EESC6353: A Novel Technique for High-Accuracy Phase-Based Ranging with Bluetooth Low Energy (BLE)

Jayson Van Marter
Email: jayson.vanmarter@utdallas.edu

Josiah W. Smith
Email: josiah.smith@utdallas.edu

***Abstract*—In this project, we examine the work of Zand *et al.* [1] for high-accuracy ranging with Bluetooth Low Energy (BLE). First, a frequency-hopping phase-based ranging solution conforming to the BLE standard in the 2.4 GHz frequency band is presented. Next, we investigate the impact of crystal offset and phase-noise on the ranging accuracy. We extend the work in [1] by mathematically examining the error introduced by device mobility in BLE ranging and investigating a multi-path and Rayleigh fading channel model. Further, we introduce a novel residual learning technique for synthetically improving signal-to-noise ratio (SNR) before ranging via subspace decomposition methods (MUSIC). Our proposed deep neural network (DNN) approach reduces localization error on both AWGN fixed fading and AWGN Rayleigh fading channels in multi-path scenarios. Simulation results for the various techniques and scenarios are presented and discussed demonstrating the superior performance of our novel methods.**

I. Introduction

In recent years, accurate estimation of the distance between low energy devices has increased in interest for many applications with the emergence of the Internet-of-Things (IoT). In many IoT applications, this distance estimation, or ranging, is used for localization, security, privacy, etc. IoT devices often communicate wirelessly using Bluetooth Low Energy (BLE), although other technologies are also used. Ranging with BLE is becoming increasingly essential for many applications relying on accurate range information between low power devices.

To perform ranging, three common approaches are adopted: Received Signal Strength Indicator (RSSI)-based, time-based, and phase-based. This project examines the phase-based ranging solution, although extensive work on RSSI-based [2] and time-based [3]–[5] ranging has been investigated elsewhere. To perform ranging, the signal phase shift between the transmitter and receiver is leveraged to estimate the distance between them. In phase-based solutions, multi-path effects are mitigated by measuring the phase changes at two devices over multiple frequencies in a process known as Multi-Carrier Phase Difference (MCPD). In this project, we investigate a BLE multi-carrier approach for ranging and the impact of crystal offset and mobility.

Estimating the range from the MCPD channel frequency response (CFR) is a spectral estimation problem and common Fourier and subspace decomposition solutions are often applied with varying results. Subspace methods, e.g., MUlitple SIgnal Classification (MUSIC), are typically advantageous over Fourier-based approaches due to higher resolution. However, the performance of these approaches suffers in the presence of noise. Recently, data-driven denoising techniques have been of significant interest. Instead of relying on knowledge of the signal structure, deep neural networks (DNNs) can be trained to capture the signal structure [6]. Recent work posits that when the noisy signal consists of a highly structured signal in the presence of of unstructured noise, a deep network can more efficiently learn to remove the structured signal than directly remove the noise [7]. This concept is known as *residual learning* [8]. Jiang and Rangaswamy employ a DNN to denoise time-domain signals to improve subsequent spectral estimation using MUSIC [8]. We propose a novel residual learning technique employing a DNN by leveraging the structure of the two-way MCPD CFR. By denoising the CFR, the DNN offers improved signal/noise subspace separation yielding superior spectral estimation by MUSIC. The architecture and training methodology of the DNN is detailed and numerical results demonstrate an improvement in ranging accuracy.

The rest of this report is organized as follows. In Section II, we present the MCPD ranging technique for range estimation. Section III provides a mathematical investigation of the ranging error introduced by a crystal offset. In Section IV, we explore the impact of mobility on the ranging accuracy and we extend our analysis to include a multi-path Rayleigh fading channel in Section V. We present a novel approach to deep residual learning considering the two-way channel frequency response of a typical BLE ranging system in Section VI. Section VII overviews our analytical results followed finally by conclusions.

II. Phase-Based Ranging Technique

It is well known that the phase-shift produced by a pure Line-of-Sight (LOS) channel is a linear function of both frequency (f) and range (r) as

$$\phi(f, r) = -2\pi f r / c_0 \pmod{2\pi}, \tag{1}$$

where c_0 is the freespace speed of light. Thus, the slope of the phase across the frequency can be measured and used to estimate the range.

Phased-based BLE ranging using the MCPD principle employs two devices in distinct roles: initiator I and reflector

R. The initiator starts the ranging procedure and the reflector responds to the initiator. After a synchronization handshake between the initiator and reflector, the initiator transmits its Local Oscillator (LO) signal on the first channel (f_0). The reflector measures the phase of the signal and and responds with a constant tone, its LO, on the same channel for a phase measurement by the initiator. This process is repeated over each channel using K_f as the number of channels and Δ_f as the uniform channel spacing. Afterwards, the reflector provides the measurement results for all channels to the initiator for range estimation.

Thus, the phase at the reflector can be modeled as

$$\phi_R(f_k, r) = -2\pi f_k \left(\frac{r}{c_0} - \Delta_t \right) - \theta_k \pmod{2\pi}, \quad (2)$$

where Δ_t is the time difference between the initiator and reflector, f_k is the frequency of the k-th channel, and θ_k is the phase difference between the initiator and reflector LO signals which changes per channel k. After the reflector's transmission, the phase at the initiator is expressed as

$$\phi_I(f_k, r) = -2\pi f_k \left(\frac{r}{c_0} + \Delta_t \right) + \theta_k \pmod{2\pi}. \quad (3)$$

The two-way phase difference is calculated as the sum of the phase at I and R as

$$\begin{aligned} \phi_{2W}(f_k, r) &= \phi_R(f_k, r) + \phi_I(f_k, r) \\ &= -\frac{4\pi f_k}{c_0} r \pmod{2\pi}. \end{aligned} \quad (4)$$

Note that the time-offset (Δ_t) and phase difference ϕ are cancelled by the summation. By measuring the phase at multiple locations, the half-wavelength ambiguity in (4) can be combated. Defining $\hat{\Delta}_\phi[k]$ as the phase difference between the k-th and the $k-1$-th channel as

$$\begin{aligned} \hat{\Delta}_\phi[k] &= \phi_{2W}(f_k, r) - \phi_{2W}(f_{k-1}, r) \\ &= -\frac{4\pi \Delta_f}{c_0} r \pmod{2\pi}, \end{aligned} \quad (5)$$

the range, r, can be estimated on the first two channels (f_0 and f_1) as

$$\hat{r} = -\frac{c_0}{4\pi \Delta_f} \hat{\Delta}_\phi[1] \left(\text{mod } \frac{c_0}{2\Delta_f} \right). \quad (6)$$

The range-ambiguity from the estimator in (6) is $c_0/(2\Delta_f)$. Using the MCPD to leverage all K_f channels, the average phase difference can be computed by

$$\bar{\Delta}_\phi = \frac{1}{K_f - 1} \sum_{k=1}^{K_f - 1} \hat{\Delta}_\phi[k] \pmod{2\pi} \quad (7)$$

which is identical to the phase difference between the first and last channels after the phase is unwrapped. Thus, the MCPD estimator can be expressed as

$$\hat{r} = -\frac{c_0}{4\pi \Delta_f} \frac{\phi_{2W}(f_{K_f}, r) - \phi_{2W}(f_0, r)}{K_f - 1}. \quad (8)$$

Hence, the bandwidth has effectively been increased by a factor of $K_f - 1$ without decreasing the unambiguous range, resulting in an improved performance in the presence of phase noise and multi-path effects.

