Efficient Channel Estimation for Aerial Wireless Communications

Dennis Ogbe, *Student Member, IEEE,* David J. Love, *Fellow, IEEE,* Matthew Rebholz, *Member, IEEE*
Patrick Bidigare, *Member, IEEE*

***Abstract*—We present a technique to jointly estimate the channel taps and the frequency offset due to the Doppler effect of a special class of doubly-dispersive aeronautical channels. The algorithm includes the use of pulse-repetition techniques at the transmitter and the "power method" subspace iteration at the receiver. We show that transmitting CAZAC sequences for channel sounding yields estimators with low computational complexity. Numerical simulations indicate performance comparable to the estimation-theoretic lower bound.**

***Index Terms*—Doubly dispersive channel, Doppler effect, channel estimation, power method, CAZAC sequence**

I. Introduction

THE performance of modern wireless communication systems fundamentally depends on the quality of the channel estimates at either the receiver, the transmitter, or both. This statement holds of course for aerial wireless communication systems, which have recently experienced a resurgence of interest from both the academic and industrial community. Aerial communication systems, especially those based on unmanned aerial vehicle (UAV) platforms, are considered enabling technologies for future wireless networks like the consumer-oriented fifth-generation (5G) cellular networks and ad-hoc public safety networks. [1] In addition to 5G applications like UAV control and video streaming, aerial communication systems are actively being researched, developed, and deployed as part of the ongoing effort to increase the number of people connected to the Internet, lead by initiatives like Google's balloon-based "Project Loon" [2]. These efforts seek to increase the availability of low-cost Internet access by connecting a distributed set of airborne transmitters through wireless backhaul. Furthermore, apart from being considered for future communication networks, aerial wireless communication systems continue to play a significant role in other civilian applications like air traffic control as well as in a variety of military applications.

In addition to the time dispersion due to multipath propagation present in many wireless channels, the effects of vehicular motion inherent to aeronautical communication systems may induce frequency dispersion due to the Doppler effect [3]. The general class of channels exhibiting both time dispersive and frequency dispersive effects, usually referred to as *doubly dispersive* channels, is continuing to attract considerable interest from researchers and practitioners, in part due to its ubiquity in modern wireless communication systems, e.g., [4]–[10]. In the classical *channel estimation followed by data transmission* set-up, which we follow in this manuscript, it is thus desirable from a channel estimation perspective to obtain knowledge of the parameters governing both the time dispersion of the channel—assumed to be modeled as a discrete-time tapped delay line [11]—as well as the frequency dispersion, here assumed to be mainly caused by the Doppler effect due to vehicular motion.

The specific model we introduce and consider in this manuscript arises from our consideration of communication systems on high-velocity airborne vehicles. The general model of doubly dispersive channels where the frequency dispersion is due to Doppler assumes that the Doppler effect and thus the resulting frequency shift varies for each multipath component of the channel, resulting in a Doppler *spectrum*. In this manuscript however, we assume that for high-velocity airborne vehicles the Doppler spectrum is dominated by the *bulk* frequency shift due to the motion of the vehicle (see Section II), rendering the individual shifts on each multipath component negligible in our analysis.

The model and corresponding estimation problem of a single bulk frequency shift coupled with multipath transmission mirrors the well-researched problem of estimating the frequency offset in orthogonal frequency division multiplexing (OFDM) systems. The tightly-spaced subcarriers in OFDM systems lose orthogonality in the presence of any kind of frequency offset, resulting in inter-channel interference. It has been shown that the bit error rate increases significantly if those offsets are left uncompensated [12].

The popularity of OFDM in modern communication systems has lead to the development of many different techniques to estimate and compensate for the frequency offset due to the Doppler effect and/or mismatches between the transmitter and receiver local oscillators. In general, the available frequency offset estimators can be classified into two categories. Blind techniques provide estimates of the frequency offset without the need for pilot symbols. The technique in [13] exploits the cyclostationarity inherent to OFDM waveforms to extract an estimate of the frequency offset. The authors of [14] derived a kurtosis-based estimator, which generalizes their single input, single output (SISO) results to multiple input, multiple output (MIMO) OFDM systems. The other class of OFDM frequency offset techniques includes semi-blind and non-blind estimators. The common property of all of these techniques, which relates to the results of this paper, is the reliance on some sort

D. Ogbe and D. J. Love are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907, USA (e-mail: {dogbe, djlove}@purdue.edu).
M. Rebholz email: rebholz.matthew@gmail.com
P. Bidigare email: bidigare@ieee.org

of redundancy in the structure of the waveform to compute the estimates. The work in [15] presented a frequency offset estimator based on the repetition of two OFDM symbols. The work in [16] uses a similar principle, keeping the redundant data within one OFDM symbol and using a second symbol for fine estimation as well as timing synchronization. The main idea behind these two techniques is that the cross-correlation of the repeated time-domain sequence of the receiver would perfectly reproduce the frequency offset in the absence of noise. The technique presented by Schmidl and Cox in [16] exploits the fact that the inverse discrete Fourier transform (IDFT) of a subcarrier allocation where every other subcarrier is set to zero produces a time-domain sequence of two repeated half-symbols. This concept was extended in [17] and generalized to produce time-domain sequences with more than two repetitions. The complexity of the Schmidl and Cox technique was further simplified in [18], reducing the number of OFDM pilot symbols needed from two to one. Furthermore, the work in [19]–[21] exploits the inherent redundancy of the usage of a cyclic prefix in OFDM to estimate the frequency offset. Finally, the frequency offset compensation problem was studied specifically for aeronautical channels in [22] and the references therein. The authors assume a two-ray, dual Doppler shift model and compensate for each Doppler shift using the techniques from [19] after separating the signals of the two incoming paths.

Our contribution to the problem is a technique inspired by the various OFDM frequency offset estimators which jointly estimates the bulk Doppler shift and channel taps of an aeronautical channel. In contrast to the aforementioned OFDM-based techniques, our algorithm assumes single-carrier modulation, but could potentially be adapted to support multi-carrier modulations with a few modifications. More specifically, although our algorithm requires a certain time-domain structure combined with time-domain processing of the sounding signals, it places no restriction on the data transmission, allowing multi-carrier modulations to be used. The main idea behind our technique is to transmit a cyclically prefixed training sequence consisting of a repeated shorter training pulse. The receiver then computes an estimate of the Doppler shift using a combination of subspace estimation and matched filtering. After estimating and correcting for the effects of the Doppler shift, the receiver then computes the conditional maximum-likelihood (ML) estimate of the channel taps. We summarize the contributions of this paper as follows.

- We introduce a system model for high-velocity airborne wireless communication systems exhibiting time dispersion due to multipath and frequency dispersion due to a bulk Doppler shift caused by vehicular motion
- We derive the estimation-theoretic lower bound for estimating the channel taps as well as the Doppler shift of these channels
- We develop a pulse-repetition based channel estimator for these parameters
- We show that using constant amplitude, zero autocorrelation (CAZAC) sequences as pulses in our algorithm decreases the computational burden of our estimator
- We present numerical studies analyzing the performance of our estimator

The rest of this paper is structured as follows. Section II describes the system model considered throughout the paper. In Section III we derive the Cramer-Rao lower bound (CRLB) for the joint estimation of the Doppler shift and the channel taps using the observation model given in Section II. We present and discuss the joint estimation algorithm in Section IV and discuss the special case for CAZAC sequences in Section IV-D. Simulation results are presented and discussed in Section V and we provide some concluding remarks in Section VI.

Notation. We will use the following notation throughout this manuscript. Bold upper-case and lower-case letters (such as $\mathbf{A}$ and $\mathbf{a}$) denote matrices and column vectors, respectively. The operators $(\cdot)^{\mathsf{T}}$, $\overline{(\cdot)}$ and $(\cdot)^*$ denote matrix transposition, element-wise complex conjugation and matrix Hermitian transposition, respectively. $\|\cdot\|_2$ denotes the vector ℓ_2-norm and $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. $\mathcal{CN}(\mathbf{a}, \mathbf{A})$ denotes a complex Gaussian random vector with mean $\mathbf{a}$ and covariance matrix $\mathbf{A}$.

II. System Model

Our system model consists of a single-antenna transmitter and receiver pair communicating over a single-input, single-output (SISO) doubly dispersive wireless communication channel. Under these assumptions, the general discrete-time complex baseband input-output model between the transmitter and the receiver can be written as

$$y[k] = \sqrt{\rho} \sum_{\ell=0}^{L-1} h_k[\ell] s[k-\ell] + n[k], \tag{1}$$

where we denote $h_k[\ell],\ \ell \in \{0,\ \ldots,\ L-1\}$ as the ℓ-th complex channel filter tap at time index k. Furthermore, $s[k],\ k \in \{0,\ \ldots,\ L_s - 1\}$ denotes the k-th sample of an arbitrary data or sounding sequence of length L_s, $n[k] \sim \mathcal{CN}(0, \sigma_k^2)$ is a sample of an additive Gaussian noise process, and ρ is a transmit signal-to-noise ratio (SNR) term.

