Joint Localization and Information Transfer for Reconfigurable Intelligent Surface Aided Full-Duplex Systems

Zhichao Shao, Member, IEEE, Xiaojun Yuan, Senior Member, IEEE, Wei Zhang, Fellow, IEEE, and Marco Di Renzo, Fellow, IEEE

Abstract—In this work, we investigate a reconfigurable intelligent surface (RIS) aided integrated sensing and communication scenario, where a base station (BS) communicates with multiple devices in a full-duplex mode, and senses the positions of these devices simultaneously. An RIS is assumed to be mounted on each device to enhance the reflected echoes. Meanwhile, the information of each device is passively transferred to the BS via reflection modulation. We aim to tackle the problem of joint localization and information retrieval at the BS. A grid based parametric model is constructed and the joint estimation problem is formulated as a compressive sensing problem. We propose a novel message-passing algorithm to solve the considered problem, and a progressive approximation method to reduce the computational complexity involved in the message passing. Moreover, an expectation-maximization (EM) algorithm is applied for tuning the grid parameters, hence mitigating the model mismatch problem. Finally, we analyze the efficacy of the proposed algorithm through the Bayesian Cramér-Rao bound. Numerical results demonstrate the feasibility of the proposed scheme and the superior performance of the proposed EM-based message-passing algorithm.

Index Terms—Reconfigurable intelligent surface, integrated sensing and communication, full-duplex system, message-passing, bayesian Cramér-Rao bound.

I. INTRODUCTION

RECONFIGURABLE intelligent surface (RIS), also known as intelligent reflecting surface, or large intelligent surface, has been recognized as a promising next-generation wireless communication technology [1], [2], [3]. An RIS consists of a large number of software-controlled meta-atoms, each of which can independently impose the required phase shift on the incident electromagnetic waves. By carefully adjusting the phase shifts of all the meta-atoms, the radiated electromagnetic waves can be shaped to propagate towards desired directions, thereby significantly enhancing the communication quality. In wireless communications, RISs have many potential applications. For example, in RIS-enhanced cellular networks, RISs are deployed to establish favorable non-line-of-sight links between base stations (BSs) and users. Thus, additional degrees of freedom are introduced by RISs for improving the system performance [4]. Other applications include the beamforming design in simultaneous wireless information and power transfer networks [5], [6], unmanned aerial vehicle aided wireless networks [7], and internet of things networks [8].

Moreover, an RIS, as a large nearly-passive scattering array, has remarkable capabilities to sense the environment, which has motivated research activities towards the integration of sensing and communications (ISAC) in RIS-aided systems. The main objective of an ISAC system is to enable coexistence between the communication and sensing functionalities by ensuring reliable communications with the users while using the same spectrum for sensing and localizing targets [9], [10], [11]. The authors of [12] employed one RIS for both sensing and communication where the direct path between the dual function radar and communication BS and the target exists. The goal was to maximize the radar signal-to-interference-plus-noise ratio (SINR) under communication SINR constraints. The authors of [13] proposed an RIS that is adaptively partitioned into two parts for communication and localization, respectively, when no direct path exists.

Motivated by the aforementioned considerations, in this work we investigate a RIS-assisted communication system by designing a configuration with integrated sensing capabilities. Possible applications of the proposed approach include perceptive mobile networks [14], [15], which evolve from current communication-only mobile networks and are expected...
to serve as ubiquitous radar-sensing networks, whilst providing uncompromising mobile communication services. Specifically, we propose a new RIS-aided ISAC scenario, where a BS not only communicates with multiple devices (such as vehicles) in full-duplex, but also senses the positions of these devices simultaneously. We assume that an RIS is mounted on each device to enhance the echoes reflected by the device. Meanwhile, the information of each device is passively transferred to the BS via reflection modulation [16], [17], in which the information is encoded into the phase adjustments of RIS meta-atoms. The proposed ISAC scenario is advantageous in two aspects. First, the communication system is full-duplex in frequency and time: the information delivered by the BS can be received by the devices using conventional receivers, and, on the same time/frequency slot, the information of the devices is passively delivered to the BS via reflection modulation thanks to the RISs. Second, the devices are “green” since they do not emit any additional electromagnetic signals during the whole process.

We consider a MIMO orthogonal frequency division multiplexing (OFDM) full-duplex ISAC system, and focus on the receiver design of the BS to jointly locating the positions of the devices and retrieving the data of targets. OFDM, which transmits data symbols over orthogonal subcarriers, is popular for implementing ISAC systems. The authors of [18] allocated different subcarriers to the communication and sensing functions, but this reduces the available bandwidth of each function. In [19], a high-resolution compressed sensing (CS) algorithm is proposed to solve the problem of joint delay-Doppler estimation of moving targets under the assumption that the signal is sparse in the Fourier-domain, but the computational cost is high while solving the semidefinite program. The authors of [20] derived a low-complexity subspace-based algorithm by applying a smoothing approach for joint estimation of range and Doppler shift in OFDM-based radar systems, but the estimation accuracy for multiple targets is not sufficient enough. In [18], [19], and [20], the receiver can be regarded as a radar receiver, which estimates the localization parameters, such as transmission delay, through the OFDM signal reflected from the targets.

In this work we propose a new RIS-aided full-duplex MIMO-OFDM system, where the receiver simultaneously estimates the localization parameters and retrieves the data information from the received echoes. The main contributions are summarized as follows.

- We establish a grid-based parametric system model, and formulate the joint estimation problem as a CS problem by exploiting the sparsity on the considered parameter grid.
- Based on a factor graph representation of the probability model, we develop a novel message-passing algorithm to efficiently solve the considered problem. During the message passing, a progressive approximation method is introduced to reduce the computational complexity.
- An expectation-maximization (EM) algorithm is applied for tuning the grid parameters and mitigating the model mismatch problem.

- The Bayesian Cramér-Rao bound (BCRB) is derived as a fundamental performance limit to evaluate the efficacy of the proposed algorithm.
- Numerical results demonstrate the feasibility of the proposed scenario, as well as the superior performance of the proposed EM-learning method.

The rest of this paper is organized as follows. In Section II, the system model for the considered RIS-aided full-duplex MIMO-OFDM system is described, and the nearly-passive beamforming and information transfer are illustrated based on the generalized Snell’s law and the reflection modulation, separately. In Section III, we propose a grid-based parametric model and solve the localization and information recovery problem through the proposed message-passing algorithm. In Section IV, an EM-based parameters learning algorithm is presented for estimating prior parameters (i.e., noise variance and sparsity) of the proposed algorithm, and the grid parameters for the possible model mismatch problem. In Section V, we derive the analytical BCRB for the proposed system. Simulations are presented and discussed in Section VI and the paper is concluded in Section VII.

Notation: Throughout the paper, bold letters indicate vectors and matrices, non-bold letters express scalars if not indicated. The operators $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^{-1}$ stand for the transpose, Hermitian transpose and inverse, respectively, of a matrix. $I_n$ denotes an $n \times n$ identity matrix and $0_{m \times n}$ is an $m \times n$ all zeros matrix. Additionally, $\text{diag}(\cdot)$ is a diagonal matrix, where the diagonal elements are the vector $\alpha$. The operators $\otimes$ and $\text{Tr}\{\cdot\}$ denote the Kronecker product and trace, respectively. $\text{Re}\{\cdot\}$, $\text{Im}\{\cdot\}$ and $\angle$ return the real part, imaginary part and phase of complex numbers, respectively. Finally, $E\{\cdot\}$ and $\text{Var}\{\cdot\}$ denote the expectation and the variance, respectively. $O(\cdot)$ refers to the big O notation.

II. SYSTEM DESCRIPTION

A. RIS-Aided Full-Duplex System

We consider a MIMO-OFDM full-duplex system, where a BS with $N_t$ transmit and $N_r$ receive antennas communicates with multiple devices, each equipped with an RIS consisting of $L$ reflecting elements and a conventional receiver. The antenna/element spacing is $\frac{\lambda}{2}$, where $\lambda$ is the wavelength. The OFDM signal has $N$ orthogonal subcarriers, and the frequency spacing of adjacent subcarriers is $f_\Delta = 1/T$, where $T$ denotes the duration of one OFDM symbol. At the beginning of each OFDM symbol, a cyclic prefix (CP) of length $T_{cp}$ is inserted to avoid the intersymbol interference, and the duration of each OFDM block is $T_b = T + T_{cp}$.

