OFDM-Based Positioning with Unknown Data Payloads: Bounds and Applications to LEO PNT

Andrew M. Graff, *Graduate Student Member, IEEE* and Todd E. Humphreys, *Senior Member, IEEE*

***Abstract*—This paper presents bounds, estimators, and signal design strategies for exploiting both known pilot resources and unknown data payload resources in time-of-arrival (TOA)-based positioning systems with orthogonal frequency-division multiplexing (OFDM) signals. It is the first to derive the Ziv-Zakai bound (ZZB) on TOA estimation for OFDM signals containing both known pilot and unknown data resources. In comparison to the Cramer-Rao bounds (CRBs) derived in prior work, this ZZB captures the low-signal-to-noise ratio (SNR) thresholding effects in TOA estimation and accounts for an unknown carrier phase. The derived ZZB is evaluated against CRBs and empirical TOA error variances. It is then evaluated on signals with resource allocations optimized for pilot-only TOA estimation, quantifying the performance gain over the best-case pilot-only signal designs. Finally, the positioning accuracy of maximum-likelihood and decision-directed estimators is evaluated on simulated low-Earth-orbit non-terrestrial-network channels and compared against their respective ZZBs.**

***Index Terms*—OFDM; positioning; Ziv-Zakai bound; NTN**

I. Introduction

POSITIONING services within existing wireless communications networks are becoming an increasingly important source of accurate localization. Such services can provide exceptional accuracy due to their access to large bandwidths and widespread deployment, making them an attractive alternative to traditional global navigation satellite system (GNSS) positioning. Low-Earth-orbit (LEO) non-terrestrial-networks (NTNs) in particular are a prime candidate for accurate user positioning services because of their ability to provide near-global coverage and strong signals with exceptionally large bandwidths [1], [2]. These LEO NTNs are undergoing a massive expansion in satellite deployment [3], further increasing their viability as the new standard for global positioning while simultaneously providing high-throughput communications.

An overwhelming majority of communications networks, including the 5G NTN standard [3], operate through orthogonal frequency-division multiplexing (OFDM), which divides the spectrum into time and frequency resource elements that may be allocated with either known pilot resources or unknown data resources. Since the structure of these pilot resources is known to users in the network, users may obtain time-of-arrival (TOA) estimates by correlating their received signal against the known pilot resources. These TOA estimates may then be used in positioning protocols such as pseudorange multilateration, time-difference-of-arrival (TDOA), or round-trip-time (RTT) [4]. But this scheme creates a challenging design tradeoff: allocating additional pilot resources improves positioning accuracy at the expense of data throughput, since OFDM resources must be diverted from data [5]. This tradeoff becomes especially complex for satellites in NTNs, which must additionally manage their power consumption when balancing the tradeoff between communications and positioning services [2]. Such a "zero-sum game" poses problems for rapidly expanding NTNs tasked with handling an ever-growing demand for both increased positioning accuracy and higher data rates. However, a new and enticing scheme emerges if users exploit both pilot resources and data resources in their TOA estimation.

Two approaches exist for users to exploit data resources in their TOA estimation: decision-directed (DD) and maximum likelihood (ML) estimation. The decision-directed estimator makes hard decoding decisions on the unknown data and then correlates the received signal against both the pilot resources and decoded data to improve estimation accuracy. While DD estimators are efficient and effective at high signal-to-noise ratios (SNRs), they are prone to decoding errors. Although these decoding errors may be mitigated through error correcting codes, such codes may not be usable for data resources intended for other users since networks may intentionally obfuscate the coding from other users; witness the scrambling by the Radio Network Temporary Identifier in long term evolution (LTE) and 5G new radio (NR). In contrast to DD estimators, ML estimators, commonly referred to as non-data-aided (NDA) estimators in prior work when no known pilots are used [6], [7], do not make hard decoding decisions but instead evaluate the likelihood of data resources over all symbols in the constellation. Since an overwhelming majority of the spectrum in communications networks is allocated to data resources, these DD and ML estimators can harness a much greater amount of signal power compared to the pilot-only estimator, resulting in significantly reduced TOA estimation errors.

When evaluating signals for TOA estimation, it is crucial to consider the impact that resource allocation has on the TOA likelihood function, which exhibits a mainlobe centered around the true TOA, sidelobes located away from the true TOA outside the mainlobe, and grating lobes caused by aliasing. At high SNR, TOA estimation errors will be concentrated near the true TOA within the mainlobe. At low SNR, however, TOA estimates may latch on to sidelobes, significantly increasing estimation error variance [8]–[10]. Bounds such as

A. Graff is with the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712, USA (e-mail: andrewgraff@utexas.edu).

T. Humphreys is with the Department of Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin, Austin, TX 78712, USA (e-mail: todd.humphreys@utexas.edu).

the Barankin bound [11], [12] and Ziv-Zakai bound (ZZB) [13] were derived to capture the behavior of this thresholding effect, which is ignored by the simpler Cramer-Rao bound (CRB). When only pilot resources are used in TOA estimation, the likelihood function is closely related to the autocorrelation function. However, the inclusion of unknown data resources in the likelihood function alters its shape in a complex manner, potentially introducing new sidelobes and sharpening the mainlobe. A bound that captures both the effects of unknown data and low SNR is needed to accurately characterize the positioning performance of DD and ML estimators.

LEO NTNs are especially amenable to positioning with unknown data resources. The large satellite constellation sizes increase the likelihood of a line-of-sight (LOS) path between users and satellites. Furthermore, phased arrays at both ends provide exceptional multipath mitigation. Finally, LEO satellites may be able to pre-compensate for Doppler due to their highly directive beams and small cell size. As a result, the fading in the post-beamforming channels remains exceptionally flat across wide bandwidths — as large as $240\,\mathrm{MHz}$ in the case of Starlink [14]. Under these favorable channels conditions, users only need to estimate and compensate for the TOA and carrier phase to equalize the received signal and begin decoding data.

This paper derives the ZZB on TOA estimation error variance for OFDM signals containing both pilot resources and unknown data resources. The derived ZZB is then compared against CRBs derived in prior work and against empirical TOA estimation error variance. Empirical TOA estimates are obtained with both ML and DD estimators. Three variants of ML estimators are considered: one that exploits only pilot resources, another based on only data resources, and a third that harnesses both pilot and data resources. These will be referred to as the pilot-only, data-only, and pilot-plus-data ML estimators, respectively. Candidate resource allocations are then generated over a range of SNRs that optimize the placement of positioning reference signals (PRSs) in the frequency domain to minimize the pilot-only TOA ZZB. The pilot-plus-data TOA ZZB is then evaluated on these optimized signals to quantify the reduction in TOA estimation error that can be achieved over the optimal allocations for pilot-only estimation. Finally, LEO NTN channels are simulated for a satellite constellation servicing a single cell. The positioning accuracy of a user in the serviced cell is evaluated using Monte Carlo methods for the pilot-only ML, data-only ML, pilot-plus-data ML, and DD estimators. These results are then compared against the derived ZZB.

A. Prior Work

Prior work has studied several TOA-based positioning algorithms with OFDM signals. Algorithms such as TDOA, pseudorange multilateration, and RTT are supported within the existing 5G NR standards [4]. More advanced approaches have demonstrated accurate positioning with 5G NR signals using both TOA and angle-of-arrival (AOA) measurements in an extended Kalman filter [15] and using multipath parameter estimates obtained from signals with optimized beam power allocations [16]. Alternatively, LTE [17] and 5G NR [18], [19] signals have been used as signals-of-opportunity (SOPs) for positioning, a paradigm that requires no cooperation between the user and the network. While exceptional positioning accuracy is demonstrated in this prior work, both the network-supported and SOP methods only obtain position estimates from known reference signals embedded in the OFDM signal and do not exploit the vast quantity of unknown data resources that are present in typical LTE and 5G NR downlink signals. The presence of OFDM reference signals has also been detected in a cognitive manner for SOP positioning [19], [20], but this approach still relies on the allocation of reference signals by the network.

The analysis of positioning within communications networks has been extended to the context of LEO NTNs. Scheduling for LEO constellations providing both communications and positioning services has been analyzed in [2]. The authors in [21] optimized LEO beamforming and beam scheduling to minimize the user positioning CRB. This work demonstrates the potential improvements LEO positioning services may provide over existing GNSS solutions. However, prior work has not yet analyzed OFDM resource allocation for LEO positioning nor explored the potential for exploiting data resources in TOA estimation.

Outside of positioning, DD and NDA estimation have been thoroughly studied in the context of communications. Prior work has analyzed DD estimators for OFDM frequency-offset estimation [22], OFDM channel estimation in high-velocity channels [23], and multiple-input multiple-output (MIMO) channel tracking [24]. The authors in [25] propose a DD channel estimator for overcoming pilot contamination in cell-free massive MIMO networks. This body of work demonstrates the potential improvements in estimator accuracy that may be gained through hard decoding decisions on unknown data. Similarly, prior work has analyzed NDA estimators for OFDM frequency-offset estimation [26], OFDM timing recovery [27], and OFDM SNR estimation [28]. Prior work has also studied the CRBs of NDA estimation. Bellili et al. derived a CRB for NDA time [29] and frequency [6] estimation for square-QAM constellations that is tighter than the simpler Modified Cramer-Rao bound (MCRB). The authors in [30] also propose a NDA TOA estimator that uses importance sampling to reduce computational costs, and they compare its performance against the MCRB.

Whereas prior work has extensively studied NDA and DD estimators for communications purposes, only a limited body of work has studied their applicability to positioning. The authors in [31] proposed a NDA AOA estimator for positioning with Gaussian frequency-shift-keying signals. Wang et al. proposed a semiblind OFDM range tracker which, after initialization with known pilot resources, tracks the multipath components of the channel using DD decoding decisions and a Kalman filter [32]. Similarly, the authors in [33] proposed a semiblind channel estimator for positioning that improves its estimates of the multipath components of the channel using DD decoding, comparing performance against the MCRB. Mensing et al. proposed a DD TOA estimator for TDOA positioning with intercell interference [34]. This work demon-

strates how unknown data payloads can be exploited, either through NDA or DD estimation, to improve positioning accuracy within communications networks. However, these papers do not compare NDA and DD estimators against one another to understand how each estimator's errors change with SNR. Furthermore, [31] and [33] did not consider OFDM signals, and [31] did not consider TOA-based positioning. Finally, [31], [33], [34] compared estimator accuracy only against the CRB, which is unable to capture low-SNR thresholding effects.

To overcome the limitations of the CRB, several studies have used the ZZB for analyzing positioning performance in OFDM systems with pilot-only TOA estimation. Prior work has used the ZZB to characterize the TOA precision of different parameterizations of OFDM pilot resource allocations [5], [35]–[37]. Furthermore, the ZZB has been used as an optimization criteria to solve for OFDM pilot resource allocations that minimize TOA estimation errors [38]. The ZZB on direct position estimation has also been derived in [39]. However, prior work has not derived the ZZB in the context of unknown data or for NDA estimation.

