Multiband Delay Estimation for Localization Using a Two-Stage Global Estimation Scheme

Yubo Wan\textsuperscript{,} Graduate Student Member, IEEE, An Liu\textsuperscript{,} Senior Member, IEEE, Qiyu Hu\textsuperscript{,} Graduate Student Member, IEEE, Mianyi Zhang\textsuperscript{,} and Yunlong Cai\textsuperscript{,} Senior Member, IEEE

Abstract—The time of arrival (TOA)-based localization techniques, which need to estimate the delay of the line-of-sight (LoS) path, have been widely employed in location-aware networks. To achieve a high-accuracy delay estimation, a number of multiband-based algorithms have been proposed recently, which exploit the channel state information (CSI) measurements over multiple non-contiguous frequency bands. However, to the best of our knowledge, there still lacks an efficient scheme that fully exploits the multiband gains when the phase distortion factors caused by hardware imperfections are considered, due to that the associated multi-parameter estimation problem contains many local optiumns and the existing algorithms can easily get stuck in a “bad” local optimum. To address these issues, we propose a novel two-stage global estimation (TSGE) scheme for multiband delay estimation. In the coarse stage, we exploit the group sparsity structure of the multiband channel and propose a Turbo Bayesian inference ( Turbo-BI) algorithm to achieve a good initial delay estimation based on a coarse signal model, which is transformed from the original multiband signal model by absorbing the carrier frequency terms. The estimation problem derived from the coarse signal model contains fewer local optiumns and thus a more stable estimation can be achieved than directly using the original signal model. Then in the refined stage, with the help of coarse estimation results to narrow down the search range, we perform a global delay estimation using a particle swarm optimization-least square (PSO-LS) algorithm based on a refined multiband signal model to exploit the multiband gains to further improve the estimation accuracy. Simulation results show that the proposed TSGE significantly outperforms the benchmarks with comparative computational complexity.

Index Terms—Delay estimation, multiband, Bayesian inference, particle swarm optimization, two-stage global estimation.

I. INTRODUCTION

Recently, the wireless-based localization has attracted great attention due to its adaptability to the existing wireless infrastructure and capacity for assisting communications [1], [2], [3], [4], [5]. It has been widely employed in multisensory extended reality (XR) [6], smart transportation [7], and connected robotics and autonomous systems (CRAS) [8]. Initially, the global navigation satellite system (GNSS) has been employed to provide location services, but it has low accuracy in indoor environments [9]. Thus, the cellular-based localization and wireless local area networks (WLAN)-based localization have been developed as alternatives to GNSS [10], [11], [12]. These approaches estimate the positions of agents by utilizing the information such as time of arrival (TOA), angle of arrival (AoA), angle of departure (AoD), time difference of arrival (TDOA), and received signal strength (RSS). In particular, the TOA-based localization is a widely studied wireless localization method [11], [12], [13], [14], which estimates the distance by multiplying the delay of the line-of-sight (LoS) path with the light speed.

However, the estimation accuracy of the multipath channel delays is limited by the bandwidth of the transmitted signal. To address this issue, the multiband delay estimation schemes have been proposed in [15], [16], [17], [18], [19], [20], [21], and [22], which make use of the channel state information (CSI) measurements across multiple frequency bands to obtain high accuracy delay estimation. Compared to delay estimation problem with only a single wideband contiguous spectrum [13], [14], [23], multiband delay estimation can obtain extra multiband gains, which consist of two parts: (i) Multiband CSI samples lead to more subcarrier apertures gain; (ii) Frequency band apertures gain brought by the difference of carrier frequency between subbands [20]. The subcarrier apertures and the frequency band apertures are shown in Fig. 1. As can be seen, the spectrum resource used for localization consists of a number of non-contiguous subbands. The green regions are the frequency subbands allocated to other applications (e.g., wireless communication) and thus are unable to utilize for localization.

Despite the existence of apertures gains, the multiband delay estimation method faces new challenges as compared to single band based delay estimation. One challenge is the phase distortion in the multiband channel frequency response (CFR) samples caused by hardware imperfections has a severe effect on delay estimation [24], [25], [26], which differs...
from band to band and needs to be calibrated. The other challenge is that the high frequency carrier in the CFR samples leads to a violent oscillation phenomenon of the likelihood function, which causes numerous “bad” local optimums [19], [20]. Therefore, it is very difficult for algorithms designed based on the multiband signal model to exploit the frequency band apertures gain, which requires a strong global search ability and the ability of overcoming the phase distortions with acceptable computation costs. Some related works are summarized below.

A. Maximum Likelihood (ML) Based Methods

In [22], a multiband TOA estimation method has been carried out for long term evolution (LTE) downlink orthogonal frequency division multiplexing (OFDM) systems, which provides a reduced standard deviation for the delay estimation. In [15], the authors proposed a low-complexity approach, which recovers the channel impulse response (CIR) with equally spaced taps and approximately estimates the first path by mitigating the energy leakage.

B. Subspace Based Estimation Methods

There have been works attempting to solve the multiband delay estimation problems by using subspace based methods [19], [20], [21]. In [21], the classical multiple signal classification (MUSIC) algorithm has been applied to delay estimation. The authors in [19] and [20] employed the multiple shift-invariance structure in the multiband channel and achieve a high accuracy delay estimation. Nevertheless, the subcarrier spacing of different subbands is assumed to be equivalent, which restricts its applications for practical systems, e.g., IEEE 802.11be (WiFi 7) standard enables multiband operations across 2.4 GHz, 5 GHz, and 6 GHz subbands to achieve a maximum throughput, where an equivalent subcarrier spacing among different subbands may not be guaranteed [27]. Moreover, the subspace based methods generally require multiple snapshots of OFDM pilot symbols to guarantee the performance [28], which consumes lots of pilot resources.

C. Compressed Sensing (CS) Based Methods

In an indoor environment, the CIR is sparse since it consists of a small number of paths. Motivated by the sparsity of CIR over the delay domain, many state-of-the-art delay estimation algorithms have been proposed based on CS methods [16], [17], [18]. In [16] and [17], the authors formulated $l_1$-norm minimization problems for capturing the signal sparsity. In [18], orthogonal matching pursuit (OMP) methods have been used for recovering the sparse CIR. However, these approaches cause energy leakage resulting from basis mismatch [29], which require dense grids and have high computational complexity.

In the aforementioned studies, on one hand, the accuracy of most algorithms is limited, since the associated multi-parameter estimation problem contains many “bad” local optimums caused by high frequency carrier terms and frequency band apertures. Though some other works have eliminated the oscillation, they only exploit the subcarrier apertures gain due to the difficulty of exploiting the frequency band apertures gain [16], [21]. On the other hand, many works have not considered the effect of phase distortion in practical systems [15], [19], [21], [30]. Though some studies in [16], [18], and [20] have considered the phase distortion, the calibration of the phase distortions with extra information is required via a handshaking procedure under the assumption of perfect channel reciprocity. However, this assumption is restrictive and the transceiver needs to have the ability of Tx/Rx switching. Furthermore, in the handshaking procedure, the phase lock loop (PLL) must keep inlock to ensure that the phase offset remains unchanged, which is difficult to achieve in practice.

In this paper, we consider a TOA-based localization system using OFDM training signals over multiple frequency subbands. Then, a novel two-stage global estimation (TSGE) scheme is proposed to fully exploit all the multiband gains, where we consider all the phase distortion factors and calibrate them implicitly without using extra handshaking procedures. Different from other existing two-stage estimation schemes, whose two estimation stages are based on different estimation algorithms but the same signal model, our proposed TSGE scheme considers different signal models in the two estimation stages and is able to find the global optimum of the associated multi-parameter estimation problem with high probability.

Specifically, in Stage 1, we build a coarse signal model, in which the high frequency carriers are all absorbed for eliminating the oscillation of the likelihood function. Although the coarse estimation algorithm derived from the coarse signal model can only exploit the subcarrier apertures gain, it does not get stuck in “bad” local optimums and thus can provide a much more stable delay estimation to narrow down the search range for the global delay estimation in the refined stage. By doing this, numerous “bad” local optimums can be excluded from the global search space in the refined stage. Then, we provide a sparse representation for multiband CFR, where we adopt a common support based sparse vector to capture the group sparsity structure in the multiband channel over the delay domain. Based on this model, a Turbo Bayesian inference (Turbo-BI) algorithm is proposed for channel parameter estimation (including the delay parameter). Compared to the CS-based delay estimation methods in [16], [17], and [18], our proposed algorithm achieves higher estimation accuracy with lower computational complexity. It is because we adopt a dynamic grid adjustment strategy and we do not need a very dense grid.

