}}, \dots, T_{\mathrm{sym}}F_{\mathrm{s}}$. Let the OFDM symbol support function be defined as

$$g_{\mathrm{s}}(t) = \begin{cases} 2t/T_{\mathrm{g}}, & 0 \leq t < T_{\mathrm{g}}/2 \\ 1, & T_{\mathrm{g}}/2 \leq t < T + T_{\mathrm{g}}/2 \\ 1 - 2(t - T - T_{\mathrm{g}}/2)/T_{\mathrm{g}}, & T + T_{\mathrm{g}}/2 \leq t < T_{\mathrm{sym}} \\ 0, & \text{otherwise} \end{cases}$$

Then, assuming a start time $t_0 = 0$, the baseband time-domain signal over the mth frame may be written as

$$x_m(t) = \sum_{i \in \mathcal{I}} x_{mi}(t - iT_{\mathrm{sym}}) g_{\mathrm{s}}(t - iT_{\mathrm{sym}}), \quad 0 \leq t < T_{\mathrm{f}} \tag{16}$$

Starlink OFDM signals are transmitted over one of eight channels spanning the 10.7–12.7 GHz band [20]. In what follows, F_{c} represents an arbitrary channel center frequency selected from among the eight. Assume $x_m(t)$ undergoes digital-to-analog conversion driven by $t_{\mathrm{s}}(t)$, and mixing from baseband to center frequency F_{c} driven by $t_{\mathrm{c}}(t)$. Then, assuming $t_0 = 0$, we arrive at the transmitted mth-frame RF signal model

$$\begin{aligned} \tilde{x}_m(t) &= x_m[t_{\mathrm{s}}(t)] \exp\left[j2\pi F_{\mathrm{c}} t_{\mathrm{c}}(t)\right] \\ &= x_m\Big[(1 + \delta\dot{t}_{\mathrm{s}})t + \delta t_{\mathrm{s}0}\Big] \\ &\quad \times \exp\Big\{j2\pi F_{\mathrm{c}}\Big[(1 + \delta\dot{t}_{\mathrm{c}})t + \delta t_{\mathrm{c}0}\Big]\Big\} \end{aligned} \tag{17}$$

C. Received Signal Model

This section develops a discrete-time received signal model for the LEO-to-Earth channel that includes sample frequency offset (SFO) and carrier frequency offset (CFO) parameters.

Let v_{los} be the scalar velocity between the SV and receiver along the line-of-sight, defined such that $v_{\mathrm{los}} < 0$ when the SV is moving towards the receiver. Define the Doppler parameter $\beta \triangleq v_{\mathrm{los}}/c$, where c is the free-space speed of light. For signals from a LEO SV traveling at an orbital altitude of 550 km to a stationary ground receiver, $|\beta|$ is bounded below 25 ppm. In practice, Starlink SVs typically do not transmit below an elevation of ∼40°, which limits $|\beta|$ to below ∼15 ppm. As the frame interval $T_{\mathrm{f}} \approx 1.33$ ms is short compared to the timescale over which the SV-receiver geometry evolves, β is modeled as constant over each frame.

Let τ_{los} be the line-of-sight propagation delay for the leading edge of the frame and $w(t)$ be receiver noise, modeled as complex-valued zero-mean additive white Gaussian noise (AWGN) with two-sided power spectral density $N_0/2$ per dimension and uncorrelated with the transmitted signal. Assume that $\beta, \delta\dot{t}_{\mathrm{s}}$, and $\delta\dot{t}_{\mathrm{c}}$ are small enough that their second-order products are negligible. Then the analog RF received signal may be modeled as

$$\begin{aligned} \tilde{y}_m(t) &= \tilde{x}_m\left[(1 - \beta)(t - \tau_{\mathrm{los}})\right] + w(t) \\ &= x_m\Big[(1 + \delta\dot{t}_{\mathrm{s}} - \beta)(t - \tau_{\mathrm{los}}) + \delta t_{\mathrm{s}0}\Big] \\ &\quad \times \exp\Big\{j2\pi(1 + \delta\dot{t}_{\mathrm{c}} - \beta)F_{\mathrm{c}}(t - \tau_{\mathrm{los}}) + j2\pi F_{\mathrm{c}}\delta t_{\mathrm{c}0}\Big\} \\ &\quad + w(t) \end{aligned} \tag{18}$$

Let the effects of τ_{los}, $\delta\dot{t}_{\mathrm{s}}$, and $\delta t_{\mathrm{s}0}$ be combined into an equivalent timing offset τ_m, and those of τ_{los}, $\delta\dot{t}_{\mathrm{c}}$, $\delta t_{\mathrm{c}0}$ be combined into an equivalent phase offset ϕ_m:

$$\begin{aligned} \tau_m &= \tau_{\mathrm{los}} - \frac{\delta t_{\mathrm{s}0}}{1 + \delta\dot{t}_{\mathrm{s}} - \beta} \\ \phi_m &= 2\pi F_{\mathrm{c}}\delta t_{\mathrm{c}0} - 2\pi(1 + \delta\dot{t}_{\mathrm{c}} - \beta)F_{\mathrm{c}}\,\tau_{\mathrm{los}} \end{aligned}$$

Then the received signal model simplifies to

$$\begin{aligned} \tilde{y}_m(t) &= x_m\Big[(1 + \delta\dot{t}_{\mathrm{s}} - \beta)(t - \tau_m)\Big] \\ &\quad \times \exp\Big\{j\Big[2\pi(1 + \delta\dot{t}_{\mathrm{c}} - \beta)F_{\mathrm{c}}t + \phi_m\Big]\Big\} + w(t) \end{aligned}$$

For the data processing described in the sequel, signals were captured using a highly directional antenna and sampled at a complex sampling rate of 250 Msps after mixing to baseband, then resampled at F_{s} such that all subsequent operations may proceed at the native Starlink information symbol rate. The receiver clock driving our capture equipment was a GNSS-disciplined OCXO whose time can be taken as synonymous with true time t. To reduce complexity, we aggregate the clock drift and Doppler parameters into an effective SFO factor $\beta_{\mathrm{s}} \triangleq -\delta\dot{t}_{\mathrm{s}} + \beta$ and an effective CFO factor $\beta_{\mathrm{c}} \triangleq -\delta\dot{t}_{\mathrm{c}} + \beta$. Let t_m be the true start time of the mth frame as received, which is governed by Doppler and by the satellite frame clock. Then, for a sampling interval $T_{\mathrm{s}} = 1/F_{\mathrm{s}}$ and receiver sample index $n \in \mathbb{Z}$ on $t_m \leq nT_{\mathrm{s}} < t_m + T_{\mathrm{f}}$, we arrive at the baseband discrete-time signal model

$$\begin{aligned} \tilde{y}_m[n] &= x_m[(1 - \beta_{\mathrm{s}})\,(n - n_m)T_{\mathrm{s}}] \\ &\quad \times \exp(-j2\pi\,\beta_{\mathrm{c}}F_{\mathrm{c}}\,nT_{\mathrm{s}} + j\phi_m) + w(nT_{\mathrm{s}}) \end{aligned} \tag{19}$$

where $n_m \in \mathbb{R}$ satisfies $n_mT_{\mathrm{s}} = \tau_m$. A continuous sequence of received samples across both occupied and unoccupied frame slots is given by

$$\tilde{y}[n] = \sum_{m \in \mathbb{Z}} \tilde{y}_m[n - t_m/T_{\mathrm{s}}]g_{\mathrm{f}}(nT_{\mathrm{s}} - t_m), \quad n \in \mathbb{Z} \tag{20}$$

where

$$g_{\mathrm{f}}(t) = \begin{cases} 1, & 0 \leq t < T_{\mathrm{f}} \\ 0, & \text{otherwise} \end{cases}$$

is the frame support function. Note that, as regards the sample index and frame support function limits, the wide frame guard interval T_{fg}—during which no signal is present—allows us to ignore small deviations from T_{f} in the received signal's frame period.

VI. DEMODULATION

This section describes the demodulation procedure used to correct residual synchronization errors and recover the transmitted information symbols from a single received frame. Here, we follow the indexing convention established in Section V: the received, equalized, phase-aligned, and hard-decoded counterparts to X_{mik} are denoted $\bar{Y}_{mik}$, $\tilde{Y}_{mik}$, Y_{mik}, and $\hat{X}_{mik}$, respectively, with corresponding OFDM-symbol-level vectors $\bar{\boldsymbol{Y}}_{mi}$, $\tilde{\boldsymbol{Y}}_{mi}$, $\boldsymbol{Y}_{mi}$, $\hat{\boldsymbol{X}}_{mi} \in \mathbb{C}^{|\mathcal{K}|}$.

In brief overview, the demodulation process proceeds as follows. Frame m is detected and coarsely acquired by matched filtering against a Doppler-adjusted PSS+SSS local replica, whereupon standard OFDM processing yields, for each symbol index $i \in \mathcal{I}$, a noisy received vector $\bar{\boldsymbol{Y}}_{mi}$. Residual SFO and CFO errors are then estimated and compensated. Finally, hard-decoding decisions are made to produce the recovered information symbol vectors $\hat{\boldsymbol{X}}_{mi}$.