III. Impact of Crystal Offset on BLE Ranging

To derive the impact of crystal offset on BLE ranging, we first define the timeline of measurements as in Fig. 1. For most BLE-transceivers, a single crystal is used for both LO and clock-generation. Therefore, the carrier frequencies and timings previously defined in Section II will be offset by a factor of η, the crystal offset, typically expressed in part-per-million (ppm). The crystal offset varies per BLE device, and we define the crystal offset of the initiator and reflector as η_i and η_r respectively.

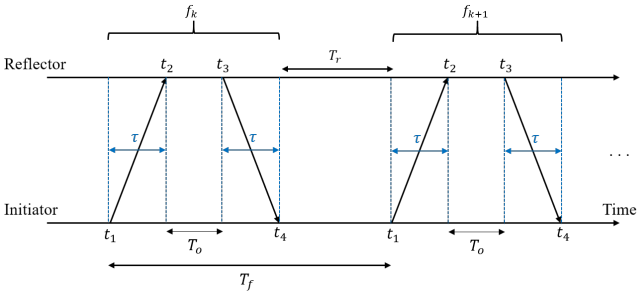

Fig. 1. BLE MCPD measurement timeline. The initiator starts the measurement process for channel k by sending its LO at t_1. From t_2 to t_3, CFR at the reflector is measured. At t_3, the reflector sends its LO to the initiator, and the initiator measures the CFR received at t_4. T_o is the defined as the intra-delay or reflector delay and T_r is defined as the inter-delay or channel hopping delay.

Including the effect of crystal offset, it can be shown that

$$\begin{aligned} \phi_{2W}(f_k, r) = &- 2\pi \left(f_k^i + f_k^r \right) \frac{r}{c_0} \\ &+ 2\pi \left(f_k^i + f_k^r \right) \left(T_k^r - T_k^i \right), \end{aligned} \quad (9)$$

from (2), (3), and (4), where f_k^i and f_k^r are the actual carrier frequencies generated by the initiator and reflector respectively, and T_k^i, T_k^r are the actual times at which CFR is measured at initiator and reflector respectively. The time difference between initiator and reflector is represented as $\Delta_t = T_k^r - T_k^i$.

Defining the actual carrier frequencies in terms of the crystal offsets η_i and η_r, we obtain $f_k^i = (1 + \eta_i) f_k$ and $f_k^r = (1 + \eta_r) f_k$ where $f_k = f_0 + k\Delta_f$ is the carrier frequency of the k-th channel. Similarly, the measurement timings are defined as $T_k^r = (1 + \eta_r) t_k$ and $T_k^i = (1 + \eta_i)(t_k + T_o)$ where T_o is the intra-delay, the time between when the reflector receives a signal from the initiator and transmits back to the initiator for each frequency. The total time per channel measurement is defined as $T_f = t_{k+1} - t_k$ which includes the time of flight, τ, twice, the intra-delay, T_o, and the inter-delay T_r as in Fig. 1. For this analysis, we assume the time of flight τ is much less than T_o, T_r, and T_r.

Substituting the actual carrier frequencies and measurement timings into (9) and (8), the ranging error is

$$e_r = \tilde{r} - r = \alpha_f T_f + \alpha_o T_o, \tag{10}$$

where

$$\alpha_f = -\frac{c_0}{2}(\eta_r - \eta_i)^2 \left(\frac{f_0}{\Delta_f} + K_f \right), \tag{11}$$

and

$$\alpha_o = -\frac{c_0}{2}(1+\eta_i)(\eta_i - \eta_r). \tag{12}$$

It is notable that the equation for α_o differs from the equation derived in [1] in that η_i is swapped with η_r which we have found to be the correct form in our own derivation.

Using practical values in (11) and (12), a simplified ranging error can be defined. Setting $\eta_i - \eta_r = \pm 20$ ppm at most, $f_0 = 2.404$ GHz, $\Delta_f = 1$ MHz, and $K_f = 75$, it can be observed that α_o is a factor of 20 times higher than α_f. Also observing $(1+\eta_i) \approx 1$, α_f and α_o can be approximated as

$$\tilde{\alpha_f} = 0, \tag{13}$$

and

$$\tilde{\alpha_o} = -\frac{c_0}{2}(\eta_i - \eta_r). \tag{14}$$

The ranging error is then approximated as

$$\tilde{e_r} = \tilde{\alpha_o} T_o = -\frac{c_0}{2}(\eta_i - \eta_r) T_o. \tag{15}$$

IV. Slow Mobility and the Doppler Effect

The Doppler effect is defined as the receiver frequency offset due to mobility. Doppler shift can be quantified as

$$f_r = \left(1 + \frac{\Delta_v}{c_0 + v_t}\right) f_t \approx \left(1 + \frac{\Delta_v}{c_0}\right) f_t, \tag{16}$$

where f_r is the received frequency, $\Delta v = v_r - v_t$ is the difference in velocity of the receiver v_r and velocity of the transmitter v_r, and f_t is the transmitted frequency. The Doppler frequency shift is then defined as

$$f_d = \frac{f_t}{c_0}\Delta_v = \frac{\Delta_v}{\lambda_t}, \tag{17}$$

where $\lambda_t = c_0/f_t$ is the wavelength of the transmitted frequency.

A. Effect of Mobility for BLE MCPD Ranging

Following the model in Section III, the Doppler frequency shift adds to the received carrier frequencies as $\hat{f}_k^i = (1+\eta_i)f_k + f_{d,k}^i$ and $\hat{f}_k^r = (1+\eta_r)f_k + f_{d,k}^r$ for the initiator and reflector respectively where $f_{d,k}^i = (f_k^r/c_0)(v_i - v_r)$ and $f_{d,k}^r = (f_k^i/c_0)(v_r - v_i)$ with initiator and reflector velocities of v_i and v_r respectively. The distance estimation error can then be re-evaluated in the case of mobility through (8) and (9) to obtain

$$\hat{e}_r = \tilde{r} - r = \hat{\alpha}_f T_f + \hat{\alpha}_o T_o, \tag{18}$$

where

$$\begin{aligned} \hat{\alpha}_f = &-\frac{c_0}{2\Delta_f}[(\eta_i - \eta_r)^2(f_o + K_f\Delta_f) \\ &- (f_{d,K_f}^i - f_{d,K_f}^r)(\eta_r - \eta_i)], \end{aligned} \tag{19}$$

and

$$\begin{aligned} \hat{\alpha}_o = &-\frac{c_0}{2}(1+\eta_i)(\eta_i - \eta_r) \\ &-\frac{1}{2}\left[(1+\eta_i)^2 + (1+\eta_i)(1+\eta_r)\right](v_i - v_r). \end{aligned} \tag{20}$$

Once again, we utilize practical values for BLE to evaluate the added distance estimation error due to mobility. Using $f_0 = 2.404$ GHz, $\Delta_f = 1$ MHz, $K_f = 75$, $\eta_i - \eta_r = \pm 20$ ppm at most, and $v_i - v_r = 6$ m/s at most for pedestrians. The change for both α_f and α_o in (11) and (12) is then only 0.2%. Therefore, the added effect of mobility for MCPD BLE ranging is negligible.