B. Contributions

The main contributions of this paper are as follows:

- A novel derivation of the ZZB on TOA estimation error variance for OFDM signals with unknown data resources. This novel bound is compared against the CRB, MCRB, and empirical errors from Monte-Carlo simulation.
- A comparison of the pilot-only and pilot-plus-data ZZBs for OFDM signals with resources optimized for pilot-only TOA estimation. This comparison provides insights into the potential gains achieved by exploiting unknown data and informs how pilot resources can be allocated to minimize overhead while meeting TOA accuracy requirements.
- Evaluation of the empirical positioning errors achieved by both ML and DD estimators on simulated LEO satellite channels in comparison to the ZZB.

The remainder of this paper is organized as follows. Section II introduces the signal model. Section III defines the TOA CRBs and derives the TOA ZZB for OFDM signals with payloads containing unknown data. Section IV defines the ML estimators for the pilot-only, data-only, and pilot-plus-data cases as well as the DD estimator. Section V-A compares the derived ZZB against the CRBs and Monte Carlo TOA errors. Section V-B evaluates the pilot-plus-data ZZB on signals with resource allocations optimized for pilot-only estimation. Section V-C evaluates the positioning accuracy of the pilot-only ML, data-only ML, pilot-plus-data ML, and DD estimators on simulated LEO NTN channels, comparing them against the derived ZZB. Finally, Section VI closes the paper by drawing conclusions from the results.

Notation: Column vectors are denoted with lowercase bold, e.g., $\boldsymbol{x}$. Matrices are denoted with uppercase bold, e.g., $\boldsymbol{X}$. Scalars are denoted without bold, e.g., x. The ith entry of a vector $\boldsymbol{x}$ is denoted $x[i]$ or in shorthand as x_i. The Euclidean norm is denoted $||\boldsymbol{x}||$. The cardinality of a set $\mathcal{S}$ is denoted $|\mathcal{S}|$. Real transpose is represented by the superscript T and conjugate transpose by the superscript H. The Q-function is denoted as $Q(\cdot)$. Zero-based indexing is used throughout the paper; e.g., $x[0]$ refers to the first element of $\boldsymbol{x}$.

II. Signal Model

Consider an OFDM signal with K subcarriers, N_{sym} symbols, a subcarrier spacing of Δ_{f} Hz, and a payload $x_m[k]$ for symbol indices $m \in \mathcal{M}$ and subcarrier indices $k \in \mathcal{K}$, where $\mathcal{M} = \{0, 1, \ldots, N_{\text{sym}}-1\}$ and $\mathcal{K} = \{0, 1, \ldots, K-1\}$. Let $d[k]$ be the mapping from subcarrier indices to offsets in frequency from the carrier in units of subcarriers. This map is defined as $d[k] = k$ for $k = 0, 1, \ldots, \frac{K}{2} - 1$ and $d[k] = k - K$ for $k = \frac{K}{2}, \frac{K}{2} + 1, \ldots, K - 1$. This signal propagates through a doubly-selective channel at a carrier frequency f_{c} with baseband frequency-domain channel coefficients $h_m[k]$ and experiences a LOS time delay τ, phase shift ϕ, and additive white Gaussian noise (AWGN) $v_m[k] \sim \mathcal{CN}(0, \sigma^2)$. Assuming negligible intercarrier interference due to small Doppler and negligible intersymbol interference due to a sufficiently long cyclic-prefix, the baseband received signal $y_m[k]$ is modeled in the frequency domain as

$$y_m[k] = \alpha_m[k] x_m[k] + v_m[k], \tag{1}$$

$$\alpha_m[k] = h_m[k] \exp\left(-j2\pi d[k]\Delta_{\text{f}}\tau + j\phi\right). \tag{2}$$

If the channel coefficients are constant across frequency and time, the complex gain $\alpha_m[k]$ can be instead modeled as

$$\alpha_m[k] = \sqrt{g} \exp\left(-j2\pi d[k]\Delta_{\text{f}}\tau + j\phi\right), \tag{3}$$

where g is the channel gain. The bounds and estimators of this paper are derived under the frequency-flat time-invariant model in (3), while the simulated channels in the results use the frequency-selective time-varying model in (2).

The payload $x_m[k]$ may contain either pilot resources, data resources, or be empty. Define $\mathcal{K}_m^{\text{P}}$ as the set of subcarrier indices containing pilot resources, $\mathcal{K}_m^{\text{D}}$ as the set of subcarrier indices containing data resources, and $\tilde{\mathcal{K}}_m \triangleq \mathcal{K}_m^{\text{P}} \bigcup \mathcal{K}_m^{\text{D}}$ as their union, during symbol m. Each data resource is modeled as randomly selected from a symbol constellation $\mathcal{C}$ with uniform probability and statistical independence from all other resource elements. This is expressed as $x_m[k] = c_m[k]$ and $P(c_m[k] = c_{\text{sym}}) = \frac{1}{|\mathcal{C}|}$ for $m \in \mathcal{M}$, $k \in \mathcal{K}_m^{\text{D}}$, and $c_{\text{sym}} \in \mathcal{C}$. Furthermore, the constellation is modeled as having unit average power such that $\mathbb{E}\left[|x_m[k]|^2\right] = 1$.

III. TOA Estimation Error Bounds

Let $\hat{\tau}$ be an unbiased estimate of the true TOA τ and define $\gamma_m[k] = \frac{|\alpha_m[k]|^2}{\sigma^2}$ as the SNR at subcarrier k during symbol m. This section will define bounds on the TOA estimation error variance $\mathbb{E}\left[(\hat{\tau} - \tau)^2\right]$ under the simplified signal model in (3).

A. Cramer-Rao Bounds

The simplest bound is the CRB when $\hat{\tau}$ is estimated using only pilot resources, which takes the form [40]

$$\mathbb{E}\left[(\hat{\tau} - \tau)^2\right] \geq \sigma^2_{\text{CRLB,P}} = I^{-1}_{\text{CRLB,P}} \tag{4}$$

$$I_{\text{CRLB,P}} \triangleq 8\pi^2 \Delta_{\text{f}}^2 \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{K}_m^{\text{P}}} d^2[k] \gamma_m[k]. \tag{5}$$

In comparison to the pilot-only CRB, the derivation of the CRB for the data-only estimator is more difficult since the likelihood function for each resource element is a Gaussian mixture distribution function with each Gaussian centered at each symbol in the constellation $\mathcal{C}$. One simplifying approach is to derive the MCRB [41] by conditioning on the unknown symbols. This simplifies to a form similar to the pilot-only CRB

$$\mathbb{E}\left[(\hat{\tau}-\tau)^2\right] \geq \sigma^2_{\text{MCRLB}} = I^{-1}_{\text{MCRLB}} \tag{6}$$

$$I_{\text{MCRLB}} \triangleq 8\pi^2\Delta_{\text{f}}^2 \sum_{m\in\mathcal{M}} \sum_{k\in\mathcal{K}_m^{\text{D}}} d^2[k]\gamma_m[k]. \tag{7}$$

A tighter CRB for unknown data is derived in [6] without using the MCRB. While this bound was derived for frequency estimation, it is easily mapped to TOA estimation with OFDM signals. This CRB takes the form

$$\mathbb{E}\left[(\hat{\tau}-\tau)^2\right] \geq \sigma^2_{\text{CRLB,D}} = I^{-1}_{\text{CRLB,D}}, \tag{8}$$

where

$$I_{\text{CRLB,D}} \triangleq 8\pi^2\Delta_{\text{f}}^2 \sum_{\substack{m\in\mathcal{M}\\ k\in\mathcal{K}_m^{\text{D}}}} \left(\frac{1+\gamma_m[k]}{A_2}\psi(\gamma_m[k]) - \gamma_m[k]\right) \times d^2[k]\gamma_m[k], \tag{9}$$

and $\psi(\cdot)$ and A_2 are defined in [6, Eqs. (37)-(38)].

These CRBs are useful for analyzing the high-SNR precision of TOA estimators, and the CRB from [6] importantly captures the increased errors caused by uncertainty in the symbol selected from the constellation, making the bound in (9) tighter than the MCRB in (7). These fundamental error bounds can provide valuable insights into the efficiency of estimators and can serve as an optimization criteria for evaluating different signal designs and resource allocations. However, the CRBs ignore the impact that sidelobes in the likelihood function have on estimation error, making these bounds inapt at lower SNRs [5].

B. Ziv-Zakai Bounds

The ZZB is superior to the CRB for TOA estimation analysis because it captures the low-SNR thresholding effects caused by sidelobes. The bound considers a binary detection problem with equally-likely hypotheses: (1) the received signal experienced delay τ_0 and phase ϕ_0, and (2) the received signal experienced delay $\tau_0+\tau_1$ and phase $\phi_0+\phi_1 \pmod{2\pi}$. Define $\boldsymbol{\theta} \triangleq [\tau,\ \phi]^T$, $\boldsymbol{\theta}_0 \triangleq [\tau_0,\ \phi_0]^T$, and $\boldsymbol{\theta}_1 \triangleq [\tau_0+\tau_1,\ \phi_0+\phi_1 \pmod{2\pi}]^T$. $\boldsymbol{\theta}$ is treated as a random variable with a known *a priori* distribution. The hypothesis test can be expressed as

$$\frac{p(\boldsymbol{y}|\boldsymbol{\theta}=\boldsymbol{\theta}_0)}{p(\boldsymbol{y}|\boldsymbol{\theta}=\boldsymbol{\theta}_1)} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \nu, \qquad \begin{matrix}\mathcal{H}_0 : \boldsymbol{\theta}=\boldsymbol{\theta}_0\\ \mathcal{H}_1 : \boldsymbol{\theta}=\boldsymbol{\theta}_1\end{matrix}, \tag{10}$$

where $\boldsymbol{y}$ is the vector containing all $y_m[k]$ for $m\in\mathcal{M}$ and $k\in\tilde{\mathcal{K}}_m$, and ν is the detection threshold. The ZZB is concerned with the minimum error probability of this hypothesis test, which corresponds to a threshold of $\nu=1$ when $\mathcal{H}_0$ and $\mathcal{H}_1$ are assumed equally-likely.