A. Coarse Frame Acquisition

Individual Starlink frame detection and acquisition is accomplished through matched filtering with a Doppler-adjusted local baseband replica. Define the coherent concatenation of the time-domain PSS and SSS functions from [20] as

$$x_{m01}(t) \triangleq \begin{cases} x_{m0}(t), & 0 \leq t < T_{\text{sym}} \\ x_{m1}(t - T_{\text{sym}}), & T_{\text{sym}} \leq t < 2T_{\text{sym}} \\ 0, & \text{otherwise} \end{cases}$$

Because the PSS and SSS are identical for all $m \in \mathbb{Z}$ and all SVs, $x_{m01}(t)$ can be used for signal detection on all captured data. The full discrete-time local replica used for correlation is the sampled product of the Doppler-adjusted $x_{m01}(t)$ and a complex exponential:

$$\tilde{y}_{01}[n, \beta] \triangleq x_{m01}\left[nT_{\text{s}}(1-\beta)\right] \exp\left(-j2\pi\beta F_{\text{c}} nT_{\text{s}}\right)$$

Sampling of the local replica is limited to $n \in [0, M(\beta) - 1]$ with $M(\beta) = \lceil 2T_{\text{sym}}/(1-\beta)T_{\text{s}} \rceil$. For $\tilde{y}[n]$ given as in (20), the complex ambiguity function is formed as

$$R[k, \beta] = \sum_{n=0}^{M(\beta)-1} \tilde{y}[n+k]\, y_{01}^{*}[n, \beta] \tag{21}$$

The nearest-sample timing offset estimate $\hat{n}_m$ and initial Doppler estimate $\hat{\beta}_m$ are obtained via a two-dimensional search over lag and Doppler hypotheses:

$$(\hat{n}_m, \hat{\beta}_m) = \underset{k,\beta}{\operatorname{argmax}}\, |R[k, \beta]|$$

These estimates are accepted if $|R[\hat{n}_m, \hat{\beta}_m]|$ exceeds a detection threshold, whereupon the mth Starlink frame is extracted from the continuous stream as

$$\tilde{y}_m[n] = \tilde{y}[n + \hat{n}_m], \quad n = 0, 1, \ldots, N_{\text{sf}}(N_{\text{s}} + N_{\text{g}}) - 1$$

Using the initial Doppler estimate $\hat{\beta}_m$, the extracted frame is coarsely Doppler compensated via time resampling and carrier frequency compensation. The resulting signal can be modeled in terms of residual synchronization errors as

$$\begin{aligned} \bar{y}_m[n] = {} & x_m\left[(1-\delta\beta_{\text{s}})\,(n-\delta n)T_{\text{s}}\right] \\ & \times \exp\left[-j2\pi\,\delta\beta_{\text{c}} F_{\text{c}}\, nT_{\text{s}} + j\phi_m\right] + w(nT_{\text{s}}) \end{aligned} \tag{22}$$

where $\delta n = n_m - \hat{n}_m$ denotes the fractional residual timing offset, $\delta\beta_{\text{s}} = \beta_{\text{s}} - \hat{\beta}_m$ the residual SFO factor, and $\delta\beta_{\text{c}} = \beta_{\text{c}} - \hat{\beta}_m$ the residual CFO factor. We assume $|\delta n|\, T_{\text{s}} < T_{\text{g}}$ and $|\delta\beta_{\text{c}}|\, F_{\text{c}} \ll F$ so that CP removal and frequency transformation yield negligible inter-symbol interference (ISI) and inter-carrier interference (ICI). Next, $\tilde{y}_m[n]$ is segmented into N_{sf} blocks of $N_{\text{s}} + N_{\text{g}}$ samples each. Each block undergoes CP removal and conversion to the frequency domain via the FFT. In the frequency domain, the ith OFDM symbol's kth received data symbol $\bar{Y}_{mik}$ can be modeled as the following function of the transmitted data symbol X_{mik}:

$$\bar{Y}_{mik} = \alpha_{mik} X_{mik} + V_{mik} \tag{23}$$

$$\alpha_{mik} = Z_m\, H_k\, \exp(-j2\pi d[k] F \tau_{mi} + j\phi_{mi}) \tag{24}$$

Here, V_{mik} denotes frequency-domain AWGN, H_k is the frequency-selective channel response at subcarrier k normalized such that $H_0 = 1$, and $Z_m = \sqrt{g}\exp(j\theta_m)$ is a frame-specific complex constant, independent of k, with channel gain g and initial phase offset θ_m. The phase offset θ_m does not appear to be arbitrary from frame to frame, but our efforts at modeling it have so far not successfully generalized to all frames and all Starlink SVs. The per-symbol time delay τ_{mi} and phase offset ϕ_{mi} relate to $\delta\beta_{\text{s}}$ and $\delta\beta_{\text{c}}$ as follows:

$$\tau_{mi} = \delta n T_{\text{s}} + iT_{\text{sym}}\left(\frac{\delta\beta_{\text{s}}}{1-\delta\beta_{\text{s}}}\right) \approx \delta n T_{\text{s}} + iT_{\text{sym}}\delta\beta_{\text{s}} \tag{25}$$

$$\phi_{mi} = \phi_m - 2\pi i T_{\text{sym}} \delta\beta_{\text{c}} F_{\text{c}} \tag{26}$$

B. Channel Equalization

The channel transfer function H_k may not have a flat amplitude response or a linear phase response over the entire 240-MHz band of a Starlink channel, in which case channel equalization is helpful before proceeding with demodulation. As a native time-domain signal, the PSS is excluded from this process; thus, our concern is the symbol index set $\mathcal{I}_1 \triangleq \mathcal{I} \setminus 0$.

Let $\mathcal{K}_{\text{l}} \triangleq \mathcal{K} \setminus \mathcal{K}_{\text{g}}$ denote the set of loaded (non-gutter) subcarriers. The channel transfer function can be identified by using the known transmitted SSS symbols X_{m1k} for all $k \in \mathcal{K}_{\text{l}}$. Suppose that the corresponding received demodulated values $\bar{Y}_{m1k}$ have been determined using the techniques from the preceding subsection, and suppose we define $\phi_{m1} \triangleq 0$. Then the following relationship holds true for all $k \in \mathcal{K}_{\text{l}}$:

$$\bar{Y}_{m1k} = Z_m H_k \exp(-j2\pi d[k] F \tau_{m1}) X_{m1k} + V_{m1k} \tag{27}$$

Under the constraint $H_0 = 1$, a pre-equalization maximum-likelihood estimate of τ_{m1} can be obtained from $\bar{Y}_{m1k} X_{m1k}^{*}$, $k \in \mathcal{K}_{\text{l}}$ in the same way that frequency is estimated in [25]. This allows the exponential factor in (27) to be eliminated, yielding

$$\bar{Y}_{m1k} = Z_m H_k X_{m1k} + V_{m1k} \tag{28}$$

One can then estimate H_k (to within the constant Z_m) by computing $\breve{H}_k = \bar{Y}_{m1k}/X_{m1k}$ for all $k \in \mathcal{K}_{\text{l}}$. To mitigate the

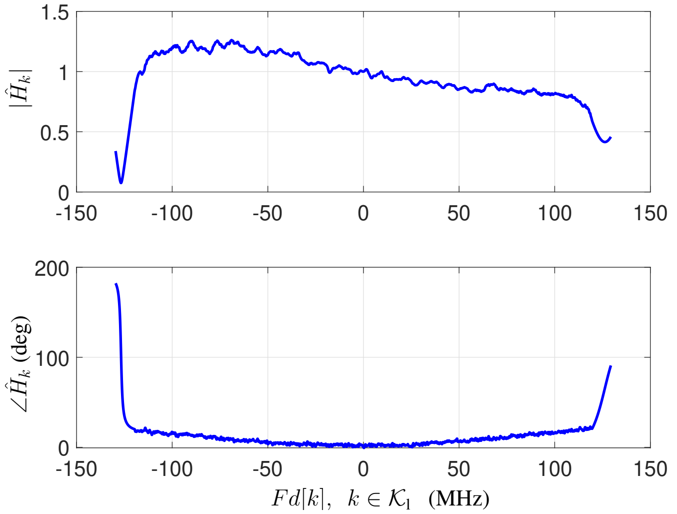

Fig. 4: Example estimated Starlink channel transfer function $\hat{H}_k$ as a function of offset frequency $Fd[k]$ for $k \in \mathcal{K}_{\mathrm{l}}$. This transfer function is for a data recording system that uses an Ettus X410 USRP with native sampling at $250\,\mathrm{MHz}$ and a 1.2-m steered-dish antenna.

effects of noise, one may apply a filter (e.g., a Savitzky–Golay filter) or a spline fit to the $\check{H}_k$ for all $k \in \mathcal{K}_{\mathrm{l}}$ to produce a smoothed function $\tilde{H}_k$. One obtains the value $\tilde{H}_0$ from the smoothed function and uses this to compute a normalized smoothed estimate of the transfer function, $\hat{H}_k = \tilde{H}_k/\tilde{H}_0$. This technique can be used for multiple successive frames of a given receiver/satellite/band combination to refine the channel transfer function estimate.

Fig. 4 shows an example estimated transfer function. Note how the amplitude falls off going from the low-frequency section to the high-frequency section. It also falls off rapidly near the two band edges. The phase stays relatively flat through most of the band, but changes rapidly near the band edges at the same frequencies that experience sharp amplitude roll-off.