In Stage 2, with the help of prior information passed from Stage 1, we perform a finer estimation based on a refined signal model. A higher estimation accuracy than that in Stage 1 can be guaranteed since this signal model contains the structure of frequency band apertures. In addition,
the proposed refined signal model significantly reduces the difficulty of getting the global optimum for estimation algorithms. However, the refined signal model still leads to a multi-dimensional likelihood function that has a few “bad” local optima, which makes it difficult to find the global optimum and fully exploit the frequency band apertures gain. For utilizing this aperture gain and the prior information properly, we adopt a global search algorithm based on the particle swarm optimization (PSO) to find a global optimum solution with high probability for the non-convex optimization of the multi-dimensional likelihood function. In particular, the coarse estimation results from Stage 1 can be utilized for determining the particle search range, which can reduce the search complexity significantly. For further reducing the search space and improving the estimation accuracy by finding the global optimum with higher probability, we employ primal-decomposition theory to decouple the objective function and get a least square (LS) solution for channel coefficients. Then, the dimension of the search space can be reduced by eliminating the channel coefficients in the primal optimization problem. The main contributions are summarized below.

- We propose a novel two-stage signal model that helps the estimation algorithm to exploit extra multiband gains and simultaneously overcome the inherent difficulties of the multiband delay estimation problem.
- Based on the properties of the multiband signal model, we propose two novel delay estimation algorithms. In Stage 1, we set up a common support based probability model and propose a Turbo-BI algorithm to exploit the group sparsity structure of the multiband channel. In Stage 2, we propose a PSO-LS algorithm, which is able to exploit multiband gains in multiband channels to improve the estimation accuracy.

The rest of this paper is organized as follows. In Section II, we describe the system model and introduce the phase distortion factors. In Section III, we formulate the two-stage signal model and outline the TSGE scheme. Sections IV and V present the Turbo-BI algorithm and the PSO-LS algorithm in Stage 1 and Stage 2, respectively. In Section VI, numerical results are presented and finally Section VII concludes the paper.

Notations: Bold upper (lower)-case letters are used to define matrices (column vectors). I denotes an identity matrix, \( \delta(\cdot) \) denotes the Dirac’s delta function, \( \text{diag}(\cdot) \) constructs a diagonal matrix from its vector argument, and \( \| \cdot \| \) denotes the Euclidean norm of a complex vector. For a matrix \( A \), \( A^T, A^H, A^{-1}, \text{tr}(A) \) represent a transpose, complex conjugate transpose, inverse, and trace of a matrix, respectively. For a scalar \( a \), \( a^* \) denotes the conjugate of a scalar. The notation \( \mathbb{R}^+ \) represents the strictly positive real number and \( \mathcal{CN}(x; \mu, \Sigma) \) denotes a complex Gaussian normal distribution corresponding to variable \( x \) with mean \( \mu \) and covariance matrix \( \Sigma \).

II. SYSTEM MODEL

As shown in Fig. 2, we consider a single-input single-output (SISO) multiband system which employs OFDM training signals over \( M \) frequency subbands. The multiband system consists of \( M \) non-overlapping single band OFDM subsystems, where the \( m \)-th subband is allocated to the \( m \)-th operator. Assume that each frequency subband has \( N_m \) orthogonal subcarriers with subcarrier spacing \( f_{s,m} \) and the carrier frequency of subband \( m \) is denoted as \( f_{c,m} \). Then, the continuous-time CIR \( h(t) \) can be written as

\[
h(t) = \sum_{k=1}^{K} \alpha_k \delta(t - \tau_k),
\]

where \( K \) denotes the number of multipath components between the transmitter and the receiver, \( \alpha_k \in \mathbb{C} \) and \( \tau_k \in \mathbb{R}^+ \) denote the complex path gain and the delay of the \( k \)-th path, respectively. The delays are sorted in an increasing order, i.e., \( \tau_{k-1} < \tau_k, k = 2, \ldots, K \), and \( \tau_1 \) is the LoS path which needs to be estimated for localization. We assume that the complex path gain and delay parameters are independent of the frequency subbands. Then, via a Fourier transform of the CIR as in [17] and [20], the CFR samples can be expressed as

\[
\tilde{h}_{m,n} = \sum_{k=1}^{K} \alpha_k e^{-j2\pi f_{m,n}\tau_k},
\]

where \( f_{m,n} = f_{c,m} + n f_{s,m}, m = 1, \ldots, M, n \in \mathbb{N}_m \triangleq \{-\frac{N_m}{2}, \ldots, \frac{N_m}{2} - 1\} \). With a slight abuse of notation, we use \( n \) instead of \( n_m \) in the following equations. We assume that \( N_m, \forall m \) is an even number without loss of generality, and denote \( N = N_1 + \ldots + N_M \) as the number of CFR samples over all subbands.

Apparently, the CFR exhibits sparsity over delay domain when \( K \) is small, which will be exploited in our proposed Turbo-BI algorithm. Then, during the period of a single OFDM symbol, the discrete-time received signal model can be written as [17] and [20]

\[
y_{m,n} = \sum_{k=1}^{K} \alpha_k e^{-j2\pi(f_{c,m}+nf_{s,m})\tau_k} e^{-j2\pi n f_{s,m} \delta_m} e^{j \varphi_m} s_{m,n} + w_{m,n},
\]

where \( w_{m,n} \) is the \( n \)-th element of the additive white Gaussian noise (AWGN) vector \( \mathbf{w}_m \in \mathbb{C}^{N_m \times 1} \), following the distribution \( \mathcal{CN}(0, \sigma_p^2 I) \). \( s_{m,n} \) denotes a known training symbol over the \( n \)-th subcarrier of subband \( m \) and we assume \( s_{m,n} = 1, \forall m, n \) for simplicity. The parameter \( \varphi_m \) and \( \delta_m \) represent the phase distortion factors caused by random phase offset and receiver timing offset [24], [25], [26], respectively. In practice, the receiver timing offset \( \delta_m \) is often within a small range and thus we assume that \( \delta_m, \forall m \) follows a prior distribution \( p(\delta_m) \sim \mathcal{N}(0, \sigma_p^2) \) [26], where \( \sigma_p \) is the timing synchronization error.
However, the global delay estimation problem will be intractable if we directly use the signal model (3) due to the huge search space of the multi-dimensional parameters and the existence of many local optima in the likelihood function, which thus motivates the set up of the novel two-stage signal model and the associated two-stage global estimation scheme.

III. TWO-STAGE GLOBAL ESTIMATION SCHEME FOR MULTIBAND DELAY ESTIMATION

In this section, we first explain why we cannot use the original signal model for delay estimation. On one hand, the original model (3) leads to a multimodal non-convex delay estimation.

A. Two-Stage Signal Model

It is difficult to directly use the received signal model (3) for delay estimation. On one hand, the original model (3) exists inherent ambiguity. Specifically, for an arbitrary constant $c \neq 0$, if we substitute two sets of variables $((\alpha_k, e^{j(\varphi_k + c)}, \varphi_m))$ and $((\alpha_k, e^{j(\varphi_k + c)}, \varphi_m - c))$ into equation (3), the equivalent observation result will be obtained. It indicates that the parameters $(\alpha_k, \varphi_m)$ are ambiguous, which leads to the difficulty of delay estimation.

On the other hand, due to the high frequency carrier $f_{c,m}$, the original signal model (3) leads to a multimodal non-convex likelihood function that has many sidelobes. Specifically, the likelihood function based on original signal model (3) can be expressed as

$$ p(y; \tau) \propto \exp\left(-\frac{||y - h(\tau)||^2}{\sigma_n^2}\right), $$

where $y = [y_1, 0, \ldots, y_{m,n}, \ldots, y_{M,N,M-1}] \in \mathbb{C}^{N \times 1}$, $h(\tau)$ is a vector consisting of the element

$$ h_{m,n} = \sum_{k=1}^{K} \alpha_k e^{-j2\pi(f_{c,m} + n f_{c,m})\tau_k} e^{-j2\pi n f_{c,m} \delta_m} e^{j\varphi_m} s_{m,n}. $$

For simplicity, we assume that $K = 1$ and the delay $\tau$ is the only unknown parameter needs to be estimated. Then, the log-likelihood function can be simplified as

$$ \ln p(y; \tau) \propto -2\text{Re}\{y^H h(\tau)\}. $$

As can be seen in the expression, the delay $\tau_k$ is multiplied by a huge carrier frequency $f_{c,m}$, implying that a slight variation of $\tau_k$ may lead to a significant phase shift. Hence, when we search for the global optimum in the likelihood function, this property will be reflected by the violent oscillation phenomena of the likelihood function, leading to numerous sidelobes. Consequently, it is extremely difficult to find the global optimum. In the worst case, the point of true values may fall into sidelobes, which will inevitably result in a large estimation error. To demonstrate the relation between true values point and the global optimum point, and the difficulty level of finding the global optimum, we plot a likelihood function curve for the one-dimensional problem of estimating the unknown LoS path delay only in Fig. 3, where the red circle marks the point of the true values of LoS path delay and the red star marks the point of global optimum. We set $M = 2$, $K = 2$, $f_{c,1} = 1.8$ GHz, $f_{c,2} = 2.6$ GHz, $N_1 = N_2 = 1024$, $f_{s,1} = f_{s,2} = 60$ KHz, the true delays $\tau_1 = 30$ ns, $\tau_2 = 59$ ns, and signal-to-noise ratio (SNR) is 5 dB. As can be seen, the likelihood function fluctuates frequently, which makes it intractable to find the global optimum. Moreover, the point of true values is not in the mainlobe, at which the global optimum locates. Even though we try our best to find the global optimum, an absolute estimation error about 27 ns is still inevitable.