As in [38], the time delays will be normalized by the OFDM sampling period $T_{\text{s}} = \frac{1}{K\Delta_{\text{f}}}$, creating $z \triangleq \frac{\tau}{T_{\text{s}}}$, $z_0 \triangleq \frac{\tau_0}{T_{\text{s}}}$ and $z_1 \triangleq \frac{\tau_1}{T_{\text{s}}}$. The minimum error probability of this detection problem is assumed to be shift-invariant, allowing the hypotheses to be simplified without loss of generality by assuming $\tau_0=0$ and $\phi_0=0$. Then the log of the likelihood ratio in (10) can be denoted

$$\log\Lambda(\boldsymbol{y},\boldsymbol{\theta}_1) \triangleq \log p(\boldsymbol{y}|\boldsymbol{\theta}=\boldsymbol{\theta}_0) - \log p(\boldsymbol{y}|\boldsymbol{\theta}=\boldsymbol{\theta}_1), \tag{11}$$

and the minimum error probability of the hypothesis test is defined as the probability that the log-likelihood ratio in (11), conditioned on $\boldsymbol{\theta}=\boldsymbol{\theta}_0$, is less than zero:

$$P_{\text{min}}(z_1,\phi_1) \triangleq P\left(\log\Lambda(\boldsymbol{y},\boldsymbol{\theta}_1) < 0 \middle|\, \boldsymbol{\theta}=\boldsymbol{\theta}_0\right). \tag{12}$$

Define $\boldsymbol{R}_{\boldsymbol{\theta}}$ as the estimation error covariance of $\boldsymbol{\theta}$ and $\mathcal{V}\{\cdot\}$ as the valley-filling function [42]. Assuming *a priori* knowledge that the TOA is uniformly distributed on $[0,T_{\text{a}}]$ and the phase is uniformly distributed on $[0,2\pi]$, and noting the scale-invariance of the valley-filling function, the ZZB on TOA error variance can be defined as [43]

$$\begin{aligned}\mathbb{E}\left[(\hat{\tau}-\tau)^2\right] &= \boldsymbol{a}^T\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{a} \\ &\geq \sigma^2_{\text{ZZB}} \\ &\triangleq \frac{1}{T_{\text{a}}}\int_0^{T_{\text{a}}} \tau_1\mathcal{V}\{(T_{\text{a}}-\tau_1)\max_{\phi_1}[P_{\text{min}}(\tau_1/T_s,\phi_1)]\}d\tau_1 \\ &= \frac{T_{\text{s}}^2}{N_{\text{a}}}\int_0^{N_{\text{a}}} z_1\mathcal{V}\{(N_{\text{a}}-z_1)\max_{\phi_1}[P_{\text{min}}(z_1,\phi_1)]\}dz_1,\end{aligned} \tag{13}$$

where $N_{\text{a}} \triangleq \frac{T_{\text{a}}}{T_{\text{s}}}$.

Expressions for $P_{\text{min}}(z_1,\phi_1)$ will now be derived. The likelihood of $\boldsymbol{y}$ is

$$p(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{m\in\mathcal{M}}\prod_{k\in\tilde{\mathcal{K}}_m} p(y_m[k]|\boldsymbol{\theta}), \tag{14}$$

which follows from the independence of each resource element. The log-likelihood ratio in (11) can be expressed as

$$\begin{aligned}&\log\Lambda(\boldsymbol{y},\boldsymbol{\theta}_1) \\ &= \sum_{m\in\mathcal{M}}\sum_{k\in\tilde{\mathcal{K}}_m} \log p(y_m[k]|\boldsymbol{\theta}=\boldsymbol{\theta}_0) - \log p(y_m[k]|\boldsymbol{\theta}=\boldsymbol{\theta}_1) \\ &= \sum_{m\in\mathcal{M}}\sum_{k\in\tilde{\mathcal{K}}_m} \log\Lambda(y_m[k],\boldsymbol{\theta}_1),\end{aligned} \tag{15}$$

where $\log\Lambda(y_m[k],\boldsymbol{\theta}_1) \triangleq \log p(y_m[k]|\boldsymbol{\theta}=\boldsymbol{\theta}_0) - \log p(y_m[k]|\boldsymbol{\theta}=\boldsymbol{\theta}_1)$ is the log-likelihood ratio for the resource at symbol m and subcarrier k. Similarly, the minimum error probability in (12) can be expressed as

$$\begin{aligned}&P_{\text{min}}(z_1,\phi_1) \\ &= P\left(\sum_{m\in\mathcal{M}}\sum_{k\in\tilde{\mathcal{K}}_m} \log\Lambda(y_m[k],\boldsymbol{\theta}_1) < 0 \middle| \boldsymbol{\theta}=\boldsymbol{\theta}_0\right).\end{aligned} \tag{16}$$

To derive this probability, the distribution of $\log\Lambda$ conditioned on the parameter vector $\boldsymbol{\theta}=\boldsymbol{\theta}_0$ must be analyzed. Accordingly, all expectations through the remainder of this section are conditioned on $\boldsymbol{\theta}=\boldsymbol{\theta}_0$.

Unknown Data Resources: First consider the case of unknown data resources. The likelihood of $y_m[k]$ conditioned on the parameter vector $\boldsymbol{\theta}$ and conditioned on knowledge of the symbol $c_m[k]$ is

$$\begin{aligned} &p(y_m[k]|c_m[k],\boldsymbol{\theta}) \\ &= \frac{1}{\pi\sigma^2}\exp\left(\frac{-1}{\sigma^2}\left|y_m[k]-\mu_m[k]\nu_k(\boldsymbol{\theta})\right|^2\right), \end{aligned} \tag{17}$$

where $\nu_k(\boldsymbol{\theta}) \triangleq \exp\left(-j2\pi z d[k]/K + j\phi\right)$ and $\mu_m[k] \triangleq \sqrt{g}c_m[k]$. Assuming equally-likely symbols in the constellation, the likelihood of $y_m[k]$ is given by

$$p(y_m[k]|\boldsymbol{\theta}) = \frac{1}{|\mathcal{C}|}\sum_{c_{\text{sym}}\in\mathcal{C}} p(y_m[k]|c_m[k]=c_{\text{sym}},\boldsymbol{\theta}), \tag{18}$$

which is a Gaussian mixture distribution function. The log-likelihood of $y_m[k]$ then becomes

$$\begin{aligned} &\log p(y_m[k]|\boldsymbol{\theta}) \\ &= \log\frac{1}{|\mathcal{C}|} + \log\sum_{c_{\text{sym}}\in\mathcal{C}} p(y_m[k]|c_m[k]=c_{\text{sym}},\boldsymbol{\theta}) \\ &= \log\frac{1}{\pi\sigma^2|\mathcal{C}|} + \log\sum_{c_{\text{sym}}\in\mathcal{C}}\exp\left(\frac{-1}{\sigma^2}\left|y_m[k]-\mu_m[k]\nu_k(\boldsymbol{\theta})\right|^2\right) \\ &= \log\frac{1}{\pi\sigma^2|\mathcal{C}|} - \frac{1}{\sigma^2}|y_m[k]|^2 \\ &+ \log\sum_{c_{\text{sym}}\in\mathcal{C}}\exp\left(\frac{1}{\sigma^2}\left(2\mathbb{R}\{y_m^*[k]\mu_m[k]\nu_k(\boldsymbol{\theta})\} - g|c_{\text{sym}}|^2\right)\right). \end{aligned} \tag{19}$$

With this log-likelihood defined, $\log p(y_m[k]|\boldsymbol{\theta}=\boldsymbol{\theta}_0)$ and $\log p(y_m[k]|\boldsymbol{\theta}=\boldsymbol{\theta}_1)$ can be substituted into (15), resulting in

$$\begin{aligned} &\log\Lambda(y_m[k],\boldsymbol{\theta}_1) \\ &= \log\sum_{c_{\text{sym}}\in\mathcal{C}}\exp\left(\frac{1}{\sigma^2}\left(2\mathbb{R}\{y_m^*[k]\mu_m[k]\} - g|c_{\text{sym}}|^2\right)\right) \\ &- \log\sum_{c_{\text{sym}}\in\mathcal{C}}\exp\left(\frac{1}{\sigma^2}\left(2\mathbb{R}\{y_m^*[k]\mu_m[k]\nu_k(\boldsymbol{\theta}_1)\} - g|c_{\text{sym}}|^2\right)\right). \end{aligned} \tag{20}$$

The distribution of this log-sum-exp form in (20) is difficult to analyze. Conditioned on a specific symbol $c_m[k]$, $y_m[k]$ becomes Gaussian distributed and (20) becomes the difference of the log of two lognormal sums. Prior work has approximated the log of lognormal sums as Gaussian-distributed [44]. Likewise, the log-likelihood ratio in (20) will be approximated as a Gaussian distribution matching the first and second moments.

Since the log-likelihood is a function of a Gaussian random variable conditioned on knowledge of the symbol $c_m[k]$, its moment generating function is easily expressed as

$$\begin{aligned} &M_{\log\Lambda}(t) \triangleq \mathbb{E}\left[\exp\left(t\log\Lambda(y_m[k],\boldsymbol{\theta}_1)\right)|\boldsymbol{\theta}=\boldsymbol{\theta}_0\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\exp\left(t\log\Lambda(y_m[k],\boldsymbol{\theta}_1)\right)|c_m[k],\boldsymbol{\theta}=\boldsymbol{\theta}_0\right]\right] \\ &= \frac{1}{|\mathcal{C}|}\sum_{c_{\text{sym}}\in\mathcal{C}}\mathbb{E}\left[\exp\left(t\log\Lambda(y_m[k],\boldsymbol{\theta}_1)\right)|c_m[k]=c_{\text{sym}},\boldsymbol{\theta}=\boldsymbol{\theta}_0\right], \end{aligned} \tag{21}$$

where the smoothing property allows the expectation to be conditioned on the symbol $c_m[k]$. The inner expectation is taken over the noise $v_m[k]$. The first moment can then be computed as

$$\begin{aligned} &\mathbb{E}\left[\log\Lambda(y_m[k],\boldsymbol{\theta}_1)|\boldsymbol{\theta}=\boldsymbol{\theta}_0\right] = \frac{\partial}{\partial t}M_{\log\Lambda}(t)\Big|_{t=0} \\ &= \frac{1}{|\mathcal{C}|}\sum_{c_{\text{sym}}\in\mathcal{C}}\mathbb{E}\left[\log\Lambda(y_m[k],\boldsymbol{\theta}_1)|c_m[k]=c_{\text{sym}},\boldsymbol{\theta}=\boldsymbol{\theta}_0\right], \end{aligned} \tag{22}$$

and the second moment can be computed as

$$\begin{aligned} &\mathbb{E}\left[(\log\Lambda(y_m[k],\boldsymbol{\theta}_1))^2|\boldsymbol{\theta}=\boldsymbol{\theta}_0\right] = \frac{\partial^2}{\partial t^2}M_{\log\Lambda}(t)\Big|_{t=0} \\ &= \frac{1}{|\mathcal{C}|}\sum_{c_{\text{sym}}\in\mathcal{C}}\mathbb{E}\left[(\log\Lambda(y_m[k],\boldsymbol{\theta}_1))^2|c_m[k]=c_{\text{sym}},\boldsymbol{\theta}=\boldsymbol{\theta}_0\right]. \end{aligned} \tag{23}$$