The frame-specific constant Z_m is estimated as $\hat{Z}_m = \sqrt{\hat{g}}\exp(j\hat{\theta}_m)$, where $\hat{g}$ is the estimated channel gain and $\hat{\theta}_m$ is the estimated initial carrier phase:

$$\hat{g} = \frac{1}{|\mathcal{K}_{\mathrm{l}}|}\sum_{k\in\mathcal{K}_{\mathrm{l}}} |\bar{Y}_{m1k}/\hat{H}_k|^2, \quad \hat{\theta}_m = \angle\sum_{k\in\mathcal{K}_{\mathrm{l}}}(\bar{Y}_{m1k}/\hat{H}_k)X^*_{m1k}$$

Normalization by $\hat{H}_k$ and $\hat{Z}_m$ equalizes the received symbols for all $k \in \mathcal{K}_{\mathrm{l}}$ and for all $i \in \mathcal{I}_1$:

$$\tilde{Y}_{mik} = \bar{Y}_{mik}/(\hat{Z}_m\hat{H}_k)$$

The equalized symbols can be modeled by the following simplification of (23)

$$\tilde{Y}_{mik} = \exp(-j2\pi d[k]F\tau_{mi} + j\phi_{mi})X_{mik} + \tilde{V}_{mik} \tag{29}$$

with $\tilde{V}_{mik}$ being complex AWGN and τ_{mi} and ϕ_{mi} given as in (25) and (26).

C. Residual Synchronization Parameter Estimation

As is clear from (29), each $\tilde{\mathbf{Y}}_{mi}$, $i \in \mathcal{I}_1$ remains distorted by residual synchronization errors and phase misalignment. We remedy this via the multi-stage process described in [26] and summarized below.

First, each constellation $\mathcal{C}_{mi}$ must be identified for $i \in \mathcal{I}_1$. Starlink employs several different quadrature amplitude modulation (QAM) schemes to encode a sequence of digital data onto the N_{s} complex information symbols that modulate the subcarriers of each OFDM symbol. We distinguish 4QAM from QPSK constellations according to the convention

$$\mathcal{C}^{\mathrm{QPSK}} = \{1, j, -1, -j\}, \quad \mathcal{C}^{\mathrm{4QAM}} = \mathcal{C}^{\mathrm{QPSK}}\exp(j\pi/4)$$

Thus, $\mathcal{C}^{\mathrm{4QAM}}$ is equivalent to $\mathcal{C}^{\mathrm{QPSK}}$ modulo a 45° rotation. We exclude the edge pilots (formally introduced in Section VII) from the determination of $\mathcal{C}_{mi}$ because edge pilot symbols are uniformly 4QAM. Let $\mathcal{K}_{\mathrm{p}}$ be the subcarrier indices corresponding to the edge pilots, and let $\mathcal{K}_{\mathrm{lnp}} \triangleq \mathcal{K}_{\mathrm{l}} \setminus \mathcal{K}_{\mathrm{p}}$ be the set of loaded non-pilot subcarrier indices. Only subcarriers $k \in \mathcal{K}_{\mathrm{lnp}}$ are involved in constellation identification.

We identify $\mathcal{C}_{mi}$ by estimating its cardinality $|\mathcal{C}_{mi}|$ from $\tilde{\mathbf{Y}}_{mi}$ via k-means clustering. If the estimated cardinality is 16 or 32 we select $\hat{\mathcal{C}}_{mi} = \mathcal{C}^{\mathrm{16QAM}}$ or $\hat{\mathcal{C}}_{mi} = \mathcal{C}^{\mathrm{32QAM}}$ as appropriate; for cardinality 4, we initially assume $\hat{\mathcal{C}}_{mi} = \mathcal{C}^{\mathrm{QPSK}}$ and defer the QPSK/4QAM disambiguation to after residual synchronization compensation. For rare cases of other estimated cardinality (e.g., 8), we declare the OFDM symbol to have composite modulation within $\mathcal{K}_{\mathrm{lnp}}$ and discard it. Let $\tilde{\mathcal{I}}_m \subseteq \mathcal{I}_1$ be the set of all OFDM symbols retained after this process as applied to the mth frame.

Given $\hat{\mathcal{C}}_{mi}$, we estimate the residual delay and phase parameters τ_{mi} and ϕ_{mi} from (29) using a symbol-marginalized maximum-likelihood (ML) formulation. Since channel equalization has already normalized the symbols to approximately unit power, the ML formulation focuses solely on phase and delay estimation. Let $\boldsymbol{\theta}_{mi} \triangleq [\tau_{mi}, \phi_{mi}]$ and let

$$\Lambda_{\tilde{\mathbf{Y}}_{mi}}(\boldsymbol{\theta}_{mi} \mid \hat{\mathcal{C}}_{mi}) = \prod_{k\in\mathcal{K}_{\mathrm{l}}} p(\tilde{Y}_{mik} \mid \boldsymbol{\theta}_{mi}, \hat{\mathcal{C}}_{mi})$$

be the likelihood function of the vector measurement $\tilde{\mathbf{Y}}_{mi}$, where $p(\tilde{Y}_{mik} \mid \boldsymbol{\theta}_{mi}, \hat{\mathcal{C}}_{mi})$ is the likelihood of $\tilde{Y}_{mik}$ for a given $\boldsymbol{\theta}_{mi}$ assuming $\mathcal{C}_{mi} = \hat{\mathcal{C}}_{mi}$. For data-bearing information symbols, the likelihood function is obtained by marginalizing over the possible values for the unknown symbol X_{mik}, assuming complex Gaussian noise under a uniform prior on $\hat{\mathcal{C}}_{mi}$:

$$p(\tilde{Y}_{mik} \mid \boldsymbol{\theta}_{mi}, \hat{\mathcal{C}}_{mi}) = \frac{1}{|\hat{\mathcal{C}}_{mi}|}\sum_{c\in\hat{\mathcal{C}}_{mi}} p(\tilde{Y}_{mik} \mid X_{mik} = c, \boldsymbol{\theta}_{mi}) \tag{30}$$

Pilot information symbols are handled as the special case where X_{mik} is known and no marginalization is required. The ML estimate of $\boldsymbol{\theta}_{mi}$ may be found by

$$\hat{\boldsymbol{\theta}}_{mi} = \underset{\boldsymbol{\theta}_{mi}}{\operatorname{argmax}} \log \Lambda_{\tilde{\mathbf{Y}}_{mi}}(\boldsymbol{\theta}_{mi} \mid \hat{\mathcal{C}}_{mi}) \tag{31}$$

This process yields a sequence of per-OFDM-symbol delay and phase estimates $\{\hat{\tau}_{mi}, \hat{\phi}_{mi} \mid i \in \tilde{\mathcal{I}}_m\}$. While these

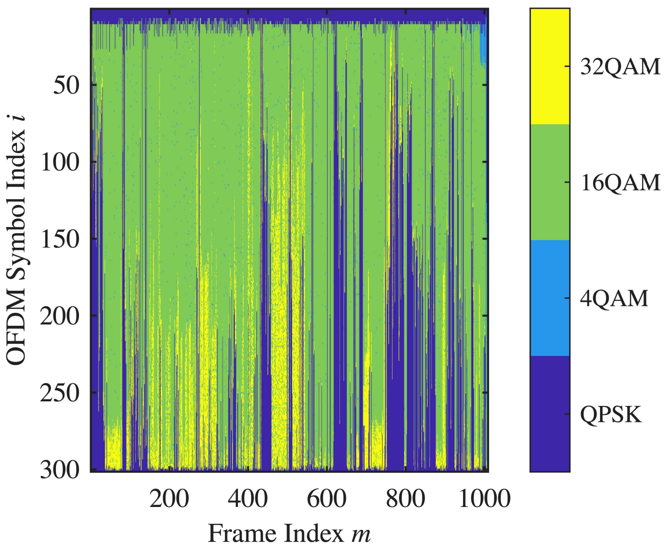

Fig. 5: Modulation scheme estimates for the set of 1009 exemplar frames, with each frame comprising $|\mathcal{I}_1| = 301$ OFDM symbols. Frames and symbols are ordered along the horizontal and vertical axes, respectively. The color of each cell indicates the estimated $\mathcal{C}_{mi}$ type for $m \in \{1, \dots, 1009\}$ and $i \in \mathcal{I}_1$. Note how the first few symbols (starting with the SSS) of every frame are QPSK-modulated. These symbols constitute the variable-length frame header.

estimates are statistically efficient on a per-symbol basis, they are noisy and do not obey the temporal structure imposed by (25) and (26) wherein τ_{mi} and ϕ_{mi} vary linearly with OFDM symbol index i. To exploit this structure and obtain a unified set of frame-level residual synchronization parameters, the per-symbol ML estimates are combined by a joint ML fit as follows. Let $\hat{\boldsymbol{\phi}}_m = [\hat{\phi}_{mi} \mid i \in \tilde{\mathcal{I}}_m]$ and $\hat{\boldsymbol{\tau}}_m = [\hat{\tau}_{mi} \mid i \in \tilde{\mathcal{I}}_m]$ respectively denote column vectors of per-OFDM-symbol phase and delay estimates obtained from the preceding ML stage for frame m. Likewise, let the vector $\boldsymbol{i} = [i \mid i \in \tilde{\mathcal{I}}_m]$ contain the relevant symbol indices, and $\mathbf{1}$ be the all-ones vector of matching dimension. Let $\boldsymbol{\psi}_m \triangleq [\phi_{m0}, \delta\beta_\text{c}, \tau_{m0}, \delta\beta_\text{s}]$ contain the parameters to be estimated, and define $\boldsymbol{J}(\boldsymbol{\psi}_m)$ as

$$\boldsymbol{J}(\boldsymbol{\psi}_m) \triangleq \begin{bmatrix} \hat{\boldsymbol{\phi}}_m - \left(\phi_{m0}\mathbf{1} - 2\pi T_\text{sym} F_\text{c} \delta\beta_\text{c}\, \boldsymbol{i}\right) \\ \hat{\boldsymbol{\tau}}_m - \left(\tau_{m0}\mathbf{1} - T_\text{sym}\, \delta\beta_\text{s}\, \boldsymbol{i}\right) \end{bmatrix}$$