The transmission protocol illustrated in Fig. 1 is described as follows. The BS first broadcasts the OFDM signal towards a given region via beamforming. The devices in the considered region receive the transmitted signal and recover the message from the BS by using traditional signal processing techniques. Meanwhile, the RIS deployed on each device applies reflection modulation onto the incident electromagnetic wave, and reflects it back to the BS. Unlike the incident wave, the

---

1We omit the details here since the receiver design at the device is standard and is not the focus of this paper.
The Doppler effects caused by the moving users can be ignored with such slow moving speeds. The transmit and receive antennas needs to be considered on the same time-frequency slot, the interference between the moving device can be ignored in such a short period of time. Since the duration of each transmission frame is typically an angle of propagation and a delay. Considering that there are signal reflected by each device as a multipath characterized by the channel between the BS and each device. In fact, the con-

\[ H_{DL,k} = \mathbf{a}_{\text{RIS}}(\phi_{i,k}, \gamma_{i,k}) \mathbf{a}_{\text{BS}}(\vartheta_k)^H \]

\[ H_{UL,k} = \mathbf{b}_{\text{BS}}(-\vartheta_k) \mathbf{a}_{\text{RIS}}(\phi_{i,k}, \gamma_{i,k})^H, \]

where \( \mathbf{a}_{\text{RIS}}, \mathbf{a}_{\text{BS}} \) and \( \mathbf{b}_{\text{BS}} \) are the steering vectors defined in what follows. We assume that the elements of each RIS are arranged in a uniform rectangular array whose size is \( \times \) \( \times \) \( L \) \( \times \) \( N_y \) \( = \) \( \mathbb{C} \). The transmitted signal in the \( m \)-th OFDM block at the BS is given by

\[ x_m(t) = \sum_{n=0}^{N-1} x_m[n] e^{j2\pi nf_0 t} \Xi(t - mT_b), \]

where \( x_m[n] = [x_{m,1}[n], \ldots, x_{m,N_y}[n]]^T \in \mathbb{C}^{N_y \times 1} \) contains the transmitted symbols at the \( n \)-th subcarrier in the \( m \)-th block; and

\[ \Xi(t) = \begin{cases} 1, & t \in [0, T_b], \\ 0, & \text{otherwise}. \end{cases} \]

The received signals at the BS are contaminated by thermal noise \( \mathbf{w}_{\text{Th},m}(t) \in \mathbb{C}^{N_y \times 1} \) can be expressed as Gaussian random variables [22]. The echos received at the BS are by the moving users can be ignored with such slow moving speeds.

\[ y_m(t) = \sum_{k=0}^{K-1} \beta_k \mathbf{H}_{UL,k} \mathbf{\Lambda}_{k,m} \mathbf{H}_{DL,k} x_m(t - \tau_k) + \mathbf{w}_{\text{Th},m}(t) \]

Fig. 1. The considered RIS-aided full-duplex system.
Fig. 2. An illustration of the generalized Snell’s law.

metasurface for backscattering enhancement), the AoAs and
the AoDs at the BS are opposite to each other. Defining
\[
\alpha_{k,m} = \beta_k a_{\text{RIS}}(\phi_{r,k}, \gamma_{r,k})^H A_{k,m} a_{\text{RIS}}(\phi_{i,k}, \gamma_{i,k}),
\]
we can rewrite (2) as
\[
y_m(t) = \sum_{k=0}^{K-1} \alpha_{k,m} H(\theta_k) x_m(t - \tau_k) + w_{\text{Th},m}(t) + w_{\text{St},m}(t),
\]
where \( H(\theta_k) = b_{\text{BS}}(-\psi_k)a_{\text{BS}}(\psi_k)^H \in \mathbb{C}^{N_x \times N_t} \). For the
\( m \)-th block, the demodulated signal at the \( n \)-th subcarrier is
\[
y_m[n] = \frac{1}{T} \int_{mT_b + T_{sp}}^{(m+1)T_b} e^{-j2\pi fm\omega t} y_m(t)dt.
\]
Plugging (1) and (9) into (10), we have
\[
y_m[n] = \sum_{k=0}^{K-1} \alpha_{k,m} e^{-j2\pi \Delta f \omega t} H(\theta_k) x_m[n] + w_m[n],
\]
where
\[
w_m[n] = \frac{1}{T} \int_{mT_b + T_{sp}}^{(m+1)T_b} e^{-j2\pi \Delta f \omega t} (w_{\text{Th},m}(t) + w_{\text{St},m}(t))dt
\]
with \( w_m[n] \in \mathbb{C}^{N_t \times 1} \) being the equivalent complex Gaussian noise vector with mean zero and covariance \( \sigma^2 I_{N_t} \).

B. Nearly-Passive Beamforming and Information Transfer of RIS

We first illustrate an anomalous reflection scenario in Fig. 2, where \( \phi_i \in [-\frac{\pi}{2}, \frac{\pi}{2}] \) and \( \gamma_i \in [0, 2\pi] \) are the elevation and
the azimuth angles of the incident wave, respectively; \( \phi_r \in [-\frac{\pi}{2}, \frac{\pi}{2}] \) and \( \gamma_r \in [0, 2\pi] \) are the elevation and the azimuth
angles of the reflected wave, respectively. According to the
generalized Snell’s law [23], to reflect an incident plane wave
into a desired direction by breaking the specular reflection law
(i.e., \( \phi_r = \phi_i, \gamma_r = \gamma_i + \pi \)), the reflection phase of each element
needs to be set linearly along the \( x \)- and \( y \)-axes. Specifically, the phase shift of the \( (i, j) \)-th element \( \theta_{k,m,i,j} \) of the \( k \)-th RIS is obtained by [24]
\[
\theta_{k,m,i,j} = \pi(i-1)q_{x,k} + \pi(j-1)q_{y,k} + \varphi_{k,m},
\]
where \( \varphi_{k,m} \) is the reference phase at the origin of the coordinates; \( q_{x,k} \) and \( q_{y,k} \) denote the phase gradients of the \( x \)- and \( y \)-axes, which can be expressed as
\[
\begin{align*}
q_{x,k} &= \frac{\sin(\phi_{r,k}) \cos(\gamma_{r,k}) + \sin(\phi_{i,k}) \cos(\gamma_{i,k})}{\sin(\phi_{r,k}) \sin(\gamma_{r,k}) + \sin(\phi_{i,k}) \sin(\gamma_{i,k})}, \\
q_{y,k} &= \frac{\sin(\phi_{r,k}) \sin(\gamma_{r,k}) - \sin(\phi_{i,k}) \sin(\gamma_{i,k})}{\sin(\phi_{r,k}) \sin(\gamma_{r,k}) + \sin(\phi_{i,k}) \sin(\gamma_{i,k})}.
\end{align*}
\]
Equation (13) reveals that the anomalous reflection from an
arbitrary pair of angles \( (\phi_{r,k}, \gamma_{r,k}) \) to an arbitrary pair of angles \( (\phi_{i,k}, \gamma_{i,k}) \) can be obtained by setting \( q_{x,k} \) and \( q_{y,k} \) accordingly. Besides, another degree of freedom provided in
(12) is the reference phase \( \varphi_{k,m} \) of the \( k \)-th RIS, which
determines the wavefront phase of the reflected beam in the
\( m \)-th OFDM block. Particularly, for the scenario of retro-
reflection (i.e., \( \phi_{r,k} = -\phi_{i,k}, \gamma_{r,k} = \gamma_{i,k} \)), we have
\[
\begin{align*}
q_{x,k} &= \frac{-\sin(\phi_{i,k}) \cos(\gamma_{i,k}) + \sin(\phi_{r,k}) \cos(\gamma_{r,k})}{\sin(\phi_{r,k}) \sin(\gamma_{r,k}) + \sin(\phi_{i,k}) \sin(\gamma_{i,k})}, \\
q_{y,k} &= \frac{-\sin(\phi_{r,k}) \sin(\gamma_{r,k}) - \sin(\phi_{i,k}) \sin(\gamma_{i,k})}{\sin(\phi_{r,k}) \sin(\gamma_{r,k}) + \sin(\phi_{i,k}) \sin(\gamma_{i,k})}.
\end{align*}
\]
In practice, the angles \( \phi_{r,k} \) and \( \gamma_{r,k} \) can be estimated through
sensors deployed at the \( k \)-th device, e.g., by using the MUSIC
algorithm [25]. Plugging (12) and (14) into (8), we have
\[
\alpha_{k,m} = \frac{\beta_k N_x \sum_{i=1}^{N_x} e^{j2\pi(i-1)\sin(\phi_{r,k}) \cos(\gamma_{r,k})} e^{j\varphi_{k,m}}}{\sum_{i=1}^{N_x} e^{j2\pi(i-1)\sin(\phi_{r,k}) \cos(\gamma_{r,k})} e^{j\varphi_{k,m}}}.
\]
where \( \varphi_{k,m} \) can be easily controlled by the \( k \)-th RIS. This
motivates us to apply reflection modulation on the incident
wave [26]. The reflected wave, or more specifically, \( \varphi_{k,m} \), can be
utilized to implicitly transmit new information generated at
the device to the BS.