Since these expectations are taken over a complex Gaussian distribution, they can be approximated using a Gauss-Hermite quadrature [45, 25.4.46]. Consider a Gauss-Hermite quadrature of size N with weights $w[n]$ and nodes $\delta_v[n]$ for $n\in 0,1,\ldots,N-1$. Then define $\tilde{\delta}_v[n_{\text{I}},n_{\text{Q}}] \triangleq \delta_v[n_{\text{I}}]+j\delta_v[n_{\text{Q}}]$. Note that conditioned on $c_m[k]=c_{\text{sym}}$, the expected value of $y_m[k]$ is $\mu_m[k]=\sqrt{g}c_{\text{sym}}$. The expectation in (22) can then be expressed as

$$\begin{aligned} &\mathbb{E}\left[\log\Lambda(y_m[k],\boldsymbol{\theta}_1)|c_m[k]=c_{\text{sym}},\boldsymbol{\theta}=\boldsymbol{\theta}_0\right] \\ &\approx \sum_{n_{\text{I}}=0}^{N-1}\sum_{n_{\text{Q}}=0}^{N-1} w[n_{\text{I}}]w[n_{\text{Q}}]\log\Lambda\left(\mu_m[k]+\tilde{\delta}_v[n_{\text{I}},n_{\text{Q}}],\boldsymbol{\theta}_1\right). \end{aligned} \tag{24}$$

Similarly, the expectation in (23) can be expressed as

$$\begin{aligned} &\mathbb{E}\left[(\log\Lambda(y_m[k],\boldsymbol{\theta}_1))^2|c_m[k]=c_{\text{sym}},\boldsymbol{\theta}=\boldsymbol{\theta}_0\right] \\ &\approx \sum_{n_{\text{I}}=0}^{N-1}\sum_{n_{\text{Q}}=0}^{N-1} w[n_{\text{I}}]w[n_{\text{Q}}]\left(\log\Lambda\left(\mu_m[k]+\tilde{\delta}_v[n_{\text{I}},n_{\text{Q}}],\boldsymbol{\theta}_1\right)\right)^2. \end{aligned} \tag{25}$$

Finally, the variance of the log-likelihood ratio for the resource at symbol m and subcarrier k is

$$\begin{aligned} &\text{Var}\left(\log\Lambda(y_m[k],\boldsymbol{\theta}_1)|\boldsymbol{\theta}=\boldsymbol{\theta}_0\right) = \\ &\mathbb{E}\left[(\log\Lambda(y_m[k],\boldsymbol{\theta}_1))^2|\boldsymbol{\theta}=\boldsymbol{\theta}_0\right] - \mathbb{E}\left[\log\Lambda(y_m[k],\boldsymbol{\theta}_1)|\boldsymbol{\theta}=\boldsymbol{\theta}_0\right]^2. \end{aligned} \tag{26}$$

The mean and variance of the log-likelihood ratio for unknown data resources is now quantified.

Known Pilot Resources: Now consider the case of known pilot resources. The log-likelihood ratio takes the form

$$\begin{aligned} &\log\Lambda(y_m[k],\boldsymbol{\theta}_1) \\ &= \log p(y_m[k]|\boldsymbol{\theta}=\boldsymbol{\theta}_0) - \log p(y_m[k]|\boldsymbol{\theta}=\boldsymbol{\theta}_1) \\ &= \frac{1}{\sigma^2}\left(\left|y_m[k]-\mu_m[k]\nu_k(\boldsymbol{\theta}_1)\right|^2 - \left|y_m[k]-\mu_m[k]\right|^2\right) \\ &= \frac{2}{\sigma^2}\left(\mathbb{R}\{y_m^*[k]\mu_m[k]\} - \mathbb{R}\{y_m^*[k]\mu_m[k]\nu_k(\boldsymbol{\theta}_1)\}\right) \\ &= \frac{2}{\sigma^2}\mathbb{R}\{y_m^*[k]\mu_m[k]\left(1-\nu_k(\boldsymbol{\theta}_1)\right)\}. \end{aligned} \tag{27}$$

Since (27) is a linear function of a Gaussian random variable $y_m[k]$, the log-likelihood ratio is itself Gaussian distributed. The mean of the log-likelihood ratio can be expressed as

$$\begin{aligned}&\mathbb{E}\left[\log\Lambda(y_m[k],\boldsymbol{\theta}_1)|\boldsymbol{\theta}{=}\boldsymbol{\theta}_0\right] \\ &= \frac{2}{\sigma^2}|\mu_m[k]|^2\left(1-\cos\left(2\pi z_1 d[k]/K+\phi_1\right)\right) \\ &= \frac{2\alpha_m^2[k]}{\sigma^2}|c_m[k]|^2\left(1-\cos\left(2\pi z_1 d[k]/K+\phi_1\right)\right),\end{aligned} \tag{28}$$

and the variance of the log-likelihood ratio can be expressed as

$$\begin{aligned}&\mathrm{Var}\left(\log\Lambda(y_m[k],\boldsymbol{\theta}_1)|\boldsymbol{\theta}{=}\boldsymbol{\theta}_0\right) \\ &= \frac{4}{\sigma^2}|\mu_m[k]|^2\left(1-\cos\left(2\pi z_1 d[k]/K+\phi_1\right)\right).\end{aligned} \tag{29}$$

The probability of error using only pilot resources simplifies to the form seen in [46]. However, expressing the mean and variance of the log-likelihood ratio allows pilot resources and data resources to be combined together in the ZZB.

Unified Expression for $P_{min}(z,\phi)$: Now that expressions have been derived for the mean and variance of the log-likelihood ratio at each resource element for both unknown data resources and known pilot resources, the distribution of the sum log-likelihood ratio $\log\Lambda(\boldsymbol{y},\boldsymbol{\theta}_1)$ can be quantified. By approximating the log-likelihood ratio of the data resources as Gaussian, it follows that the sum log-likelihood is also approximately Gaussian. Furthermore, for large numbers of subcarriers, this approximation will improve by the central limit theorem. The mean of the sum log-likelihood ratio $\log\Lambda(\boldsymbol{y},\boldsymbol{\theta}_1)$ is

$$\begin{aligned}&\mathbb{E}\left[\log\Lambda(\boldsymbol{y},\boldsymbol{\theta}_1)|\boldsymbol{\theta}{=}\boldsymbol{\theta}_0\right] \\ &= \sum_{m\in\mathcal{M}}\sum_{k\in\tilde{\mathcal{K}}_m}\mathbb{E}\left[\log\Lambda(y_m[k],\boldsymbol{\theta}_1)|\boldsymbol{\theta}{=}\boldsymbol{\theta}_0\right],\end{aligned} \tag{30}$$

and its variance is

$$\begin{aligned}&\mathrm{Var}\left(\log\Lambda(\boldsymbol{y},\boldsymbol{\theta}_1)|\boldsymbol{\theta}{=}\boldsymbol{\theta}_0\right) \\ &= \sum_{m\in\mathcal{M}}\sum_{k\in\tilde{\mathcal{K}}_m}\mathrm{Var}\left(\log\Lambda(y_m[k],\boldsymbol{\theta}_1)|\boldsymbol{\theta}{=}\boldsymbol{\theta}_0\right).\end{aligned} \tag{31}$$

Finally, the probability of error in (16) can be expressed as

$$\begin{aligned}&P_{\min}(z_1,\phi_1) \\ &\approx Q\left(\mathbb{E}\left[\log\Lambda(\boldsymbol{y},\boldsymbol{\theta}_1)|\boldsymbol{\theta}{=}\boldsymbol{\theta}_0\right]/\sqrt{\mathrm{Var}\left(\log\Lambda(\boldsymbol{y},\boldsymbol{\theta}_1)|\boldsymbol{\theta}{=}\boldsymbol{\theta}_0\right)}\right),\end{aligned} \tag{32}$$

which can be substituted into (13) to compute the ZZB. When only pilot resources are used, the probability of error in (32) simplifies to the form in [46].

IV. Maximum Likelihood Estimation

In the results presented in the following section, the bounds in Section III will be evaluated against ML estimators and a DD estimator. Recall that three variants of ML estimator are considered: the pilot-only, data-only, and pilot-plus-data ML estimators. The pilot-only ML estimator can be expressed as

$$\hat{\boldsymbol{\theta}}_{\mathrm{P}} = \underset{\boldsymbol{\theta}_1}{\operatorname{argmax}}\sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}_m^{\mathrm{P}}}\log p(y_m[k]|\boldsymbol{\theta}{=}\boldsymbol{\theta}_1). \tag{33}$$

The data-only ML estimator is expressed as

$$\hat{\boldsymbol{\theta}}_{\mathrm{D}} = \underset{\boldsymbol{\theta}_1}{\operatorname{argmax}}\sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}_m^{\mathrm{D}}}\log p(y_m[k]|\boldsymbol{\theta}{=}\boldsymbol{\theta}_1). \tag{34}$$

And the pilot-plus-data ML estimator is expressed as

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}_1}{\operatorname{argmax}}\sum_{m\in\mathcal{M}}\sum_{k\in\tilde{\mathcal{K}}_m}\log p(y_m[k]|\boldsymbol{\theta}{=}\boldsymbol{\theta}_1). \tag{35}$$

Note that (33)-(35) differ in the sets of subcarriers in the summation.

The DD estimator requires an initial estimate of the signal parameters $\boldsymbol{\theta}$ and therefore will only be applied when both pilot and data resources are present in the signal. After obtaining the pilot-only ML parameter estimate $\hat{\boldsymbol{\theta}}_{\mathrm{P}}$, ML decoding decisions are made:

$$\hat{c}_m[k] = \underset{c_{\mathrm{sym}}\in\mathcal{C}}{\operatorname{argmax}}\log p(y_m[k]|c_m[k]=c_{\mathrm{sym}},\boldsymbol{\theta}{=}\hat{\boldsymbol{\theta}}_{\mathrm{P}}). \tag{36}$$

It is important to note that this DD estimator uses no error correction codes, which, as mentioned in Section I, cannot be assumed to be of benefit in the current context because they are user-specific whereas this paper's scheme is designed to exploit all data. After decoding the data resources, another ML parameter estimate is obtained using both pilot and data resources by treating the decoded data resources as known symbols:

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{\mathrm{DD}} = \underset{\boldsymbol{\theta}_1}{\operatorname{argmax}}&\sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}_m^{\mathrm{P}}}\log p(y_m[k]|\boldsymbol{\theta}{=}\boldsymbol{\theta}_1) \\ &+\sum_{m\in\mathcal{M}}\sum_{k\in\mathcal{K}_m^{\mathrm{D}}}\log p(y_m[k]|c_m[k]=\hat{c}_m[k],\boldsymbol{\theta}{=}\boldsymbol{\theta}_1).\end{aligned} \tag{37}$$

The argmax operations in (33)-(35) and (37) are evaluated using a grid search with quadratic peak interpolation. The TOA z has a discretized grid spanning from 0 to N_a in intervals of size Δ_z, while the phase ϕ has a discretized grid spanning from 0 to 2π in intervals of size Δ_ϕ.