Therefore, we build new signal models without inherent ambiguity to help the estimation algorithms get the global optimum solution with high probability and improve the probability that the point of true values is in the region of the mainlobe. Moreover, the signal models should reserve the structure of frequency band apertures and subcarrier apertures. Motivated by the above facts, we propose a two-stage signal model, which is transformed from the original signal model (3) by absorbing different phase terms into the complex gain.

1) Coarse Signal Model:

$$ y_{m,n} = \sum_{k=1}^{K} \alpha_k e^{-j2\pi n f_{c,m} \tau_k} e^{-j2\pi n f_{c,m} \delta_m} s_{m,n} + w_{m,n}, $$

where $\alpha_k, m = \alpha_k e^{j\varphi_m} e^{-j2\pi f_{c,m} \tau_k}, \forall k, m$. In signal model (4), we absorb the terms of random phase offset $e^{j\varphi_m}$ and carrier phase $e^{-j2\pi f_{c,m} \tau_k}$ into $\alpha_k$ and they will be estimated as a whole denoted as $\alpha_k, m$. Apparently, the original ambiguity existing in signal model (4) between $\alpha_k$ and $e^{j\varphi_m}$ has been eliminated according to this equivalent transformation and all subbands share the common sparse delay domain. As shown in Fig. 4, we depict the likelihood function based on signal model (4) with the same parameters setup as Fig. 3. The point of true values locates at the mainlobe region and the likelihood function no longer frequently fluctuates since the carrier frequency terms are reduced to zero, i.e., absorbed into $\alpha_k$. In this case, we can exploit the subcarrier apertures gain and estimate delay parameters without ambiguity. This helps to achieve a much more stable delay estimation in the coarse estimation stage whose primary purpose is to narrow down the search range to a relatively small region with high stability (probability). However, the estimation accuracy is limited, since we have...
absorbed the carrier phase term and thus cannot exploit the frequency band apertures gain, which motivates the refined signal model (5) in Stage 2.

2) Refined Signal Model:

\[ y_{m,n} = \sum_{k=1}^{K} \alpha_k' e^{-j2\pi f_c'e_c\tau_k} e^{-j2\pi n f_c\tau_k} e^{-j2\pi n s m} e^{j\phi_m} + w_{m,n}, \]

(5)

where \( f_c'e_c = f_c - f_c, e_k' = \alpha_k e^{j\phi_1} e^{-j2\pi f_c\tau_k}, \phi_m = \phi_m - \phi_1, k, m \). In the signal model (5), we absorb the random phase offset, \( e^{j\phi_1} \), and carrier phase of the first frequency subband, \( e^{-j2\pi f_c\tau_k} \), into \( \alpha_k' \) and reserve the residual term, e.g., \( e^{-j2\pi (f_c - f_c')\tau_k} \) and \( e^{j(\phi_2 - \phi_1)} \) when \( m = 2 \), in (5). After that, if we substitute the sets of variables \((|\alpha_k'| e^{j(\phi_2 - \phi_1)}, e^{j(\phi_m - \phi_1)}), \forall c, m = 2, \ldots, M \) into the term \((\alpha_k', e^{j\phi_m})\) of the refined signal model (5), i.e., \( |\alpha_k'| e^{j(\phi_2 - \phi_1)} - |\alpha_k'| e^{j(\phi_m - \phi_1)}, e^{j\phi_m} \leftarrow e^{j(\phi_m - \phi_1)}, m = 2, \ldots, M \). (Note that when \( m = 1, \phi_m = 0 \) is a constant and thus cannot be substituted), we will observe that the observation results \( y_{m,n} \) remain unchanged for \( m = 2, \ldots, M \) but vary with \( c \) for \( m = 1 \). It means that the ambiguity between \( \alpha_k \) and \( e^{j\phi_m} \) is eliminated and the refined signal model (5) is unambiguous. Compared to (4), signal model (5) has the extra structure of frequency band apertures, \( e^{-j2\pi f_c'e_c\tau_k} \), and residual random phase offset, \( e^{j\phi_m} \). Fig. 5 illustrates the likelihood function based on signal model (5) with the same parameters setup as Fig. 3, where the likelihood function fluctuates less than that in Fig. 3 since the carrier frequency terms are reduced from \( f_c, c \) to \( f_c'e_c, c \) and the point of true values is now in the mainlobe region. Moreover, we observe that the mainlobe is sharper than that in Fig. 4 due to the existence of frequency band apertures, which leads to a potential performance improvement. However, there are still numerous local optimums in the likelihood functions, because the local optimums not only depend on the carrier frequency term, but also the frequency band apertures. How to find the global optimum by exploiting the frequency band apertures gain and simultaneously overcoming the oscillation phenomenon of the likelihood function with low-complexity is a challenging problem. To solve this problem, we propose the PSO-LS algorithm later.

In summary, both two signal models are essential and we are able to find the global optimum with much higher probability based on the proposed two-stage signal model. In Stage 1 based on signal model (4), we exploit the subcarrier apertures and an initial stable delay estimation result can be provided to narrow down the search range for the global delay estimation in Stage 2. By doing this, numerous bad local optimums are excluded from the global search space in Stage 2. Then, in Stage 2, signal model (5) is used for providing a refined estimation by exploiting both the subcarrier apertures gain and the frequency band apertures gain. The proposed refined signal model (5) significantly attenuates the oscillation by reducing the carrier frequency term and therefore reduces the difficulty of getting the global optimum.

B. Outline of the TSGE Scheme

Based on the two-stage signal model, the TSGE scheme is depicted as follows:

- Stage 1: We build coarse signal model (4) and perform an initial delay estimation using the proposed Turbo-BI algorithm. By doing this, we exploit channel sparsity over delay domain and the subcarrier apertures gain. Then, we provide the estimation result to Stage 2.
- Stage 2: Based on the coarse estimation result from Stage 1 and the refined signal model (5), a more refined delay estimation is performed. To fully exploit subcarrier and frequency band apertures gains and overcome the difficulty of finding the global optimum as shown in Fig. 5, we propose the PSO-LS algorithm.

The overall algorithm is summarized as in Algorithm 1. The details of the Turbo-BI and PSO-LS algorithms are presented in Section IV and V, respectively.

IV. Turbo-BI Algorithm in Stage 1

A. Common Support Based Sparse Representation

We first describe the sparse representation over delay domain for the coarse signal model (4), which is a necessary step before employing the sparse recovery methods, e.g., Turbo-BI algorithm. One commonly used method is to define a uniform grid \( D = \{\bar{d}_1, \ldots, \bar{d}_L\} \) of \( L (L \gg K) \) delay.
that denotes the linear steering vector,\( \tau \) example, if the whose non-zero elements correspond to the true delays. For and the corresponding true delay is \( D \) make the equation (6) hold approximately, which leads to \( L \) in [16], [17], and [18] employ a dense grid (\( \Delta \) grid vector leakage caused by delay mismatch and high computational problem (27). Given the channel support vector, \( s \), the conditional prior distribution of the elements of \( x_m, \forall m \) is independent and can be written as
\[
 p(x_{m,l} | s_l) = (1 - s_l) \delta (x_{m,l}) + s_l \mathcal{CN}(x_{m,l}; 0, \sigma^2_{m_l}). \quad (8)
\]

\section{B. Turbo-BI Algorithm}

To achieve the goal of delay estimation, we need to estimate the sparse vector \( x \), support vector \( s \), and the uncertain parameters \( \xi \equiv (\delta^T, (\Delta \tau)^T)^T \), where \( \delta \equiv [\delta_1, \ldots, \delta_M]^T \), given the observations \( y = [y_1; \ldots; y_M] \in \mathbb{C}^{N \times 1} \). In particular, for given \( \xi \), we are interested in computing the conditional marginal posterior \( p(x|y, \xi) \). For the uncertain parameters \( \xi \), we adopt the maximum a posteriori (MAP) estimation method as
\[
 \xi^* = \arg\max_{\xi} p(x|y, \xi) \quad (9)
\]

\section{B. Turbo-BI Algorithm}

To achieve the goal of delay estimation, we need to estimate the sparse vector \( x \), support vector \( s \), and the uncertain parameters \( \xi \equiv (\delta^T, (\Delta \tau)^T)^T \), where \( \delta \equiv [\delta_1, \ldots, \delta_M]^T \), given the observations \( y = [y_1; \ldots; y_M] \in \mathbb{C}^{N \times 1} \). In particular, for given \( \xi \), we are interested in computing the conditional marginal posterior \( p(x|y, \xi) \). For the uncertain parameters \( \xi \), we adopt the maximum a posteriori (MAP) estimation method as
\[
 \xi^* = \arg\max_{\xi} p(x|y, \xi) \quad (9)
\]

It is a high-dimensional non-convex objective function and we cannot obtain a closed-form expression due to the multi-dimensional integration over \( x \) and \( s \), making it difficult to directly maximize \( \ln p(y, \xi) \). To handle this issue, we adopt majorization-minimization (MM) method to construct a surrogate function and then use alternating optimization (AO) method to find a stationary point of (9). Inspired by the expectation-maximization (EM) method [33], we propose a Turbo-BI algorithm performing iterations between following two steps until convergence.