Since the localization parameters are assumed to be constant
in \( M \) consecutive OFDM blocks, the complex path coefficient
\( \beta_k \) is a common constant. Then, (15) is extended to
\[
[\alpha_{k,1}, \ldots, \alpha_{k,M}] = \beta_k [e^{j\varphi_{k,1}}, \ldots, e^{j\varphi_{k,m}}].
\]
To retrieve information from the phase of \( \alpha_{k,m} \), i.e., \( \angle \alpha_{k,m} = \angle \beta_k + \varphi_{k,m} \), the differential phase shift keying (DPSK)
modulation is used to cancel out the unknown common phase
\( \angle \beta_k \) caused by the environment during the signal transmission.

The modulation process is described as
\[
\varphi_{k,m} = \varphi_{k,m-1} + S_{k,m} + S_{\text{ref}}, \quad m \in [2, M]
\]
where \( S_{k,m} \in \{S_1, \ldots, S_V\} \) is the modulated phase difference with \( \{S_1, \ldots, S_V\} \) being the set of \( V \) possible phase differences in DPSK, and \( S_{\text{ref}} \) is the known common reference phase to ensure \( \varphi_{k,m} \in \{S_1, \ldots, S_V\} \). For example, when
differential quadrature phase shift keying (DQPSK) is applied,
\( S_{k,m} \in \{\pi, \frac{3\pi}{4}, \frac{\pi}{4}, \frac{\pi}{2}\} \) and \( S_{\text{ref}} = \pi \).

Recall that \( \{x_m[n]\} \) is the self-information known by the
BS. Then, the problem of interest is to jointly estimate the
unknown parameters \( \{\theta_k\}, \{\tau_k\} \) and \( \{\alpha_{k,m}\} \) from the noisy
observations \( \{y_m[n]\} \) in (11). After obtaining the estimated
\( \hat{\alpha}_{k,m} \), the estimated phase difference \( \hat{S}_{k,m} \) is calculated as
\[
\hat{S}_{k,m} = \angle \hat{\alpha}_{k,m} - \angle \hat{\alpha}_{k,m-1} - S_{\text{ref}}, \quad m \in [2, M]
\]
where \( \angle \beta_k \) is canceled during the subtraction.
III. JOINT LOCALIZATION AND INFORMATION RECOVERY

In this section, we first introduce the grid-based parametric model to formulate the problem stated in the preceding section, and we then present a novel message-passing algorithm to solve the problem.

A. Parametric Grid Model

The service ranges of angles and delays are first quantized to into $Q$ and $U$ discrete values [27], [28] as

$$\vartheta = [\vartheta_0, \ldots, \vartheta_{Q-1}]^T, \quad \vartheta_q \in [\vartheta_{\min}, \vartheta_{\max}],$$

$$\tau = [\tau_0, \ldots, \tau_{U-1}]^T, \quad \tau_u \in [\tau_{\min}, \tau_{\max}],$$

(19a)

(19b)

where $\vartheta_0 = \vartheta_{\min}, \vartheta_{Q-1} = \vartheta_{\max}, \vartheta_0 = \vartheta_{\min}, \tau_{U-1} = \tau_{\max}$. A uniform grid is employed, and the differences between adjacent angles and delays are $\frac{\vartheta_{\max}-\vartheta_{\min}}{Q}$ and $\frac{\tau_{\max}-\tau_{\min}}{U}$, respectively, with $Q \gg K$ and $U \gg K$ to ensure the desired resolution. Based on this grid model, a device with delay $\vartheta_q$ and angle $\tau_u$ is denoted as a virtual device. Then, there are $QU$ virtual devices in total, and among them only $K (K \ll QU)$ devices are real. From (11), the grid-based signal model is rewritten as

$$y_m[n] = \sum_{u=0}^{U-1} \sum_{q=0}^{Q-1} e^{-j2\pi n f_d \tau_u} H(\vartheta_q, \vartheta_u) x_m[n] \zeta_q, u, m + w_m[n],$$

(20)

where $\zeta_q, u, m$ denotes a possible device with the angle $\vartheta_q$ and the delay $\tau_u$ in the $m$-th OFDM block, and

$$\zeta_q, u, m = \begin{cases} 0, & \text{if it is a virtual device,} \\ \alpha_{k, m}, & \text{if it is a real device.} \end{cases}$$

Equation (20) is further simplified as

$$y_m[n] = Z_m[n] \zeta_m + w_m[n],$$

(21)

with

$$Z_m[n] = (e^{-j2\pi n f_d \tau})^T \otimes H(\vartheta_q, \vartheta_u) I_Q \otimes x_m[n]),$$

(22)

where $e^{-j2\pi n f_d \tau} = [e^{-j2\pi n f_d \tau_0}, \ldots, e^{-j2\pi n f_d \tau_{U-1}}] \in C^{U \times 1}$, $H(\vartheta_q, \vartheta_u) = [H(\vartheta_0), \ldots, H(\vartheta_{Q-1})] \in C^{N \times Q N}$, and

$$\zeta_m = [\zeta_1, m, \ldots, \zeta_{Q, U, m}]^T \in C^{QU \times 1}. $$

Since there are only $K$ devices in the considered region, $\zeta_m$ is a sparse vector, where the non-zero elements are given by $\alpha_{k, m} [k \in [1, K]]$ from (20). Collecting the received signals on $N$ subcarriers, we obtain

$$y_m = [y_m[1]^T, \ldots, y_m[N]^T]^T = Z_m \zeta_m + w_m,$$

(23)

where $Z_m = [Z_m[1]^T, \ldots, Z_m[N]^T]^T \in C^{N \times N \times U \times Q}$ and $w_m = [w_m[1]^T, \ldots, w_m[N]^T]^T \in C^{N \times N \times 1}$. Considering all the demodulated symbols in one transmission frame, we have

$$Y = \begin{bmatrix} y_1, \ldots, y_M \end{bmatrix} = \begin{bmatrix} Z_1 \zeta_1, \ldots, Z_M \zeta_M \end{bmatrix} + W,$$

(24)

where $W = [w_1, \ldots, w_M] \in C^{N \times N \times M}$. With the knowledge of the transmitted symbols $x_m[n]$ and the grids in (19), the matrix $Z_m$ is known by the BS. The estimation problem is converted into the problem of estimating $\zeta = [\zeta_1, \ldots, \zeta_M] \in C^{U \times Q \times M}$ based on $Y$ in (24).

B. Message-Passing Algorithm

The minimum mean square error (MMSE) estimator of $\zeta$ given $Y$ is given by

$$\hat{\zeta} = E\{\zeta | Y\}.$$  

(25)

However, the calculation of the conditional mean $E\{\zeta | Y\}$ is computationally involving. In the following, we develop an efficient iterative algorithm to approximately calculate the MMSE estimator (25) based on the message-passing rule.