Our channel model assumes an airborne platform where the Doppler shift is approximated as being of the same magnitude on each multipath component of $h_k[\ell]$. In general, the expression for a channel tap $h_k[\ell]$ can be written as [3]

$$h_k[\ell] = \frac{1}{\sqrt{N_p}} \sum_{n=1}^{N_p} e^{j\theta_n} e^{j2\pi f_{d_n} k T_s} g_{\text{total}}(\ell T_s - \tau_n), \tag{2}$$

where N_p denotes the number of paths, θ_n and τ_n denote the phase shift and time delay of path n, g_{total} denotes the convolution of the transmitter and receiver pulse shaping filters, T_s denotes the sampling period of the system, and f_{d_n} represents the Doppler shift on the n-th path from transmitter to receiver. As mentioned in Section I, we make the assumption that the Doppler shift is equal across all paths, i.e., $f_{d_1} = f_{d_2} = \cdots = f_{d_{N_p}} = f_d$. This is a reasonable assumption when considering a ring of scatterers close to the transmitter traveling at the same velocity as the transmitter, for example the surface of an airplane. In this case, if we let v denote the relative velocity between the transmitter and receiver, we can define the Doppler shift as $f_d = f_c v/c$, f_c being the carrier frequency of the communication system. Figure 1 illustrates two possible scenarios of our system model. In either case (air-to-air or air-to-ground), the Doppler shift creating a frequency offset at the receiver is due to non-zero relative velocities between the transmitter and the receiver. To further illustrate the near-equal Doppler shift assumption, consider the case of scattering off the airframe. In this case, the Doppler *spread*, i.e., the variance of the Doppler shifts on the different paths, will be most significantly affected by the roll, pitch, and yaw rates of the aircraft. Suppose for example scattering off the engines of a commercial two-engine airliner. The yaw rate for a "dutch roll" maneuver is 2.2 rad/second [23], resulting in a maximum Doppler frequency of $\frac{f_c}{c} \cdot 12.4$ meters per second for an engine offset by 6 meters from the fuselage. However, realistic cruising speeds for commercial jet airliners are around 250 meters per second, resulting in a bulk Doppler shift about 20 times larger than the Doppler spread due to maneuvering. Similar arguments can be made for scattering off the ground in the air-to-ground case. Here, the ratio of the Doppler spread due to the scatterers and the Bulk Doppler shift is bounded by the ratio of the diameter of the ring of scatterers and the distance from the transmitter to the receiver. Scatterer ring diameters in the single-kilometer range and standoff distances on the order of tens of kilometers (expected standoff ranges for tactical distributed beamforming applications, for example [24]) give bulk/scatterer ratios comparable to the aforementioned air-to-air setting. Considering, for example, a distance of 50 km and a scatterer ring diameter of 1 km, the bulk Doppler shift is approximately 50 times larger than the shifts due to scattering.

To simplify notation, we denote the sampled baseband frequency offset as $\alpha = 2\pi T_s f_d$. We can thus rewrite (2) as

$$h_k[\ell] = e^{j\alpha k} \sum_{n=1}^{N_p} \frac{1}{\sqrt{N_p}} e^{j\theta_n} g_{\text{total}}(\ell T_s - \tau_n), \tag{3}$$

where, if the number of paths N_p grows large, we can approximate the summation term as a circularly symmetric complex Gaussian random variable [11]. We can thus write the time-varying channel impulse response as

$$h_k[\ell] = e^{j\alpha k} h[\ell], \tag{4}$$

where $h[\ell]$ can be approximated as $\mathcal{CN}(0, 1)$. Our doubly dispersive channel model thus consists of two separate components. We model the frequency dispersive part due to the Doppler shift with the complex exponential $e^{j\alpha k}$. We model the time dispersive part as a purely feed-forward tapped delay line with L taps. For the remainder of this manuscript, we will refer to these taps as finite impulse response (FIR) taps of the channel and denote them as $h[\ell],\ \ell \in \{0,\ \ldots,\ L-1\}$. Substituting (4) in (1) yields the input-output relationship that is considered in this paper:

$$y[k] = e^{j\alpha k} \sqrt{\rho} \sum_{\ell=0}^{L-1} h[\ell] s[k-\ell] + n[k]. \tag{5}$$

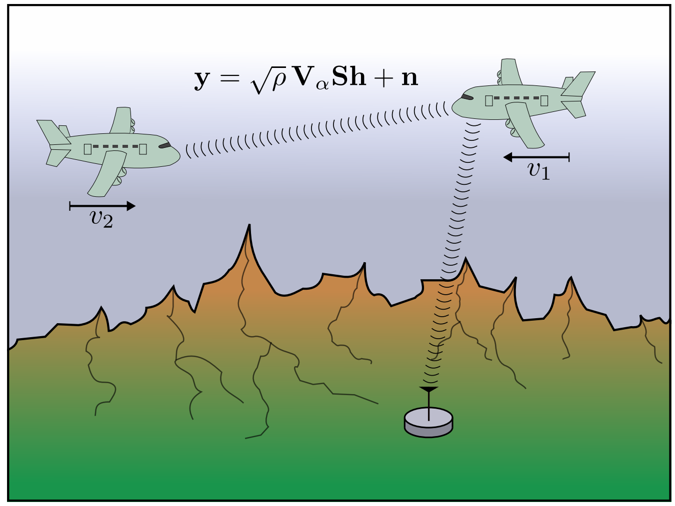


Fig. 1: An aircraft wishes to estimate the channel taps $\mathbf{h}$ of an air-to-ground or air-to-air channel with a single dominant Doppler component.

In our model, the receiver wishes to estimate a vector consisting of the L FIR taps of the channel impulse response $h[\ell]$ using the received samples of a known training sequence $s[k]$. We assume the that the training sequence $s[k]$ contains L_s samples. The received signal $y[k]$ thus consists of $N = L_s + L - 1$ samples after the convolution with the channel impulse response. Using matrix-vector notation, and after defining the vectors $\mathbf{s} = [s[0],\ \cdots,\ s[L_s-1]]^{\mathsf{T}}$, $\mathbf{h} = [h[0],\ \cdots,\ h[L-1]]^{\mathsf{T}}$, and $\mathbf{y} = [y[0],\ \cdots,\ y[N-1]]^{\mathsf{T}}$, as the training sequence, received samples, and channel impulse response vectors, respectively, the vector of received samples $\mathbf{y}$ can be written as

$$\mathbf{y} = \sqrt{\rho}\,\mathbf{V}_\alpha \mathbf{S}\mathbf{h} + \mathbf{n}, \tag{6}$$

where $\mathbf{V}_\alpha = \operatorname{diag}\left(\left[1\ e^{j\alpha}\ \cdots\ e^{j\alpha(N-1)}\right]\right)$ represents the Doppler shift matrix, $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C})$ is a vector of complex Gaussian noise samples with covariance matrix $\mathbf{C}$, and $\mathbf{S} \in \mathbb{C}^{N \times L}$ is the Toeplitz matrix obtained by linearly shifting the samples of the training sequence $\mathbf{s}$ for each column.

The estimation problem that the receiver is seeking to solve is a joint Doppler/channel estimation problem, since the Doppler shift α is unknown. More specifically, using our proposed technique, the receiver will use its estimate of α to correct for the effects of the Doppler shift matrix $\mathbf{V}_\alpha$ when obtaining an estimate of the vector of FIR taps $\mathbf{h}$. In the next section, we derive theoretical bounds on the variance of the estimators for the Doppler shift α and the channel taps $\mathbf{h}$ in this joint estimation framework.