- Turbo-BI-E Step: For given \( \xi \), approximately calculate the posterior \( p(x|y, \xi) \) by combining the message passing and linear minimum mean square error (LMMSE) approaches via a turbo framework, as will be elaborated in Subsection IV-B.1.
- Turbo-BI-M Step: Given \( p(x|y, \xi) \), construct a surrogate function for \( \ln p(y, \xi) \) based on the MM method, partition \( \xi \) into \( B \) blocks \( \xi = (\xi_1, \ldots, \xi_B) \), then alternatively
maximize the surrogate function with respect to $\xi_j$, $j = 1, \ldots, B$, as will be elaborated in Subsection IV-B.2.

1) Turbo-BI-E Step: The Turbo-BI-E step contains two modules, as illustrated in Fig. 6. Module A is a LMMSE estimator based on the observation $y_m, \forall m$ and Module B is a sparsity combiner that utilizes the sparsity information of $x_m, \forall m$ to further refine the estimation results. The extrinsic estimation of one module will be treated as a prior mean for the other module in the next iteration. Based on the iterations between these two modules, the channel prior information and observations information are combined to be exploited. Specifically, in Module A, we assume that $x_m, \forall m$ follows a Gaussian distribution $CN\left(x_m; x_{A,m}^{pri}, \sigma_{A,m}^2 I\right)$, where $x_{A,m}^{pri}$ and $v_{A,m}^{pri}$ are the extrinsic message output from Module B. We define $\Phi_m = S_m A_m (\Delta \tau) \in \mathbb{C}^{N_m \times L}$ as the measurement matrix and we can obtain the conditional distribution $p(y_m | x_m) = CN\left(\Phi_m x_m, \sigma_n^2 I\right)$. Then, the posterior distribution of $x_m$ is given by $p(x_m | y_m) = CN\left(x_{A,m}^{post}, V_{A,m}^{post}\right)$, where

$$V_{A,m}^{post} = \left(\frac{\Phi_m^H \Phi_m}{\sigma_n^2} + \frac{1}{v_{A,m}^{pri}} I\right)^{-1}, \quad x_{A,m}^{post} = V_{A,m}^{post} \frac{x_{A,m}^{pri}}{v_{A,m}^{pri}} + \frac{\Phi_m^H y_m}{\sigma_n^2}.$$

Note that in [34], the singular value decomposition (SVD) of $\Phi_m$ is utilized to reduce the computational complexity of $V_{A,m}^{post}$, which involves a matrix inverse operation. However, the complexity of the matrix inverse operation is acceptable. It is because $L \ll N_m$, the complexity of the matrix inversion operation, $O\left(L^3\right)$, is relatively low. Finally, we calculate the extrinsic message passing:

$$v_{A,m,i}^{ext} = \left(\frac{1}{v_{A,m,i}^{post}} - \frac{1}{v_{A,m,i}^{pri}}\right)^{-1},$$

$$x_{A,m,i}^{ext} = v_{A,m,i}^{ext} x_{A,m,i}^{post} + \frac{x_{A,m,i}^{pri}}{v_{A,m,i}^{pri}},$$

where $v_{A,m,i}^{post}$ is the $i$-th diagonal element of $V_{A,m}^{post}$.

In Module B, we assume that $x_{B,m}^{pri}$ is modeled as an AWGN observation of $x_m$ [35], [36]:

$$x_{B,m}^{pri} = x_m + z_m,$$

where $z_m \sim CN\left(0, v_{B,m}^{pri} I\right)$ is independent of $x_m$, $x_{B,m}^{ext} = x_{A,m}^{ext}$ and $v_{B,m}^{ext} = \frac{1}{L} \sum_{l=1}^{L} v_{A,m,i}^{ext}$ are the extrinsic message from Module A. Based on (14), we combine the sparsity prior information of $x_m$ and the extrinsic messages from Module A, aiming at calculating the posterior distributions $p\left(x_{m,l} | x_{B,m}^{pri}\right)$ by performing the sum-product message passing (SPMP) [37] over the factor graph, where $x_{B,m}^{pri} = \left(x_{B,1,m}^{pri} T, \ldots, x_{B,M,m}^{pri} T\right)$. Particularly, the factor graph of the joint distribution $p\left(x_{B,m}^{pri} : x_m\right)$ is shown in Fig. 7, where the function expression of each factor node is listed in Table I. At subband $m$, the factor graph is denoted by $G_m$. As can be seen, factor graphs $G_m$’s share the common support vector $s$.

We now outline the message passing scheme on graph $G$. The details are elaborated in Appendix A. According to the sum-product rule, the message passing over the path $x_{m,l} \rightarrow f_{m,l} \rightarrow s_l$ are given by (33) and (34). Then the message is passed back over the path $s_l \rightarrow f_{m,l} \rightarrow x_{m,l}$ using (35) and (36). After calculating the updated messages $\{v_{f_{m,l} \rightarrow x_{m,l}}^{ext}\}$, the approximate posterior distributions are given by

$$p\left(x_{m,l} | x_{B,m}^{pri}\right) \approx v_{f_{m,l} \rightarrow x_{m,l}}^{ext} p_{gm_{l} \rightarrow x_{m,l}}^{ext} p_{gm_{l} \rightarrow x_{m,l}},$$

where $v_{gm_{l} \rightarrow x_{m,l}}^{ext} (x_{m,l}) = CN\left(x_{m,l}; x_{B,m}^{pri}, v_{B,m}^{pri}\right)$. Then the posterior mean and variance are given by

$$x_{B,m,l}^{post} = \mathbb{E}\left(x_{m,l} | x_{B,m,l}^{pri}\right) = \int x_{m,l} p\left(x_{m,l} | x_{B,m,l}^{pri}\right) dx_{m,l},$$

$$v_{B,m,l}^{ext} = \frac{1}{L} \sum_{l=1}^{L} \text{Var}\left(x_{m,l} | x_{B,m,l}^{pri}\right) = \frac{1}{L} \sum_{l=1}^{L} \int \left| x_{m,l} - \mathbb{E}\left(x_{m,l} | x_{B,m,l}^{pri}\right) \right|^2 p\left(x_{m,l} | x_{B,m,l}^{pri}\right) dx_{m,l}. \quad (17)$$

Finally, based on the derivation in [38], the extrinsic update for Module A can be calculated as

$$x_{A,m}^{pri} = x_{A,m}^{ext} = v_{A,m}^{ext} \left(\frac{x_{B,m}^{post}}{v_{B,m}^{post}} - \frac{x_{B,m}^{pri}}{v_{B,m}^{pri}}\right),$$

$$v_{A,m}^{ext} = \left(\frac{1}{v_{B,m}^{post}} - \frac{1}{v_{B,m}^{pri}}\right)^{-1}. \quad (19)$$
2) Turbo-BI-M Step: In the M-step, we construct a surrogate function at fixed point $\xi$ for the objective function (9) based on the MM method as

$$u(\xi; \hat{\xi}) = \int p(x | y; \hat{\xi}) \ln \frac{p(x, y; \xi)}{p(x | y; \hat{\xi})} d\xi,$$

which satisfies basic properties $u(\xi; \hat{\xi}) \leq \ln p(y; \xi), \forall \xi$;

$$u(\xi; \hat{\xi}) = \ln p(y; \hat{\xi}), \forall \xi; \frac{\partial u(\xi; \hat{\xi})}{\partial \xi} = \frac{\partial \ln p(y; \xi)}{\partial \xi} = \forall \xi.$$  