From (16), each non-zero row of $\zeta$ is the multiplication of a common constant $\beta_q$ and phase shifts in different OFDM blocks $[e^{j\varphi_{q,k}}, \ldots, e^{j\varphi_{q,M}}]$. Then we define $\Xi = \text{diag}(\varphi) \chi$, where $\beta_k$ and $[e^{j\varphi_{k,1}}, \ldots, e^{j\varphi_{k,M}}]$ can be found as the non-zero elements in $\nu \in C^{U \times 1}$ and non-zero rows in $\chi \in C^{U \times Q \times M}$, respectively. The posterior probability $p(\zeta | Y)$ is factorized as

$$p(\zeta | Y) = \frac{1}{p(Y)} p(Y | \zeta) p(\zeta | \nu, \chi) p(\nu) p(\chi),$$

(26)

where

$$p(Y | \zeta) = \prod_{m=1}^M CN(y_m; z_m \zeta_m, \sigma^2 I_N, \nu),$$

(27)

$$p(\zeta | \nu, \chi) = \delta(\zeta - \text{diag}(\nu) \chi),$$

(28)

$$p(\nu) = \prod_{i=1}^{UQ} \left( (1 - \rho) \delta(\nu_i) + \rho CN(\nu_i; \bar{\beta}, \nu^2) \right),$$

(29)

$$p(\chi) = \prod_{i=1}^{UQ} \prod_{m=1}^M \left( \sum_{l=1}^V \delta(\chi_{i,m} - e^{jS_l}) \right).$$

(30)

In (28)–(30), $\delta(\cdot)$ is the Dirac delta function; $\rho = \frac{\kappa}{\nu^2}$ is the sparsity of $\nu$ (or the row sparsity of $\zeta$); $p(\beta_k)$ is approximated as a Gaussian distribution with mean $\bar{\beta}$ and variance $\nu^2$; $X_{i,m}$ is the element of $\chi$ in the $i$-th row and the $m$-th column; $S_l$ denotes the $l$-th phase shift within the set of $V$ possible phases in DPSK, where each phase has probability equal to $\frac{1}{V}$. The factor graph representation of $p(\zeta | Y)$ is shown in Fig. 3. The hollow circles and the solid squares represent the variable nodes and the factor nodes, respectively. Detailed message passings are described as follows.

1) Forward Message Passing: Given the message

$$M_{\zeta_i,m} \rightarrow p(\zeta_i,m | \nu, \chi, \zeta_i,m) \sim CN(\zeta_i,m; \bar{\nu}, \nu^2),$$

where $\bar{\nu_i}$ and $\nu'_i$ are the outputs of the input linear step in the generalized approximate message passing (GAMP) algorithm [29] (more details can be found at the end of this subsection), the message from the factor node $p(\zeta_i,m | \nu_i,X_i,m)$ to the variable node $\nu_i$ is calculated from the sum-product rule as

$$M_{\zeta_i,m} | \nu_i, X_i,m) \rightarrow p(\nu_i) = \int_{X_i,m} \delta(\zeta_i,m - \nu_i, \chi, \zeta_i,m) CN(\zeta_i,m; \bar{\nu}, \nu^2),$$

(31)

Inserting (28) and the prior probability (30) into (31), we have

$$M_{\zeta_i,m} | \nu_i, X_i,m) \rightarrow p(\nu_i) = \int_{X_i,m} \prod_{i=1}^{UQ} \delta(\chi_{i,m} - \nu_i, \chi, \zeta_i,m) CN(\zeta_i,m; \bar{\nu}, \nu^2).$$

(32)

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Gaussian mixture in (33) is calculated as the product of Gaussian mixtures as follows.

We propose a progressive approximation method to simplify the number of OFDM blocks of overall Gaussian components grows exponentially with the product of multiple Gaussian mixtures. The number Note that the exact form of (33) is difficult to calculate due to the factor node

\[ \zeta \]

Fig. 3. The factor graph of the proposed algorithm.

2) Backward Message Passing: With the Bernoulli Gaussian prior in (29) and the Gaussian mixture in (32), the message from the variable node \( \nu_i \) to the factor node \( p(\zeta_i | \nu_i, \chi_i, m) \) is calculated as

\[
\mathcal{M}_{\nu_i} = p(\zeta_i | \nu_i, \chi_i, m) = \prod_{m' = 1, m' \neq m}^{M} \mathcal{M}_{p(\zeta_i | \nu_i, \chi_i, m') - \nu_i} (\nu_i) = \left(1 - \beta\right) \delta(\nu_i) + \rho \mathcal{N}(\nu_i; \beta, \nu^3) \times \prod_{m'=1, m' \neq m}^{M} \left(\sum_{l=1}^{V} \frac{1}{V} \mathcal{N}(\nu_i; e^{-jS_l \hat{r}_{i,m'}}, v_{r,i,m'}^r)\right).
\]

Note that the exact form of (33) is difficult to calculate due to the product of multiple Gaussian mixtures. The number of overall Gaussian components grows exponentially with the number of OFDM blocks \( M \). To alleviate this problem, we propose a progressive approximation method to simplify the product of Gaussian mixtures as follows.

The multiplication between the \( m' \)-th and the \( (m' + 1) \)-th Gaussian mixture in (33) is calculated as

\[
\left(\sum_{l=1}^{V} \frac{1}{V} \mathcal{N}(\nu_i; e^{-jS_l \hat{r}_{i,m'}}, v_{r,i,m'}^r)\right) \times \left(\sum_{q=1}^{V} \frac{1}{V} \mathcal{N}(\nu_i; e^{-jS_q \hat{r}_{i,m'+1}}, v_{r,i,m'+1}^r)\right)
\]

\[= \sum_{l=1}^{V} \sum_{q=1}^{V} \xi_{l,q} \mathcal{N}(\nu_i; \hat{\nu}_{i,m'+1,l,q}, v_{r,i,m'+1}^r).
\]

where, from the Gaussian message combining property,\(^3\) the variance \( v_{r,i,m'+1}^r \), the mean \( \hat{\nu}_{i,m'+1,l,q} \) and the weight \( \xi_{l,q} \) are given by (35a)–(35c), respectively, and are shown at the bottom of the next page. Since each Gaussian component in (34) shares the same variance \( v_{r,i,m'+1}^r \), we cluster the mean values into \( V \) groups and approximate the values of each group as one mean value. Clustering algorithms, such as the K-means [30] and expectation maximization (EM) [31], can be used for this purpose. Specifically, we group the mean values \( \{\hat{\nu}_{i,m'+1, l,q}\} \) as follows.

We assume \( |\hat{r}_{i,m'}| > e^{-jS_l \hat{r}_{i,m'}} \) and \( \frac{1}{V} \mathcal{N}(\nu_i; e^{-jS_l \hat{r}_{i,m'}}, v_{r,i,m'}^r) (S_l \in \{S_1, \cdots, S_V\}) \) is multiplied with \( \sum_{q=1}^{V} \xi_{l,q} \mathcal{N}(\nu_i; \hat{\nu}_{i,m'+1,l,q}, v_{r,i,m'+1}^r) \) to obtain one group of Gaussian functions \( \sum_{q=1}^{V} \xi_{l,q} \mathcal{N}(\nu_i; \hat{\nu}_{i,m'+1,l,q}, v_{r,i,m'+1}^r) \).