III. Cramer-Rao Lower Bound

The derivation of the Cramer-Rao lower bounds for the joint estimator of the Doppler shift and the channel taps follows. The receiver wishes to estimate the parameter vector

$$\boldsymbol{\theta} = \left[\alpha\ \mathbf{h}^{\mathsf{T}}\right]^{\mathsf{T}} \tag{7}$$

from the observation given in (6). In order to derive the CRLB for any estimator $\widehat{\boldsymbol{\theta}}$, we construct the Fisher information matrix $\mathcal{I}(\boldsymbol{\theta})$, where its elements are defined as [25]

$$\mathcal{I}_{k,\ell}\left(\boldsymbol{\theta}\right) = \left[\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_k}\right]^* \mathbf{C}^{-1} \left[\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_\ell}\right] \tag{8}$$

and $\boldsymbol{\mu}(\boldsymbol{\theta})$ denotes the expectation of $\boldsymbol{\theta}$. A lower bound on the variance of the i-th element of $\boldsymbol{\theta}$ is then given by

$$\mathrm{var}(\widehat{\theta}_i) \geq \left[\mathcal{I}^{-1}(\boldsymbol{\theta})\right]_{ii}. \tag{9}$$

Lemma 1. *For $\boldsymbol{\theta}$ as defined in* (7) *and the observation model given in* (6)*, the Fisher information matrix is given by*

$$\mathcal{I}(\boldsymbol{\theta}) = \rho \begin{bmatrix} -\mathbf{h}^*\mathbf{S}^*\mathbf{V}_\alpha^*\mathbf{A}^*\mathbf{C}^{-1}\mathbf{A}\mathbf{V}_\alpha\mathbf{S}\mathbf{h} & -j\mathbf{h}^*\mathbf{S}^*\mathbf{V}_\alpha^*\mathbf{A}^*\mathbf{C}^{-1}\mathbf{V}_\alpha\mathbf{S} \\ j\mathbf{S}^*\mathbf{V}_\alpha^*\mathbf{C}^{-1}\mathbf{A}\mathbf{V}_\alpha\mathbf{S}\mathbf{h} & \mathbf{S}^*\mathbf{V}_\alpha^*\mathbf{C}^{-1}\mathbf{V}_\alpha\mathbf{S} \end{bmatrix} \tag{10}$$

where $\mathbf{A} = \mathrm{diag}\left([0 \ \cdots \ N-1]\right)$.

Proof: By inspection of (6), we have $\boldsymbol{\mu}(\boldsymbol{\theta}) = \sqrt{\rho}\,\mathbf{V}_\alpha\mathbf{S}\mathbf{h}$. It can then be shown that

$$\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_1} = \sqrt{\rho}\,\frac{\partial \mathbf{V}_\alpha}{\partial \alpha}\mathbf{S}\mathbf{h} = j\sqrt{\rho}\,\mathbf{A}\mathbf{V}_\alpha\mathbf{S}\mathbf{h}, \tag{11}$$

where $\mathbf{A} = \mathrm{diag}\left([0 \ \cdots \ N-1]\right)$. It can further be shown that

$$\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_2} = \sqrt{\rho}\,\mathbf{V}_\alpha\mathbf{S}. \tag{12}$$

After application of (8), we arrive at (10). ∎

Theorem 1. *The Cramer-Rao Lower Bound for any estimate of the Doppler frequency α, denoted by $\widehat{\alpha}$, is given by*

$$\begin{aligned} \mathrm{var}(\widehat{\alpha}) \geq (1/\rho)\cdot \\ \big(&\mathbf{h}^*\mathbf{S}^*\mathbf{V}_\alpha^*\mathbf{A}^*\mathbf{C}^{-1}\mathbf{A}\mathbf{V}_\alpha\mathbf{S}\mathbf{h} \\ &+\mathbf{h}^*\mathbf{S}^*\mathbf{V}_\alpha^*\mathbf{A}^*\mathbf{C}^{-1}\mathbf{V}_\alpha\mathbf{S}\left(\mathbf{S}^*\mathbf{V}_\alpha^*\mathbf{C}^{-1}\mathbf{V}_\alpha\mathbf{S}\right)^{-1}\cdot \\ &\mathbf{S}^*\mathbf{V}_\alpha^*\mathbf{C}^{-1}\mathbf{A}\mathbf{V}_\alpha\mathbf{S}\mathbf{h}\big)^{-1}. \end{aligned} \tag{13}$$

Furthermore, the Cramer-Rao Lower Bound for any estimate of the channel $\mathbf{h}$, denoted by $\widehat{\mathbf{h}}$, is given by

$$\begin{aligned} \mathrm{cov}(\widehat{\mathbf{h}}) \ \geq (1/\rho)\cdot \\ \big(&\mathbf{S}^*\mathbf{V}_\alpha^*\mathbf{C}^{-1}\mathbf{V}_\alpha\mathbf{S} \\ &+\mathbf{S}^*\mathbf{V}_\alpha^*\mathbf{C}^{-1}\mathbf{A}\mathbf{V}_\alpha\mathbf{S}\mathbf{h}\cdot \\ &\left(\mathbf{h}^*\mathbf{S}^*\mathbf{V}_\alpha^*\mathbf{A}^*\mathbf{C}^{-1}\mathbf{A}\mathbf{V}_\alpha\mathbf{S}\mathbf{h}\right)^{-1}\cdot \\ &\mathbf{h}^*\mathbf{S}^*\mathbf{V}_\alpha^*\mathbf{A}^*\mathbf{C}^{-1}\mathbf{V}_\alpha\mathbf{S}\big)^{-1}, \end{aligned} \tag{14}$$

where $\mathbf{A} \geq \mathbf{B}$ for two compatible matrices $\mathbf{A}$ and $\mathbf{B}$ means $\mathbf{A}-\mathbf{B}$ is a positive semi-definite matrix, or, equivalently, the product $\mathbf{x}^*\left(\mathbf{A}-\mathbf{B}\right)\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{C}^n$.

Proof: We invert (10) using the Schur complement formula [26] and apply (9) to arrive at (13) and (14). ∎

We note that (13) and (14) hold for arbitrary noise covariance matrices $\mathbf{C}$. In the case of white noise, the expressions for the CRLB simplify significantly, a fact shown in Corollary 1.

Corollary 1. *If the additive noise in* (6) *is i.i.d. Gaussian distributed, i.e.,* $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$*, the Cramer-Rao Lower Bounds are given by*

$$\begin{aligned} &\mathrm{var}(\widehat{\alpha}) \geq \\ &\frac{1}{\rho}\left(\mathbf{h}^*\mathbf{S}^*\mathbf{A}^*\mathbf{A}\mathbf{S}\mathbf{h} + \mathbf{h}^*\mathbf{S}^*\mathbf{A}^*\mathbf{S}(\mathbf{S}^*\mathbf{S})^{-1}\mathbf{S}^*\mathbf{A}\mathbf{S}\mathbf{h}\right)^{-1} \end{aligned} \tag{15}$$

and

$$\begin{aligned} &\mathrm{cov}(\widehat{\mathbf{h}}) \geq \\ &\frac{1}{\rho}\left(\mathbf{S}^*\mathbf{S} + \mathbf{S}^*\mathbf{A}\mathbf{S}\mathbf{h}(\mathbf{h}^*\mathbf{S}^*\mathbf{A}^*\mathbf{A}\mathbf{S}\mathbf{h})^{-1}\mathbf{h}^*\mathbf{S}^*\mathbf{A}^*\mathbf{S}\right)^{-1}. \end{aligned} \tag{16}$$

The expressions for the CRLB for both the Doppler and the channel estimates are crucial tools for analyzing the performance of our proposed algorithms. The simulation results in Section V will verify that our proposed techniques fall within a desirably small margin to the theoretical bounds on the estimator performance.

IV. Estimation Procedure

This section provides a detailed description of our proposed block-based estimation algorithm. Subsection IV-A covers the details of the block structure of the training signal, which is required in Subsection IV-B to extract information about the Doppler shift at the receiver. The Doppler information is then used in Subsection IV-C to compute an estimate of the channel coefficients. We close this section by providing an overview of the estimation steps in Algorithm 2.

A. Block-based processing

Our proposed algorithm demands that the training sequence consists of M repetitions of an arbitrary training pulse $\mathbf{x}$ and a cyclic prefix $\mathbf{x}_{CP}$. Specifically, in the notation of Section II, the training sequence is structured as

$$\mathbf{s} = [\ \underbrace{\mathbf{x}_{CP}^\mathsf{T}}_{\text{cyclic prefix}} \ |\ \underbrace{\mathbf{x}^\mathsf{T} \quad \mathbf{x}^\mathsf{T} \quad \cdots \quad \mathbf{x}^\mathsf{T}}_{M \text{ training pulses}}\ ]^\mathsf{T}, \tag{17}$$

where the cyclic prefix samples $\mathbf{x}_{CP} = [x[K-L+1] \ \cdots \ x[K-1]]^\mathsf{T} \in \mathbb{C}^{L-1\times 1}$ consist of a block of $L-1$ data symbols rotated cyclically [11]. In order to mitigate the effects of inter-symbol interference (ISI) induced by the time dispersion due to the L channel taps of $\mathbf{h}$, the receiver processes only the samples in the time interval $k \in [L, KM+L-1]$ and discards the rest. This results in a modified input/output model, which can be expressed in terms of the cyclic convolution of concatenation of M training pulses and the channel vector $\mathbf{h}$ [11]. In addition to the cyclic prefix removal operation, the receiver also reorders the M segments of the received signal corresponding to the M transmitted training pulses into columns of a matrix of received samples denoted as $\mathbf{Y} \in \mathbb{C}^{K\times M}$. The process of cyclic prefix removal and sample reordering is visualized in Figure 2. The goal of this block processing at the receiver is to write an expression for the individual columns of $\mathbf{Y}$ in terms of the cyclic convolution of one sequence pulse $\mathbf{x}$ with the channel vector $\mathbf{h}$, which is given as

$$\mathbf{y}_\ell = \sqrt{\rho}\,\widetilde{\mathbf{V}}_\ell\widetilde{\mathbf{X}}\mathbf{h} + \mathbf{n}_\ell, \quad \ell \in 0,\ \ldots,\ M-1. \tag{18}$$

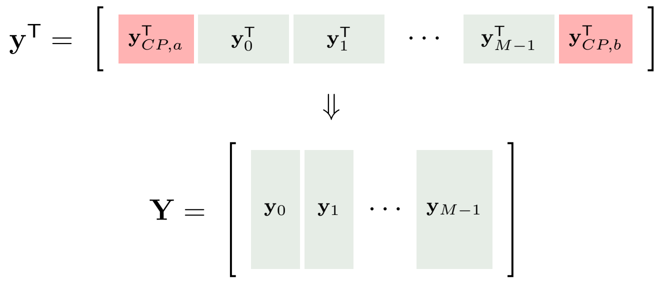


Fig. 2: The receiver drops the cyclic prefix components (red), extracts M receive pulses (green, K samples each), and re-orders them into the matrix of received pulses $\mathbf{Y}$.