Then, we partition $\xi$ into $B = 2$ blocks with $\xi_1 = \delta$, $\xi_2 = \Delta \tau$ based on their distinct physical meaning, and alternatively update $\xi_1$ and $\xi_2$ as

$$\delta^{(i+1)} = \arg\max_{\delta} u(\delta, \Delta \tau^{(i)}; \delta^{(i)}, \Delta \tau^{(i)}),$$

$$\Delta \tau^{(i+1)} = \arg\max_{\Delta \tau} u(\delta^{(i+1)}, \Delta \tau; \delta^{(i)}, \Delta \tau^{(i)}).$$

Since the optimization problems (21) and (22) are non-convex and it is hard to find their optimal solutions, we apply a one-step gradient update for $\delta$ and $\Delta \tau$ as follows:

$$\delta^{(i+1)} = \delta^{(i)} + \gamma_\delta \cdots \zeta^{(i)} \cdots,$$

$$\Delta \tau^{(i+1)} = \Delta \tau^{(i)} + \gamma_{\Delta \tau} \cdots \zeta^{(i)} \cdots,$$

where $\gamma_\delta$ and $\gamma_{\Delta \tau}$ are the step size determined by the Armijo rule [39], $\zeta^{(i)}$ and $\zeta_{\Delta \tau}$ are the gradients of the objective function in (21) and (22) with respect to $\delta$ and $\Delta \tau$, respectively. The detailed derivations for $\zeta^{(i)}$ and $\zeta_{\Delta \tau}$ are presented in Appendix B. Moreover, the convergence of this in-exact MM algorithm to a stationary point can be guaranteed [40, Theorem 1].

C. Summary of the Turbo-BI Algorithm and Complexity Analysis

The Turbo-BI algorithm is summarized in Algorithm 2. Finally, we analyze the computational complexity of the proposed Turbo-BI algorithm. The computational complexity of the Turbo-BI-E step is dominated by the matrix multiplication $\Phi_m^H \Phi_m$ in (10), which is $O(N_m L^2)$.

In the Turbo-BI-M step, the computational complexity of choosing the right step size mainly depends on calculating the cost function. We denote the number of calculating the cost function for $\gamma_\delta$ and $\gamma_{\Delta \tau}$ in every backtracking line search as $R_{\delta,1}$ and $R_{\Delta \tau,2}$, respectively. Then, the complexity of choosing $\gamma_\delta$ and $\gamma_{\Delta \tau}$ is $O(N L^2 R_{\delta,1})$ and $O(N L^2 R_{\Delta \tau,2})$, respectively. Besides, the complexity in calculating $\zeta^{(i)}$ and $\zeta_{\Delta \tau}$ are $O(NL^2)$ and $O(NL)$ based on matrix multiplication, respectively. Hence, the overall computational complexity of Turbo-BI algorithm is $O(NL^2 I_{in} + NL^2 (R_{\delta,1} + R_{\Delta \tau,2}))$ per iteration, where $I_{in}$ denotes the number of Turbo iterations for convergence. Note that though the complexity of Turbo-BI algorithm is proportional to the square of $L$, the grid size $L$ is not the dominate factor impacting the computational complexity. In fact, $L$ can be small in our problem and we do not need a very dense grid to guarantee the estimation accuracy. It is because we adopt a dynamic grid adjustment strategy in Turbo-BI algorithm, where the grids can be updated in the M-step to obtain a high-accuracy delay estimation. Moreover, the maximum delay spread of the channel is limited compared to the cyclic prefix (CP). Hence, we can use some prior knowledge about the maximum delay spread to predetermined the grid size within a given range.

| Factor | Distribution | Functional form |
|--------|--------------|-----------------|
| $g_{m,l}(x_{B,m,l}, x_{m,l})$ | $p(x_{B,m,l} | x_{m,l})$ | $\mathcal{C}N(x_{m,l}; x_{B,m,l}^*, v_{B,m,l}^*)$ |
| $f_{m,l}(s_l, x_{m,l})$ | $p(x_{m,l} | s_l)$ | $(1 - s_l) \delta(x_{m,l}) + s_l \mathcal{C}N(x_{m,l}; 0, \sigma^2_{m,l})$ |
| $d_i(s_l)$ | $p(s_l)$ | $p_s$ |
V. PSO-LS ALGORITHM IN STAGE 2

In Stage 2, we aim to fully exploit the frequency band apertures gain and perform a refined delay estimation based on the estimation result from Stage 1. First, we reformulate the refined estimation signal model (5) as a linear form:

\[ y = H(\theta)x + w, \]

where

\[ H(\theta) = \begin{bmatrix} h_{11} & \cdots & h_{1K} \\ \vdots & \ddots & \vdots \\ h_{M1} & \cdots & h_{MK} \end{bmatrix} \in \mathbb{C}^{N \times K}, \]

\[ h_{mk}(n) = e^{-j2\pi(f_m \tau + n f_s) e^{-j2\pi n f_s \delta_m e^{j\phi_m}}, \]

\[ \phi_m \] denotes the \( n \)-th element of the vector \( \phi \), \( \theta = [\tau_1, \ldots, \tau_K, \delta_1, \ldots, \delta_M, \phi^T_1, \ldots, \phi^T_M] \in \mathbb{R}^{(K+2M-1) \times 1} \)

denotes the vector consisting of unknown parameters, \( x = [\alpha_1, \ldots, \alpha_K] \in \mathbb{C}^{K \times 1} \), and \( w \sim \mathcal{CN}(0, \sigma^2 I) \in \mathbb{C}^{N \times 1} \).

We adopt the MAP method for estimation, that takes the prior information of \( \delta_m \) into consideration. Then, the optimization problem can be formulated as

\[ \mathcal{P}_1: \max_{\theta, x} \ln p(y|x) + \sum_{m=1}^{M} \ln p(\delta_m), \]

s.t. \( 0 \leq \phi_m \leq 2\pi, \forall m \in \{2, \ldots, M\} \),

where \( p(y|x) \propto \exp\left(-\frac{|y-H(\theta)x|^2}{\sigma^2_n}\right) \) is the likelihood function. After an equivalent transformation, \( \mathcal{P}_1 \) can be reformulated as

\[ \mathcal{P}_2: \min_{\theta, x} \frac{|y - H(\theta)x|^2}{\sigma^2_n} + \sum_{m=1}^{M} \frac{\delta_m^2}{2\sigma^2_p} \]

s.t. \( 0 \leq \phi_m \leq 2\pi, \forall m \in \{2, \ldots, M\} \).

The non-convex problem \( \mathcal{P}_2 \) has a multimodal and multi-dimensional objective function, which is extremely challenging to solve due to the existence of numerous local optima. In this case, conventional algorithms, such as gradient descent and exhaustive search algorithms, have high-computational complexity or are easily trapped into local optima. To overcome these drawbacks, we adopt the PSO method to find a good solution for the non-convex optimization problem \( \mathcal{P}_2 \). Though the PSO-based algorithms may also trap into local optima and cannot guarantee the global optimum, they have been shown to have low computational complexity and a strong optimization ability for complex multimodal optimization problems that are able to readily escape from local optima [41], [42], [43]. In addition, when the number of particles in PSO goes to infinity, the convergence to the global optimum can be guaranteed [44], which implies that PSO with sufficient particles is able to find the global optimum with high probability.

In the PSO method, we employ a number of particles which are potential solutions, to find the optimal solution by iterations, where the search space is bounded by the constraints of the target optimization problem [45]. Generally, PSO starts with a random initialization of the particles’ locations in a large search space. In our problem \( \mathcal{P}_2 \), however, we can narrow down the search space based on coarse estimation results from Stage 1. Specifically, we set the search space as

\[ S = [\beta - e, \beta + e] \in \mathbb{R}^{D \times 1}, \]

where \( \beta \) has the values consisting of the coarse estimation results, \( e \) is the search range obtained by evaluating the mean squared error (MSE) of coarse estimation results based on offline training or evaluating the Cramér-Rao bound (CRB) values based on coarse signal model (4), and \( D \) denotes the dimension of the particles (search space). Since there is no estimation for \( \phi_m \) in Stage 1, the search space for \( \phi_m \) is set as \([0, 2\pi]\). Note that \( S \) is a real set with dimension \( D = 3K + 2M - 1 \), since the \( K \)-dimension complex vector \( x \) can be seen as a \( 2K \)-dimension real vector in a real domain search space. Then, for the \( i \)-th iteration, each particle comes to a better location by updating its velocity and position based on the equations:

\[ V_{q,d}^{(i+1)} = \omega V_{q,d}^{(i)} + c_1 r_1 \cdot q,d \cdot p_{best}^{(i)} - X_{q,d}^{(i)} \]

\[ + c_2 r_2 \cdot q,d \cdot (g_{best}^{(i)} - X_{q,d}^{(i)}), \]

\[ X_{q,d}^{(i+1)} = X_{q,d}^{(i)} + V_{q,d}^{(i+1)}, \]

where \( c_1 \) and \( c_2 \) are acceleration coefficients, and \( \omega \) is the inertia factor, which is proposed to balance the global and local searchabilities [42], [45]. Note that \( X_{q,d}^{(i)} \) and \( V_{q,d}^{(i)} \) are the position and velocity of the \( d \)-th element of the \( q \)-th particle, respectively, where \( q \in \{1, 2, \ldots, Q_p\}, d \in \{1, 2, \ldots, D\} \) with \( Q_p \) denoting the number of particles. \( r_1 \cdot q,d \) and \( r_2 \cdot q,d \) are two independent random numbers uniformly distributed within \([0, 1]\). Besides, \( g_{best} \) denotes the best position in the whole swarm and \( p_{best}^{(i)} \) denotes the best position values at the \( d \)-th dimension that the \( q \)-th particle has come across so far at the \( i \)-th iteration. The goodness of the particle position is measured by the objective function value in \( \mathcal{P}_2 \) (also called fitness in PSO algorithms). In our problem, a smaller fitness indicates a better position.