Minimizing the squared magnitude of this function yields the ML estimate for $\boldsymbol{\psi}_m$:

$$\hat{\boldsymbol{\psi}}_m = \underset{\boldsymbol{\psi}_m}{\operatorname{argmin}} \left\| \boldsymbol{J}(\boldsymbol{\psi}_m) \right\|_2^2 \tag{32}$$

Compensation of the residual synchronization errors yields the phase-aligned symbols for all $k \in \mathcal{K}_\text{l}$ and $i \in \tilde{\mathcal{I}}_m$:

$$Y_{mik} = \tilde{Y}_{mik} \exp\left(-j\hat{\phi}_{mi} + j2\pi d[k] F \hat{\tau}_{mi}\right) \tag{33}$$

D. Information Symbol Detection

With the phase-aligned information symbol vector $\boldsymbol{Y}_{mi}$ available for all $i \in \tilde{\mathcal{I}}_m$, the QPSK/4QAM ambiguity for cardinality-4 symbols is resolved: For each $\boldsymbol{Y}_{mi}$, we select $\hat{\mathcal{C}}_{mi} = \mathcal{C}^\text{QPSK}$ or $\hat{\mathcal{C}}_{mi} = \mathcal{C}^\text{4QAM}$ depending on the relative alignment against $\boldsymbol{Y}_{m(i-1)}$, with $\boldsymbol{Y}_{m1}$, the phase-aligned SSS, providing an absolute reference.

Each hard-decoded symbol estimate $\hat{X}_{mik}$ is then obtained by mapping Y_{mik} to its nearest constellation point in $\hat{\mathcal{C}}_{mi}$:

$$\hat{X}_{mik} = \underset{c \in \hat{\mathcal{C}}_{mi}}{\operatorname{argmin}} \; |Y_{mik} - c|$$

This process produces the frame-level information symbol matrix

$$\hat{\boldsymbol{X}}_m \triangleq \left[\hat{\boldsymbol{X}}_{mi} \mid i \in \tilde{\mathcal{I}}_m\right] \in \mathbb{C}^{|\mathcal{K}| \times |\tilde{\mathcal{I}}_m|}$$

whose columns correspond to hard-decoded frequency-domain OFDM symbols. Fig. 5 illustrates the estimated modulation scheme for each OFDM symbol across the exemplar frames.

E. Alternate Residual Synchronization Parameter Estimation

A more nearly optimal method was developed to remove the residual carrier Doppler shift error $\delta\beta_\text{c}$ and phase ϕ_{m0} when determining the edge pilot symbols. In addition to marginalizing over $\mathcal{C}_{mi}$, as in (30), this method marginalized the probability distribution for $\tilde{\boldsymbol{Y}}_{mi}$ over the 4 possibilities for the ith symbol interval, namely QPSK, 4QAM, 16QAM, and 32QAM. This marginalization procedure assumed that each of these 4 possibilities is equally likely, despite evidence to the contrary. The marginalized probability density function for $\tilde{\boldsymbol{Y}}_{mi}$ was conditioned on the residual phase angles of its $\tilde{Y}_{mik}$ components

$$\phi_{mik} = \phi_{m0} - 2\pi(\delta\beta_\text{c} F_\text{c} + \delta\beta_\text{s} F d[k])(T_\text{sym}/[1 - \hat{\beta}_m])i$$

The negative-log-likelihood functions for all of the $\tilde{\boldsymbol{Y}}_{mi}$ symbol vectors were summed for $i \in \mathcal{I}_2$, where $\mathcal{I}_2 \triangleq \{2, 3, \dots, 301\}$ denotes OFDM symbol indices excluding the PSS ($i = 0$) and SSS ($i = 1$). Two additional terms were added to penalize non-zero $\delta\beta_\text{c}$ and the difference between ϕ_{m0} and its estimate based purely on the PSS and SSS symbols. The resulting negative-log-likelihood cost function was minimized to obtain estimates of $\delta\beta_\text{c}$ and ϕ_{m0}. This minimization applied the approximation $\delta\beta_\text{s} = \delta\beta_\text{c}$ in the ϕ_{mik} formula above. It started with a brute-force search over an extended $\delta\beta_\text{c}$ range in order to get near the global minimum instead of a local minimum. Afterwards, it applied Newton's method to find the optimal estimates of $\delta\beta_\text{c}$ and ϕ_{m0}. These estimates were then used to compute the phase-aligned, demodulated symbols $\{Y_{mik} \mid i \in \mathcal{I}_2, k \in \mathcal{K}_\text{l}\}$.

All loaded subcarriers—including those for the edge pilots—were included in this data fitting procedure so that the resulting Y_{mik} values could be used generally; e.g., to determine whether or not there exist edge pilots and other predictable symbols.

VII. EDGE PILOTS

We developed a tailored averaging procedure to search for information symbols that repeat exactly from frame to frame in the Starlink downlink. The procedure computed $|\mathcal{I}_2| \times |\mathcal{K}_\text{l}| = 306000$ averages of the received, equalized, and phase-aligned symbols Y_{mik}—one average per unique combination of $i \in \mathcal{I}_2$ and $k \in \mathcal{K}_\text{l}$. The averaging was performed across frames (i.e., over multiple m values) from multiple time intervals, multiple Starlink frequency bands, and multiple Starlink satellites (see the Supplementary Material for details). If the same coefficient X_{mik} is present for all frames, then the resulting average approaches X_{mik}. If, instead, X_{mik} varies with m, then the average approaches zero. Combinations (i,k) that produced large absolute values of the corresponding averages were examined to ensure that the individual Y_{mik} values used to produce the given average appeared to be noisy versions of the same frame-to-frame-repeated symbol X_{ik}.

Application of this averaging procedure revealed two bands of 4QAM-modulated pilot symbols, one near each edge of a Starlink channel's allocated bandwidth. The existence of these edge pilots was reported in [8], [23], but not their symbol values, which are published here. Each of the two edge pilot bands occupies 8 subcarriers. The relevant indices of both bands are $\mathcal{K}_\text{p} \triangleq \{488, 489, ..., 495, 528, 529, ..., 535\}$. A 4QAM pilot coefficient is modulated onto each $k \in \mathcal{K}_\text{p}$ for each $i \in \mathcal{I}_2$.

Notably, we found that the edge pilots appear to be identical, not only across all frames from the same Starlink SV, but also across all SVs in the constellation. The pilot coefficient X_{mik} takes the following form for OFDM symbol $i \in \mathcal{I}_2$ and subcarrier $k \in \mathcal{K}_\text{p}$:

$$X_{mik} = \exp\left[j\frac{\pi}{2}\left(s_{\text{p}ik} + \frac{1}{2}\right)\right] \tag{34}$$

$$s_{\text{p}ik} = \left\lfloor \frac{q_{\text{p}k}}{4^{(301-i)}} \right\rfloor \bmod 4 \tag{35}$$

The quantity $s_{\text{p}ik} \in \{0,1,2,3\}$ for $i \in \mathcal{I}_2$ and $k \in \mathcal{K}_\text{p}$ is a function of $q_{\text{p}k}$, a 150-digit hexadecimal number given in Appendix A for each $k \in \mathcal{K}_\text{p}$. The foregoing results are identical to those in [27] except for slight notational adjustments and for the basis of representation (base-4 vs. hexadecimal).

VIII. LOW-ENTROPY QPSK SYMBOLS

This section reveals the surprising fact that nearly all QPSK-modulated information symbols in the Starlink Ku-band downlink are not high-entropy user data. Rather, outside a short header region, these symbols exhibit a highly repetitive structure across the full capture corpus: deviations relative to a reference set of information symbols called the *reference template* are BPSK-valued and follow a simple periodicity with deterministic shifts across OFDM symbols. This structural repetition enables construction of long, near noise-free local replicas for matched filtering, boosting processing gain.

A. Reference Template

Our analysis of low-entropy structure (aside from the edge pilots) in decoded frame data is based on the exemplar frames mentioned in Section IV, with discovered patterns verified against data from the larger corpus. The exemplar frames, denoted $\{\hat{\boldsymbol{X}}_m \mid m \in \mathcal{M}_\text{e}\}$ with $\mathcal{M}_\text{e} = \{1, 2, \ldots, 1009\}$, all come from a single capture of data from one SV. They are ordered chronologically, with no frames from occupied slots discarded. The pre-correlation SNR for each exemplar frame is greater than 9 dB and the average over the exemplar set is 13.8 dB.