Here, the additive noise vector $\mathbf{n}_\ell$ is drawn from the complex Gaussian distribution with zero mean and covariance matrix $\widetilde{\mathbf{C}}$, which is a truncated version of $\mathbf{C}$ from 6. The diagonal matrix $\widetilde{\mathbf{V}}_\ell \in \mathbb{C}^{K\times K}$ accounts for the Doppler shift of the ℓ-th block and is given by

$$\widetilde{\mathbf{V}}_\ell = e^{j\ell K\alpha}\mathrm{diag}\left(\begin{bmatrix} e^{j(L-1)\alpha} & e^{jL\alpha} & \cdots & e^{j(K+L-2)\alpha}\end{bmatrix}\right) \tag{19}$$

Furthermore, $\widetilde{\mathbf{X}} \in \mathbb{C}^{K\times L}$ is the cyclic convolution matrix obtained from the training pulse $\mathbf{x}$, truncated to its first L columns. The $K \times K$ cyclic convolution matrix obtained from a vector $\mathbf{x} \in \mathbb{C}^{K\times 1}$ is given by the matrix of cyclic shifts of $\mathbf{x}$ and is defined as

$$\mathbf{X} = \begin{bmatrix} x[0] & x[K-1] & \cdots & x[1] \\ x[1] & x[0] & \cdots & x[2] \\ x[2] & x[1] & \cdots & \vdots \\ \vdots & x[2] & \cdots & x[K-2] \\ x[K-2] & \vdots & \vdots & x[K-1] \\ x[K-1] & x[K-2] & \cdots & x[0] \end{bmatrix}. \tag{20}$$

Upon further inspection of (19), we note that the effects of the Doppler shift α can be separated into the Doppler shift internal to each block and the frequency offset between the blocks. Specifically, if we let $\widetilde{\mathbf{V}}_0$ represent the Doppler shift on each block ($\widetilde{\mathbf{V}}_\ell$ from (19) with $\ell = 0$) and define the inter-block Doppler offset vector $\mathbf{d}^*$ as

$$\mathbf{d}^* = \begin{bmatrix} 1 & e^{jK\alpha} & \cdots & e^{j(M-1)K\alpha}\end{bmatrix}, \tag{21}$$

we can write the expression for $\mathbf{Y}$ as

$$\mathbf{Y} = \sqrt{\rho}\,\widetilde{\mathbf{V}}_0\widetilde{\mathbf{X}}\mathbf{h}\mathbf{d}^* + \mathbf{N}, \tag{22}$$

where $\widetilde{\mathbf{X}}$ is defined as above and $\mathbf{N} = [\mathbf{n}_0 \cdots \mathbf{n}_{M-1}]$ is the matrix of additive noise vectors. Writing the input/output model like (22) is a desirable step, since it lets us break the estimation algorithm into two distinct parts: The first step computes an estimate of the Doppler frequency, denoted $\widehat{\alpha}$, which is then used to cancel out the effects of $\mathbf{d}^*$ and $\widetilde{\mathbf{V}}_0$, effectively turning the estimation of the channel coefficients into a straightforward linear Gaussian estimation problem.

B. Doppler Estimation

This estimation step directly exploits the block structure of the training sequence. We observe that at high SNRs, i.e., for large ρ, the received signal matrix $\mathbf{Y}$ can be approximated as the rank-1 outer product of the vectors $\sqrt{\rho}\,\widetilde{\mathbf{V}}_0\widetilde{\mathbf{X}}\mathbf{h}$ and $\mathbf{d}^*$. It is a well known fact that the singular value decomposition (SVD) can be used to construct low-rank approximations to matrices of any dimension [27]. In order to extract an estimate of $\mathbf{d}^*$ from (22), we can therefore use the SVD of $\mathbf{Y}$, defined as

$$\begin{aligned}\mathbf{Y} &= \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^* \\ &= \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_K \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_K \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_M \end{bmatrix}^* \end{aligned} \tag{23}$$

If $\mathbf{Y}$ were truly rank one, we would have $\sigma_2 = \sigma_3 \ldots = \sigma_K = 0$ and could therefore write the SVD as the outer product $\mathbf{Y} = \sigma_1\mathbf{u}_1\mathbf{v}_1^*$. Comparing with the model given in (22), it becomes clear that in high SNRs, $\mathbf{d}^*$ must be some scaled version of the vector $\mathbf{v}_1^*$, which is often referred to as the dominant right singular vector. This observation is the key to the Doppler estimation step, and we define the estimate of the Doppler offset vector, denoted as $\widehat{\mathbf{d}}$, as

$$\widehat{\mathbf{d}} = \sqrt{M}\mathbf{v}_1, \tag{24}$$

where the scale factor $\sqrt{M}$ compensates for the fact that $\mathbf{v}_1$ is usually obtained with unit norm. The simulation results from Section V will show that this estimator provides acceptable results even in low-SNR regimes, where $\mathbf{Y}$ is unlikely to be of rank one due to noise.

Since the estimator for $\mathbf{d}$ only requires the dominant right singular vector of $\mathbf{Y}$, computing the full SVD of $\mathbf{Y}$ can be a waste of computational resources. Fortunately, there exist well-known iterative algorithms in numerical linear algebra to compute the dominant eigenvectors and singular vectors directly, with the simplest one being the power method [28]. The power method belongs to a larger class of general eigenvalue or "power" iterations, which have been applied to problems in the space of multiple-input, multiple-output (MIMO) communication systems in recent works [29,30]. The basic idea behind the power method is that the repeated multiplication of a randomly selected vector $\mathbf{x}$ with a matrix $\mathbf{A}$ converges to a scalar multiple of the dominant eigenvector of $\mathbf{A}$. Normalization between the iterations of the power method produces the unit norm dominant eigenvector of $\mathbf{A}$. The algorithm is defined as

Algorithm 1 Power Method

Input: A diagonalizable matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$
Let $\mathbf{q}^0$ be a randomly chosen unit vector $\in \mathbb{C}^n$
for $k = 1, 2, \ldots$ **do**
  $\mathbf{z}^{(k)} = \mathbf{A}\mathbf{q}^{(k-1)}$
  $\mathbf{q}^{(k)} = \mathbf{z}^{(k)}/\|\mathbf{z}^{(k)}\|_2$
end for

In Lemma 2, we briefly show the convergence of $\mathbf{q}^{(k)}$ to the dominant eigenvector of $\mathbf{A}$.

Lemma 2. *The power method given in Algorithm 1 converges to a scalar multiple of the dominant eigenvector of* $\mathbf{A}$.