For getting a better solution in the PSO algorithm, we aim to reduce the dimension of the search space in problem \( \mathcal{P}_2 \) and thus propose the PSO-LS algorithm. We first employ primal-decomposition theory to decouple the problem \( \mathcal{P}_2 \) as:

\[ \mathcal{P}_3:\min_\theta \frac{|y - H(\theta)g^*(\theta)|^2}{\sigma^2_n} + \sum_{m=1}^{M} \frac{\delta_m^2}{2\sigma^2_p} \]

s.t. \( 0 \leq \phi_m \leq 2\pi, \forall m \in \{2, \ldots, M\} \),

where \( g^*(\theta) = \text{argmin}_{\theta} |y - H(\theta)x|^2 \). Then we can obtain a LS solution of \( g^*(\theta) \) as

\[ g^*(\theta) = (H(\theta)^H H(\theta))^{-1} H(\theta)^H y. \]

Using this method, the dimension of search space is reduced from \( \mathbb{R}^{3K + 2M - 1} \) to \( \mathbb{R}^{K + 2M - 1} \). Therefore, for given the number of particles, the global optimum is able to be found with higher probability. To verify this, we compare the performance with the pseudo global optimum found by running the PSO-LS with a large number of particles \((Q_p = 1000)\) around the true delays. The simulation results show that the proposed
Algorithm 3 PSO-LS Algorithm

Input: \( y \), search space \( S \), \( Q_p \), maximum iteration number \( I_{PSO} \), threshold \( \epsilon \).
Output: \( \text{gbest} \).

1: Initialize \( X_q^{(1)} \) within the search space \( S \) in (28), \( V_q^{(1)} \), \( \text{gbest}_q^{(1)} \), \( \text{pbest}_q^{(1)} \), \( \forall q, d \).
2: for \( i = 1, \ldots, I_{PSO} \) do
3:   for \( q = 1, \ldots, Q_p \) do
4:     Update \( q \) using (29).
5:   end for
6:   Update the position of particle \( q \) based on the objective in (31).
7:   Update \( \text{gbest}_q^{(i+1)} \) and \( \text{pbest}_q^{(i+1)} \), \( \forall q, d \).
8:   end for
9:   if \( \| \text{gbest}^{(i+1)} - \text{gbest}^{(i)} \| \leq \epsilon \) then
10:      break
11:   end if
12: end for

PSO-LS with only a moderate number of particles \((Q_p = 100)\) can already achieve a performance close to the pseudo global optimum. Then, we analyze the computational complexity of the PSO-LS algorithm. The complexity mainly depends on the calculations of the objective function in (31), which is \( O(NK^2Q_p) \) per iteration. Finally, the proposed PSO-LS algorithm is presented in Algorithm 3.

VI. Simulation Results

In this section, we provide numerical results to evaluate the performance of the proposed schemes and draw useful insights. In the default setup, we consider that CSI samples are collected using one OFDM training symbol with subcarrier spacing \( f_{s1} = f_{s2} = 60 \) KHz and subband bandwidth \( B_1 = B_2 = 40 \) MHz at \( M = 2 \) subbands, with central frequencies \( f_{c1} = 1.80 \) GHz and \( f_{c2} = 2.02 \) GHz. \( \varphi_m, \forall m \) and \( \delta_m, \forall m \) are generated following a uniform distribution within \([0, 2\pi]\) and a Gaussian distribution \( N(0, \sigma^2) \), respectively. We use the Quadriga platform to generate multiband CSI samples in an indoor factory (InF) scenario, which is depicted in 3GPP R16 [46]. Other system parameters are set as follows unless otherwise specified: the grid size \( L = 26 \), the initial grid location \( \{ \vec{d}_1, \ldots, \vec{d}_L \} \) follows a uniformly-spaced location within \([-125,500] \) ns and the adjacent spacing is 25 ns, SNR is 7 dB, \( I_{EM} = 50, I_{PSO} = 500, Q_p = 100, \epsilon = 10^{-5}, c_1 = 2.5, c_2 = 0.5, \) and \( \omega = 0.99 - \frac{0.79}{I_{PSO}} t \) for the \( t \)-th PSO iteration. To assess the performance of the schemes, we adopt the empirical cumulative distribution function (CDF) of the LoS path delay estimation errors using 200 Monte Carlo trials.

For comparison, we consider the following four benchmark schemes:

- **Turbo-BI algorithm**: We adopt the coarse estimation results of the Turbo-BI algorithm in Stage 1 as one of the benchmarks.
- **Multiband weighted delay estimation (MBWDE) algorithm [20]**: It adopts a weighted subspace fitting algorithm for delay estimation, which is able to exploit the frequency band apertures gain.

- **Two-stage gradient descent (TSGD) scheme**: This scheme has the same coarse estimation implementation as TSGE in Stage 1 and then the gradient descent method is employed to perform refined estimation in Stage 2.

- **Ideal gradient descent (IGD) scheme**: This scheme adopts gradient descent method based on problem \( \mathcal{P}_2 \), where we set the initial point to be the true values. In fact, this scheme is infeasible because we cannot know the true values of the estimation parameters in practical environment. In the simulations, however, we can regard this scheme as a performance upper bound for practical delay estimation algorithms that consider phase distortion factors.

A. Convergence of the Proposed Algorithms

We first illustrate the convergence behavior of the proposed TSGE scheme, i.e., Turbo-BI and PSO-LS algorithms. As illustrated in Fig. 8, Turbo-BI converges within 50 iterations and PSO-LS converges within about 200 iterations (up to a small convergence error). Due to the fact that the objective function of PSO-LS algorithm contains numerous local optimums, the algorithm at each iteration may get stuck and then get out from local optimums and thus the curve in Fig. 8b is not strictly decreasing. However, from the overall trend, we observe that the algorithm is able to escape from local optimums, making the objective value gradually decrease and finally converges.

B. Impact of the Factor \( \delta \)

We study the impact of the receiver timing offset \( \delta \). Since the InF scenario does not consider the effect of \( \delta \), we construct a two-path channel model with Rayleigh distributed magnitudes. The delays are set to follow a uniform distribution within \([20, 200] \) ns. Fig. 9 depicts the CDF of LoS path delay estimation errors for different standard deviation \( \sigma_p \) achieved by TSGE. As can be seen, the factor \( \delta \) has significantly degraded the estimation performance. Besides, when \( \delta \) is considered, the estimation performance mainly depends on the prior standard deviation \( \sigma_p \), where the delay estimation errors increase with \( \sigma_p \). It is reasonable since a larger \( \sigma_p \) means less prior information of \( \delta \) we have. Therefore, in the following simulations, we assume that \( \sigma_p = 0 \) ns in order to avoid its effect on the estimation performance.

C. Performance of TSGE Scheme

In Fig. 10, we compare the delay estimation performance of the TSGE scheme with benchmarks. First, we observe that the CS-based schemes (i.e., TSGE, Turbo-BI, and TSGD) and the IGD scheme achieve a better performance than the subspace-based algorithm, i.e., MBWDE. On one hand, it is mainly because that the received CSI samples are collected using only a single OFDM training symbol, which leads to a limited ability of subspace-based algorithms to suppress the noise interference. In contrast, CS-based schemes have
a strong ability of reducing the effects of noise by signal sparse reconstruction, thus they achieve a better performance. On the other hand, the MBWDE algorithm is proposed based on an ideal multiband channel model, where the phase distortion factors $\phi_m$ are assumed to be perfectly eliminated via a handshaking procedure under the assumption of perfect channel reciprocity. Hence, in our simulation scenarios with $\phi_m$ explicitly considered, MBWDE does not perform well though it is able to exploit the frequency band apertures gains. As illustrated in Fig. 10, MBWDE in an ideal scenario, i.e., $\phi_m = 0$, $\forall m$, has a significant performance improvement.