V. BLE Ranging in a Multi-path Rayleigh Fading Channel

In typical BLE environments indoors, multi-path is abundant and can corrupt the LOS signal measurements. To evaluate a multi-path channel with two-way MCPD BLE ranging, consider a two path environment CFR

$$\begin{aligned} H_{2W}(f_k) &= \left(a_1 e^{-j2\pi f_k \frac{r_1}{c_0}} + a_2 e^{-j2\pi f_k \frac{r_2}{c_0}}\right)^2 \\ &= a_1^2 e^{-j2\pi f_k \frac{2r_1}{c_0}} + a_1 a_2 e^{-j2\pi f_k \frac{r_1+r_2}{c_0}} \\ &\quad + a_2^2 e^{-j2\pi f_k \frac{2r_2}{c_0}}, \end{aligned} \tag{21}$$

with one-way distances r_1 and r_2 and complex attenuations a_1 and a_2. It can be observed that for that for two one-way paths, there are three paths that show up in the two-way CFR: $2r_1$, $r_1 + r_2$, and $2r_2$. This principle can be extrapolated for any number of signal paths.

A. Spectral Analysis Methods for LOS Distance Estimation

To recover the line of sight (LOS) distance for ranging from a multi-path channel frequency response as in (21), a spectral estimation algorithm such as the Fourier transform or MUSIC can be employed. The Fourier transform approach can be expressed as

$$h_{2W}(n) = \sum_{k=0}^{K-1} H_{2W}(f_k) e^{-j\frac{2\pi}{K}kn}, \tag{22}$$

where $-\pi \leq n < \pi$ with spacing $2\pi/K$, and the set of distance estimates is $d = -(n \times c_0)/(2\pi\Delta_f)$. The estimated distance is then peak in $h_{2W}(n)$ with lowest distance d. For two-way BLE ranging, the distance corresponding to the LOS peak, d, is divided by two to obtain the one-way distance. The resolution of the Fourier transform approach is a function of the Fourier transform size K. The Fourier transform approach has low complexity and can be employed by using the fast Fourier transform (FFT), but it suffers from lack of resolution between two paths with similar distances.

A more popular approach for time of flight or distance estimations using channel frequency response measurements is MUSIC (MUltiple SIgnal Classification) which achieves

super-resolution at the cost of higher complexity [9]. To define the MUSIC distance estimation method, first define

$$H(f_k) = \sum_{i=0}^{L_p-1} a_i e^{-j2\pi f_k \tau_i}, \tag{23}$$

as the CFR with L_p paths with propagation delays $\tau_i = r_i \times c_0$. The channel frequency response measurements can then be expressed in the form

$$x(k) = H(f_k) + w(k) = \sum_{i=0}^{L_p-1} a_i e^{-j2\pi f_k \tau_i} + w(k), \tag{24}$$

where $k = 0, 1, ..., L-1$ is the set of frequency measurement indices and $w(k)$ denotes additive white Gaussian noise with mean zero and variance σ_w^2. The CFR signal model can thus be expressed in matrix form as

$$\mathbf{x} = \mathbf{H} + \mathbf{w} = \mathbf{V}\mathbf{a} + \mathbf{w}, \tag{25}$$

where

$$\begin{aligned}
\mathbf{x} &= [x(0)\ x(1)\ \cdots\ x(L-1)]^T, \\
\mathbf{H} &= [H(f_0)\ H(f_1)\ \cdots\ H(f_{L-1})]^T, \\
\mathbf{w} &= [w(0)\ w(1)\ \cdots\ w(L-1)]^T, \\
\mathbf{V} &= \left[\mathbf{v}(\tau_0)\ \mathbf{v}(\tau_1)\ \cdots\ \mathbf{v}(\tau_{L_p-1})\right], \\
\mathbf{v}(\tau_i) &= \left[1\ e^{-j2\pi\Delta_f\tau_i}\ \cdots\ e^{-j2\pi(L-1)\Delta_f\tau_i}\right]^T, \\
\mathbf{a} &= \left[a_0'\ a_1'\ \cdots\ a_{L_p-1}'\right]^T, \\
a_i' &= a_i e^{-j2\pi f_0 \tau_i},
\end{aligned}$$

and the T superscript denotes the matrix transpose operation. MUSIC super-resolution techniques utilize eigen-decomposition of the autocorrelation matrix for the signal model in (25). From the eigen-decomposition, MUSIC utilizes the noise subspace for spectral analysis. The autocorrelation of $\mathbf{x}$ is defined as

$$\mathbf{R}_{xx} = E\left[\mathbf{x}\mathbf{x}^H\right] = \mathbf{V}\mathbf{A}\mathbf{V}^H + \sigma_w^2\mathbf{I}, \tag{26}$$

where $\mathbf{A} = E\left[\mathbf{a}\mathbf{a}^H\right]$ and the superscript H denotes the conjugate transpose operation. Following the assumptions:

1) The propagation delays τ_i are all different, so the matrix $\mathbf{V}$ has full column rank.
2) The magnitude of the parameters a_i is constant.
3) The phase is a uniform random variable in $[0, 2\pi]$.

Then, the $L_p \times L_p$ covariance matrix $\mathbf{A}$ is nonsingular. It follows that assuming $L > L_p$, the rank of the matrix $\mathbf{V}\mathbf{A}\mathbf{V}^H$ is L_p. Equivalently, the $L - L_p$ smallest eigenvalues of the signal autocorrelation matrix $\mathbf{R}_{xx}$ are all equal to σ_w^2. The eigenvectors (EVs) corresponding to $L - L_p$ smallest eigenvalues of $\mathbf{R}_{xx}$ are called noise EVs, while the EVs corresponding to L_p largest eigenvalues are called signal EVs. Thus, the L-dimensional subspace that contains the signal vector $\mathbf{x}$ can be split into two orthogonal subspaces, the signal subspace and the noise subspace, by the signal EVs and noise EVs respectively.