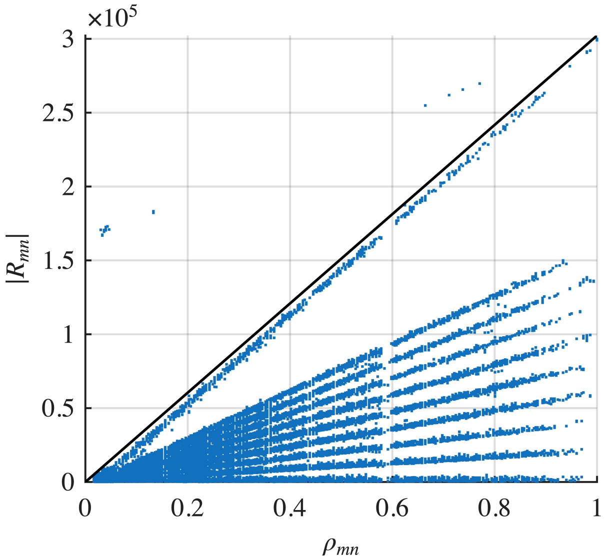

Fig. 6: $|R_{mn}|$ vs. ρ_{mn} for $m, n \in \mathcal{M}_\text{e}$, with $n < m$ to avoid duplicate pairwise comparisons. The black line shows the upper limit of $|R_{mn}|$, which obtains when all QPSK-modulated OFDM symbols carry precisely the same information. A small number of points lie above this theoretical limit. Investigation revealed that these correspond to consecutive frames (i.e., $\{m, m+1\}$) where the SV appears to have retransmitted duplicate data.

To identify portions of the Starlink frame that repeat from frame to frame, we compute pairwise cross-correlations. We begin by defining several useful sets and measures. Let the set of non-SSS QPSK-modulated symbol indices for frame m be

$$\mathcal{I}_{\text{Q}m} \triangleq \{\, i \in \mathcal{I}_2 \mid \hat{\mathcal{C}}_{mi} = \mathcal{C}^{\text{QPSK}} \,\}$$

and the corresponding QPSK ratio be $\rho_m \triangleq |\mathcal{I}_{\text{Q}m}|/|\mathcal{I}_2|$. For a pair of frames $\{m, n\}$, define the indices of symbols that are QPSK in both frames as

$$\mathcal{I}_{\text{Q}mn} \triangleq \{\, i \in \mathcal{I}_2 \mid \hat{\mathcal{C}}_{mi} = \hat{\mathcal{C}}_{ni} = \mathcal{C}^{\text{QPSK}} \,\}$$

The corresponding shared QPSK ratio is $\rho_{mn} \triangleq |\mathcal{I}_{\text{Q}mn}|/|\mathcal{I}_2|$. Define the cross-correlation between frames m and n, restricted to their shared QPSK OFDM symbols and loaded non-pilot subcarriers, as

$$R_{mn} \triangleq \sum_{i \in \mathcal{I}_{\text{Q}mn}} \sum_{k \in \mathcal{K}_\text{lnp}} \hat{X}^*_{mik} \hat{X}_{nik} \tag{36}$$

If QPSK-modulated OFDM symbols carried high-entropy user data from frame to frame, then $|R_{mn}|$ would concentrate near zero regardless of ρ_{mn}. An empirical analysis of $|R_{mn}|$ instead shows a clear dependence on ρ_{mn}, as illustrated in Fig. 6: the data form a structured fan with vertically quantized bands. This behavior indicates that QPSK-modulated OFDM symbols exhibit various degrees of correlation across frames and therefore warrant focused analysis.

Further study revealed that pure-QPSK frames, those for which $\rho_m = 1$, exhibit striking consistency: Nearly all information symbols agree across pure-QPSK frames in the exemplar set. Exceptions are a mix of rare hard-decoding errors and some genuine frame-to-frame variation within the frame header. Let the indices of pure-QPSK frames be denoted $\mathcal{M}_{\text{ep}} \triangleq \{m \in \mathcal{M}_{\text{e}} \mid \rho_m = 1\}$. We define the *reference template* $\boldsymbol{T} \in \mathbb{C}^{|\mathcal{K}_{\text{lnp}}| \times |\mathcal{I}_2|}$ as the element-wise mode of the hard-decoded symbols $\{\hat{\boldsymbol{X}}_m \mid m \in \mathcal{M}_{\text{ep}}\}$. In other words, the (i, k)th element of $\boldsymbol{T}$ for $i \in \mathcal{I}_2$ and $k \in \mathcal{K}_{\text{lnp}}$ is defined as

$$T_{ik} \triangleq \operatorname*{argmax}_{c \in \mathcal{C}^{\text{QPSK}}} \big|\{\, m \in \mathcal{M}_{\text{ep}} \mid \hat{X}_{mik} = c \,\}\big| \tag{37}$$

We provide the full reference template $\boldsymbol{T}$ in the Supplementary Material.

B. Tessellation Codes

This section leverages the reference template $\boldsymbol{T}$ to expose the highly regular structure of QPSK-modulated OFDM symbols, then introduces the repeating bit patterns that underlie this structure.

To quantify deviations from the reference template, we restrict $\{\hat{\boldsymbol{X}}_m \mid m \in \mathcal{M}_{\text{e}}\}$ to QPSK-modulated symbols, then compute the element-wise product between the decoded information symbol and the conjugate of the corresponding reference template element. Let D_{mik} denote the (i, k)th element of the deviation, with corresponding OFDM-symbol-level vector $\boldsymbol{D}_{mi} \in \mathbb{C}^{|\mathcal{K}_{\text{lnp}}|}$ and frame-level matrix $\boldsymbol{D}_m \in \mathbb{C}^{|\mathcal{K}_{\text{lnp}}| \times |\mathcal{I}_{\text{Q}m}|}$. The element is defined for $m \in \mathcal{M}_{\text{e}}$, $i \in \mathcal{I}_{\text{Q}m}$, and $k \in \mathcal{K}_{\text{lnp}}$ as

$$D_{mik} \triangleq \hat{X}_{mik} T^*_{ik} \tag{38}$$

Inspection of $\boldsymbol{D}_m$ reveals regular patterns, as follows.

BPSK Deviations: A histogram analysis of $\boldsymbol{T}$, and of the set of all QPSK-valued OFDM symbols in the exemplar frames $\{\hat{\boldsymbol{X}}_{mi} \mid m \in \mathcal{M}_{\text{e}}, i \in \mathcal{I}_{\text{Q}m}\}$, and of the set of all deviation matrices $\{\boldsymbol{D}_m \mid m \in \mathcal{M}_{\text{e}}\}$ revealed that, whereas the former two are uniformly distributed on $\mathcal{C}^{\text{QPSK}}$, the deviation matrices are in fact BPSK-modulated; in particular, $D_{mik} \in \{1, -1\}$. Thus, QPSK-modulated OFDM symbols can be expressed as the product of a BPSK-valued sequence and the corresponding column of $\boldsymbol{T}$. Analysis of the sequence of deviation vectors $\{\boldsymbol{D}_{mi} \mid m \in \mathcal{M}_{\text{e}}, i \in \mathcal{I}_{\text{Q}m}\}$ reveals that the QPSK-modulated portion of each frame comprises two distinct temporal regions: a header region and a T-code region.

Header: After the PSS and SSS, each frame begins with a sequence of QPSK-modulated OFDM symbols that form a header of variable length whose information content differs from the highly regular structure observed later in the frame. In this paper we do not attempt to interpret the header's content semantically; however, preliminary analysis shows multi-modal frame-to-frame repetition.

Let $\mathcal{I}_{\text{H}m} \triangleq \{i \in \mathcal{I}_2 \mid i \leq i_{\text{h}m}\}$ denote the header-symbol indices for the mth frame, where $i_{\text{h}m}$ is the index of the final header symbol. Fig. 7 shows the distribution of $i_{\text{h}m}$ across the exemplar frames. Within the header region, individual OFDM symbols often manifest long sequences of value -1. These can be seen as vertical blue strips on close inspection of Fig. 8. Perhaps certain header regions are reserved for data only transmitted intermittently and are otherwise set to -1. In any case, these runs of constant symbols and other obvious header structure clearly indicate that the header data are not encrypted.

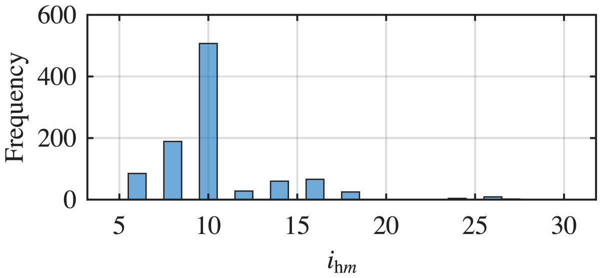

Fig. 7: Distribution of the header length $i_{\text{h}m}$ across the exemplar frames.

T-Codes: Let $\mathcal{I}_{\text{T}m} \triangleq \{i \in \mathcal{I}_{\text{Q}m} \mid i > i_{\text{h}m}\}$ denote the indices of post-header QPSK-modulated OFDM symbols. Inspection of $\{\boldsymbol{D}_{mi} \mid i \in \mathcal{I}_{\text{T}m}\}$ reveals a tessellated pattern, an example of which is observable in Fig. 8. Between successive OFDM symbols, the deviation's BPSK sequence is circularly shifted by 16 subcarriers from each OFDM symbol to the next. Within each OFDM symbol, the deviation's sequence repeats at 60-subcarrier intervals for $k \in \mathcal{K}_{\text{lnp}}$. We refer to the 60-bit BPSK sequence as a tessellation code, or *T-code*. Whereas the tessellation periodicity is consistent across frames, the T-codes themselves vary from frame to frame. If non-QPSK OFDM symbols interrupt a continuous sequence of QPSK-modulated OFDM symbols, the pattern continues as if there were no interruption. Continuity is verified by observing that the circular-shift phase accumulates correctly across such gaps. Let $N_{\text{T}} = 60$ be the T-code length and $d_{\text{T}} = 16$ the circular shift period. Then for all $i \in \mathcal{I}_{\text{T}m}$, $k \in \mathcal{K}_{\text{lnp}}$ we may write

$$D_{mik} = D_{mi(k+N_{\text{T}}) \bmod |\mathcal{K}_{\text{lnp}}|} \tag{39}$$

$$\boldsymbol{D}_{m(i+\Delta i)} = \texttt{circshift}(\boldsymbol{D}_{mi},\ d_{\text{T}} \cdot \Delta i) \tag{40}$$

where Δi is the OFDM symbol index difference and $\texttt{circshift}(\boldsymbol{x}, d)$ is a function that circularly shifts elements of the vector $\boldsymbol{x}$ by d positions. We provide all unique T-codes within the exemplar frames in the Supplementary Material.