Second, TSGE significantly outperforms Turbo-BI and TSGD. This is because TSGE exploits the extra frequency band apertures gain compared to Turbo-BI. Moreover, note that the likelihood function based on the refined signal model (5) has numerous local optima, which requires a strong global search ability for the algorithms in Stage 2. Compared to the gradient descent algorithm in the TSGD scheme, the PSO-LS algorithm in the TSGE scheme has stronger global search ability, thus achieves higher estimation accuracy than TSGD.

Finally, it is observed that the CDF curve of TSGE is nearly close to that of IGD, which indicates a negligible performance loss and validates the effectiveness of the proposed algorithms. In Table II and III, we present the computing cost and the estimation accuracy of the considered algorithms, i.e., the proposed TSGE combined with Turbo-BI and PSO-LS algorithms with different iteration numbers, TSGD, MBWDE, and the primal PSO algorithm based on the problem $P_2$. The computing cost is characterized by the spent CPU time when running the algorithms with the Intel Xeon 6248R CPU. It is observed that our proposed TSGE scheme is able to achieve better estimation accuracy than TSGD with comparable computing cost (TSGE $46.4s$ vs TSGD $44.8s$). Although the CPU time of MBWDE is smaller, it performs poorly in practical scenarios with phase distortion factors. In addition, we can see that TSGE is able to achieve a flexible trade-off between the computing cost and estimation performance. In particular, when we have sufficient computing resource, we can use a large number of particles and iterations to pursue the highest estimation accuracy. Conversely, when the computing resource is limited, we can use a small number of particles and iterations to achieve a relatively accurate estimation.

In Table III, it can be seen that PSO-LS is able to achieve higher estimation accuracy with less computing cost than primal PSO, which validates the effectiveness of the proposed PSO-LS algorithms. Moreover, the computing cost of Turbo-BI is acceptable with relatively high estimation accuracy compared to other benchmarks. In fact, if the computing resource is limited, we can further release the computing resource by reducing the iteration numbers of Turbo-BI for the implementation of the PSO-LS algorithm.

Hence, we conclude that the TSGE scheme combined with Turbo-BI and PSO-LS algorithm is able to achieve a better trade-off between performance and complexity.
TABLE II
CPU TIME AND RMSE COMPARISON

|                | TSGE$^1$ | TSGE$^2$ | TSGD  | MBWDE |
|----------------|----------|----------|-------|-------|
| CPU time (s)   | 46.4     | 10.1     | 44.8  | 0.8   |
| RMSE (ns)      | 0.2876   | 0.5425   | 0.4338| 1.671 |

TABLE III
CPU TIME AND RMSE FOR TURBO-BI AND PSO BASED ALGORITHMS

|                | Turbo-BI | PSO-LS | Turbo-BI | PSO-LS | Turbo-BI | Primal PSO |
|----------------|----------|--------|----------|--------|----------|------------|
|                | ($I_{EM} = 50$) | ($I_{PSO} = 500$) | ($I_{EM} = 30$) | ($I_{PSO} = 100$) | ($I_{EM} = 30$) | ($I_{PSO} = 500$) |
| CPU time (s)   | 13.6     | 32.8   | 8.8      | 1.3    | 8.8      | 20.3       |
| RMSE (ns)      | 0.3733   | 0.2876 | 0.7156   | 0.5425 | 0.7156   | 0.5517     |

D. Impact of Frequency Band Spacing

Fig. 11 illustrates the root mean square error (RMSE) of the LoS delay estimation for different schemes versus frequency band spacing, when SNR = 10 dB and bandwidth $B_1 = B_2 = 60$ MHz. We define $\text{RMSE}(\hat{\tau}) = \sqrt{\mathbb{E}\{ (\hat{\tau} - \tau)^2 \}}$, where $\hat{\tau}$ is the estimated LoS delay and $\tau$ is the true LoS delay. In particular, we fix $f_{c,1}$ and change $f_{c,2}$ for different frequency band spacings. It can be seen that the RMSE decreases with the increase of frequency band spacing for MBWDE, TSGE, and IGD. This is because these schemes are able to exploit frequency band apertures gain, which enlarges as frequency band spacing increases. Furthermore, the TSGD scheme has a better performance than the Turbo-BI algorithm when frequency band spacing is narrow, but the performance gap decreases as the frequency band spacing increases. When the frequency band spacing is 260 MHz, the performance of TSGD is even worse than Turbo-BI. This is reasonable since the oscillation of the likelihood function becomes more violent with the increase of frequency band spacing, which causes more bad local optima and makes it more difficult to fully exploit the frequency band apertures gain. However, TSGD has a limited global search ability, which results in a high probability of being stuck into local optima. Hence, TSGD can only exploit part of frequency band apertures gain and has a poor delay estimation performance. Besides, TSGE has more stable performance than TSGD due to its strong global optimization ability. Finally, we observe that the performance of Turbo-BI is irrelevant to the frequency band spacing since it only exploits subcarrier apertures gain in Stage 1.

E. Impact of SNR

In Fig. 12, we show the impact of SNR on the delay estimation performance. As can be seen, the RMSE of all schemes decreases with the increase of SNR because of the reduction of noise interference to delay estimation. Besides, the performance gap of TSGE and IGD over Turbo-BI becomes larger as SNR increases.

F. Impact of the Bandwidth

In Fig. 13, we investigate the RMSE of the LoS delay estimation versus the bandwidth with SNR = 10 dB. It is observed that the delay estimation accuracy increases as the bandwidth, which is mainly because larger bandwidth leads to more subcarrier apertures gain. Furthermore, we observe that the performance gain of TSGE over MBWDE increases with bandwidth.

G. Impact of Multipath Environment

We first investigate the effect of LoS power ratio on the LoS path delay estimation performance with different numbers of frequency subbands, where the LoS power ratio is defined as the power of the LoS path over the total channel power given by $\rho = \frac{|\alpha_1|^2}{\sum_{k=1}^{K} |\alpha_k|^2}$. The simulation scenario is the same as
For the case $M$ the LoS path delay estimation performance as illustrated in Fig. 14a, the performance of TSGE outperforms other schemes for all $\rho$ regimes. Besides, the performance of all schemes improves as $\rho$ increases since: (i) The effective SNR for the LoS path increases; (ii) The multipath interference from the NLoS path reduces when $\rho$ increases. Furthermore, it is observed that increasing the number of frequency subbands leads to a significant performance improvement due to more available multiband gains.

Then, we investigate the effect of delay separation on the LoS path delay estimation performance as illustrated in Fig. 14b. As can be seen, the RMSE decreases rapidly and finally converges with the increase of delay separation. Moreover, it is observed that TSGE has a significant performance improvement as compared to the Turbo-BI and MBWDE algorithms.

VII. Conclusion

In this paper, we studied a delay estimation problem in the multiband OFDM system with phase distortion factors considered. We proposed a novel two-stage global estimation scheme that fully exploits the multiband gains to improve the delay estimation performance. In particular, in Stage 1, we build a coarse signal model and then achieve a coarse delay estimation using the Turbo-BI algorithm. Then, with the help of the coarse estimation results, we further conducted a more refined delay estimation by employing the proposed PSO-LS algorithm based on the refined signal model in Stage 2. Finally, simulation results validated the effectiveness of our proposed TSGE scheme. Future work can consider a dynamic multiband based localization problem in a multiple-input multiple-output (MIMO) system with Doppler effect considered.

Appendix

A. Message Passing for Module B of Turbo-BI

The message from variable node $x_{m,l}$ to factor node $f_{m,l}$ is given by

$$v_{x_{m,l} \rightarrow f_{m,l}}(x_{m,l}) = \mathcal{CN}(x_{m,l}; x_{B,m,l}^{pri}, v_{B,m}^{pri}),$$

(33)

The message from factor node $f_{m,l}$ to variable node $s_l$ is given by

$$v_{f_{m,l} \rightarrow s_l}(s_l) \propto \int f_{m,l}(x_{m,l}, s_l) v_{x_{m,l} \rightarrow f_{m,l}}(x_{m,l}) dx_{m,l} = \pi_l^{m}\delta(s_l-1) + \left(1 - \pi_l^{m}\right)\delta(s_l),$$

(34)

where

$$\pi_l^{m} = \left(1 + \frac{\mathcal{CN}(0; x_{B,m,l}^{pri}, v_{B,m}^{pri})}{\mathcal{CN}(0; x_{B,m,l}^{pri}, v_{B,m}^{pri} + \sigma_{m,l}^2)}\right)^{-1}.$$