Proof: Without loss of generality, we suppose that the eigenvalues of $\mathbf{A}$ are ordered as

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \cdots \geq |\lambda_n|, \tag{25}$$

where we denote λ_1 as the dominant eigenvalue and the corresponding eigenvector $\mathbf{x_1}$ as the dominant eigenvector. Then, since $\mathbf{A}$ is diagonalizable, we can write $\mathbf{q}^0$ as linear combination of the eigenvectors of $\mathbf{A}$, i.e.,

$$\mathbf{q}^0 = \sum_{i=1}^{n} q_i \mathbf{x}_i. \tag{26}$$

We can thus write $\mathbf{A}^k\mathbf{q}^{(0)}$ as

$$\mathbf{A}^k\mathbf{q}^{(0)} = q_1\lambda_1^k \left(\mathbf{x}_1 + \sum_{i=2}^{n} \frac{q_i}{q_1}\left(\frac{\lambda_i}{\lambda_1}\right)^k \mathbf{x}_i\right) \tag{27}$$

It is straightforward to show that $\mathbf{q}^{(k)}$ is a scalar multiple of $\mathbf{A}^k\mathbf{q}^{(0)}$. Furthermore, since the ratio $(\lambda_i/\lambda_1)^k$ approaches zero for all $i \neq 1$, all components of $\mathbf{q}^{(k)}$ corresponding to eigenvectors other than $\mathbf{x}_1$ vanish as $k \to \infty$. ■

Since the set of right singular vectors of $\mathbf{Y}$ is equal to the set of eigenvectors of $\mathbf{Y}^*\mathbf{Y}$, we can use the power method to compute the dominant right singular vector $\mathbf{v}_1$, from which the estimator $\widehat{\mathbf{d}}$ can be obtained. Note that due to the block structure of $\mathbf{Y}$ we can safely assume that (25) holds for moderate to high SNRs, since the signal component of $\mathbf{Y}$ is rank 1 and all higher-rank components are due to additive noise. In practice, this means that a sufficiently accurate estimate of the dominant right singular vector can be obtained with a moderate number of iterations.

Recall that in order to mitigate the effects of the Doppler shift in the model from (22), the receiver must ideally cancel both the block-based Doppler offset due to $\mathbf{d}^*$ as well as the intra-block Doppler shift due to $\widetilde{\mathbf{V}}_0$. With perfect knowledge of α, the receiver could simply construct the inverse of $\widetilde{\mathbf{V}}_0$ for this. However, since α is not known, the receiver has to extract an estimate of it from the previously computed estimate of the Doppler offset vector $\mathbf{d}^*$. This estimate can be constructed using a simple correlation operation, and is given by

$$\widehat{\alpha} = \arg\max_{\beta \in \mathbb{R}^+} \left|\sum_{\ell=1}^{M} e^{j(\ell-1)K\beta}\widehat{d}_\ell\right|, \tag{28}$$

where $\widehat{d}_\ell$ is the ℓ-th element of $\widehat{\mathbf{d}}$. To minimize computational cost, the receiver could have a precomputed set of vectors to correlate against stored, resulting in a single matrix multiplication per estimation.

C. Channel estimation

To arrive at the expression for the estimates of the L channel taps, we observe that with perfect knowledge of α, the channel estimation problem would reduce to the well known linear Gaussian estimation problem, for which we could write the estimator as [25]

$$\widehat{\mathbf{h}}_\alpha = \frac{1}{M\sqrt{\rho}}\left(\widetilde{\mathbf{X}}^*\widetilde{\mathbf{V}}_0^*\widetilde{\mathbf{C}}^{-1}\widetilde{\mathbf{V}}_0\widetilde{\mathbf{X}}\right)^{-1}\widetilde{\mathbf{X}}^*\widetilde{\mathbf{V}}_0^*\widetilde{\mathbf{C}}^{-1}\mathbf{Y}\mathbf{d}. \tag{29}$$

Since in our system setup the Doppler shift α is unknown, the receiver has to resort to using the estimates of $\widetilde{\mathbf{V}}_0$ and $\mathbf{d}$, which are obtained as described in the previous subsection. More specifically, the receiver computes the estimate as

$$\widehat{\mathbf{h}} = \frac{1}{M\sqrt{\rho}}\left(\widetilde{\mathbf{X}}^*\widehat{\mathbf{V}}_0^*\widetilde{\mathbf{C}}^{-1}\widehat{\mathbf{V}}_0\widetilde{\mathbf{X}}\right)^{-1}\widetilde{\mathbf{X}}^*\widehat{\mathbf{V}}_0^*\widetilde{\mathbf{C}}^{-1}\mathbf{Y}\widehat{\mathbf{d}}, \tag{30}$$

where $\widehat{\mathbf{d}}$ was derived in the previous section and $\widehat{\mathbf{V}}_0$ is constructed using the Doppler shift estimate $\widehat{\alpha}$ as

$$\widehat{\mathbf{V}}_0 = \operatorname{diag}\left(\left[e^{j(L-1)\widehat{\alpha}}\ e^{jL\widehat{\alpha}}\ \cdots\ e^{j(K+L-2)\widehat{\alpha}}\right]\right). \tag{31}$$

To conclude our discussion on the specifics of our proposed estimation algorithm, the steps outlined in 2 provides a short summary of the necessary steps at the receiver with our proposed technique.

Algorithm 2 Doubly dispersive channel estimation (Summary)

Input: Received estimation sequence $\mathbf{y}$	▷ (6)
1. Drop cyclic prefix and reshape	▷ (22), Fig. 2
2. Estimate Doppler shift vector $\mathbf{d}$	▷ (24), Alg. 1
3. Estimate Doppler frequency using correlation	▷ (28)
4. Compute channel estimate $\widehat{\mathbf{h}}$	▷ (30)

D. Simplified channel estimation using CAZAC sequences

Until now, the discussion of our channel estimation algorithm has remained independent of the choice of base sequence for our training pulses, denoted earlier as $\mathbf{x}$. While there are many potential choices of pseudo-random (PN) training sequences, we have exclusively considered the class of Zadoff-Chu sequences [31] in this work. Zadoff-Chu sequences belong to the class of constant-amplitude, zero autocorrelation (CAZAC) waveforms and have most recently found use in various applications in the LTE physical layer [32]. The k-th symbol of a Zadoff-Chu sequence of length K is given by

$$x_k = \begin{cases} e^{-j\pi uk(k+2q)/K} & \text{if } K \text{ is even} \\ e^{-j\pi uk(k+1+2q)/K} & \text{if } K \text{ is odd,} \end{cases} \tag{32}$$

where the parameter q is any positive integer or zero and the parameter u is some positive integer relatively prime to K. For the remainder of this text, we used the values $q = 0$ and $u = 1$. Zadoff-Chu sequences have various beneficial properties, the most interesting for this application being the fact that cyclic shifts of the same sequence are orthogonal. Recall that our estimator, given in (30), utilizes the circular convolution matrix obtained from a training pulse, denoted as $\widetilde{\mathbf{X}}$. If the training pulse $\mathbf{x}$ is a Zadoff-Chu sequence, it can be shown that the circular convolution matrix $\widetilde{\mathbf{X}}$ satisfies

$$\widetilde{\mathbf{X}}^*\widetilde{\mathbf{X}} = \widetilde{\mathbf{X}}\widetilde{\mathbf{X}}^* = K\mathbf{I}, \tag{33}$$

which greatly simplifies the computation of (30), since it is easily shown that under white noise this expression simplifies to

$$\widehat{\mathbf{h}} = \frac{1}{M\sqrt{\rho}}\widetilde{\mathbf{X}}^{*}\widehat{\mathbf{V}}_{0}^{*}\mathbf{Y}\widehat{\mathbf{d}}. \tag{34}$$

The switch from (30) to (34) obviates the need to compute expensive high-dimensional matrix inverses on-the-fly in the white noise regime, resulting in a substantially decreased computational load. Given these favorable properties, the remainder of this paper will assume that every training pulse is a Zadoff-Chu sequence of appropriate length. An added benefit to using Zadoff-Chu sequences in practical systems are their constant-cross-correlation properties [31], which are leveraged in the LTE physical random access channel (PRACH) to resolve collisions using initial access. In multi-user systems employing our proposed estimation technique, different users could be assigned different root parameters u and q to minimize conflicts during channel estimation.

V. Numerical Studies

In this section, we evaluate the performance of our proposed algorithm using Monte-Carlo methods and synthetic data. Our quantities of interest for all simulations are the sample mean squared error of the channel estimates $\widehat{\mathbf{h}}$ and the sample error variance of the Doppler frequency estimate $\widehat{\alpha}$. In all results of this section, the L channel taps were drawn independently from a complex Gaussian distribution, i.e., $h[\ell] \sim \mathcal{N}(0,\mathbf{I}), \ell \in 0,\ldots,L-1$. In all simulations, the length of the cyclic prefix was held constant at $L-1$ symbols, the minimum required length to write the input-output model using circular convolution. Furthermore, for each Monte-Carlo iteration a channel realization, noise realization, and a initial seed vector for the power method iteration was generated independently.

A. Comparison to theoretical optimum

Figures 3 and 4 study the performance of our proposed estimator in relation to the theoretical optimum derived in Section III. Specifically, in Fig. 3 we examine the difference between the sample mean squared error of the channel estimate and the CRLB given in (14) for signal-to-noise ratios ranging from -10 dB to 20 dB. Correspondingly, Fig. 4 examines the difference between the sample variance of the estimation error of $\widehat{\alpha}$ and the CRLB given in (13) for the same range of SNRs. We compare these performance metrics in both figures for three different cases.