The project matrix of the noise subspace is then determined as

$$\mathbf{P}_w = \mathbf{Q}_w\mathbf{Q}_w^H, \tag{27}$$

where $\mathbf{Q}_w = \left[\mathbf{q}_{L_p}\ \mathbf{q}_{L_p+1}\ \cdots\ \mathbf{q}_{L-1}\right]$ and $\mathbf{q}_i$, $L_p \le i \le L-1$ are the noise EVs. Since the vector $\mathbf{v}(\tau_i)$, $0 \le i \le L_p-1$ must be in the signal subspace and the noise and signal subspaces are orthogonal, then

$$\mathbf{P}_w\mathbf{v}(\tau_k) = 0. \tag{28}$$

Thus, the multi-path propagation delays τ_i can be determined by finding the delay values at which the following MUSIC pseudospectrum peaks:

$$\begin{aligned}
S_{MUSIC}(\tau) &= \frac{1}{||\mathbf{P}_w\mathbf{v}(\tau)||^2} = \frac{1}{\mathbf{v}^H(\tau)\mathbf{P}_w\mathbf{v}(\tau)} \\
&= \frac{1}{||\mathbf{Q}_w^H\mathbf{v}(\tau)||^2} = \frac{1}{\sum_{i=L_p}^{L-1}|\mathbf{q}_i^H\mathbf{v}(\tau)|^2},
\end{aligned} \tag{29}$$

where τ is the set of search delays with a defined time resolution and size. The set of τ defines the steering vectors $\mathbf{v}(\tau)$.

Thus far, the MUSIC method defined has assumed knowledge of the true correlation matrix $\mathbf{R}_{xx}$. In practice, the correlation matrix must be estimated from a limited number of samples (i.e. around 75 measurements for MCPD BLE). For a single CFR snapshot of length N, the data sequence is divided into M consecutive segments of length L, and the correlation matrix is estimated as

$$\hat{\mathbf{R}}_{xx} = \frac{1}{M}\sum_{k=1}^{M-1}\mathbf{x}(k)\mathbf{x}(k)^H, \tag{30}$$

where $M = N-L+1$ and $\mathbf{x}(k) = [x(k)\ \cdots\ x(k+L-1)]^T$. This technique of dividing the measurement snapshot into multiple segments for correlation matrix estimation is commonly referred to as spatial smoothing.

The correlation matrix calculated as in (30) can be defined as a forward approach. It has been well-studied that the estimated correlation matrix in (30) can be improved using a forward-backward approach [10]. The forward-backward correlation matrix is calculated as

$$\hat{\mathbf{R}}_{xx}^{(FB)} = \frac{1}{2}\left(\hat{\mathbf{R}}_{xx} + \mathbf{J}\hat{\mathbf{R}}_{xx}^*\mathbf{J}\right), \tag{31}$$

where the superscript $*$ denotes conjugate, and $\mathbf{J}$ is the $L \times L$ reversal matrix whose antidiagonal is all ones with zeros everywhere else. For analysis of BLE MCPD ranging performance using MUSIC, we utilize the forward-backward correlation matrix.

The remaining unknown parameter for MUSIC distance estimation is the number of paths L_p. While there exists methods to estimation L_p such as the minimum descriptive length criteria (MDL) [11], this analysis simply assumes a fixed $L_p = 5$, which is greater than the number of paths in the two-way MCPD CFRs analyzed.

B. Rayleigh Fading Channel

To evaluate the performance of MCPD BLE ranging under a worst-case fading scenario, with no guaranteed line of sight, we employ the Rayleigh fading model. Under Rayleigh fading, each channel tap is assumed to be complex zero-mean Gaussian. Therefore, the LOS channel tap is not guaranteed to be recoverable. To further refine the analysis, we assume a two tap model which translates to 2 paths for one-way ranging or 3 paths for three way ranging.

A two tap Rayleigh fading channel can be expressed as

$$h(t) = a_1\delta(t-\tau_1) + a_2\delta(t-\tau_2), \tag{32}$$

where a_1 and a_2 are complex Gaussian random variables with mean zero. Expressing this channel response in the frequency domain,

$$H(f_k) = a_1 e^{-j2\pi f_k \tau_1} + a_2 e^{-j2\pi f_k \tau_2}, \tag{33}$$

from which the two-way BLE MCPD channel response is expressed as in (21) where $\tau_1 = r_1/c_0$ and $\tau_2 = r_2/c_0$. Following the analysis in section IV, a_1, a_2, τ_1, and τ_1, can be assumed to be time-invariant random variables within a single CFR snapshot.

VI. Residual Learning Technique for Improved BLE Ranging

In this section, we present a novel technique for improving localization accuracy by employing a residual learning technique for noise estimation and removal. For the simple, one-way, additive noise channel model, as examined by [8], the signal consists of a sum of sinusoidal signals in the presence of noise. This scenario can be modeled as

$$z_n = y_n + e_n, \quad n = 0, 1, \ldots, N-1, \tag{34}$$

$$y_n = \sum_{m=0}^{M-1} \alpha_m e^{-j2\pi f_m n}, \quad n = 0, 1, \ldots, N-1, \tag{35}$$

where $f_m \in [-0.5, 0.5)$ and α_m are the normalized frequency and complex amplitude of the k-th component, respectively, M is the model order, and e_n is the complex noise.

Written in vector form with $\mathbf{z} = [z_0, z_1, \ldots, z_{N-1}]^T$, $\mathbf{y} = [y_0, y_1, \ldots, y_{N-1}]^T$, and $\mathbf{e} = [e_0, e_1, \ldots, e_{N-1}]^T$,

$$\mathbf{z} = \mathbf{y} + \mathbf{e}. \tag{36}$$

For this scenario, addressed in [8], the noise vector $\mathbf{e}$ can be estimated as $\hat{\mathbf{e}}$ using the residual learning technique and removed to estimate the noiseless signal $\mathbf{y}$ as

$$\hat{\mathbf{y}} = \mathbf{z} - \hat{\mathbf{e}}. \tag{37}$$

On the other hand, in BLE ranging, the two-way CFR is employed to cancel undesired terms and attain the two-way phase difference (4). Assuming an additive noise model on both propagation paths, the two-way CFR can be modeled as

$$s_{2W,k} = (s_{R,k} + n_{R,k})(s_{I,k} + n_{I,k}), \tag{38}$$

$$s_{R,k} = \sum_{m=0}^{M-1} \alpha_{R,m} e^{j\phi_R(f_k^r, r_m)}, \tag{39}$$

$$s_{I,k} = \sum_{m=0}^{M-1} \alpha_{I,m} e^{j\phi_I(f_k^i, r_m)}, \tag{40}$$

where $\alpha_{R,m}$, $\alpha_{I,m}$ are the complex gains for each channel in each propagation direction and $\phi_R(f_k^r, r_m)$, $\phi_I(f_k^i, r_m)$ are defined in (2) and (3), respectively. Defining

$$\begin{aligned}
\mathbf{s}_{2W} &= [s_{2W,1}\ s_{2W,2}\ \cdots\ s_{2W,K_f}]^T, \\
\mathbf{s}_R &= [s_{R,1}\ s_{R,2}\ \cdots\ s_{R,K_f}]^T, \\
\mathbf{s}_I &= [s_{I,1}\ s_{I,2}\ \cdots\ s_{I,K_f}]^T, \\
\mathbf{n}_R &= [n_{R,1}\ n_{R,2}\ \cdots\ n_{R,K_f}]^T, \\
\mathbf{n}_I &= [n_{I,1}\ n_{I,2}\ \cdots\ n_{I,K_f}]^T,
\end{aligned}$$

and using $\odot$ as the Hadamard product, we can rewrite (38) in vector notation as

$$\mathbf{s}_{2W} = \underbrace{\mathbf{s}_R \odot \mathbf{s}_I}_{\boldsymbol{\psi}} + \underbrace{(\mathbf{s}_R \odot \mathbf{n}_I + \mathbf{s}_I \odot \mathbf{n}_R + \mathbf{n}_R \odot \mathbf{n}_I)}_{\boldsymbol{\eta}}. \tag{41}$$