V. Results

A. Bounds

Fig. 1 compares the derived ZZB in (13), the CRB in (9), and the MCRB in (7) on TOA estimation against Monte-Carlo-simulation-based RMSEs assuming the data-only ML estimator in (34). The TOA RMSEs are scaled by the speed of light and plotted in units of meters. The OFDM signal consists of $K = 64$ subcarriers, $N_{\mathrm{sym}} = 1$ symbols, a subcarrier spacing of $\Delta_{\mathrm{f}} = 15\,\mathrm{kHz}$, and an *a priori* TOA duration of $T_{\mathrm{a}} = 6.25\,\mu\mathrm{s}$. All 64 resource elements are allocated as data resources. Empirical TOA RMSEs were estimated over 20000 Monte Carlo iterations of random noise at each SNR. The grid search was conducted with intervals of $\Delta_z = 1/8$ sample and $\Delta_\phi = 15^\circ$.

The MCRB is the loosest bound, only converging with the empirical RMSEs at high SNRs above 10 dB for QPSK and 17 dB for 16QAM. The CRB in (9) remains tighter than the MCRB over a larger range of SNRs, capturing the slight deviation from the MCRB above 7 dB SNR for QPSK and 14 dB SNR for 16QAM. Below these SNRs, the empirical

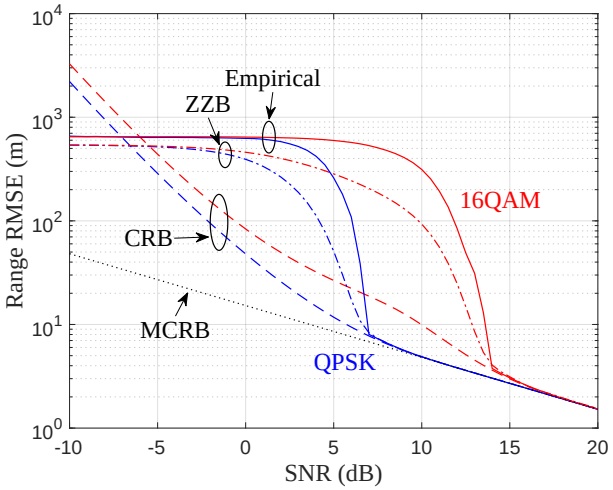

Fig. 1: A comparison of the Monte Carlo empirical root mean square errors (RMSEs), MCRB, CRB, and ZZB on TOA estimation using *a priori* unknown data. Results are shown for both QPSK and 16QAM constellations.

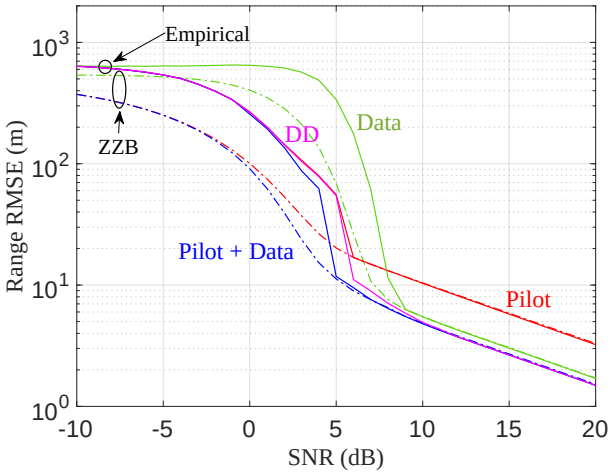

Fig. 2: A comparison of the Monte Carlo empirical RMSEs and bounds for pilot-only, data-only, and pilot-plus-data estimation. Results are shown for a single OFDM symbol with 64 subcarriers, QPSK modulation, and 8 sparsely-placed pilot resources.

TOA errors experience the low-SNR thresholding effect and RMSEs increase suddenly, a phenomenon not captured by the CRB. The ZZB provides a much tighter bound in this thresholding regime than the MCRB and CRB. Asymptotically as SNR decreases, the ZZB and empirical RMSEs converge to different values since the empirical TOA estimator is ML, not minimum mean square error. Accordingly, the ZZB RMSE converges to $\sqrt{T_\mathrm{a}/12}$ s, the standard deviation of a uniform distribution with duration T_a s [47].

Fig. 2 provides a different perspective and compares the empirical RMSEs against their ZZBs for the four types of estimators: pilot-only ML, data-only ML, pilot-plus-data ML, and DD. The OFDM signal is parameterized identically to the signal in Fig. 1 but is additionally allocated with a sparse placement of 8 pilot resources. The pilot resource placement

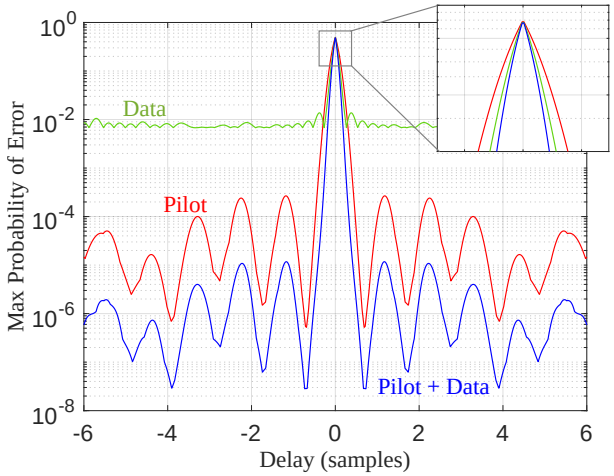

Fig. 3: A plot of the maximin probability of error $\max_\phi [P_{\min}(z,\phi)]$ against delay z computed as part of the ZZB in (13) for pilot-only, data-only, and pilot-plus-data estimation. Results are shown for the same signal in Fig. 2 at $5\,\mathrm{dB}$ SNR.

is optimized to minimize TOA error at $0\,\mathrm{dB}$ SNR using the integer-optimization routines in [38]. Power is allocated equally across all resources.

Below $8\,\mathrm{dB}$ SNR, the pilot-only estimator reduces error compared to the data-only estimator. Above this SNR, however, the data-only estimator reduces TOA errors significantly over the pilot-only estimator, ultimately achieving an RMSE of $3.0\,\mathrm{m}$ compared to the pilot-only estimator's RMSE of $5.8\,\mathrm{m}$ at $15\,\mathrm{dB}$ SNR. The pilot-plus-data estimator achieves the lowest RMSE across all SNRs. This is most notable in the thresholding regime, with the pilot-plus-data estimator achieving an RMSE of $11.8\,\mathrm{m}$ compared to pilot-only RMSE of $54.5\,\mathrm{m}$ and data-only RMSE of $338.2\,\mathrm{m}$ at $5\,\mathrm{dB}$ SNR. Meanwhile, the DD estimator is only capable of improving upon the pilot-only estimator at and above $6\,\mathrm{dB}$ SNR, after which it plateaus near the same RMSE as the pilot-plus-data estimator.

Fig. 3 provides insight into how the different types of estimation affect the probability of error in (16), thereby changing the ZZB in (13). As seen in Fig. 2, all three ZZBs at $5\,\mathrm{dB}$ SNR are experiencing the low-SNR thresholding effect to varying degrees. At this SNR, Fig. 3 shows that the pilot-only probability of error exhibits high sidelobes and a wide peak near the true delay, resulting in the large RMSE in Fig. 2. Meanwhile, the data-only probability of error exhibits a sharper peak but higher sidelobes that remain relatively flat across all delays. This elevated sidelobe presence significantly increases the likelihood of TOA estimates occurring outside of the mainlobe. As a result, data-only estimation has the highest RMSE in Fig. 2 at $5\,\mathrm{dB}$ SNR despite the sharpened mainlobe. In contrast, the pilot-plus-data probability of error exhibits both the sharpest mainlobe peak and the lowest sidelobe probability, allowing pilot-plus-data estimation to mitigate the low-SNR thresholding effect significantly and achieve a lower RMSE than the pilot-only and data-only estimators.

B. Sparse Resource Optimization

Now consider the problem of allocating PRSs in an OFDM signal designed for both positioning and communications. To minimize the reduction in data rate, the PRSs will be allocated sparsely throughout the bandwidth of the signal. Consider an OFDM signal consisting of $K = 240$ subcarriers, $N_{\text{sym}} = 4$ symbols, a subcarrier spacing of $\Delta_{\text{f}} = 240\,\text{kHz}$, and an *a priori* TOA duration of $T_{\text{a}} = 156.25\,\text{ns}$. Assume the subcarriers are divided into 20 resource blocks of 12 subcarriers each, where each resource block is restricted to containing either a PRS block or a data block filled with QPSK data resources. Letting N_{PRS} denote the number of resource blocks that are allocated with a PRS block, the allocation problem is to determine the best placement of these N_{PRS} PRS blocks among the 20 available resource blocks.

Each PRS block has been arbitrarily chosen to consist of pilot resources placed in a size 4 comb pattern, similarly to the PRS in 5G NR, which is visualized in Fig. 4. All non-pilot resource elements in each PRS block are allocated as data resources.

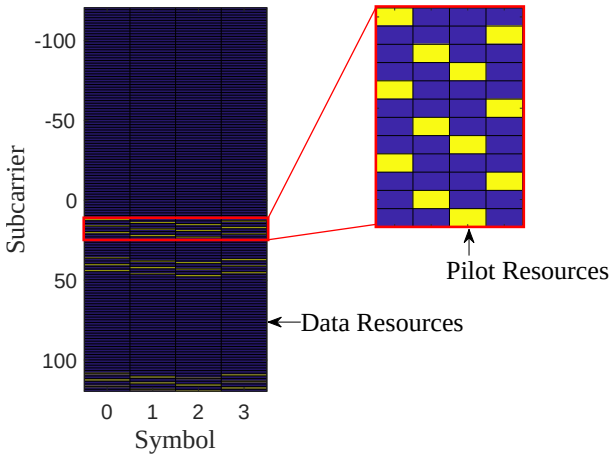

Fig. 4: A visualization of the optimized OFDM resource allocation for $N_{\text{PRS}} = 3$ at $0\,\text{dB}$ SNR.

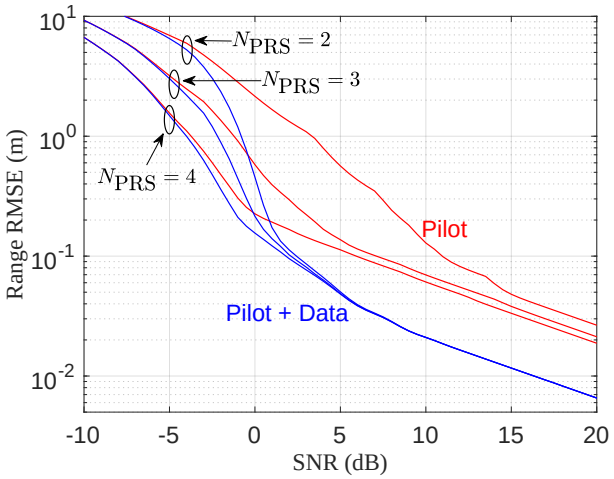

Fig. 5: The ZZB on TOA RMSE for the PRS-optimized resource allocations. Results are shown for $N_{\text{PRS}} \in \{2, 3, 4\}$.