Then, the term \( \sum_{q=1}^{V} \xi_{l,q} \mathcal{N}(\nu_i; \hat{\nu}_{i,m'+1,l,q}, v_{r,i,m'+1}^r) \) is approximated as \( \xi_{l} \mathcal{N}(\nu_i; \hat{\nu}_{i,m'+1,l}, v_{r,i,m'+1}^r) \), where the approximated mean \( \hat{\nu}_{i,m'+1, l} \) variance \( v_{r,i,m'+1}^r \) and weight \( \xi_{l} \) are calculated, respectively, as

\[
\hat{\nu}_{i,m'+1, l} = \int_{-\infty}^{\infty} \nu_{i,m'+1, l,q} d\nu_{i,m'+1, l,q}, \quad \xi_{l} = \sum_{q=1}^{V} \xi_{l,q}.
\]

The approximated Gaussian function for the \( l \)-th group is \( \xi_{l} \mathcal{N}(\nu_i; \hat{\nu}_{i,m'+1, l}, v_{r,i,m'+1}^r) \), and the approximated Gaussian functions for the other groups can be directly obtained, where the mean values are calculated by adding additional phase shifts \( q \frac{2\pi}{V} (q \in \{1, \cdots, V - 1\}) \) to \( \hat{\nu}_{i,m'+1, l} \) (i.e., \( e^{j\pi q \frac{2\pi}{V}} \hat{\nu}_{i,m'+1, l} \)), and the variances are the same as \( v_{r,i,m'+1}^r \). Dropping the common group index \( l \) of \( \hat{\nu}_{i,m'+1, l} \) and \( v_{r,i,m'+1}^r \), (34) is simplified as

\[
\left(\sum_{l=1}^{V} \frac{1}{V} \mathcal{N}(\nu_i; e^{-jS_l \hat{r}_{i,m'}}, v_{r,i,m'}^r)\right) \times \left(\sum_{q=1}^{V} \frac{1}{V} \mathcal{N}(\nu_i; e^{-jS_q \hat{r}_{i,m'+1}}, v_{r,i,m'+1}^r)\right)
\]

\[\approx \sum_{l=1}^{V} \xi_{l} \mathcal{N}(\nu_i; e^{j(1-l) \frac{2\pi}{V}} \hat{\nu}_{i,m'+1, l}, v_{r,i,m'+1}^r).
\]

\(^3\)The multiplication of two Gaussian functions is another Gaussian function: \( \mathcal{N}(x; a, A) \mathcal{N}(x; b, B) = \mathcal{N}(x; c, C) \) with \( d = \mathcal{N}(0; a-b, A+B) \), \( C = (A^{-1} + B^{-1})^{-1} \) and \( c = C(a/B + b/A) \).
One visual example of the clustering process is shown in Fig. 4, and further details about the approximation can be found in the appendix. When (37) is further multiplied with the \((m'+2)\)-th Gaussian mixture, (34)-(37) is repeated until all the \(M-1\) Gaussian mixtures are multiplied together. Then (33) is simplified as

\[
\mathcal{M}_{\nu_i}^{-p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)}(\nu_i) \\
\approx \left(1 - \frac{1}{\beta} \right) \delta(\nu_i) + \rho \mathcal{N}(\nu_i; \hat{\beta}, \nu_i) \\
+ \sum_{l=1}^{V} \xi_l \mathcal{N}(\nu_i; \psi_i, \nu_i) \\
= \left(1 - \pi_{\nu_i}^{-p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)} \right) \delta(\nu_i) \\
+ \pi_{\nu_i}^{-p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)} \sum_{l=1}^{V} \xi_l \mathcal{N}(\nu_i; \tilde{\nu}_i, \nu_i),
\]

where \(\tilde{\nu}_i, M-1\) and \(\nu_i, M-1\) are the approximated mean and variance of the products of \(M-1\) Gaussian mixtures, as shown in (39a)-(39d) given at the bottom of the next page.

With the prior probability in (30) and the Bernoulli Gaussian mixture in (38), the message from factor node \(p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)\) to variable node \(\mathbf{v}_i, \mathbf{m}\) is calculated as

\[
\mathcal{M}_{p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)}^{-\mathbf{v}_i, \mathbf{m}}(\mathbf{v}_i) \\
= \int \mathcal{M}_{\nu_i}^{-p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)}(\nu_i) \\
\times \mathcal{M}_{\mathbf{v}_1, \mathbf{v}_m}^{-p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)}(\nu_1) \\
\times \mathcal{M}_{\mathbf{v}_m}^{-p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)}(\nu_m) \\
\times \mathcal{M}_{\mathbf{v}_m}^{-p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)}(\nu_m) \\
\times \delta(\mathbf{v}_m - \mathbf{v}_m) \\
\times \sum_{q=1}^{V} \frac{1}{V} \delta(\mathbf{v}_m - e^{jS_q} \tilde{\mathbf{v}}_i, \nu_i) \\
\times \left(1 - \pi_{\nu_i}^{-p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)} \right) \delta(\mathbf{v}_i) \\
+ \pi_{\nu_i}^{-p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)} \sum_{l=1}^{V} \xi_l \mathcal{N}(\mathbf{v}_i, \nu_i),
\]

where (a) follows from the proposed progressive approximation detailed in (34)-(37), and \(\nu_i, \mathbf{m}\) is the approximated mean value for the \(l\)-th cluster.

3) Message Update at \(\zeta_{i,m}\): By combining the message \(\mathcal{M}_{\zeta_{i,m}}^{-p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)}(\zeta_{i,m}) \sim \mathcal{N}(\zeta_{i,m}; \hat{\zeta}_{i,m}, \nu_i, \mathbf{m})\) and (40), the belief of \(\zeta_{i,m}\) is expressed as

\[
\mathcal{M}_{\zeta_{i,m}} = \mathcal{M}_{\zeta_{i,m}}^{-p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)}(\zeta_{i,m}) \mathcal{M}_{p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)}^{-\zeta_{i,m}}(\mathbf{v}_i, \mathbf{m}) \\
= \mathcal{N}(\zeta_{i,m}; \hat{\zeta}_{i,m}, \nu_i, \mathbf{m}) \left(1 - \pi_{\zeta_{i,m}}^{-p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)} \right) \delta(\zeta_{i,m}) \\
+ \pi_{\zeta_{i,m}}^{-p(\mathbf{v}_i, \mathbf{v}_1, \mathbf{v}_m, X_m)} \sum_{l=1}^{V} \xi_l \mathcal{N}(\zeta_{i,m}; \hat{\zeta}_{i,m}, \nu_i, \mathbf{m}),
\]

where the variance \(\zeta_i, \mathbf{m}\), the mean \(\hat{\zeta}_i, \mathbf{m}\), the weight \(\xi_i, \mathbf{m}\) and the probability \(\pi_i, \mathbf{m}\) are given by (42a)-(42d), respectively, shown at the bottom of the next page. With (41), the mean and variance of \(\zeta_{i,m}\) are calculated as

\[
\hat{\zeta}_{i,m} = \int_{-\infty}^{\infty} \zeta_{i,m} \pi_{\zeta_{i,m}} \sum_{l=1}^{V} \xi_l \mathcal{N}(\zeta_{i,m}; \hat{\zeta}_{i,m}, \nu_i, \mathbf{m}) d\zeta_{i,m} \\
= \pi_{\zeta_{i,m}} \sum_{l=1}^{V} \xi_l \mathcal{N}(\zeta_{i,m}; \hat{\zeta}_{i,m}, \nu_i, \mathbf{m})
\]

(43a)
\[
v_{i,m}^c = \int_{-\infty}^{\infty} (\xi_{i,m} - \hat{\xi}_{i,m})^2 ((1 - \pi_{\xi_{i,m}}) \delta(\xi_{i,m}) + \pi_{\xi_{i,m}}) \frac{d\xi_{i,m}}{} \nonumber = (1 - \pi_{\xi_{i,m}}) \hat{\xi}_{i,m}^2 + \pi_{\xi_{i,m}} \sum_{l=1}^V \xi_{i,m,l} \mathcal{CN}(\xi_{i,m,l}; \hat{\xi}_{i,m,l}, v_{i,m,l}^c)) \nonumber \times \left( v_{i,m}^c + (\hat{\xi}_{i,m,l} - \xi_{i,m,l})^2 \right) . \quad (43b) \]

The overall algorithm is summarized in Algorithm 1, where steps 3-10 are based on the GAMP described in [29], the operator \( \text{Ave}\{\mathbf{A}\} \) averages all the elements in \( \mathbf{A} \) and \( \varepsilon \) is a predefined small value. Steps 3-4 are the output linear step to obtain the quantities \( \{\hat{\eta}_{j,m}\} \) and \( \{v_{j,m}^p\} \). Using these quantities, steps 5-6 compute the marginal posterior means \( \{\hat{\xi}_{j,m}\} \) and variances \( \{v_{j,m}^2\} \). Steps 7-8 then use these posterior moments to compute the scaled residuals \( \{\hat{\varphi}_{j,m}\} \) and the inverse residual variances \( \{v_{j,m}^p\} \). Steps 9-10 correspond to the input linear step to compute \( \{v_{i,m}^c\} \) and \( \{\hat{\varphi}_{i,m}\} \), where \( \hat{\varphi}_{i,m} \) can be interpreted as an observation of \( \xi_{i,m} \) under an additive white Gaussian noise (AWGN) channel with zero mean and variance \( v_{i,m}^c \).