IX. DISCUSSION

A. T-Codes: Significance, Persistence, and Diversity

1) Significance: The repeating T-code structure implies that the information content of large blocks within Starlink frames is only $N_{\text{T}} = 60$ bits. If these bits can be determined (e.g., by circularly shifting, stacking, and averaging), then the QPSK-modulated data within each frame can be coherently accumulated. As shown in Section III, such extended coherent processing directly translates into increased processing gain. Note that T-code symbols appear to be fairly common: across the exemplar frames, 21.5% of all post-header OFDM symbols were T-code symbols.

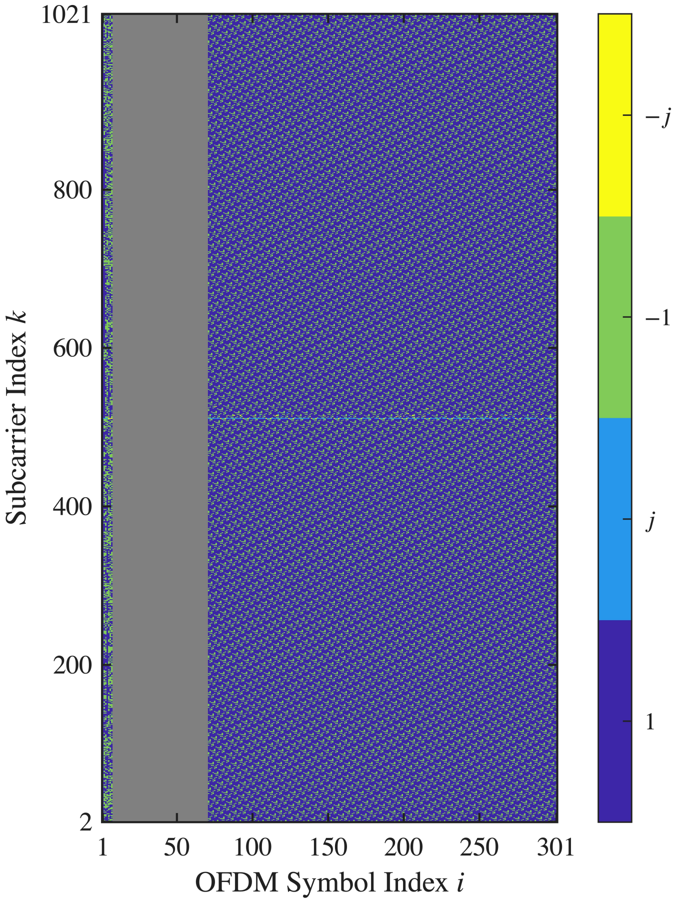

Fig. 8: A color-coded plot of $\boldsymbol{D}_m$ for $m = 814 \in \mathcal{M}_{\mathrm{e}}$. Readers are encouraged to magnify the plot to appreciate its fine structure. The horizontal and vertical axes are ordered respectively by OFDM symbol index $i \in \mathcal{I}_1$ and subcarrier index $k \in \mathcal{K}_{\mathrm{lnp}}$. The narrow vertical strip to the far left is the deviation within the frame's header, $\{\boldsymbol{D}_{mi} \mid i \in \mathcal{I}_{\mathrm{H}m}\}$. The gray region spans the non-QPSK symbols for this frame, with indices $i \in \{\mathcal{I}_2 \backslash \mathcal{I}_{\mathrm{Q}m}\}$. The frame's T-code tiles the remainder of the frame with a tessellated pattern. The thin horizontal line and speckles near the center ($k \approx 511$) are decoding artifacts caused by increased phase noise at the far edges of the Starlink channel. Note how the T-code and the header deviation exhibit only green and blue colors, consistent with $D_{mik} \in \{1, -1\}$.

2) Persistence: Whether any given signal feature is useful for opportunistic PNT depends on its reliability and persistence. We therefore offer several observations and conjectures regarding the purpose and durability of the low-entropy structures discovered in Section VIII-B. SpaceX's 2024 patent [23] indicates that the Starlink SV-side downlink signal processing chain includes both a single-tap whitening scrambler and a low-density parity-check (LDPC) encoder. Whitening scramblers disperse signal energy to ensure a uniform power spectral density, thereby maintaining DC balance and reducing the peak-to-average power ratio of the transmitted OFDM waveform. The use of a single-tap architecture is significant: by applying a deterministic complex-domain phase rotation specific to each subcarrier independently, the transmitter preserves inter-carrier orthogonality and enables computationally efficient, memoryless descrambling at the receiver.

The presence of a constant reference template is consistent with the output of a single-tap scrambler, while the observed tessellation pattern is consistent with the numerological interaction between a 16-wide vectorized hardware processing architecture and a 60-bit periodicity introduced by structured channel coding (e.g., an LDPC encoder). The observed frame-to-frame T-code variations suggest that the scrambler state is re-seeded between frames.

We conjecture that the tesselation pattern shown in Fig. 8, which is present in some form in the majority of exemplar frames, is the result of a single-tap whitening scrambler and an LDPC encoder operating on *unladen* payload allocations—OFDM symbols that are not bearing any user data. Such empty resource blocks arise due to time-varying user demand and unavoidable inefficiency in resource scheduling. Starlink SVs likely populate empty payload allocations with a deterministic fill pattern (e.g., an all-ones vector) prior to scrambling and LDPC encoding. The resulting QPSK symbols thus appear to be user data upon demodulation, yet are in fact low-entropy (60-bit) transformations of the same reference template.

These observations suggest that T-codes are neither arbitrary artifacts nor transient byproducts of a particular software release, but are instead tied to the transmitter's physical-layer architecture. As such, they are likely to remain a durable feature of the Starlink downlink, even if the fraction of OFDM symbols exhibiting T-code structure varies with per-beam user demand.

3) Diversity: Although a 60-bit-sequence set could theoretically include 2^{60} members, analysis of T-codes extracted from the exemplar frames suggests a much smaller active set. Of the 1009 exemplar frames, 530 contained T-codes, among which there were only 40 unique sequences. Each unique sequence appeared on average 13.25 times, the least frequent 3 times and the most frequent 24 times. These results suggest that the active T-code family comprises far fewer than 2^{60} unique T-codes.

B. Processing Gain and TOA Bounds

To understand the implications of these discoveries for PNT, we approximate the potential processing gain now available via exploitation of all low-entropy structure in the Starlink frame known to date, relying on empirical data obtained from the exemplar frames. For this analysis, we make several simplifying assumptions. First, we conservatively ignore the CP on all OFDM symbols other than the PSS and the SSS. Second, we treat the header as unknown data. Third, we assume that T-code symbols can be distinguished from non-T-code symbols even at low SNR.

To calculate the frame-level processing gain, let $\bar{N}$ denote the average number of information symbols per frame to which we can assign a non-uniform distribution. Since symbols with a uniform prior do not contribute to the metric L, we may substitute $\bar{N}$ for N in (5). We count in $\bar{N}$ all PSS, SSS, and edge pilot symbols, as these are fully known. For frames containing T-codes, we assume that the corresponding T-code sequence is hard-decoded by stacking, shifting, and combining the frame's T-code symbols. Let $|\mathcal{I}_{\mathrm{T}m}|$ denote the number of T-code OFDM symbols in frame m, and define the stacking factor for frame m as $M_m = |\mathcal{I}_{\mathrm{T}m}||\mathcal{K}_{\mathrm{lnp}}|/N_{\mathrm{T}}$. Define $\bar{M}$ as the average of M_m over the exemplar frames $m \in \mathcal{M}_{\mathrm{e}}$. Thus, T-code information symbols are included in $\bar{N}$ as the hard-decoded bits that permit the assignment of a non-uniform distribution to each symbol.

Next, we estimate the average (across frames) of the correlation coefficient μ from (4). For each subcarrier $k \in \mathcal{K}_{\mathrm{l}}$, we construct the local replica element l_k as follows. For pilot symbols, l_k is set to the known information symbol value, and since the local replica is exact, $|\mu| = 1$. For T-code symbols, l_k is reconstructed from the hard-decoded T-code bits and the known tesselation pattern. Assume that l_k agrees with the true symbol with probability $(1 - p_{\mathrm{e}})$ and disagrees with probability p_{e}. Then $|\mu| = |1 - 2p_{\mathrm{e}}|$, where bit-error probability for coherent BPSK demodulation is

$$p_{\mathrm{e}} = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\bar{M}\,\mathrm{SNR}_{\mathrm{pre}}}\right)$$

with $\operatorname{erfc}(\cdot)$ being the complementary error function. The average correlation coefficient $\bar{\mu}$ is calculated as the weighted average of $|\mu|$ across all exploitable symbols.