The pilot-only ZZB was evaluated for all permutations of PRS allocation for $N_{\text{PRS}} \in \{2, 3, 4\}$ at each SNR. The pilot-optimal allocation at each SNR was then chosen as the allocation that minimized the ZZB. The pilot-plus-data ZZB was then evaluated on the pilot-optimal allocations. Fig. 5 plots both the pilot-only and pilot-plus-data ZZBs of the pilot-optimal allocations against SNR. Increasing the number of PRS resource blocks reduces the ZZB for both pilot-only and pilot-plus-data estimation, yielding the greatest improvement at lower SNRs where the additional pilot resources can mitigate the low-SNR thresholding effect. However, this improvement becomes negligible for the pilot-plus-data ZZBs above approximately $6\,\text{dB}$ SNR where the bounds converge. In this high SNR regime, the pilot-plus-data ZZBs show significant reductions in error compared to the pilot-only ZZBs. At $10\,\text{dB}$ SNR, all three pilot-plus-data ZZBs have a RMSE of $2.1\,\text{cm}$ compared to RMSEs of $6.1\,\text{cm}$, $7.0\,\text{cm}$, and $12.9\,\text{cm}$ for the pilot-only ZZBs. Into low SNRs, the pilot-plus-data ZZBs still show notable improvements over their respective pilot-only ZZBs even though the PRS allocations are optimized for pilot-only estimation at every SNR.

C. LEO Satellite Positioning

Positioning with LEO satellite downlink signals is a particularly apt application for the ML and DD TOA estimators discussed in Section IV. LEO channels can span wide bandwidths, enabling highly-accurate TOA estimates and therefore accurate user positioning services. Furthermore, LEO channels experience minimal fading especially when combined with highly-directional phased arrays at both the transmitter and receiver, resulting in flat channel responses across the wide signal bandwidths. Finally, LEO satellites provide communication services to large cells which are likely to contain many network users, increasing the amount of downlink data resources that will need to be allocated. If the LEO satellites transmit in bursts to manage power consumption [2], downlink resources are likely to be fully allocated to maximize throughput. One example is the downlink Starlink signal which consists of frames with fully allocated data [14]. In such a fully allocated burst, the allocation of PRSs comes at the cost of decreased throughput. Therefore, it may be beneficial to allocate fewer resources to positioning services and instead have receivers exploit the unknown data resources to obtain accurate positioning. In this section, the positioning performance of the ML and DD TOA estimators is evaluated on a simulated LEO downlink channel.

Setup: The simulated environment consists of four LEO satellites and one receiver. The receiver's cell receives downlink service from each satellite in a time-duplexed manner with the interval between downlink bursts from each satellite denoted as T_{burst}. During each downlink burst, the servicing satellite directs its beam to the center of the cell and pre-compensates for the expected Doppler shift experienced by a stationary user at the center of the cell. For simplicity, the receiver is stationary and located at the center of its cell. The receiver also has perfect knowledge of each satellite's relative position and applies conventional beamforming weights to its

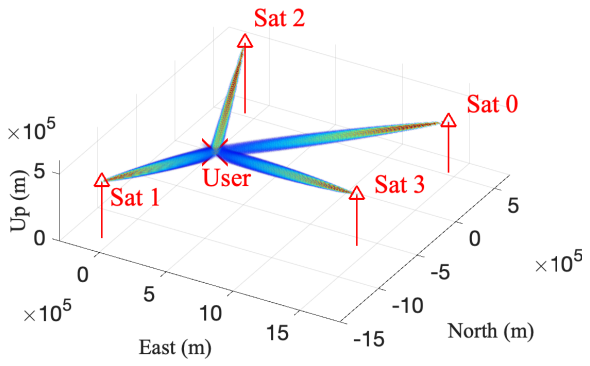

Fig. 6: A visualization of the four satellites in the simulated LEO constellation and each satellite's downlink beam.

phased array elements to direct its beam in the known direction of the servicing satellite. Fig. 6 shows a graphical depiction of the simulation setup and the downlink beams.

Each burst consists of $N_{\text{sym}} = 4$ OFDM symbols. The OFDM signal parameters and simulation parameters are listed in Table I. Simulated frequency-selective, time-varying channels are generated using QuaDRiGa [48]. Two channel models are considered: *QuaDRiGa_NTN_DenseUrban_LOS* [49] and *5G-ALLSTAR_DenseUrban_LOS* [50].

For each channel model, 1000 channel realizations were generated. For each channel realization, 1000 realizations of AWGN were generated. Pseudorange estimates were obtained for each channel and AWGN realization using each of the estimators Additionally, the ZZB was computed for each channel realization.

Let $\boldsymbol{r}_i$ be the east-north-up (ENU) coordinates of satellite i for $i \in \{0, 1, 2, 3\}$, $\boldsymbol{r}$ be the ENU coordinates of the receiver, and δ_t be the clock offset between the receiver and the satellite constellation. Define c as the speed of light, $\boldsymbol{\epsilon}_\rho$ as pseudorange estimation error, and $\boldsymbol{\theta}_\rho \triangleq [\boldsymbol{r}^T,\ c\delta_t]^T$. Additionally define

$$h_i(\boldsymbol{\theta}_\rho) \triangleq c||\boldsymbol{r}_i - \boldsymbol{r}|| + c\delta_t, \tag{38}$$

which is expressed in vector form as $\boldsymbol{h}(\boldsymbol{\theta}_\rho) \triangleq [h_0(\boldsymbol{\theta}_\rho),\ h_1(\boldsymbol{\theta}_\rho),\ h_2(\boldsymbol{\theta}_\rho),\ h_3(\boldsymbol{\theta}_\rho)]^T$. Then the vectorized pseudorange measurement equation for all satellites becomes

$$\boldsymbol{\rho} = \boldsymbol{h}(\boldsymbol{\theta}_\rho) + \boldsymbol{\epsilon}_\rho. \tag{39}$$

Letting $\boldsymbol{\Sigma}_{\boldsymbol{\rho}}$ be the covariance of the pseudorange error $\boldsymbol{\epsilon}_\rho$, a positioning solution can then be obtained by solving the weighted nonlinear least-squares problem:

$$\hat{\boldsymbol{\theta}}_\rho = \underset{\boldsymbol{\theta}_\rho}{\operatorname{argmin}} \left|\left|\boldsymbol{\rho} - \boldsymbol{h}(\boldsymbol{\theta}_\rho)\right|\right|^2_{\boldsymbol{\Sigma}_{\boldsymbol{\rho}}^{-1}}. \tag{40}$$

The error covariance of this estimate can be approximated by linearizing the pseudorange residuals at the true values of $\boldsymbol{\theta}_\rho$. Defining the matrix

$$\boldsymbol{A} \triangleq \begin{bmatrix} \frac{\boldsymbol{r}_0 - \boldsymbol{r}}{||\boldsymbol{r}_0 - \boldsymbol{r}||^2} & \frac{\boldsymbol{r}_1 - \boldsymbol{r}}{||\boldsymbol{r}_1 - \boldsymbol{r}||^2} & \frac{\boldsymbol{r}_2 - \boldsymbol{r}}{||\boldsymbol{r}_2 - \boldsymbol{r}||^2} & \frac{\boldsymbol{r}_3 - \boldsymbol{r}}{||\boldsymbol{r}_3 - \boldsymbol{r}||^2} \\ 1 & 1 & 1 & 1 \end{bmatrix}^T, \tag{41}$$

TABLE I: LEO Simulation Parameters

OFDM Parameters	
K	240 subcarriers
Δ_f (Subcarrier Spacing)	240 kHz
Symbol Duration	4.167 µs
Cyclic Prefix Duration	0.521 µs
Bandwidth	57.60 MHz
Symbol Constellation	QPSK
Simulation Parameters	
f_c (Carrier Frequency)	10.7 GHz
Polarization	LHCP @ RX & TX
TX Gain	34 dB
TX Beamwidth	3.67°
RX Array	32×32 URA ($\lambda/2$ spacing)
RX Gain	30 dB
RX Beamwidth	3.58°
EIRP	-15 dBW/4kHz
σ^2 (Resource Noise Power)	$-173.8 + 10\log_{10}(\Delta_f)$ dBm
T_a (*a priori* TOA Duration)	156.25 ns
T_{burst} (Burst Interval)	1 ms
Satellite Altitude	550 km
Satellite Constellation	Walker-Delta 53°:1584/22/39
Elevation Mask	30°
Channel Models	*QuaDRiGa_NTN_DenseUrban_LOS* *5G-ALLSTAR_DenseUrban_LOS*

the error covariance of $\hat{\boldsymbol{\theta}}_\rho$ can then be described as $\boldsymbol{Q} \triangleq \left(\boldsymbol{A}^T \boldsymbol{\Sigma}_{\boldsymbol{\rho}}^{-1} \boldsymbol{A}\right)^{-1}$. The diagonal elements of this error covariance are defined as $[\sigma_x^2,\ \sigma_y^2,\ \sigma_z^2,\ \sigma_{c\delta t}^2]^T \triangleq \text{diag}(\boldsymbol{Q})$.

Since each pseudorange is obtained independently, $\boldsymbol{\Sigma}_{\boldsymbol{\rho}}$ is a diagonal matrix consisting of elements $\sigma_{\rho_0}^2$, $\sigma_{\rho_1}^2$, $\sigma_{\rho_2}^2$, and $\sigma_{\rho_3}^2$. These variances are obtained either from the ZZB or from the empirical error variance of the TOA estimates over AWGN realizations for a single channel realization. It is important to note that the distribution of the TOA errors is unknown and not guaranteed to be Gaussian. However, the ZZBs and empirical error variances can be compared through this linearized model.

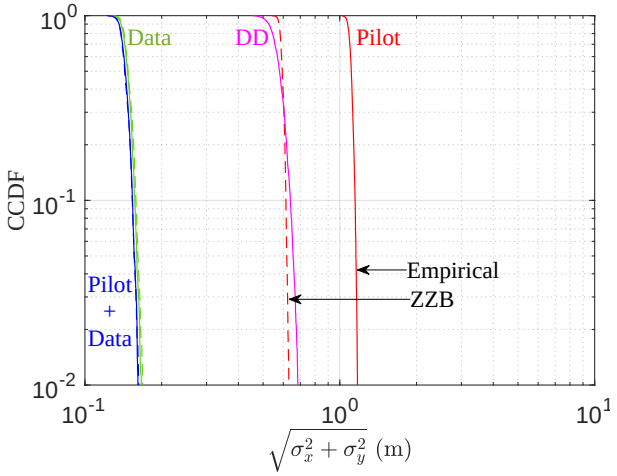

Fig. 7: The complementary cumulative distribution functions (CCDFs) of horizontal RMSE over channel realizations using the QuaDRiGa channel model, comparing the pilot-only, data-only, pilot-plus-data, and DD estimators. Results are shown for both the ZZB and empirical data.