### C. Convergence and Complexity Analysis

Algorithm 1 stems from BG-GAMP [29], since steps 3-10 are based on BG-GAMP [29]. As such, the convergence analysis of BG-GAMP in [32] can be applied to Algorithm 1. Here, we empirically show the convergence of Algorithm 1 over 50 independent trials in Fig. 5. We see that after several iterations the NMSE of \( \xi_{\text{real}} \) converges to a fixed value.

Moreover, the computational complexity of Algorithm 1 mainly comes from step 11 due to the multiple multiplications of Gaussian mixtures. Specifically, there are \( 2V^M-1 \) \((M \geq 3)\) multiplications of two complex Gaussian functions.
while calculating (33). However, with our proposed progressive approximation method, the number is reduced to $V + (M - 1)V^2$, which largely reduces the computational cost. The multiplication between two complex Gaussian functions requires $O(1)$ computational cost. Moreover, steps 3-10 require $O((N,N)NUQMM)$. Then, the overall computational cost in Algorithm 1 is $O((N,N+N+V+(M-1)V^2)UQMM)$ per iteration.

For comparison, the complexity of three state-of-the-art algorithms, namely, the Bernoulli-Gaussian generalized approximate message passing (BG-GAMP) [29], orthogonal matching pursuit (OMP) [33], turbo compressive sensing (Turbo-CS) [34] and sparse Bayesian learning (SBL) [35], are provided in Table I, where $l_{BG-GAMP}$, $l_{Turbo-CS}$, $l_{SBL}$ and $l_{Algorithm1}$ denote the number of iterations used in BG-GAMP, Turbo-CS, SBL and Algorithm 1, respectively. BG-GAMP is a simplified version of Algorithm 1, where there is no shared $v_i$ for each row and the prior information $p(\chi)$ is not considered.

In OMP, for obtaining the possible non-zero indices the main complexity comes from calculating the product between each column of $Z$ and each column of $Y$ resulting in a complexity $O((N,N)N)$ per iteration. The main computational costs in both Turbo-CS and SBL lie in the matrix inversions and matrix multiplications, which are $O((N,N)^3)$ and $O((N,N)^2UQ)$ per iteration, respectively. Moreover, the running time of different algorithms averaged over 50 independent trials are also shown in Table I, where the simulation parameters are presented in Section VI. The results are obtained by using MATLAB on a PC equipped with a Core i7-10700 CPU, 24GB RAM and one NVIDIA GeForce GTX 1050Ti graphic card.

### IV. Parameters Learning

In practice, the prior parameters $\sigma^2$ and $\rho$ are usually unknown and need to be estimated. Moreover, the true localization parameters may not be the grid values given in (19), also known as the model mismatch problem. As such, the true localization parameters are represented as

$$\theta_{\text{real},k} = \theta_q + \theta_{\Delta,q}, \quad k = 0, 1, \ldots, K - 1$$

and

$$\tau_{\text{real},k} = \tau_u + \tau_{\Delta,u}, \quad k = 0, 1, \ldots, K - 1$$

where $\theta_{\text{real},k}$ and $\tau_{\text{real},k}$ denote the true localization parameters; $\theta_q$ with $q \in \{0, 1, \ldots, Q - 1\}$ and $\tau_u$ with $u \in \{0, 1, \ldots, U - 1\}$ denote the nearest grid points; $\theta_{\Delta,q}$ and $\tau_{\Delta,u}$ correspond to the unknown off-grid gaps. To address these issues, we treat $\omega = \{\sigma^2, \rho, \theta_{\Delta}, \tau_{\Delta}\}$ as unknown parameters and utilize the EM approach [36] to learn these parameters, where $\hat{\omega} = [\theta_{\Delta,0}, \ldots, \theta_{\Delta,Q-1}]^T$ and $\tau_{\Delta} = [\tau_{\Delta,0}, \ldots, \tau_{\Delta,U-1}]^T$. The EM process is described as

$$\omega(l + 1) = \arg \max _{\omega} E_{\pi(\omega|Y,\omega(l))} \{ \ln p(Y, \omega; \omega) \},$$

where $\omega(l) = \{\sigma^2(l), \rho(l), \theta_{\Delta}(l), \tau_{\Delta}(l)\}$ is the estimate of $\omega$ in the $l$-th EM-learning iteration. $E_{\pi(\omega|Y,\omega(l))} \{ \cdot \}$ represents the expectation over the posterior distribution $p(\omega|Y,\omega(l))$. Since it is impractical to update all parameters in $\omega$ at once, we update each parameter at a time while fixing the others.

#### A. Learning $\sigma^2$ and $\rho$

We first derive the EM update for the noise variance $\sigma^2$ given the previous parameters $\omega(l)$. The joint probability density function (pdf) $p(Y, \omega; \sigma^2)$ in (45) is decoupled as

$$\sigma^2(l + 1) = \arg \max _{\sigma^2} \int _{\omega} p(\omega|Y,\omega(l)) \ln p(Y|\omega,\sigma^2)$$

$$+ \int _{\omega} p(\omega|Y,\omega(l)) \ln p(\omega;\sigma^2).$$

Setting the derivative of (46) (with respect to $\sigma^2$) to zero yields

$$\int _{\omega} p(\omega|Y,\omega(l)) \frac{\partial p(Y|\omega,\sigma^2)}{\partial \sigma^2} = 0,$$

where the last term of (46) is dropped since it is irrelevant to $\sigma^2$ by noting $p(\omega|\sigma^2) = p(\omega)$. With (27), the update of $\sigma^2$ is

$$\sigma^2(l + 1) = \frac{1}{N,N} \sum _{m=1} ^{M} \text{Tr} \{ (\tilde{y}_m$$

$$- \tilde{Z}_m \tilde{\zeta}_m(l))(\tilde{y}_m - \tilde{Z}_m \tilde{\zeta}_m(l))^H$$

$$+ \tilde{Z}_m \tilde{V}_m \tilde{\zeta}_m(l) \tilde{V}_m^H \},$$

where $\tilde{\zeta}_m(l)$ and $V_m \tilde{\zeta}_m(l)$ are the posterior mean and variance of $\zeta$ at the $m$-th column in the $l$-th EM-learning iteration, and they are calculated by (43a) and (43b), respectively.

Similar to (46), the EM update for the sparsity $\rho$ can be written as

$$\rho(l + 1) = \arg \max _{\rho} \int _{\omega} p(\omega|Y,\omega(l)) \ln p(Y|\omega,\rho)$$

$$+ \int _{\omega} p(\omega|Y,\omega(l)) \ln p(\omega;\rho).$$

Also, we set the derivatives of (49) (with respect to $\rho$) to zero, which yields

$$\int _{\omega} p(\omega|Y,\omega(l)) \frac{\partial p(\omega;\rho)}{\partial \rho} = 0.$$
where the first term is dropped since it is irrelevant to \( \rho \). With (29), the update of \( \rho \) is [29]

\[
\rho(l + 1) = \frac{1}{UQM} \sum_{i=1}^{M} \sum_{m=1}^{M} \pi_{\xi_{i,m}}(l),
\]

(51)

where \( \pi_{\xi_{i,m}}(l) \) is defined in (42d).