1) **White Noise.** The curves labeled “White Noise” present results for uncorrelated white complex Gaussian Noise, i.e., $\mathbf{C} = \mathbf{I}$. As noted earlier, the training pulses consist of Zadoff-Chu sequences, with the sequence length K fixed to 127. The equation used for the channel estimation step is thus (34).
2) **Correlated Noise.** The curves labeled “Correlated Noise” examine the performance of our estimator for correlated noise environments. To simulate such an environment, we constructed a synthetic noise correlation matrix

$$\mathbf{C} = \begin{bmatrix} 1 & \beta & \beta^2 & \cdots & \beta^N \\ \beta & 1 & \beta & \cdots & \beta^{N-1} \\ \beta^2 & \beta & 1 & \cdots & \beta^{N-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta^N & \beta^{N-1} & \beta^{N-2} & \cdots & 1 \end{bmatrix}, \tag{35}$$

where the specific correlation parameter β used for producing the curves in Figs. 3 and 4 was $\beta = 0.8$. Due to the structure of $\mathbf{C}$, the channel estimation step uses (30), including the costly matrix inverses.
3) **Grid Search.** We provide an alternative channel and Doppler estimation algorithm as a point of reference for our proposed algorithms. This estimator, denoted as “Grid Search” in Figures 3 and 4, computes an approximate maximum-likelihood estimate of the channel taps and the frequency offset α by performing the minimization

$$\widehat{\alpha} = \underset{\beta \in G}{\operatorname{argmin}} \quad \left\|\mathbf{y} - \mathbf{y}_{\beta}\right\|^2, \tag{36}$$

$$\widehat{\mathbf{h}} = \widehat{\mathbf{h}}_{\widehat{\alpha}}, \tag{37}$$

where

$$\mathbf{y}_{\beta} = \sqrt{\rho}\,\mathbf{V}_{\beta}\mathbf{S}\mathbf{h}_{\beta} \tag{38}$$

$$\widehat{\mathbf{h}}_{\beta} = \frac{1}{\sqrt{\rho}}\left(\mathbf{S}^{*}\mathbf{V}_{\beta}^{*}\mathbf{C}^{-1}\mathbf{V}_{\beta}\mathbf{S}\right)^{-1}\mathbf{S}^{*}\mathbf{V}_{\beta}^{*}\mathbf{C}^{-1}\mathbf{y} \tag{39}$$

over a pre-defined set of grid points G for the observation vector $\mathbf{y}$ (6). We note that in (36), $\mathbf{V}_{\beta} = \operatorname{diag}\left(\begin{bmatrix}1 & e^{j\beta} & \cdots & e^{j\beta(N-1)}\end{bmatrix}\right)$. In order to obtain a fair comparison, the set of grid points is chosen to be equivalent to the set of points over which our proposed estimators perform the Doppler correlation step (28), rendering this search prohibitively expensive from a computational perspective. However, this estimator helps to provide a good baseline for the performance of our proposed algorithms due to its near-optimal performance.

We show the error curves in Figures 3 and 4 for three different values of the Doppler frequency, each presented using a different line style. More specifically, the mapping is described in Table I. Although our studies suggest that the impact of the specific Doppler frequency on the estimation error is minimal, these three values were chosen to provide examples of realistic conditions and reasonable values for reference. For example, the aforementioned cruising speed of 250 m/s combined with a transmission at the center frequency $f_c = 120$ MHz (reserved for aeronautical mobile radio according to [33]) results in a bulk Doppler shift of approximately 100 Hz. The values of 50 Hz and 0 Hz were studied to provide additional points for reference. In all of these studies, one training sequence pulse consisted of a Zadoff-Chu sequence of length $K = 127$, while the entire training sequence consisted of $M = 5$ pulses with a cyclic prefix. At each iteration, $L = 16$ channel taps were generated.

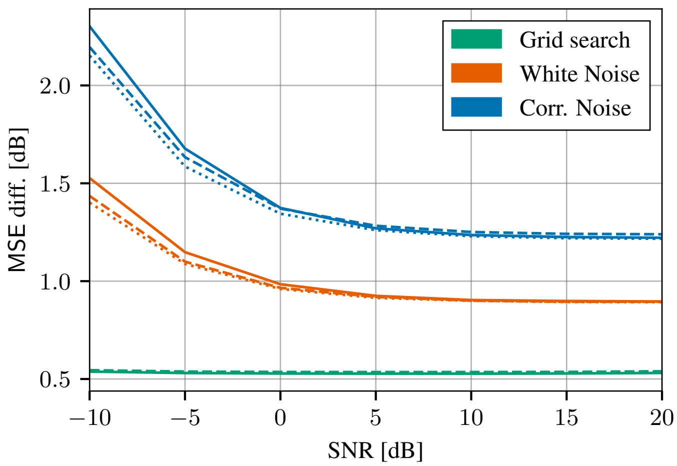


Fig. 3: Difference between sample MSE of the channel estimate $\widehat{\mathbf{h}}$ and the CRLB (14) vs. SNR of our proposed estimator for white noise and colored noise environments. Grid search based estimator pictured for reference. $K = 127$ samples, $M = 5$ pulses, $L = 16$ channel taps.

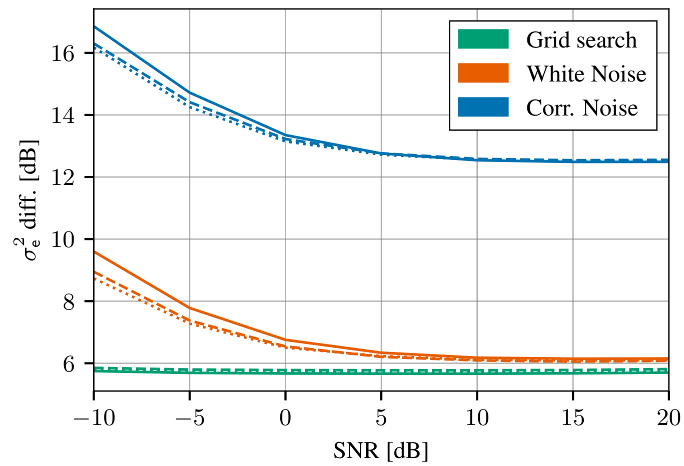


Fig. 4: Difference between sample error variance of the frequency offset estimate $\widehat{\alpha}$ and the CRLB (13) vs. SNR of our proposed estimator for white noise and colored noise environments. Grid search based estimator pictured for reference. $K = 127$ samples, $M = 5$ pulses, $L = 16$ channel taps.

$\mathbf{f_d}$	**Style**	
0 Hz	dotted	····
50 Hz	dashed	- - -
100 Hz	solid	——

TABLE I: Doppler shift to line style mapping for Figs. 3– 6

The results in Figure 3 and 4 indicate that our proposed algorithm produces desirable results largely independent of the magnitude of the Doppler shift present, especially for medium to high SNRs. Note that due to the loss of information from performing cyclic prefix removal at the receiver, our proposed estimator never fully attains the CRLB, even for high SNRs. Our algorithm performs especially well for the white noise scenario, quickly approaching a sub-1 dB difference from the theoretical optimum and a difference to the computationally expensive minimum-distance grid search estimator of less than 0.5 dB. The constant gap for the colored noise case is explained by the fact that the estimator from (30) disregards any correlation across blocks that could be present in the noise samples.

The contrast between Figure 3 and Figure 5 illustrates the importance of the intra-block Doppler correction for this algorithm. To generate the results in Figure 5, we disabled the intra-block Doppler correction from (28), resulting in $\widehat{\mathbf{V}}_0 = \mathbf{I}$ at all iterations. The severe performance penalty of only correcting for the per-block Doppler shifts is evident in the increasing loss in estimator performance as a function of increasing SNR. We note that as expected, this performance loss does not occur for the 0 Hz case.

B. Performance under model mismatch

Figure 6 studies the loss in performance of our proposed estimator when the bulk Doppler assumption introduced in Section II does not apply and the Doppler shift varies across the multipath components of the channel. In this case, the channel model from (4) becomes

$$h_k[\ell] = e^{j\alpha[\ell]k} h[\ell], \tag{40}$$

where $\alpha[\ell] = 2\pi T_s f_d[\ell]$ and $f_d[\ell]$ is sampled from a probability distribution for each channel tap. We furthermore restrict ourselves to the white noise case for this study and re-use the parameters $K = 127$, $L = 16$, and $M = 5$ from the previous subsection.