LEO Results: Fig. 7 and Fig. 8 plot the CCDF over channel realizations of the horizontal RMSEs and vertical RMSEs

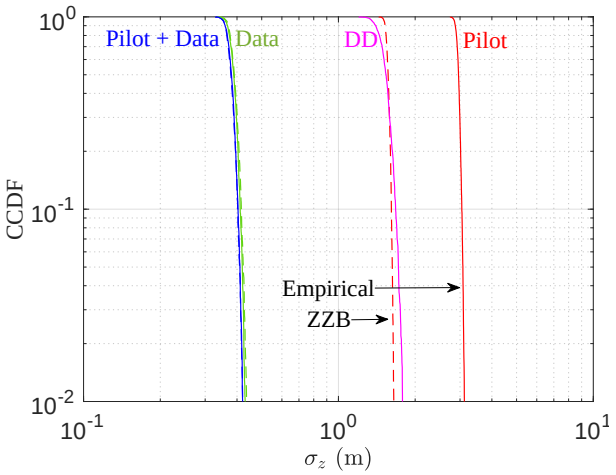

Fig. 8: As Fig. 7 but for vertical RMSE.

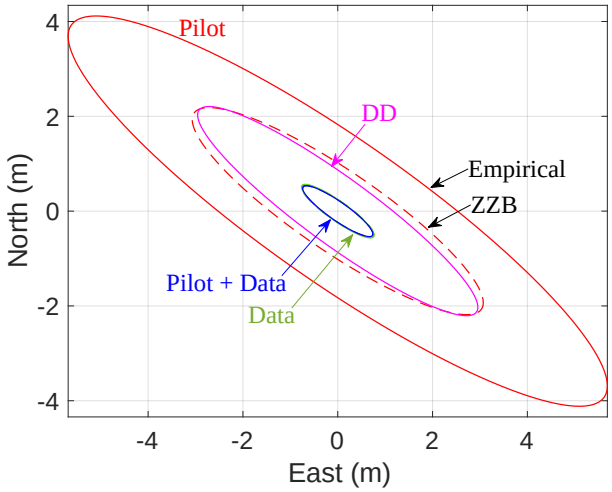

Fig. 9: The 95 % horizontal positioning error ellipses for the QuaDRiGa channel model. Results are obtained using the Chebyshev inequality with both the ZZB and empirical RMSEs over all channel realizations.

using the *QuaDRiGa_NTN_DenseUrban_LOS* channel model. Both the ZZBs and empirical RMSEs are depicted for each of the estimators. The pilot-only estimator results in the greatest errors, having a 90th percentile horizontal RMSE of 1.15 m and vertical RMSE of 3.05 m. The decision-directed estimator yields moderate improvements over the pilot-only estimator, having a 90th percentile horizontal RMSE of 0.64 m and vertical RMSE of 1.68 m. Accuracy is improved significantly with the data-only estimator, having a 90th percentile horizontal RMSE of 0.16 m and vertical RMSE of 0.41 m. Finally, the pilot-plus-data estimator achieves the best accuracy, having a 90th percentile horizontal RMSE of 0.15 m and vertical RMSE of 0.40 m.

Fig. 9 provides an alternative depiction of these results, visualizing the 95 % error ellipses using both the ZZB and empirical RMSEs over all channel realizations. Since the error distribution is not guaranteed to be Gaussian, the 95 % error ellipses are computed using the multivariate Chebyshev inequality [51], resulting in a 2D ellipse corresponding to 6.32 standard deviations. The Chebyshev inequality allows confidence intervals to be constructed for any arbitrary distribution with a finite variance. This depiction highlights how the data-only and pilot-plus-data estimators reduce positioning error significantly compared to the pilot-only estimator and even the DD estimator. The data-only and pilot-plus-data ellipses have a semi-major axis of approximately 0.9 m compared to the DD ellipse's 3.6 m and pilot-only ellipse's 6.9 m. The ZZB ellipses are nearly coincident with the empirical ellipses for data-only and pilot-plus-data estimation, while the ZZB ellipse for pilot-only estimation has a gap compared to the empirical ellipse and a semi-major axis of 3.7 m.

The pilot-plus-data and data-only estimators are capable of outperforming the pilot-only estimator by exploiting significantly more resources in the signal. Compared to the DD estimator, these estimators are not negatively impacted by errors in the hard decoding process. As a result, the DD estimator requires much higher SNRs to approach the performance of the pilot-plus-data and data-only estimators.

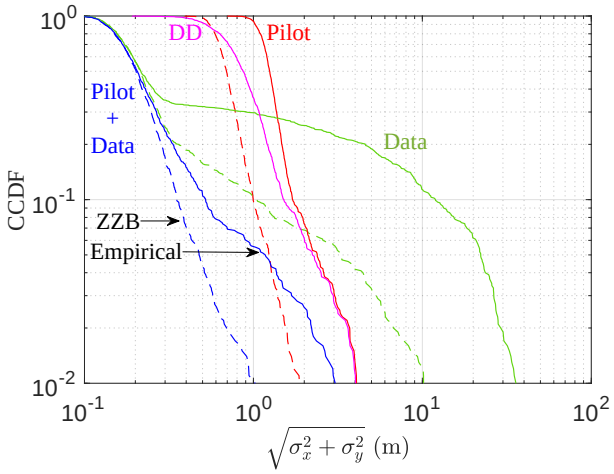

Fig. 10: The CCDFs over channel realizations of horizontal RMSE using the 5G-ALLSTAR channel model, comparing the pilot-only, data-only, pilot-plus-data, and DD estimators. Results are shown for both the ZZB and empirical data.

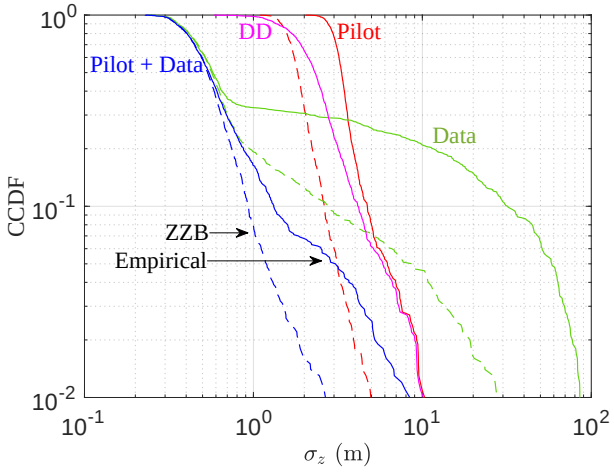

Fig. 11: As Fig. 10 but for vertical RMSE.

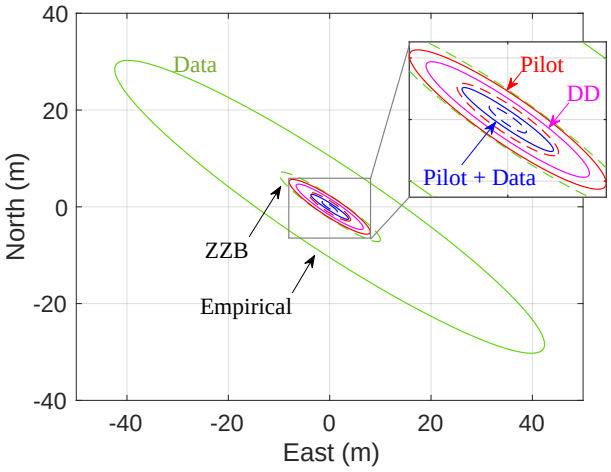

Fig. 12: The 95 % horizontal positioning error ellipses for the 5G-ALLSTAR channel model. Results are obtained using the Chebyshev inequality with both the ZZB and empirical RMSEs over all channel realizations.

Figs. 10 and 11 plot the CCDF over channel realizations of the horizontal RMSEs and vertical RMSEs using the *5G-ALLSTAR_DenseUrban_LOS* channel model, similar to Figs. 7 and 8. Different patterns emerge with this channel model, as greater fluctuations in SNR are simulated. Similar to the results in Figs. 7 and 8, the pilot-plus-data estimator achieves the greatest accuracy, having a 90th percentile horizontal RMSE of 0.52 m and vertical RMSE of 1.32 m. However, the data-only estimator only outperforms the DD and pilot-only estimators in approximately 70 % of channel realizations. In the remaining 30 % of channel realizations, the low SNRs cause the data-only estimator to enter its thresholding regime, resulting in RMSEs surpassing those achieved using only pilots. The data-only estimator has a 90th percentile horizontal RMSE of 11.96 m and vertical RMSE of 32.00 m. Meanwhile, the DD estimator exhibits less improvement over the pilot-only estimator in this channel model, having a 90th percentile horizontal RMSE of 1.52 m and vertical RMSE of 4.16 m compared to the pilot-only estimator's 90th percentile horizontal RMSE of 1.71 m and vertical RMSE of 4.73 m.

Fig. 12 visualizes the 95 % error ellipses for the 5G-ALLSTAR channel model, similar to Fig. 9. As with the results in Fig. 10 and Fig. 11, the data-only estimator exhibits significant errors due to the poor channel conditions. Meanwhile, the pilot-plus-data estimator reduces postioning error compared to both the pilot-only and DD estimators. The pilot-plus-data ellipse has a semi-major axis of approximately 4.5 m compared to the DD ellipse's 8.0 m, pilot-only ellipse's 9.6 m, and data-only ellipse's 52 m.

The large SNR fluctuations in the *5G-ALLSTAR_DenseUrban_LOS* channel model resulted in the estimators entering their low-SNR thresholding regimes. The pilot-plus-data estimator provides robustness against these thresholding effects at low SNR while simultaneously maximizing accuracy in high SNR. In comparison to Fig. 7 and Fig. 8, the ZZBs are much looser due to the bound not being as tight in the low-SNR thresholding regime as in the high-SNR regime.

VI. Conclusions

This paper has derived a novel ZZB on TOA estimation with OFDM signals containing unknown data resources. This ZZB serves as a lower bound to both ML and DD estimators that can exploit unknown data resources to improve estimation accuracy. The ZZB was shown to be tighter to empirical errors than the CRB and MCRB derived in prior work, making it a useful criterion for evaluating different OFDM resource allocations for TOA estimation. Comparisons were then made between four different types of estimators: pilot-only ML, data-only ML, pilot-plus-data ML, and DD, demonstrating that the pilot-plus-data estimator can significantly improve TOA estimation accuracy. The ZZB was then used to evaluate different allocations of PRSs within a wideband OFDM signal, which can guide resource allocations that minimize overhead while still achieving TOA accuracy requirements. Finally, the ZZBs and TOA estimators were evaluated on simulated LEO channels, quantifying the distribution of positioning RMSEs across channel realizations. These results highlight the potential for OFDM networks to significantly improve positioning performance while still prioritizing data throughput.

Acknowledgments

This work was supported by the U.S. Department of Transportation under Grant 69A3552348327 for the CARMEN+ University Transportation Center, by the U.S. Space Force under an STTR contract with Coherent Technical Services, Inc., and by affiliates of the 6G@UT center within the Wireless Networking and Communications Group at The University of Texas at Austin.