For the exemplar frames, whose mean $\mathrm{SNR}_{\mathrm{pre}}$ is $13.8\,\mathrm{dB}$, $\bar{N} \approx 69{,}500$ and $\bar{\mu} \approx 1$. Applying (5) yields an average processing gain $\bar{L} \approx 48.4\,\mathrm{dB}$, which is only $6.6\,\mathrm{dB}$ below the theoretical maximum $L \approx 55\,\mathrm{dB}$ calculated in Section III for full-frame correlation with completely known symbols.

At lower SNR, processing gain degrades due to increased T-code bit estimation errors. However, for the exemplar frames, the large stacking factor $\bar{M} \approx 1044$ provides a boost of approximately $30\,\mathrm{dB}$ to the effective SNR for T-code bit decisions. Thus, even at $\mathrm{SNR}_{\mathrm{pre}} = -25\,\mathrm{dB}$, T-code recovery remains nearly perfect, yielding $\bar{\mu} \approx 0.995$ and $\bar{L} \approx 48.4\,\mathrm{dB}$.

The gray traces in Fig. 2 show the resulting single-frame TOA precision bounds when exploiting all known low-entropy elements (LEEs)—the PSS, SSS, edge pilots, and T-codes—assuming T-code occurrence statistics as observed in the exemplar frames. The bounds are computed with the receiver's local oscillator center frequency tuned such that the edge pilots occupy subcarriers at the outer edges of the captured bandwidth, thereby maximizing the mean-square bandwidth and tightening both the ZZB and CRB. Notably, the LEEs ZZB remains coincident with its corresponding CRB across the entire pre-correlation SNR range plotted, matching the behavior of the full-frame bounds. This indicates that the discovered predictable elements are sufficient to suppress ambiguity sidelobes in the local replica correlation function.

C. TOA Based on Edge Pilots

A lower-performance but especially simple implementation of Starlink-based opportunistic LEO PNT is possible with only the PSS, SSS, and edge pilots. Fig. 9 shows TOA RMSE bounds as functions of $\mathrm{SNR}_{\mathrm{pre}}$ when the PSS+SSS local replica is augmented with the known edge pilot symbols. For full-bandwidth processing with $B_{\mathrm{f}} = 240\,\mathrm{MHz}$, adding both edge pilot bands yields only a modest change in the ZZB-CRB divergence point, but reduces the TOA RMSE bound by an approximately constant factor of 2.75 across the plotted SNR range due to additional processing gain provided by the edge pilots.

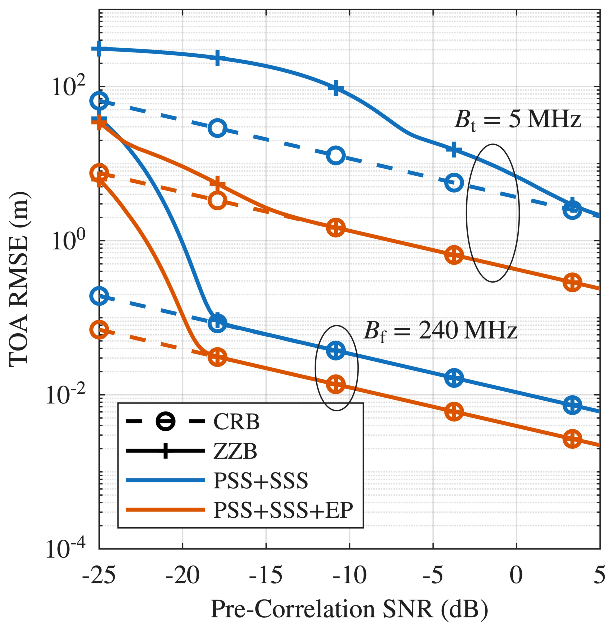

Fig. 9: Lower bounds on single-frame TOA RMSE vs. $\mathrm{SNR}_{\mathrm{pre}}$ for coherent processing with the PSS, SSS, and edge pilots (EP) under two representative capture bandwidths. From top to bottom, bounds are shown for (1) narrowband correlation with a $5\,\mathrm{MHz}$ subband of the PSS+SSS combination; (2) narrowband correlation with a coherent combination of the $5\,\mathrm{MHz}$ PSS+SSS subband concatenated in time with a single edge pilot band; (3) full-bandwidth correlation over $240\,\mathrm{MHz}$ with the PSS+SSS combination, as shown in Fig. 2; and (4) full-bandwidth correlation over $240\,\mathrm{MHz}$ with PSS+SSS and both edge pilot bands.

For narrowband processing with a tight $B_{\mathrm{t}} = 5\,\mathrm{MHz}$, augmenting the narrowly filtered PSS+SSS with a single edge pilot provides a larger benefit. Concatenating the edge pilot band with a matched $5\,\mathrm{MHz}$ PSS+SSS subband increases the effective coherent span and significantly suppresses TOA ambiguity, shifting the ZZB-CRB divergence point from approximately 6.2 to $-12.0\,\mathrm{dB}$.

D. Other Frame-Invariant Data

Reference [8] reported additional inter-frame-correlated data besides the PSS, SSS, header, and edge pilots, modeling these additional "beacon" data as appearing deterministically in all frames. Fig. 6 does indicate various levels of pairwise correlation between frames, but as we have revealed, this is due to the small-family T-code deviations from the reference template $\boldsymbol{T}$. Since T-codes vary from frame to frame and are approximately zero-mean, there is no guarantee of significant inter-frame correlation of QPSK-modulated data unless one wipes the T-code modulation from the underlying reference template $\boldsymbol{T}$. We carefully searched multiple Starlink satellite signals for the purported additional frame-to-frame repeated data using the averaging technique presented in Section VII and did not find them. We conjecture that the additional correlated data found by [8] are not actually invariant across Starlink frames but only appeared so in the cases analyzed therein.

X. CONCLUSION

We developed an acquisition and demodulation framework to hard-decode Starlink Ku-band downlink frames and applied it to a corpus of frames captured across multiple satellites and dates. Our analysis revealed key low-entropy structures whose existence or details were previously unknown to the public: edge pilots, the reference template, and T-codes. Exploiting these discoveries yields approximately $48\,\mathrm{dB}$ of processing gain, making precise TOA extraction feasible at minimum SNRs well below those required for communications. These results support robust opportunistic PNT using compact feed-horn receivers and suggest that Starlink side-beam signals may be far more usable for navigation than previously assumed.

ACKNOWLEDGMENTS

Research was supported by the U.S. Department of Transportation (DOT) under Grant 69A3552348327 for the CARMEN+ University Transportation Center, by another U.S. DOT Grant for the CARNATIONS University Transportation Center, by the Department of Defense NDSEG Fellowship program, and by affiliates of the `6G@UT` center within the Wireless Networking and Communications Group at The University of Texas at Austin.

APPENDIX A
EDGE PILOT CODES

The hexadecimal numbers $q_{\mathrm{p}k}$ used to define the edge pilot symbols $s_{\mathrm{p}ik}$ are given below for each of the 16 pilot subcarrier indices $k \in \{488, 489, ..., 495, 528, 529, ..., 535\}$. The edge pilots are also given directly in the Supplementary Material as a $|\mathcal{K}_\mathrm{p}|$-by-$|\mathcal{I}_2|$ complex-valued matrix whose (k, i)th element is X_{mik} from (34).

$q_{p488} =$

```
            76 3404 6DA4 5F89 042D 0117
E163 167D 4AE8 32D8 5751 5F3C AD90 3376
97FB 8F1C D048 EFEF 559E CD79 688B CBBF
44D2 FA9B DFAE 639D B5D7 B1DD 2DDC E4EC
9733 C0D4 DCCF 3172 A0EC 34CC 226C 530E
```

$q_{p489} =$

```
            CD 9AFA A654 147A 5FE2 B407
B51F FE15 215B 3A71 6241 3961 9628 A9C3
3E8E 3A32 E514 6C09 BD3E E902 6CA5 2032
D7FD 3896 0FFC 5259 9E9B 8A7F 6942 334B
D4C6 D99D 4331 DEF5 6745 70B2 45FB B25F
```

$q_{p490} =$

```
            02 481A 2B27 8B88 F096 C8D1
74D3 69D0 CF67 81B7 0EBD 402D 6A6F 4C98
5DA6 2658 66A8 374D C0B3 E491 7146 FE32
74CA 5D61 C3F9 A31C B812 5F29 1155 CBD4
F4F8 4E93 C0D8 54BB C54E E144 43EC 2DF8
```

$q_{p491} =$

```
            D8 DC99 C265 4265 B8C3 2450
114C 37E2 B725 A822 F105 4B46 F272 8771
22E4 7109 F113 D59E 37DF F418 FEA3 627C
7A5C C0A9 3ABA 0F94 08E9 58DF 4179 C4DE
40CE F842 D333 632B 3E77 BEB3 4B2E 6045
```

$q_{p492} =$

```
            3C C5CA 83B0 D330 89B1 4C3B
6AC3 D194 6359 726B 4966 B2E9 66BE 6112
4A5D 53E2 2A73 EDBE B383 A92F 06CA 6CAA
8A5B 1ECE 6954 6514 5E28 6EEE 1804 CD79
A00C 84FC 80C8 7DE9 DF57 2F9B 54AE 798B
```

$q_{p493} =$

```
            C7 7BD5 9D15 C2C9 17EE C97F
B479 B9F0 B2BF 5D2E CCD8 0248 D2AC 68C8
4CEA 11BA D18D 9F6F 31B6 AFD7 8334 7943
562E 2C68 32EA 7682 8FCD AB31 EFF6 A9A8
8EA4 8E3A FA62 5B2F CDA7 B99B 0295 E926
```