Then, the message passed from variable node $s_l$ to factor node $f_{m,l}$ is given by

$$v_{s_l \rightarrow f_{m,l}}(s_l) \propto v_{d_l \rightarrow s_l}(s_l) \prod_{m' \neq m} v_{f_{m',l} \rightarrow s_l}(s_l) = \hat{\pi}^{m}_{l} s_l + \left(1 - \hat{\pi}^{m}_{l}\right)(1 - s_l),$$

(35)

where

$$\hat{\pi}^{m}_{l} = \frac{\pi_l^{m}}{\pi_l^{m} + \left(1 - \pi_l^{m}\right)}.$$}

Finally, the message passed from factor node $f_{m,l}$ to variable node $x_{m,l}$ is given by

$$v_{f_{m,l} \rightarrow x_{m,l}}(x_{m,l}) \propto \sum_{s_l} f_{m,l}(x_{m,l}, s_l) v_{s_l \rightarrow f_{m,l}}(s_l) = \pi_l^{m} \mathcal{CN}(x_{m,l}; 0, \sigma_{m,l}^2) + \left(1 - \pi_l^{m}\right)\delta(x_{m,l}).$$

(36)
B. Gradient Derivation in Turbo-BI-M Step

\[
\zeta_{\delta_m} = -\frac{2}{\sigma_{ns}^2} \text{Re} \left[ \mathbf{b}_m^H \mathbf{b}_m - \mathbf{b}_m^H \mathbf{y}^m \right] \\
+ \text{tr} \left( \mathbf{A}_m \Sigma_{m} \mathbf{A}_m^H \mathbf{S} \right) \\
- \frac{1}{\sigma_{p}^2},
\]

(37)

\[
\zeta_{\Delta \tau_l} = -\frac{2}{\sigma_{ns}^2} \sum_{m=1}^{M} \mathbf{S}_m^H \mathbf{S}_m \left( \mathbf{a}_m \left( \mathbf{d}_l + \Delta \tau_l \right) \right) \\
\times \left( \left( \| \mu_{l,m} \|^2 + \Sigma_{l,m} \right) - \left( \mathbf{a}_m \left( \mathbf{d}_l + \Delta \tau_l \right) \right)^H \mathbf{S}^H \right) \\
\times \left( \mu_{l,m}^H \mathbf{y}_{m,l} - \mathbf{S}_m \sum_{j \neq l} \left( \mathbf{a}_m \left( \mathbf{d}_j + \Delta \tau_j \right) \right) \right),
\]

(38)

where \( \mathbf{b}_m = \mathbf{S}_m \mu_{l,m} \mathbf{a}_m \), \( \mathbf{b}_m = \frac{\partial \mathbf{b}_m}{\partial \mathbf{a}_m} \), \( \mathbf{y}_{m,l} = \mathbf{y}_m - \mathbf{S}_m \sum_{j \neq l} \left( \mu_{j,m} \mathbf{a}_m \left( \mathbf{d}_j + \Delta \tau_j \right) \right) \), \( \mathbf{a}_m \left( \mathbf{d}_l + \Delta \tau_l \right) = \frac{\partial \mathbf{a}_m \left( \mathbf{d}_l + \Delta \tau_l \right)}{\partial \Delta \tau_l} \), and \( \mathbf{S}_m = \frac{\partial \mathbf{S}_m}{\partial \mathbf{a}_m} \). The \( (j, l) \)-th element of the posterior mean \( \mu_{l,m} \) and covariance \( \Sigma_{l,m} \) associated with \( p(\mathbf{x} | \mathbf{y}, \xi) \), which can be approximated using the \( \mathbf{x}^{\text{post}}_{A,m} \) and \( \mathbf{V}^{\text{post}}_{A,m} \) calculated in the Turbo-BI-E step.

REFERENCES

[1] J. A. del Peral-Rosado, R. Raulefs, J. A. Lopez-Salcedo, and G. Seco-Granados, “Survey of cellular mobile radio localization methods: From 1G to 5G,” IEEE Commun. Surveys Tuts., vol. 20, no. 2, pp. 1124–1148, 2nd Quart, 2018.

[2] A. Guerra, F. Guidi, and D. Dardari, “Single-anchor localization and orientation performance limits using massive arrays: MIMO vs. beamforming,” IEEE Wireless Commun., vol. 17, no. 8, pp. 5241–5255, Aug. 2016.

[3] D. Dardari, N. Decarli, A. Guerra, and F. Guidi, “LOS/NLOS near-field localization with a large reconfigurable intelligent surface,” IEEE Trans. Wireless Commun., vol. 21, no. 6, pp. 4282–4294, Jun. 2022.

[4] J. Talvitie, M. Valkama, G. Destino, and H. Wymeersch, “Novel algorithms for high-accuracy joint position and orientation estimation in 5G mmWave systems,” in Proc. IEEE Globecom Workshops (GC Wkshps), Dec. 2017, pp. 1–7.

[5] J. Talvitie, M. Koivisto, T. Levanen, M. Valkama, G. Destino, and H. Wymeersch, “High-accuracy joint position and orientation estimation in sparse 5G mmWave channel,” in Proc. IEEE Int. Conf. Commun. (ICC), May 2019, pp. 1–7.

[6] M. Chen, W. Saad, and C. Yin, “Virtual reality over wireless networks: Quality-of-service model and learning-based resource management,” IEEE Trans. Commun., vol. 66, no. 4, pp. 2234–2241, Apr. 2018.

[7] Q. Mao, F. Hu, and Q. Hao, “Deep learning for intelligent wireless networks: A comprehensive survey,” IEEE Commun. Surveys Tuts., vol. 20, no. 4, pp. 2595–2621, 4th Quart, 2018.

[8] W. Saad, M. Bennis, and M. Chen, “A vision of 6G wireless systems: Applications, trends, technologies, and open research problems,” IEEE Netw., vol. 34, no. 3, pp. 134–142, May 2020.

[9] D. Jourdan, D. Dardari, and M. Win, “Position error bound for UWB localization in dense cluttered environments,” IEEE Trans. Aerosp. Electron. Syst., vol. 44, no. 2, pp. 613–628, Apr. 2008.
Yubo Wan (Graduate Student Member, IEEE) received the B.S. degree in information engineering from Sichuan University, Chengdu, China, in 2020. He is currently pursuing the Ph.D. degree with the College of Information Science and Electronic Engineering, Zhejiang University. His current research interests include channel estimation, integrated sensing and communication, and compressive sensing.

An Liu (Senior Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Peking University, China, in 2004 and 2011, respectively. From 2008 to 2010, he was a Visiting Scholar with the Department of ECEE, University of Colorado at Boulder. He was a Post-Doctoral Research Fellow from 2011 to 2013, a Visiting Assistant Professor in 2014, and a Research Assistant Professor with the Department of ECE, HKUST, from 2015 to 2017. He is currently a Distinguished Research Fellow with the College of Information Science and Electronic Engineering, Zhejiang University. His research interests include wireless communications, stochastic optimization, compressive sensing, and machine/deep learning for communications. He is serving an Editor for IEEE Transactions on Signal Processing, an Editor for IEEE Transactions on Wireless Communications, and a member of the Signal Processing for Communications and Networking Technical Committee (SPCOM TC) of IEEE Signal Processing Society. He served as an Editor for IEEE WIRELESS COMMUNICATIONS LETTERS from 2017 to 2022.

Qiyu Hu (Graduate Student Member, IEEE) received the B.S. degree in information engineering from Zhejiang University, Hangzhou, China, in 2018, where he is currently pursuing the Ph.D. degree with the College of Information Science and Electronic Engineering. His current research interests include wireless communications, optimization theory, machine learning for physical layer communication, and semantic communication.

Mianyi Zhang received the B.S. degree in information engineering from Zhejiang University, Hangzhou, China, in 2021, where he is currently pursuing the Ph.D. degree with the College of Information Science and Electronic Engineering. His current research interests include unmanned aerial vehicle communications, optimization theory, machine learning for physical layer signal processing.

Yunlong Cai (Senior Member, IEEE) received the B.S. degree in electronic engineering from the University of Surrey, Guildford, U.K., in 2006, and the Ph.D. degree in electronic engineering from the University of York, York, U.K., in 2010. From 2010 to 2011, he was a Post-Doctoral Fellow with the Electronics and Communications Laboratory, CNAM, Paris, France. Since February 2011, he has been with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China, where he is currently a Professor. He held visiting professor positions at various universities, including Georgia Institute of Technology, McGill University, and University of York. His research interests include transceiver design for multiple-antenna systems, cooperative and relay communications, UAV communications, and machine learning for communications. He has published over 200 journals and conference papers in these areas. He served as the General Chair for the 18th IEEE International Symposium on Wireless Communications Systems, which was held in Hangzhou, in October 2022. He regularly sits on the technical program committee boards of prominent IEEE conferences, such as ICC, GLOBECOM, and VTC. He is currently an Associate Editor of IEEE SIGNAL PROCESSING LETTERS.