Figure 6 plots these trials for normally distributed $f_d[\ell] \sim \mathcal{N}(\overline{f}_d, \sigma_f^2)$ for the three different mean values $\overline{f}_d$ given in Table I as a function of the standard deviation σ_f. Each curve is obtained by using our proposed algorithms (designed for the bulk Doppler-only model) on the channel

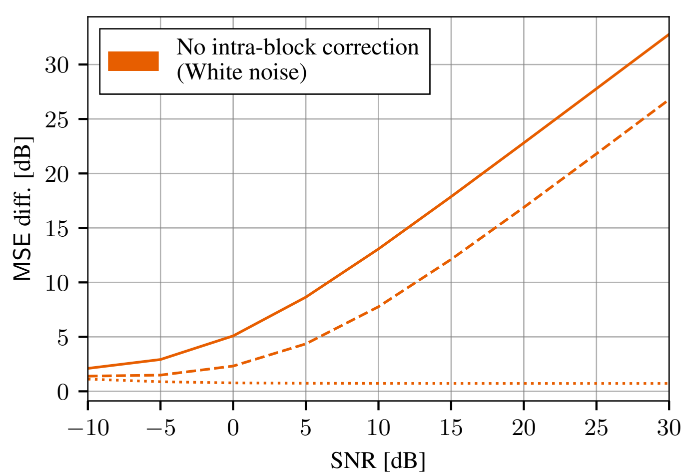


Fig. 5: Difference between sample MSE of the channel estimate $\widehat{\mathbf{h}}$ and the CRLB (14) vs. SNR of our proposed estimator for white noise without intra-block Doppler correction from (28). $K = 127$ samples, $M = 5$ pulses, $L = 16$ channel taps.

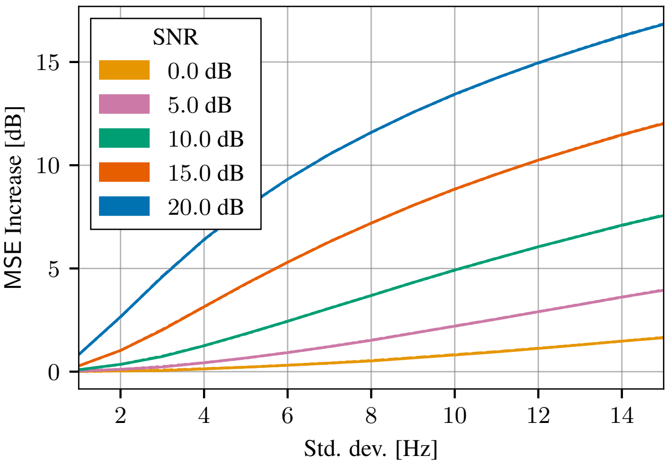


Fig. 6: Increase in mean squared error relative to the bulk-only Doppler shift model when using our proposed estimator on the mismatched random Doppler shift model form (40). “MSE Loss” denotes the difference in mean square error performance between the matched and mismatched scenario. $K = 127$ samples, $M = 5$ pulses, $L = 16$ channel taps. Pictured for various SNR regimes.

model from (40). The plot shows the difference in mean squared error between this mismatched scenario (Estimator assumes equal Doppler, but channel model generates random Doppler per multipath component) and the matched scenario from in Fig. 3 (Estimator assumes equal Doppler and channel model generates equal Doppler per multipath component) for a range of SNR regimes. As expected, the performance of our estimator degrades with increasing standard deviation of the Doppler frequency. The simulation furthermore indicates that the performance in high-SNR regimes is limited by the model mismatch, whereas it is limited by the noise power in the low to medium SNR regime. This implies that our estimators remain attractive options in the low-to medium SNR regime, especially in the 0-10 dB range. We note that the choice of normally distributed Doppler shifts is applicable especially in the air-to-ground scenario with automobiles as scatterers due to the widely accepted assumption of normally distributed velocities [34]–[37].

C. Pulse length and repetition count

The plot in Figure 7a) shows the channel estimation MSE as the length of a training pulse K increases. The results were obtained assuming white noise, i.e., $\beta = 0$, $M = 5$ training pulse repetitions, $L = 16$ channel taps, and transmit powers ranging from -10 dB to 30 dB. The training sequence pulses were again chosen to be Zadoff-Chu sequences with the length of one pulse varying from $K = 29$ to $K = 331$. For these simulations, the Doppler shift was fixed to $f_d = 50$ Hz. As we can see, performance gains can be achieved when increasing the pulse length, but the magnitude of the performance gains decreases with the length. Figure 7b) gives a similar comparison, however, whereas the number of pulses M was fixed in Figure 7a), we now fix the length of one training pulse to $K = 127$ symbols and vary the number of pulses M between $M = 2$ and $M = 20$. In this case, the MSE behaves similar as it does with increasing pulse length.

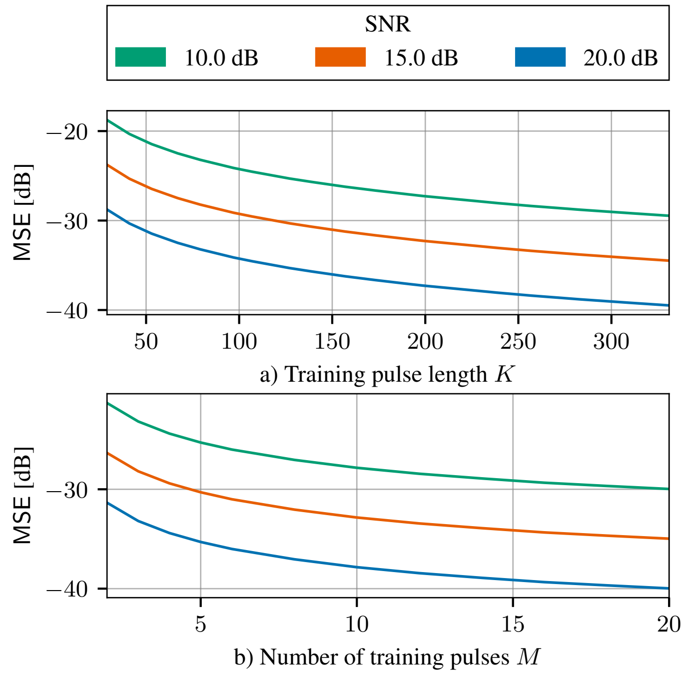


Fig. 7: Impact of sounding signal length factors on MSE for different SNRs.
a) Sample mean squared error vs. training pulse length K for fixed pulse repetition $M = 5$. White noise and $L = 16$ channel taps.
b) Sample mean squared error vs. pulse repetition M for fixed training pulse length K. White noise and $L = 16$ channel taps.

VI. Concluding Remarks

This paper studied the problem of jointly estimating the bulk Doppler shift and the channel taps of a doubly dispersive aerial channel. Work in this area is of importance and significance because channel estimation is an essential part of any modern wireless communication system, including aeronautical systems. This paper and other current and future work focusing on aeronautical wireless communication systems is in line with the recent surge in popularity of airborne platforms in consumer electronics, public safety networks, and defense applications. Our contribution to the field concerns a special class of channels in which the Doppler spread is dominated by the bulk Doppler shift due to vehicular motion. For this class of channels, we presented a simple channel model incorporating time and frequency dispersion before deriving the theoretical bounds of the resulting channel estimation problem. In addition to the theoretical analysis, we developed a novel channel estimation algorithm by combining traditional pulse-repetition techniques with simple and computationally efficient processing techniques. Our numerical studies indicate that our proposed techniques perform comparatively well to computationally expensive brute-force search methods.

The main contribution of this paper lies in our combination of prior work in estimation theory and signal processing into an efficient algorithm for our model. However, aerial channels like the one considered in this paper are still sources of many open problems. More specifically, future work in this area includes the extension of our model to multi-antenna systems and multi-user scenarios.

References

[1] S. A. R. Naqvi, S. A. Hassan, H. Pervaiz, and Q. Ni, "Drone-aided communication as a key enabler for 5G and resilient public safety networks," *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 36–42, Jan 2018.

[2] Google, "Project Loon." [Online]. Available: https://www.google.com/loon/

[3] E. Haas, "Aeronautical channel modeling," *IEEE Trans. Veh. Technol.*, vol. 51, no. 2, pp. 254–264, Mar. 2002.

[4] G. Matz and F. Hlawatsch, "Fundamentals of time-varying communication channels," in *Wireless Communications Over Rapidly Time-Varying Channels*, F. Hlawatsch and G. Matz, Eds. Oxford: Academic Press, 2011, pp. 1 – 63.

[5] G. Matz, H. Bolcskei, and F. Hlawatsch, "Time-frequency foundations of communications: Concepts and tools," *IEEE Signal Processing Mag.*, vol. 30, no. 6, pp. 87–96, 2013.

[6] G. Matz, "On non-WSSUS wireless fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2465–2478, 2005.

[7] A. Kannu and P. Schniter, "Design and analysis of MMSE pilot-aided cyclic-prefixed block transmissions for doubly selective channels," *IEEE Trans. Signal Processing*, vol. 56, no. 3, pp. 1148–1160, 2008.

[8] P. Schniter, "Low-complexity equalization of OFDM in doubly selective channels," *IEEE Trans. Signal Processing*, vol. 52, no. 4, pp. 1002–1011, 2004.