Hence, we have decomposed the two-way CFR into the sum of a structured signal term ($\boldsymbol{\psi}$) and noise term ($\boldsymbol{\eta}$). However, the noise term ($\boldsymbol{\eta}$) consists of products of the signal terms at R and I with the noise terms for each channel. Thus, if the DNN can estimate the noise as $\hat{\boldsymbol{\eta}}$ by residual learning, the structured signal term can be recovered thereby improving the SNR and subsequent spectral estimation. Then, the structured signal can be estimated as

$$\hat{\boldsymbol{\psi}} = \mathbf{s}_{2W} - \hat{\boldsymbol{\eta}}. \tag{42}$$

A. DNN Network Architecture and Training

The objective of the network is to estimate the additive noise so it can be removed to increase the SNR. To achieve this objective, we propose a regressive fully convolutional neural network (FCNN) architecture. In an FCNN, at each convolutional layer, the input and output sizes are identical by zero-padding the input signal on either size. FCNNs have been employed for residual learning in [8], and we adopt a similar approach. Instead of appending the real and imaginary parts of the noisy CFR signal to form a vector of size $2N$ from a complex-valued vector of size N, we layer the real and imaginary parts to form an array of size $N \times 2$ as the input to our DNN. As shown in Fig. 2, the network consists of 8 convolutional layers of varying kernel and layer sizes with a batch normalization (BN) and activation layer. Each convolution layer zero-pads the input data to retain the same length of the signal throughout. Additionally, from empirical investigation, the network achieves optimal convergence when the activation functions alternate between Rectified Linear Unit (ReLU) and hyperbolic tangent (tanh). Finally, before

the RMSE regression layer, we employ a convolution layer of kernel size 16 without BN or activation function to allow the network to resolve scaling and translation within the nonlinear portion prior to this layer. A detailed explanation of the mathematical form of the various layers of the DNN can be found in [8]. While our approach employs a real-valued FCNN, complex-valued FCNNs have found success in similar data-enhancement tasks [12]. A complex-valued network may achieve improved performance as phase-relationships in complex-valued data can be leveraged more efficiently. In this work, we are satisfied with the results of the real-valued network, which are further detailed in Section VII.

To train the network, we generate datasets synthetically for the multi-path AWGN channel and multi-path Rayleigh fading AWGN channel scenarios. We assume a two-tap multi-path propagation model and construct two large training datasets in simulation. Each sample consists of a LOS tap and a ghost multi-path tap, e.g., $M = 2$. The location of the LOS tap is swept across the range $r_{LOS} \in [2, 12]$ m and the second tap is placed at a farther distance than the LOS tap on the range $r_{MP} \in [r_{LOS}, r_{LOS} + 10]$ m. For the AWGN channel, the attenuation factors for the second tap, $\alpha_{R,1}$ and $\alpha_{I,1}$, are always less than the attenuation factors for the LOS tap. On the other hand, under Rayleigh fading, each tap experiences random attenuation resulting in decreased localization performance. Since the network is intended to minimize the difference between the predicted noise, $\hat{\boldsymbol{\eta}}$, and the ground-truth noise, $\boldsymbol{\eta}$, each simulated sample consists of the noisy CFR, s_{2W}, and the ground-truth noise term, $\boldsymbol{\eta}$, from (41). We train two networks, one for an AWGN channel and the other for AWGN with Rayleigh fading. Each network is trained for 1000 epochs with 16384 samples using an Adam optimizer [13] with an initial learning rate of 0.1, piecewise learning rate decay of 0.9 every 3 epochs, and a batch size of 64. No normalization is applied to the input data, and the data is shuffled every epoch. Fig. 2 shows intermediate activations at several points throughout the network, the predicted noise, and the noise-subtracted CFR. The results shown are for the AWGN with Rayleigh fading channel. As expected, the noise-subtracted CFR both closely resembles the ground-truth CFR and results in an improved spectral estimation. The sample shown in Fig. 2 consists of a LOS one-way distance of 12 m and a second tap at 22 m. The MUSIC spectrum from the noisy CFR does not contain a peak for the LOS reflection at the two-way distance of 24 m. However, the MUSIC spectrum obtained from the noise-subtracted data contains the 3 expected peaks including the LOS reflection at 24 m. Further analysis of the DNN performance is examined in Section VII.

VII. Analytical Results

To analyze the performance of BLE ranging in simulation, we first verify the effect of crystal offset as in section III. Then, we verify the effect of mobility as in section IV. After, we empirically compare the distance estimation resolution of the Fourier transform and MUSIC approaches. Finally, the potential benefit of a residual learning technique in combination with MUSIC is analyzed.

A. Simulation Analysis of the Effect of Crystal Offset and Mobility on BLE Ranging

To evaluate the effect of crystal offset and mobility on BLE ranging, we independently analyze the effect due to intra-delay T_o, inter-delay T_r (the full channel measurement delay is $T_f = T_o + T_r$), crystal offset $\epsilon = \eta_i - \eta_r$, and mobility $\Delta_v = v_i - v_r$ while keeping each of the other three parameters constant. In the simulation, the full two-way MCPD BLE phase measurement for one path is generated as in (9) with the included effect of mobility on the carrier frequencies. The distance estimations are computed using linear least squares slope estimation across the phase to verify the theoretical estimation error as in (18).

The simulated carrier frequencies are $2.404 \leq f_k \leq 2.478$ GHz with frequency measurement spacing of $\Delta_f = 1$ MHz, and the distance is fixed to $r = 10$ m (single path assumption). Four cases are evaluated to analyze the effect of changing the intra-delay, inter-delay, mobility, or crystal offset. The parameters of the four cases are listed in Table I.

TABLE I
Crystal Offset and Mobility Simulation Parameters

Case	T_o	T_r	Δ_v	ϵ
1	0 to 1 ms	0	0	-10 ppm
2	1 ms	0 to 1 ms	0	-10 ppm
3	1 ms	1 ms	0 to 30 m/s	-10 ppm
4	1 ms	1 ms	30 m/s	0 to -10 ppm

Simulation results for each case are shown in Fig. 3. From the simulation, it can be assessed that the ranging error is most sensitive to changes in intra-delay T_o and crystal offset ϵ, while being relatively insensitive to changes in inter-delay T_r and mobility Δ_v. Of particular note, changes in mobility within pedestrian speeds of 6 m/s results in distance estimation changes of less than 1 cm. This follows the previous analyses conducted in sections III and IV where the effect of inter-delay and mobility is expected to be low for realistic BLE parameters.