$q_{p494} =$

```
            61 52EF 153B 8511 0FB0 B7E2
4D83 34B1 C419 6DE8 72B5 9876 7BC3 CB4A
4827 A09D 924A A7F5 7EB9 46F1 981D 036E
3001 934B 10C9 E22A BB6A F1F0 47B3 A874
CA95 E68C BA67 063F 605F D05D 532A AD3C
```

$q_{p495} =$

```
            CD 8CAC F9DE FACD 2CB9 8114
39D8 B7E1 6F9E 09BE D473 7020 7150 A86D
FE24 EA12 98CC B090 7F5B AB67 D466 0462
C6B1 0F74 B8D9 FA7B 6F9E C139 9B30 B43A
F622 A894 B222 0B6B 509A 84AA BB58 D023
```

$q_{p528} =$

```
            CC BF3A 1692 9836 160C EC6E
B741 7AE6 C37D C1E8 28CE FB60 CE0E 6C3B
546A 76B0 AE1E 7BC0 E957 7528 B0F7 8F82
A410 4EA2 C316 B945 D385 200C 7E5A 1C5B
48F5 F9F9 AF5C 4BA9 20AC A3A5 99DB 9974
```

$q_{p529} =$

```
            9C F72F 5F5B 95CE 7342 C925
CF1A AF45 7F18 2C32 810E 2F74 8670 5D5F
A2D9 C892 3B01 73FB 206B 4604 5C6F 162B
B9FF D051 DB5E 5900 EFD2 DE24 D4BB 3FE8
7DD7 76F0 0B56 13A7 D22B 2821 E139 A599
```

$q_{p530} =$

```
                    29 6319 D723 2101 8995 3BB7
30DC 6046 E4EC 5FB4 8F97 18D5 B600 A015
78CA C315 9B58 EE8A 3066 6392 1FBE 78EE
7C1E 8E04 9B42 30A1 4EB4 9549 33AB 64F6
7B39 6DD6 DB12 BCBB 3CCA 60EA 79E0 614B
```

$q_{p531} =$

```
                    10 17FB BD3D 0398 1EE9 F442
4D47 3B8A 73E1 36C7 7795 6EAE BD4C A51E
9B70 D9F5 D106 57F2 6859 5A5C 3687 D2DD
06C9 8630 F817 CABE F3EE 6608 2235 0A70
F10A 29A8 7402 12A9 CF7E 7D81 4D60 A69C
```

$q_{p532} =$

```
                    71 2EA4 82B2 8E96 676E 65D0
9994 9655 8731 4F2B 562D 0E75 0FE5 66E8
9205 A8D4 DFED 2C4F AFFC 5ED1 EA6F B63E
C135 1344 4006 B78A DFB4 BDB6 CB05 4706
01C9 F8F4 9014 2306 9C9F BD68 D292 C16F
```

$q_{p533} =$

```
                    58 4E9F 48AC A087 84E6 9664
4C78 ED96 84FC 484F 32AA 1B4D A8E9 5457
358D F89F E8B9 D84D 47F3 0D3C A2F2 DDF0
E76E 57F1 4A44 6753 26ED CF15 052C B62B
7DF0 EBE6 2305 7605 CF24 06E2 5BD5 6B3B
```

$q_{p534} =$

```
                    4A F2EC F329 83A9 E781 852F
6E90 DC6C CE90 1863 F527 E038 DA22 C0CE
02E4 4FA0 5637 18D9 3E74 5429 3962 B435
94CC 2EE4 27FA E6F1 5C12 38D9 C85A BC4E
303F 3AEC 3404 A523 10CA C037 8665 E19A
```

$q_{p535} =$

```
                    08 4AA7 3DF9 F605 3582 9A71
6EC9 4D95 AA69 01B4 1E81 AEF2 8B03 F08C
DE7D 4542 5B11 6400 9D56 459C 4286 E269
F4B8 EBDB A8BF 6FC7 9847 B08A 69F7 9AF6
E6A7 AF05 DA50 4455 BA72 727D D7BE 7744
```

REFERENCES

[1] S. McNeil, “NATO says it is working to counter Russia’s GPS jamming after interference with EU leader’s plane,” https://apnews.com/article/europe-security-russia-gps-jamming-nato-rutte-6efed7d030da19a51dfacec39ed77325, Sept. 2025, accessed: 2025-10-24.

[2] M. Zulhusni, “UN agencies say GPS interference is getting worse,” https://techhq.com/news/un-agencies-say-gps-interference-is-getting-worse-heres-whats-at-risk, Mar. 2025, accessed: 2025-10-24.

[3] H. W. Jones, “The recent large reduction in space launch cost,” in *48th International Conference on Environmental Systems (ICES)*. Albuquerque, New Mexico: ICES, July 2018, iCES-2018-81.

[4] P. A. Iannucci and T. E. Humphreys, “Economical fused LEO GNSS,” in *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, 2020, pp. 426–443.

[5] Z. Tan, H. Qin, L. Cong, and C. Zhao, “Positioning using IRIDIUM satellite signals of opportunity in weak signal environment,” *Electronics*, vol. 9, no. 1, p. 37, 2019.

[6] M. Lei, Y. Liu, M. Gao, Z. Fang, J. Chen, and Y. Xu, “Robust Helmert variance component estimation for positioning with dual-constellation LEO satellites’ signals of opportunity,” *Electronics*, vol. 14, no. 17, p. 3437, 2025.

[7] Z. M. Kassas, N. Khairallah, and S. Kozhaya, “Ad astra: Simultaneous tracking and navigation with megaconstellation LEO satellites,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 39, no. 9, pp. 46–71, Jan. 2024.

[8] S. Kozhaya, J. Saroufim, and Z. M. Kassas, “Unveiling Starlink for PNT,” *NAVIGATION*, vol. 72, no. 1, Mar. 2025.

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

[10] 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*, vol. 42, no. 1, pp. 146–160, Jan. 2024.

[11] E. Grayver, R. Nelson, E. McDonald, E. Sorensen, and S. Romano, “Position and navigation using Starlink,” in *Proceedings of the 2024 IEEE Aerospace Conference*. Big Sky, MT, USA: IEEE, Mar. 2024, pp. 1–12.

[12] S. C. Morgan, Z. M. Komodromos, W. Qin, Z. L. Clements, A. M. Graff, W. J. Morrison, and T. E. Humphreys, “A mock implementation of fused LEO GNSS,” in *Proceedings of the IEEE/ION PLANS Meeting*, Salt Lake City, UT, 2025.

[13] Z. M. Komodromos, S. C. Morgan, Z. L. Clements, W. Qin, W. J. Morrison, and T. E. Humphreys, “Network-aided pseudorange-based LEO PNT from OneWeb,” in *Proceedings of the IEEE/ION PLANS Meeting*, Salt Lake City, UT, 2025.

[14] P. A. Iannucci and T. E. Humphreys, “Fused low-Earth-orbit GNSS,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 60, no. 4, pp. 3730–3749, Jun. 2022.

[15] M. Neinavaie, J. Khalife, and Z. M. Kassas, “Acquisition, Doppler tracking, and positioning with Starlink LEO satellites: First results,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 3, pp. 2606–2610, 2022.

[16] S. E. Kozhaya and Z. M. Kassas, “Positioning with Starlink LEO satellites: A blind Doppler spectral approach,” in *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, 2023, pp. 1–5.

[17] R. C. Johnson, Ed., *Antenna Engineering Handbook*, 3rd ed. New York, NY, USA: McGraw-Hill, 1993.

[18] C. A. Balanis, *Antenna Theory: Analysis and Design*, 3rd ed. Hoboken, NJ, USA: John Wiley & Sons, 2005.

[19] C. Ruf, J. Perez, and S. White, “EVLA memo 222: Assessment of low Earth orbit broadband constellation downlink power flux density at Earth,” National Radio Astronomy Observatory (NRAO), Socorro, NM, USA, Tech. Rep., Aug 2023. [Online]. Available: https://library.nrao.edu/public/memos/evla/EVLAM_222.pdf

[20] 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*, vol. 59, no. 5, pp. 6016–6030, Oct. 2023.

[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*, pp. 1–19, 2025.

[22] 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*, pp. 1–15, 2025.

[23] M. S. McCormick, T. Larsson, A. Tamrat, J. C. Strait, C. Chen, P. J. Treigherman, S. O. Roche, and R. Gottula, “Configurable orthogonal frequency division multiplexing (OFDM) signal and transmitter and receiver for user terminal to satellite uplink communications,” June 4 2024, US Patent 12,003,350 B1.

[24] A. M. Graff and T. E. Humphreys, “Ziv-Zakai-optimal OFDM resource allocation for time-of-arrival estimation,” *IEEE Transactions on Wireless Communications*, vol. 24, no. 8, pp. 6886–6901, 2025.

[25] D. Rife and R. Boorstyn, “Single tone parameter estimation from discrete-time observations,” *IEEE Transactions on information theory*, vol. 20, no. 5, pp. 591–598, 1974.

[26] W. Qin, Z. M. Komodromos, S. C. Morgan, and T. E. Humphreys, “Maximum likelihood time of arrival and Doppler estimation for precise Starlink-based PNT,” in *Proceedings of the IEEE/ION PLANS Meeting*, Salt Lake City, UT, 2025.

[27] M. L. Psiaki, “The why and the how of research into LEO-based signals of opportunity,” in *Proceedings of the ION GNSS+ Meeting*, Baltimore, MD, Sept. 2025.