References

[1] H. K. Dureppagari, C. Saha, H. S. Dhillon, and R. M. Buehrer, "NTN-based 6G localization: Vision, role of LEOs, and open problems," *IEEE Wireless Communications*, vol. 30, no. 6, pp. 44–51, 2023.

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

[3] X. Lin, S. Cioni, G. Charbit, N. Chuberre, S. Hellsten, and J.-F. Boutillon, "On the path to 6G: Embracing the next wave of low Earth orbit satellite access," *IEEE Communications Magazine*, vol. 59, no. 12, pp. 36–42, 2021.

[4] S. Dwivedi, R. Shreevastav, F. Munier, J. Nygren, I. Siomina, Y. Lyazidi, D. Shrestha, G. Lindmark, P. Ernström, E. Stare *et al.*, "Positioning in 5G networks," *IEEE Communications Magazine*, vol. 59, no. 11, pp. 38–44, 2021.

[5] A. M. Graff and T. E. Humphreys, "Purposeful co-design of OFDM signals for ranging and communications," *EURASIP Journal on Advances in Signal Processing*, 2024.

[6] F. Bellili, N. Atitallah, S. Affes, and A. Stéphenne, "Cramér-Rao lower bounds for frequency andphase NDA estimation from arbitrary square QAM-modulated signals," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4517–4525, 2010.

[7] A. Masmoudi, F. Bellili, S. Affes, and A. Stephenne, "A non-data-aided maximum likelihood time delay estimator using importance sampling," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4505–4515, 2011.

[8] A. Zeira and P. M. Schultheiss, "Realizable lower bounds for time delay estimation. 2. Threshold phenomena," *IEEE transactions on signal processing*, vol. 42, no. 5, pp. 1001–1007, 1994.

[9] J. A. Nanzer, M. D. Sharp, and D. Richard Brown, "Bandpass signal design for passive time delay estimation," in *2016 50th Asilomar Conference on Signals, Systems and Computers*, Nov. 2016, pp. 1086–1091.

[10] Z. Sahinoglu, S. Gezici, and I. Güvenc, *Ultra-wideband positioning systems: theoretical limits, ranging algorithms, and protocols*. Cambridge university press, 2008.

[11] E. W. Barankin, "Locally best unbiased estimates," *The Annals of Mathematical Statistics*, vol. 20, no. 4, pp. 477–501, 1949. [Online]. Available: http://www.jstor.org/stable/2236306

[12] R. McAulay and E. Hofstetter, "Barankin bounds on parameter estimation," *IEEE Transactions on Information Theory*, vol. 17, no. 6, pp. 669–676, 1971.

[13] J. Ziv and M. Zakai, "Some lower bounds on signal parameter estimation," *IEEE Transactions on Information Theory*, vol. 15, no. 3, pp. 386–391, 1969.

[14] 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.

[15] M. Koivisto, M. Costa, J. Werner, K. Heiska, J. Talvitie, K. Leppänen, V. Koivunen, and M. Valkama, "Joint device positioning and clock synchronization in 5G ultra-dense networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 2866–2881, 2017.

[16] A. Kakkavas, H. Wymeersch, G. Seco-Granados, M. H. C. García, R. A. Stirling-Gallacher, and J. A. Nossek, "Power allocation and parameter estimation for multipath-based 5G positioning," *IEEE Transactions on Wireless Communications*, vol. 20, no. 11, pp. 7302–7316, 2021.

[17] K. Shamaei and Z. M. Kassas, "LTE receiver design and multipath analysis for navigation in urban environments," *Navigation*, vol. 65, no. 4, pp. 655–675, 2018.

[18] ——, "Receiver design and time of arrival estimation for opportunistic localization with 5G signals," *IEEE Transactions on Wireless Communications*, vol. 20, no. 7, pp. 4716–4731, 2021.

[19] M. Neinavaie, J. Khalife, and Z. M. Kassas, "Cognitive opportunistic navigation in private networks with 5G signals and beyond," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 1, pp. 129–143, 2021.

[20] 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.

[21] H. Xv, Y. Sun, Y. Zhao, M. Peng, and S. Zhang, "Joint beam scheduling and beamforming design for cooperative positioning in multi-beam LEO satellite networks," *IEEE Transactions on Vehicular Technology*, 2023.

[22] K. Shi, E. Serpedin, and P. Ciblat, "Decision-directed fine synchronization in OFDM systems," *IEEE Transactions on Communications*, vol. 53, no. 3, pp. 408–412, 2005.

[23] J. Ran, R. Grunheid, H. Rohling, E. Bolinth, and R. Kern, "Decision-directed channel estimation method for OFDM systems with high velocities," in *The 57th IEEE Semiannual Vehicular Technology Conference, 2003. VTC 2003-Spring.*, vol. 4. IEEE, 2003, pp. 2358–2361.

[24] E. Karami and M. Shiva, "Decision-directed recursive least squares MIMO channels tracking," *EURASIP Journal on Wireless Communications and Networking*, vol. 2006, pp. 1–10, 2006.

[25] Y. Xiong, L. Tang, S. Sun, L. Liu, S. Mao, Z. Zhang, and N. Wei, "Data-aided channel estimation and combining for cell-free massive MIMO with low-resolution ADCs," *IEEE Communications Letters*, 2024.

[26] X. Ma, C. Tepedelenlioglu, G. B. Giannakis, and S. Barbarossa, "Non-data-aided carrier offset estimators for OFDM with null subcarriers: identifiability, algorithms, and performance," *IEEE Journal on selected areas in communications*, vol. 19, no. 12, pp. 2504–2515, 2001.

[27] A. Al-Dweik, "A novel non-data-aided symbol timing recovery technique for OFDM systems," *IEEE Transactions on Communications*, vol. 54, no. 1, pp. 37–40, 2006.

[28] F.-X. Socheleau, A. Aissa-El-Bey, and S. Houcke, "Non data-aided SNR estimation of OFDM signals," *IEEE communications letters*, vol. 12, no. 11, pp. 813–815, 2008.

[29] A. Masmoudi, F. Bellili, S. Affes, and A. Stéphenne, "Closed-form expressions for the exact Cramér–Rao bounds of timing recovery estimators from BPSK, MSK and square-QAM transmissions," *IEEE transactions on signal processing*, vol. 59, no. 6, pp. 2474–2484, 2011.

[30] A. Masmoudi, F. Bellili, S. Affes, and A. Ghrayeb, "Maximum likelihood time delay estimation from single- and multi-carrier DSSS multipath MIMO transmissions for future 5G networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 4851–4865, 2017.

[31] S. Monfared, T.-H. Nguyen, T. Van der Vorst, P. De Doncker, and F. Horlin, "Iterative NDA positioning using angle-of-arrival measurements for IoT sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11 369–11 382, 2020.

[32] W. Wang, T. Jost, C. Gentner, S. Zhang, and A. Dammann, "A semiblind tracking algorithm for joint communication and ranging with OFDM signals," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 7, pp. 5237–5250, 2015.

[33] R. Adam and P. A. Hoeher, "Semi-blind channel estimation for joint communication and positioning," in *2013 10th Workshop on Positioning, Navigation and Communication (WPNC)*. IEEE, 2013, pp. 1–5.

[34] C. Mensing, S. Sand, A. Dammann, and W. Utschick, "Data-aided location estimation in cellular OFDM communications systems," in *GLOBECOM 2009 - 2009 IEEE Global Telecommunications Conference*, 2009, pp. 1–7.

[35] T. Laas and W. Xu, "On the Ziv-Zakai bound for time difference of arrival estimation in CP-OFDM systems," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2021, pp. 1–5.

[36] A. Dammann, T. Jost, R. Raulefs, M. Walter, and S. Zhang, "Optimizing waveforms for positioning in 5G," in *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2016, pp. 1–5.

[37] E. Staudinger, M. Walter, and A. Dammann, "Optimized waveform for energy efficient ranging," in *2017 14th Workshop on Positioning, Navigation and Communications (WPNC)*. IEEE, 2017, pp. 1–6.

[38] A. M. Graff and T. E. Humphreys, "Ziv-Zakai-optimal OFDM resource alocation for time-of-arrival estimation," *IEEE Transactions on Wireless Communications*, 2024, submitted for review.

[39] A. Gusi-Amigó, P. Closas, A. Mallat, and L. Vandendorpe, "Ziv-Zakai bound for direct position estimation," *Navigation*, vol. 65, no. 3, pp. 463–475, 2018.

[40] W. Xu, M. Huang, C. Zhu, and A. Dammann, "Maximum likelihood TOA and OTDOA estimation with first arriving path detection for 3GPP LTE system," *Transactions on Emerging Telecommunications Technologies*, vol. 27, no. 3, pp. 339–356, 2016.

[41] A. N. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramer-Rao bound and its application to synchronization problems," *IEEE Transactions on Communications*, vol. 42, no. 234, pp. 1391–1399, 1994.

[42] S. Bellini and G. Tartara, "Bounds on error in signal parameter estimation," *IEEE Transactions on Communications*, vol. 22, no. 3, pp. 340–342, 1974.

[43] K. L. Bell, Y. Steinberg, Y. Ephraim, and H. L. Van Trees, "Extended Ziv-Zakai lower bound for vector parameter estimation," *IEEE Transactions on information theory*, vol. 43, no. 2, pp. 624–637, 1997.

[44] N. B. Mehta, J. Wu, A. F. Molisch, and J. Zhang, "Approximating a sum of random variables with a lognormal," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2690–2699, 2007.

[45] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. US Government printing office, 1968, vol. 55.

[46] D. Dardari and M. Z. Win, "Ziv-Zakai bound on time-of-arrival estimation with statistical channel knowledge at the receiver," in *2009 IEEE International Conference on Ultra-Wideband*. IEEE, 2009, pp. 624–629.

[47] D. Chazan, M. Zakai, and J. Ziv, "Improved lower bounds on signal parameter estimation," *IEEE transactions on Information Theory*, vol. 21, no. 1, pp. 90–93, 1975.

[48] F. Burkhardt, S. Jaeckel, E. Eberlein, and R. Prieto-Cerdeira, "QuaDRiGa: A MIMO channel model for land mobile satellite," in *The 8th European Conference on Antennas and Propagation (EuCAP 2014)*. IEEE, 2014, pp. 1274–1278.

[49] S. Jaeckel, L. Raschkowski, and L. Thieley, "A 5G-NR satellite extension for the QuaDRiGa channel model," in *2022 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*. IEEE, 2022, pp. 142–147.

[50] N. Cassiau, G. Noh, S. Jaeckel, L. Raschkowski, J.-M. Houssin, L. Combelles, M. Thary, J. Kim, J.-B. Dore, and M. Laugeois, "Satellite and terrestrial multi-connectivity for 5G: Making spectrum sharing possible," in *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. IEEE, 2020, pp. 1–6.

[51] A. W. Marshall and I. Olkin, "Multivariate chebyshev inequalities," *The Annals of Mathematical Statistics*, pp. 1001–1014, 1960.