This evaluation of ranging estimation error holds for all paths in a multi-path environment, and the rest of this analysis will assume full knowledge of the parameters in Table I to compensate for the corresponding distance estimation error.

B. Analysis of Distance Estimation Techniques: Fourier Transform (FFT) and MUSIC

To evaluate the resolution capability of the Fourier transform approach for distance estimation compared to MUSIC, we simulate using a two-path model. The simulation parameters are $2.404 \leq f_k \leq 2.478$ GHz with $\Delta_f = 1$ MHz frequency spacing, $r_1 = 2$ m, $r_2 = 6.5$ m, $a_1 = 1$, and $a_2 = 1$ in equation (21) for the two-way CFR. Since BLE MCPD ranging uses a two-way channel frequency response, the expected two-way multi-path distances are 4 m, 8.5 m, and 13 m. The

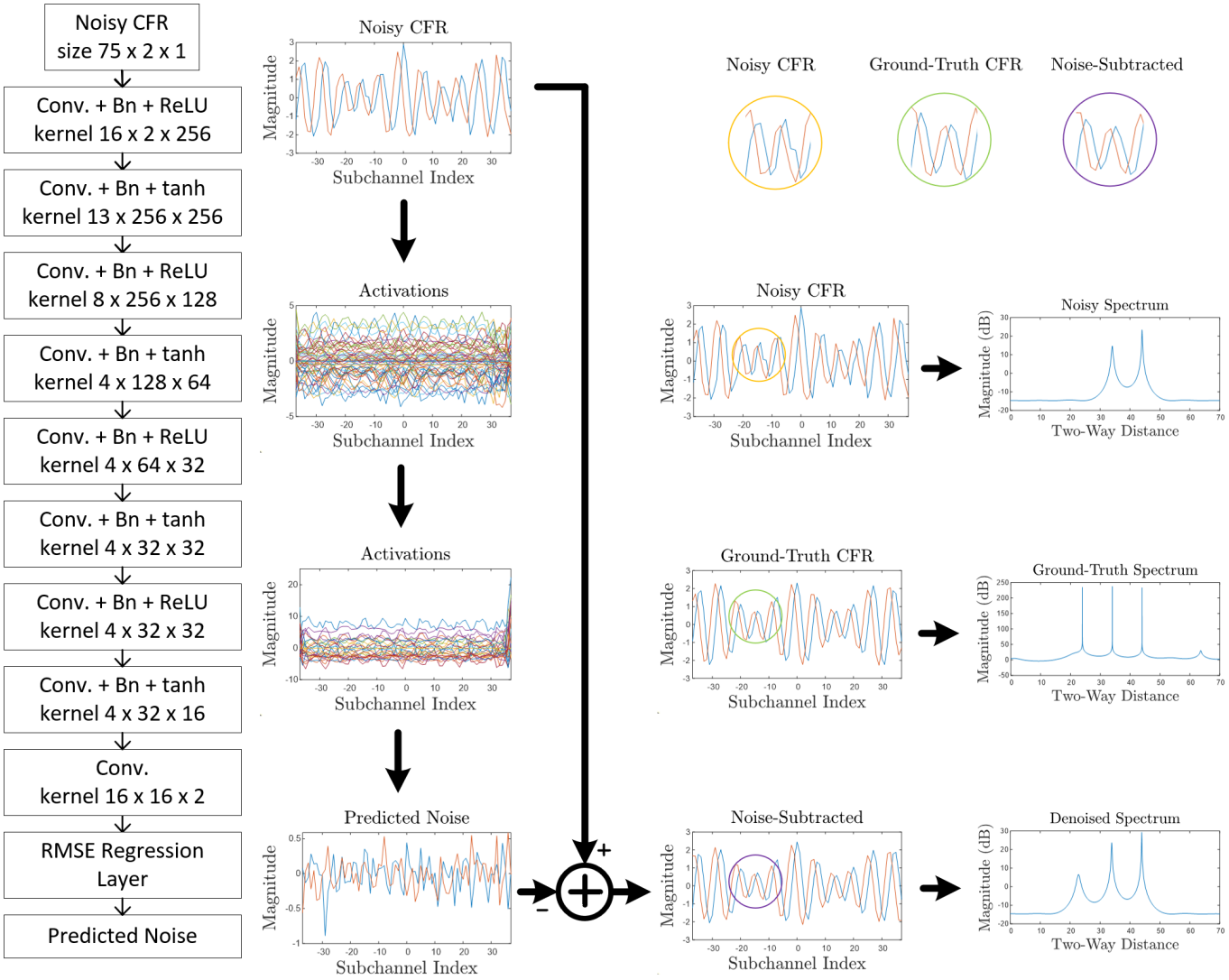

Fig. 2. Network architecture of the proposed DNN with input noisy CFR, intermediate activations, and predicted noise. Each CFR and the predicted noise shows the real (blue) and imaginary (orange) parts of the signal. The predicted noise is subtracted from the noisy CFR and the spectra of all three are computed using MUSIC. A portion of the CFR is displayed at the top for the noisy, ground-truth, and noise-subtracted cases to demonstrate performance of the DNN to learn the noise.

simulation implements AWGN with an SNR of 15 dB and an SNR of 25 dB. An example iteration for each SNR is displayed in Fig. 4. By observing the peaks in the figure, it is clear that MUSIC is capable of offering higher resolution distance estimates, especially as SNR increases.

Over 1000 iterations, the mean LOS distance estimate, standard deviation (STD), mean error, and root-mean-square error are analyzed and shown in Tables II and III for SNR = 15 dB and SNR = 25 dB respectively. Beyond observing the FFT and MUSIC spectrums, it is clear that MUSIC offers significantly lower mean error in both cases. Furthermore, as SNR is increased, MUSIC more clearly performs better in RMSE as well, whereas the bias for the FFT-based approach is relatively indifferent to the increase in SNR.

Due to the higher resolution of MUSIC over FFT, especially as SNR increases, the following residual learning technique analysis will evaluate performances using MUSIC only.