**B. Learning \( \vartheta_\Delta \) and \( \tau_\Delta \)**

Inserting (44) into (21) and following the steps in (21)-(24), the off-grid-based signal model is written as

\[
Y = [Z'_1, \cdots, Z'_M] + W,
\]

(52)

where \( \{Z'_m[n]\} \) is constructed by \( \{Z_m[n]\} \) with \( Z'_m[n] = (e^{-j2\pi n/\Delta_\tau} + \vartheta_\Delta) \mathbf{H}(\vartheta + \vartheta_\Delta) (I_Q \otimes x_m[n]) \). Based on (52), the optimization problem in (45) is calculated as

\[
\omega(l + 1) = \arg \max_{\omega} \int_p(\xi|Y; \omega(l)) \ln p(Y|\xi; \omega),
\]

where the term \( \int_p(\xi|Y; \omega(l)) \ln p(\xi; \omega) \) is dropped since it is irrelevant to \( \omega \) by noting \( p(\xi; \omega) = p(\xi) \). With (27), the objective function in (53) is calculated as

\[
\int_p(\xi|Y; \omega(l)) \ln p(Y|\xi; \omega) = \sum_{m=1}^{M} \int_{\xi_m} p(\xi_m | y_m; \omega(l)) \ln CN(y_m; Z'_m \xi_m; \sigma^2 I_{N_x,N_y}) = -M \ln(\pi \sigma^2) - \frac{1}{\sigma^2} \sum_{m=1}^{M} \int_{\xi_m} p(\xi_m | y_m; \omega(l))(y_m H y_m)
\]

\[
- y_m^H Z'_m \xi_m - (Z'_m \xi_m)^H y_m + (Z'_m \xi_m)^H (Z'_m \xi_m),
\]

(54)

where \( p(\xi_m | y_m; \omega(l)) = CN(\xi_m; \hat{\xi}_m(l), V_\xi m(l)) \) is obtained from the output of Algorithm 1. Inserting (54) into (53) and ignoring the terms independent of \( \omega \), (53) is simplified as

\[
\omega(l + 1) = \arg \max_{\omega} \sum_{m=1}^{M} \int_{\xi_m} p(\xi_m | y_m; \omega(l)) \left( y_m^H Z'_m \xi_m + (Z'_m \xi_m)^H y_m - \xi_m^H Z'_m Z'_m \xi_m \right),
\]

(55)

where

\[
\int_{\xi_m} p(\xi_m | y_m; \omega(l)) y_m^H Z'_m \xi_m = y_m^H Z'_m \int_{\xi_m} \xi_m p(\xi_m | y_m; \omega(l)) = y_m^H Z'_m \hat{\xi}_m(l),
\]

(56a)

\[
\int_{\xi_m} p(\xi_m | y_m; \omega(l)) (Z'_m \xi_m)^H y_m = (y_m^H Z'_m \hat{\xi}_m(l))^H,
\]

(56b)

\[
\int_{\xi_m} p(\xi_m | y_m; \omega(l)) \xi_m^H Z'_m Z'_m \xi_m = \text{Tr} \left\{ Z'_m^H Z'_m \int_{\xi_m} \xi_m \xi_m^H p(\xi_m | y_m; \omega(l)) \right\}
\]

\[
= \text{Tr} \left\{ Z'_m^H Z'_m \left( \hat{\xi}_m(l) \hat{\xi}_m(l)^H + V_\xi m(l) \right) \right\},
\]

(56c)

and (56c) is calculated based on the identity \( E[x^2] = E[x]^2 + \text{Var}[x] \). Based on these calculations, (55) is simplified as

\[
\omega(l + 1) = \arg \max_{\omega} \sum_{m=1}^{M} 2 \text{Re} \left\{ y_m^H Z'_m \hat{\xi}_m(l) \right\} - \text{Tr} \left\{ Z'_m^H Z'_m C_m(l) \right\},
\]

(57)

where \( C_m(l) = \hat{\xi}_m(l) \hat{\xi}_m(l)^H + V_\xi m(l) \). However, it is hard to obtain an optimal solution to (57) since \( Z'_m \) is a non-linear function of \( \vartheta_\Delta \) and \( \tau_\Delta \). We find a suboptimal solution that increases the value of the objective function step by step. Defining \( G(\omega) \triangleq \sum_{m=1}^{M} 2 \text{Re} \{ y_m^H Z'_m \hat{\xi}_m(l) \} - \text{Tr} \{ Z'_m^H Z'_m C_m(l) \} \), we update the value of \( \omega \) in the derivative direction, i.e.

\[
\omega(l + 1) = \omega(l) + \epsilon \frac{\partial G(\omega)}{\partial \omega},
\]

(58)

where \( \epsilon \) is an appropriate stepsize that is selected based on the backtracking line search method, which constantly decreases the value of \( \epsilon \) until \( \omega(l + 1) \) is found to make \( G(\omega(l + 1)) \geq G(\omega(l)) \). The gradient of \( G(\omega) \) with respect to \( \vartheta_\Delta,q \) and \( \tau_\Delta,u \) are, respectively, given by

\[
\frac{\partial G(\omega)}{\partial \vartheta_\Delta,q} = \sum_{m=1}^{M} 2 \text{Re} \left\{ y_m^H \frac{\partial Z'_m}{\partial \vartheta_\Delta,q} \hat{\xi}_m(l) \right\} - \text{Tr} \left\{ \frac{\partial Z'_m^H Z'_m + Z'_m^H \frac{\partial Z'_m}{\partial \vartheta_\Delta,q}}{\partial \vartheta_\Delta,q} C_m(l) \right\},
\]

(59a)

\[
\frac{\partial G(\omega)}{\partial \tau_\Delta,u} = \sum_{m=1}^{M} 2 \text{Re} \left\{ y_m^H \frac{\partial Z'_m}{\partial \tau_\Delta,u} \hat{\xi}_m(l) \right\} - \text{Tr} \left\{ \frac{\partial Z'_m^H Z'_m + Z'_m^H \frac{\partial Z'_m}{\partial \tau_\Delta,u}}{\partial \tau_\Delta,u} C_m(l) \right\}. \]

(59b)

From (52), the derivatives of \( Z'_m \) at the \( n \)-th subcarrier are, respectively, given by

\[
\frac{\partial Z'_m[n]}{\partial \vartheta_\Delta,q} = (e^{-j2\pi n \Delta_\tau (\tau + \vartheta_\Delta)})^T \mathbf{H}(\vartheta + \vartheta_\Delta) (I_Q \otimes x_m[n])],
\]

(60a)

\[
\frac{\partial Z'_m[n]}{\partial \tau_\Delta,u} = \frac{\partial (e^{-j2\pi n \Delta_\tau (\tau + \vartheta_\Delta)})^T \mathbf{H}(\vartheta + \vartheta_\Delta) (I_Q \otimes x_m[n])]}{\partial \tau_\Delta,u}.
\]

(60b)

where

\[
\frac{\partial \mathbf{H}(\vartheta + \vartheta_\Delta)}{\partial \vartheta_\Delta,q} = \begin{bmatrix} 0_{N_x \times N_q} & \cdots & 0_{N_x \times N_q} \frac{\partial \mathbf{H}(\vartheta + \vartheta_\Delta)}{\partial \vartheta_\Delta,q} \end{bmatrix} 0_{N_x \times N_q} & \cdots & 0_{N_x \times N_q} \end{bmatrix}^T,
\]

(60c)

\[
\frac{\partial \mathbf{H}(\vartheta + \vartheta_\Delta)}{\partial \tau_\Delta,u} = \begin{bmatrix} 0 & \cdots & 0 & j2\pi n \mathbf{f}_d \end{bmatrix}^T.
\]
Algorithm 2 The EM-Based Parameters Learning

Input: $Y$, prior distributions of $p(x)$

Output: $\hat{\zeta}$, $V'$, $\hat{\rho}$, $\hat{\phi}$, $\hat{\tau}$

1: Initialization: $\theta$, $\tau$, $\theta_{\Delta}(1)$, $\tau_{\Delta}(1)$, $\sigma^2(1)$, $\rho(1)$;
2: for $l = 1, \cdots, L_{\text{max}}$ do
3: for $l' = 1, \cdots, L_{\text{max}}$ do
4: Use Algorithm 1 to obtain $\hat{\zeta}(l')$, $V'(l')$ and $\hat{\pi}_{\zeta,m}(l')$;
5: Update $\tau$ by (48) and $\rho$ by (51);
6: if $|\sigma^2(l' + 1) - \sigma^2(l')| < \varepsilon_1$ and $|\rho(l' + 1) - \rho(l')| < \varepsilon_2$, stop
7: end for
8: Fix $\tau$, and maximize (57) with respect to $\tau_{\Delta}$ through gradient descent in (58) to obtain $\theta_{\Delta}(l + 1)$;
9: Fix $\theta_{\Delta}$, and maximize (57) with respect to $\tau_{\Delta}$ through gradient descent in (58) to obtain $\tau_{\Delta}(l + 1)$;
10: Generate $Z_m'$ based on $\theta_{\Delta}(l + 1)$ and $\tau_{\Delta}(l + 1)$ by (52);
11: if $|\text{Ave}(|\tau_{\Delta}(l + 1) - \tau_{\Delta}(l)|) < \varepsilon_3$ and $|\text{Ave}(|\theta_{\Delta}(l + 1) - \theta_{\Delta}(l)|) < \varepsilon_4$, stop
12: end for

The overall EM-based parameters learning procedure is summarized in Algorithm 2, where $\varepsilon_1$, $\varepsilon_2$, $\varepsilon_3$ and $\varepsilon_4$ are predefined small values. The computational complexity of Algorithm 2 mainly lies in steps 8-9. Since we use the gradient descent method to iteratively maximize (57), the calculation of the gradient for each localization parameter in (59) requires $O(((UQ)^3 + (UQ)^2 N + UQ N_1 N M))$ per EM-learning iteration.