[9] G. Leus, S. Zhou, and G. B. Giannakis, "Orthogonal multiple access over time-and frequency-selective channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 8, pp. 1942–1950, 2003.

[10] X. Ma, G. B. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly selective wireless fading channels," *IEEE Trans. Signal Processing*, vol. 51, no. 5, pp. 1351–1366, 2003.

[11] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, UK ; New York: Cambridge University Press, Jul. 2005.

[12] T. Pollet, M. V. Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 191–193, Feb. 1995.

[13] H. Bolcskei, "Blind estimation of symbol timing and carrier frequency offset in wireless OFDM systems," *IEEE Trans. Commun.*, vol. 49, no. 6, pp. 988–999, Jun. 2001.

[14] Y. Yao and G. B. Giannakis, "Blind carrier frequency offset estimation in SISO, MIMO, and multiuser OFDM systems," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 173–183, Jan. 2005.

[15] P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*, vol. 42, no. 10, pp. 2908–2914, Oct. 1994.

[16] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, no. 12, pp. 1613–1621, Dec. 1997.

[17] M. Morelli and U. Mengali, "An improved frequency offset estimator for OFDM applications," in *Communication Theory Mini-Conference, 1999*, Jun. 1999, pp. 106–109.

[18] D. Huang and K. B. Letaief, "Carrier frequency offset estimation for OFDM systems using subcarriers," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 813–823, May 2006.

[19] J. J. v. d. Beek, M. Sandell, and P. O. Borjesson, "ML estimation of time and frequency offset in OFDM systems," *IEEE Trans. Signal Processing*, vol. 45, no. 7, pp. 1800–1805, Jul. 1997.

[20] N. Lashkarian and S. Kiaei, "Class of cyclic-based estimators for frequency-offset estimation of OFDM systems," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2139–2149, Dec. 2000.

[21] M.-H. Hsieh and C.-H. Wei, "A low-complexity frame synchronization and frequency offset compensation scheme for OFDM systems over fading channels," *IEEE Trans. Veh. Technol.*, vol. 48, no. 5, pp. 1596–1609, Sep. 1999.

[22] M. Erturk, J. Haque, W. Moreno, and H. Arslan, "Doppler mitigation in OFDM-based aeronautical communications," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 1, pp. 120–129, Jan. 2014.

[23] D. R. Ellis and E. Seckel, "Flying qualities of small general aviation airplanes. part 1. the influence of dutch-roll frequency, dutch-roll damping, and dihedral effect," Princeton University, Tech. Rep., 1969.

[24] D. Scherber, P. Bidigare, R. ODonnell, M. Rebholz, M. Oyarzun, C. Obranovich, W. Kulp, D. Chang, and D. R. B. III, "Coherent distributed techniques for tactical radio networks: Enabling long range communications with reduced size, weight, power and cost," in *MILCOM 2013 - 2013 IEEE Military Communications Conference*, Nov 2013, pp. 655–660.

[25] S. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, 1st ed. Englewood Cliffs, N.J: Prentice Hall, Apr. 1993.

[26] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge: Cambridge University Press, Oct. 2012.

[27] G. Strang, *Linear Algebra and Its Applications, 4th Edition*, 4th ed. Belmont, CA: Cengage Learning, Jul. 2005.

[28] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 4th ed. Baltimore: Johns Hopkins University Press, Dec. 2012.

[29] D. Ogbe, D. J. Love, and V. Raghavan, "Noisy beam alignment techniques for reciprocal MIMO channels," *IEEE Trans. Signal Processing*, vol. 65, no. 19, pp. 5092–5107, Oct 2017.

[30] H. Ghauch, T. Kim, M. Skoglund, and M. Bengtsson, "Subspace estimation and decomposition in large millimeter-wave MIMO systems," *IEEE Journ. Sel. Topics in Sig. Proc.*, vol. 10, no. 3, pp. 528–542, Apr. 2016.

[31] D. Chu, "Polyphase codes with good periodic correlation properties (Corresp.)," *IEEE Trans. Inform. Theory*, vol. 18, no. 4, pp. 531–532, Jul. 1972.

[32] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation," 3rd Generation Partnership Project (3GPP), TS 36.211, Jan. 2016.

[33] Federal Communications Commission, Office of Engineering and Technology, "FCC Online Table of Frequency Allocations," June 2018.

[34] S. Yousefi, E. Altman, R. El-Azouzi, and M. Fathy, "Analytical model for connectivity in vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 6, pp. 3341–3356, Nov 2008.

[35] Z. Zhang, G. Mao, and B. D. O. Anderson, "On the information propagation process in mobile vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 5, pp. 2314–2325, Jun 2011.

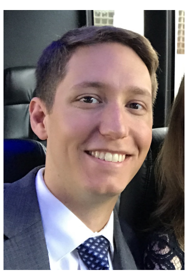
[36] ——, "Stochastic characterization of information propagation process in vehicular ad hoc networks," *IEEE Trans. Intell. Transport. Syst.*, vol. 15, no. 1, pp. 122–135, Feb 2014.

[37] R. Nagel, "The effect of vehicular distance distributions and mobility on vanet communications," in *2010 IEEE Intelligent Vehicles Symposium*, June 2010, pp. 1190–1194.

Dennis Ogbe (S'13) received the B.S. degree (with honors) in electrical engineering in 2014 from Tennessee Technological University in Cookeville, TN and is currently a PhD candidate at the School of Electrical and Computer Engineering at Purdue University in West Lafayette, IN. During the summer of 2016, he was an intern at Nokia Bell Labs in Arlington Heights, IL. His current research interests are in the areas of signal processing for wireless communications, multiple-input multiple-output (MIMO) communications, information theory, and collaborative spectrum sharing systems. Mr. Ogbe is an active member of Eta Kappa Nu.

David J. Love (S'98–M'05–SM'09–F'15) received the B.S. (with highest honors), M.S.E., and Ph.D. degrees in electrical engineering from the University of Texas at Austin in 2000, 2002, and 2004, respectively. Since 2004, he has been with the School of Electrical and Computer Engineering at Purdue University, where he is now the Nick Trbovich Professor of Electrical and Computer Engineering and leads the College of Engineering Preeminent Team on Efficient Spectrum Usage. He currently serves as a Senior Editor for IEEE Signal Processing Magazine and previously served as an Editor for the IEEE Transactions on Communications, an Associate Editor for the IEEE Transactions on Signal Processing, and a guest editor for special issues of the IEEE Journal on Selected Areas in Communications and the EURASIP Journal on Wireless Communications and Networking. He is a member of the Executive Committee for the National Spectrum Consortium. His research interests are in the design and analysis of broadband wireless communication systems, 5G wireless systems, multiple-input multiple-output (MIMO) communications, millimeter wave wireless, software defined radios and wireless networks, coding theory, and MIMO array processing. He has around 30 U.S. patent filings, 28 of which have issued. He currently co-advises a team in the DARPA Spectrum Collaboration Challenge (SC2) that finished in the top-ten in both the first and second phase events and previously co-advised the Purdue team that was a finalist in the DARPA Spectrum Challenge.

Dr. Love has been recognized as a Thomson Reuters Highly Cited Researcher (2014 and 2015). Along with his co-authors, he won best paper awards from the IEEE Communications Society (2016 IEEE Communications Society Stephen O. Rice Prize), the IEEE Signal Processing Society (2015 IEEE Signal Processing Society Best Paper Award), and the IEEE Vehicular Technology Society (2009 IEEE Transactions on Vehicular Technology Jack Neubauer Memorial Award). He has received multiple IEEE Global Communications Conference (Globecom) best paper awards. He was an invited participant to the 2011 NAE Frontiers of Engineering Education Symposium and 2016 EU-US NAE Frontiers of Engineering Symposium.



Matthew Rebholz (M'13) received his B.S. and M.S. degrees in Electrical Engineering from Virginia Tech in 2010 and 2013 respectively. Since 2011 he has worked as a Scientist at Raytheon BBN Technologies in Arlington, VA, and Raleigh, NC. While at BBN he has worked on distributed RF communications and beamforming, signal processing, and semantic technologies. His research interests include algorithm development for distributed time transfer and causal modeling.

T. Patrick Bidigare received the B.S.E.E. degree in computer engineering and the and the Ph.D. degree in mathematics from the University of Michigan, Ann Arbor, in 1992 and 1997 respectively. His doctoral work was in the area of algebraic combinatorics. He was with the Environmental Institute of Michigan (now General Dynamics) from 1991 to 2007 where he served as principal investigator on programs in areas of protection, radar responsive tagging and secure communications. Since 2007 he has been a technical director at BBN Technologies (now Raytheon) in Arlington, VA where he leads programs in distributed RF sensing and communications. His research interests include hardware and algorithm development for implementing distributed clock synchronization, network self-localization and distributed beamforming.