TABLE II
FFT VS. MUSIC FOR A TWO-PATH MODEL WITH SNR = 15 DB

Method	LOS Est. (m)	STD (m)	Mean Error (m)	RMSE (m)
FFT	1.642	0.026	-0358	0.359
MUSIC	2.091	0.335	0.091	0.347

TABLE III
FFT VS. MUSIC FOR A TWO-PATH MODEL WITH SNR = 25 DB

Method	LOS Est. (m)	STD (m)	Mean Error (m)	RMSE (m)
FFT	1.642	0.009	-0.358	0.358
MUSIC	2.002	0.066	0.002	0.066

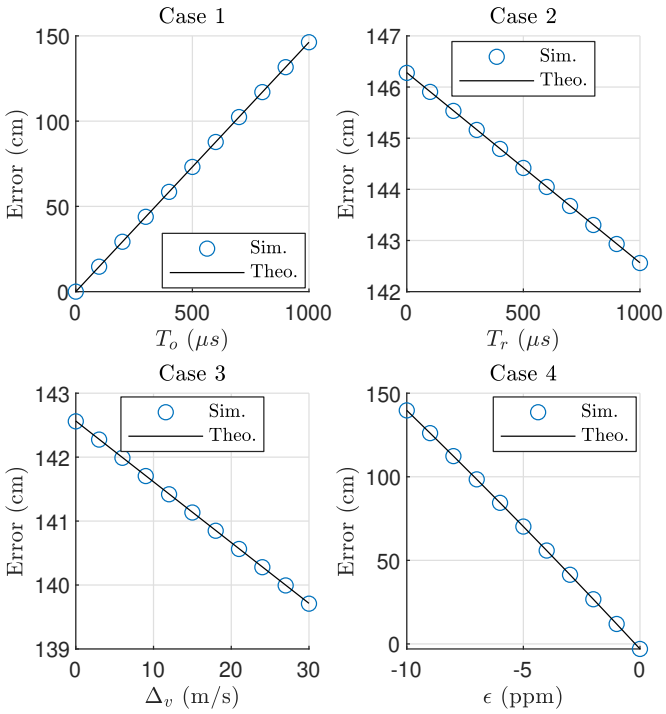

Fig. 3. Mobility and crystal offset simulations with different parameters as in Table I.

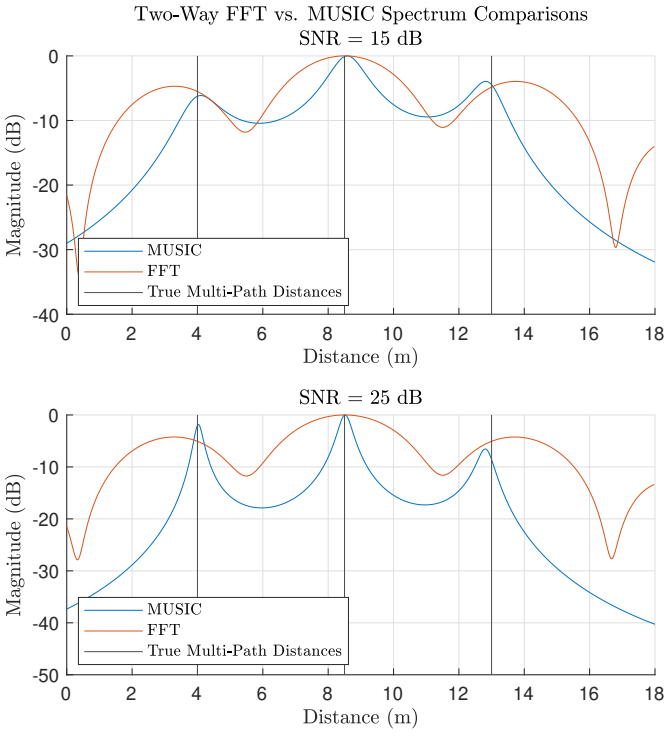

Fig. 4. Spectrum comparisons between FFT and MUSIC. The simulation uses a two-way channel frequency response with one-way paths of 2 m and 6.5 m. The corresponding two-way paths are then 4 m, 8.5 m, and 13 m as indicated in the figure.

C. Novel BLE Residual Learning Technique on AWGN and Multi-Path Rayleigh Fading Channel

To analyze the potential performance gain achieved by the proposed residual learning technique, we first train two independent networks using the methodology discussed in Section VI. The first network is trained using simulated data consisting of two-tap multi-path on an AWGN channel. Similarly, the second network is trained using data on an AWGN channel with Rayleigh fading. The first and second networks will be denoted by *A-DNN* and *AR-DNN*. To compare the ranging performance between the DNN methods and the traditional MUSIC-based approach, we simulate 160 scenarios consisting of a two-tap multi-path ranging scenario where the LOS path is held at a constant 10 m and the second distance varies from 10.05 m - 18 m with a step size of 5 cm. Each scenario is repeated 20 times with new AWGN noise and Rayleigh fading coefficients, when applicable, and the SNR is 15 dB. For range estimation, the peaks of the spectrum are estimated with a threshold of 13 dB below the highest peak. The MUSIC algorithm is applied either directly to the noisy CFR signal or the noise-subtracted CFR after applying the DNN.

First, we compare the classical MUSIC approach to the network trained on samples with AWGN noise and without Rayleigh fading. The simulated CFR data are also subject to AWGN noise and fixed channel fading. Fig. 5 compares the empirical cumulative distribution functions (CDFs) for the direct MUSIC approach against the MUSIC-*A-DNN* technique. The MUSIC-*A-DNN* shows superior performance over MUSIC alone, demonstrating the networks effectiveness in identifying the noise from the noisy CFR data. Once the noise is predicted, it can be subtracted from the raw CFR resulting in higher SNR.

Similarly, we repeat the experiment using AWGN noise and fixed channel fading with *AR-DNN*, the network trained using an AWGN and Rayleigh fading channel. Despite the fact this network was trained exclusively on data with Rayleigh fading, it shows similar performance to the *A-DNN* when applied to fixed fading data, as shown in Fig. 5. Again, the residual learning approach demonstrates a performance gain over the traditional MUSIC-based methods, resulting in lower ranging error on average. For all three approaches, the absolute error is consistently less than 150 cm.

We extend our investigation to the multi-path scenario with an AWGN channel with Rayleigh fading. To analyze the performance of both, the MUSIC-*A-DNN* and MUSIC-*AR-DNN* networks in comparison to classical MUSIC, we apply the same set of 160 two-tap multi-path scenarios and examine the ranging error across all 3200 samples. When the MUSIC-*A-DNN* method is applied to the AWGN with Rayleigh fading data, it yields a worse result than the MUSIC algorithm alone as it has only been trained on fixed fading AWGN samples and is likely removing meaningful portions of the CFR by including them in the noise prediction. As expected, we observe a slight performance gain when applying the MUSIC-*AR-DNN* technique to the AWGN with Rayleigh fading data,

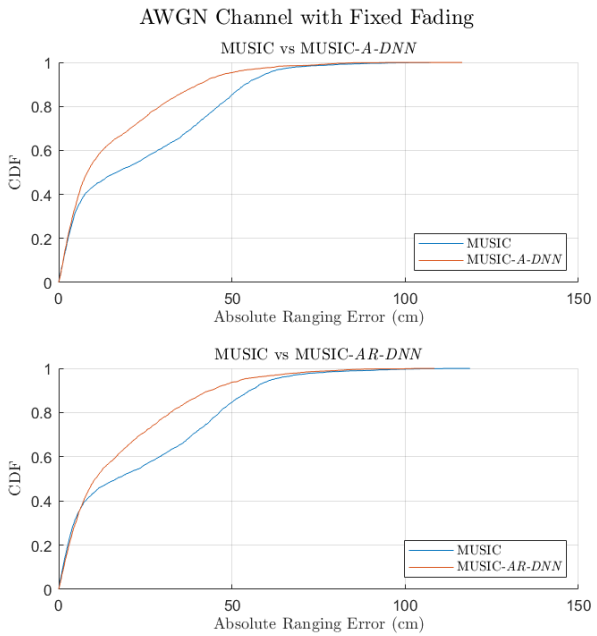

Fig. 5. Comparison of the error CDFs for the MUSIC-*A-DNN* and MUSIC-*AR-DNN* on an AWGN channel with fixed Fading.

signifying an improvement in SNR and ability to resolve the LOS peak, as shown in Fig. 6. However, the improvement over the traditional MUSIC technique on an AWGN Rayleigh fading channel is less than the improvement gained on an AWGN fixed fading channel. This is to be expected due to the nature of the Rayleigh fading data and the residual learning approach. Residual learning, at best, serves to improve SNR in additive noise scenarios, which explains its effectiveness in the AWGN case. For Rayleigh fading with AWGN, on the other hand, the channel response is random, in addition to additive noise, and the LOS is not guaranteed. In this case, the residual learning technique can still predict the additive noise with high accuracy, but the Rayleigh fading remains unaddressed and contributes heavily to the performance degradation. Hence, a residual DNN is expected to achieve a smaller performance gain on a Rayleigh fading channel than an AWGN channel.

To overcome the deficiencies of residual learning on a Rayleigh fading channel, an alternative approach can be explored employing either traditional or deep learning techniques. Intuitively, space-time diversity could be exploited to improve the performance of our proposed methods on a Rayleigh fading channel. Additionally, the sparsity of the signal in the frequency domain could be exploited using sparse neural networks [14], alternative supervised-learning FCNN techniques [12], compressed sensing, or other signal processing techniques. These methods have seen success in other arenas and may enjoy success for BLE ranging with some minor modifications. However, the field of signal processing with deep learning, specifically for spectral estimation and sinusoidal-based structured signals has only begun to be explored, and we believe it will have a tremendous impact on the future of wireless communications and signal processing.

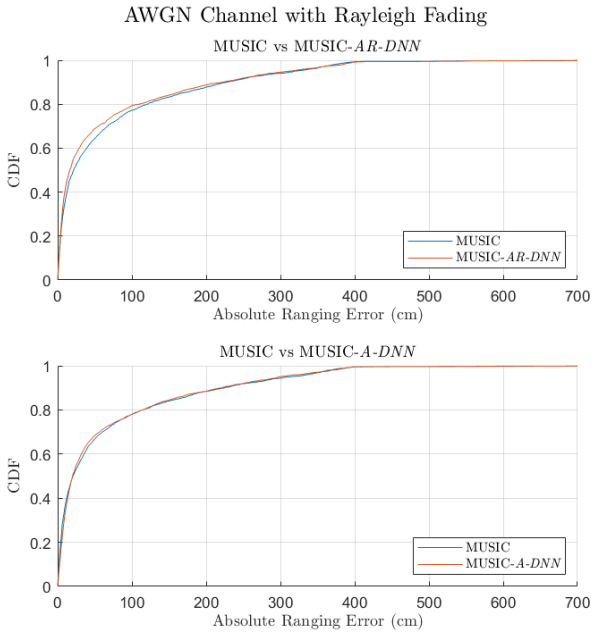

Fig. 6. Comparison of the error CDFs for the MUSIC-*AR-DNN* and MUSIC-*A-DNN* on an AWGN channel with Rayleigh Fading.

VIII. Conclusion

In this report, we investigate the impact of crystal offset, mobility, multi-path, and Rayleigh fading on BLE ranging. We extend the work of Zand *et al.* [1] by deriving an analytical expression for the error introduced by target mobility and provide simulation results to verify our analysis. On a typical BLE system, the ranging error due to mobility will likely be less than 1 cm with velocities less than 6 m/s, as shown in Fig. 3. While higher velocities result in increased error, for nearly all BLE applications, such velocities would result in lost connection in mere seconds. Further, the error introduced by mobility is dependent on the presence of crystal offset and timing mismatches and can be mitigated as these errors are eradicated. Additionally, we propose a novel residual learning technique for reducing BLE ranging error under multi-path AWGN fixed fading and Rayleigh fading channels. By extracting the additive noise from a structured signal, the proposed FCNN architecture can improve the SNR and result in superior spectral estimation. We train two DNNs for extracting noise from either AWGN fixed channel fading (MUSIC-*A-DNN*) and AWGN Rayleigh channel fading scenarios (MUSIC-*AR-DNN*), respectively. After the noise is estimated and subtracted form the noisy CFR signal, the MUSIC spectrum is computed and compared against the MUSIC spectrum taken directly from the noisy CFR. In simulation over multi-path fixed and Rayleigh fading channels, the DNN-based approaches consistently demonstrate a lower empirical ranging error than the traditional MUSIC algorithm. The performance gain using

the MUSIC-*A-DNN* on AWGN fixed fading data is greater than that of the MUSIC-*AR-DNN* on AWGN Rayleigh fading data due to the fact that the residual learning approach does not address the error introduced by the Rayleigh fading, but only extracts the additive noise. Alternative approaches, such as exploiting diversity or other deep learning techniques may result in a greater performance improvement and would be an interesting topic for further investigation.

ACKNOWLEDGMENT

The authors would like to thank Dr. Al-Dhahir for his abundant effort in teaching the EESC6353 Broadband Digital Communication course to provide a sufficient understanding of the relevant topics investigated in Zand *et al.*'s paper [1].

REFERENCES

[1] P. Zand, J. Romme, J. Govers, F. Pasveer, and G. Dolmans, "A high-accuracy phase-based ranging solution with bluetooth low energy (BLE)," in *Wireless Communications and Networking Conf.*, 2019, pp. 1–8.

[2] A. Ozer and E. John, "Improving the accuracy of bluetooth low energy indoor positioning system using Kalman filtering," in *CSCI*, 2016, pp. 180–185.

[3] B. T. Fang, "Simple solutions for hyperbolic and related position fixes," *Trans. on Aerospace and Electronic Systems*, vol. 26, no. 5, pp. 748–753, 1990.

[4] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *Trans. on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1067–1080, 2007.

[5] A. Bensky, *Wireless positioning technologies and applications*. Artech House, 2016.

[6] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising," *Trans. on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017.

[7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *CVPR*, 2016, pp. 770–778.

[8] Y. Jiang, H. Li, and M. Rangaswamy, "Deep learning denoising based line spectral estimation," *IEEE Signal Processing Letters*, vol. 26, no. 11, pp. 1573–1577, 2019.

[9] X. Li and K. Pahlavan, "Super-resolution toa estimation with diversity for indoor geolocation," *Trans. on Wireless Communications*, vol. 3, no. 1, pp. 224–234, 2004.

[10] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Pearson/Prentice Hall, 2005.

[11] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *Trans. on Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 387–392, 1985.

[12] J. Gao, B. Deng, Y. Qin, H. Wang, and X. Li, "Enhanced radar imaging using a complex-valued convolutional neural network," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 1, pp. 35–39, 2019.

[13] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[14] S. Srinivas, A. Subramanya, and R. Venkatesh Babu, "Training sparse neural networks," in *CVPR*, 2017, pp. 138–145.