V. BAYESIAN CRAMÉR-RAO BOUND

In this section, we develop a mean square error lower-bound of the considered estimation problem. Recalling the non-grid-based signal model in (11) and following the steps in (23)-(24), we have

$$Y = [Z_{\text{real},1}, \cdots, Z_{\text{real},M} \zeta_{\text{real},1}, \cdots, \zeta_{\text{real},M}] + W,$$

where $Z_{\text{real},m} \in \mathbb{C}^{N \times K}$ is similar to $Z_m$ but it is constructed by the angles $\theta_{\text{real}} = [\theta_0, \cdots, \theta_{K-1}]^T \in \mathbb{R}^{K \times 1}$ and delays $\tau_{\text{real}} = [\tau_0, \cdots, \tau_{K-1}]^T \in \mathbb{R}^{K \times 1}$; $\zeta_{\text{real},m} = [\zeta_{0,0}, \cdots, \zeta_{0,K-1}, \cdots, \zeta_{K-1,0}, \cdots, \zeta_{K-1,K-1}]^T \in \mathbb{C}^{K \times K}$, by stacking the columns of (61) sequentially on the top of one another, we have

$$y = Z_{\text{real}} \zeta_{\text{real}} + w,$$

where

$$Z_{\text{real}} = \begin{bmatrix}
Z_{\text{real},1} & 0 & \cdots & 0 \\
0 & Z_{\text{real},2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & Z_{\text{real},M}
\end{bmatrix},$$

$\zeta_{\text{real}} = [\zeta_{\text{real},1}^T, \cdots, \zeta_{\text{real},M}^T]^T \in \mathbb{C}^{KM \times 1}$ and $w \in \mathbb{C}^{N_1 N M \times 1}$. Let $\kappa = [\theta_{\text{real}}^T, \tau_{\text{real}}^T]^T$ be the parameter vector. The Bayesian information matrix (BIM) [37] is defined as

$$\text{J}_y(\kappa) = \text{J}_y^{D}(\kappa) + \text{J}_y^{P}(\kappa),$$

where

$$\text{J}_y^{D}(\kappa) = \begin{bmatrix}
J_y^{D}(\theta_{\text{real}}) & J_y^{D}(\theta_{\text{real}}, \tau_{\text{real}}) & J_y^{D}(\theta_{\text{real}}, \zeta_{\text{real}}) \\
J_y^{D}(\tau_{\text{real}}, \theta_{\text{real}}) & J_y^{D}(\tau_{\text{real}}) & J_y^{D}(\tau_{\text{real}}, \zeta_{\text{real}}) \\
J_y^{D}(\zeta_{\text{real}}, \theta_{\text{real}}) & J_y^{D}(\zeta_{\text{real}}, \tau_{\text{real}}) & J_y^{D}(\zeta_{\text{real}})
\end{bmatrix}$$

and $\text{J}_y^{P}(\kappa)$ has a similar form. Specifically, the $(i, j)$-th element of $\text{J}_y^{D}(\kappa)$ and $\text{J}_y^{P}(\kappa)$ are, respectively, calculated as

$$\text{J}^{D}(\kappa)_{i,j} = \frac{2}{\sigma^2} \text{Re} \left\{ \frac{\partial}{\partial \kappa_i} \frac{Z_{\text{real}} \zeta_{\text{real}}}{\partial \kappa_j} \right\}.$$

Based on (67), the submatrices in (64) are

$$\text{J}^{D}(\theta_{\text{real}})_{i,j} = \frac{2}{\sigma^2} \text{Re} \left\{ \frac{\partial}{\partial \theta_{i}} \frac{Z_{\text{real}} \zeta_{\text{real}}}{\partial \theta_{j}} \right\}$$

$$\text{J}^{D}(\tau_{\text{real}})_{i,j} = \frac{2}{\sigma^2} \text{Re} \left\{ \frac{\partial}{\partial \tau_{i}} \frac{Z_{\text{real}} \zeta_{\text{real}}}{\partial \tau_{j}} \right\}$$

$$\text{J}^{D}(\zeta_{\text{real}})_{i,j} = \frac{2}{\sigma^2} \text{Re} \left\{ \frac{\partial}{\partial \kappa_{i}} \frac{Z_{\text{real}} \zeta_{\text{real}}}{\partial \kappa_{j}} \right\}.$$
Furthermore, we note that $p(\zeta_{\text{real}})$ needs to be known a priori in Algorithm 1, we approximate $p(\zeta_{\text{real}})$ as a Gaussian mixture $\sum_{l=1}^{L} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\zeta_{\text{real}}-\beta_l\tau_l)^2}{2\sigma^2}}$. The first-order derivative of the log-likelihood function $\ln p(\zeta_{\text{real}})$ is

$$\frac{\partial \ln p(\zeta_{\text{real}})}{\partial \zeta_{\text{real}}} = \sum_{l=1}^{L} e^{-\frac{(\zeta_{\text{real}}-\beta_l\tau_l)^2}{2\sigma^2}} \frac{-2(\zeta_{\text{real}}-\beta_l\tau_l)}{2\sigma^2},$$

which can be inserted into (66) to obtain $[J'_{\zeta}(\zeta_{\text{real}})]_{i,j}$. Specifically, when $\beta = 0$, $[J'_{\zeta}(\zeta_{\text{real}})]_{i,j} = \frac{\zeta_i}{\tau_i}$. Since $\theta_{\text{real}}$ and $\tau_{\text{real}}$ are uniformly distributed, $[J'_{\theta}(\theta_{\text{real}})]_{i,j}$, $[J'_{\tau}(\tau_{\text{real}})]_{i,j}$, $[J'_{\zeta}(\zeta_{\text{real}}, \theta_{\text{real}})]_{i,j}$, $[J'_{\zeta}(\zeta_{\text{real}}, \tau_{\text{real}})]_{i,j}$ and $[J'_{\zeta}(\zeta_{\text{real}}, \tau_{\text{real}})]_{i,j}$ are all zeros. With (63), the variance of the estimator $\hat{\kappa}_i$ is lower bounded by

$$\text{Var}\{\hat{\kappa}_i\} \geq \text{Var}\{\kappa_i\} = [J^{-1}_{\zeta}(\kappa)]_{i,i}. \quad (70)$$

Furthermore, we note that $d_{\text{real}} = \frac{1}{2} \tau_{\text{real}}$, where $d_{\text{real}} = [d_{\text{real}}_1, \ldots, d_{\text{real}}_K]^T \in \mathbb{R}^K$ are the distances between the BS and the devices with $c$ being the speed of light. The variance of the estimator $\hat{d}_{\text{real},i}$ is lower bounded by

$$\text{Var}\{\hat{d}_{\text{real},i}\} \geq \text{Var}\{d_{\text{real},i}\} = \frac{c^2}{4} \text{Var}\{\tau_{\text{real},i}\}. \quad (71)$$

VI. NUMERICAL RESULTS

The simulation parameters are listed in Tab. II. The fading coefficient $\beta_k$ is given by [38]

$$\beta_k = \eta_k \sqrt{G_t G_r G \frac{L_\lambda^2}{64\pi^3 d_k^4}}. \quad (72)$$

Moreover, the SINR is defined as

$$10 \log \frac{\text{Tr}(\mathbf{Z}_{\text{real}} \mathbf{C}_{\text{real}})^{\dagger} (\mathbf{Z}_{\text{real}} \mathbf{C}_{\text{real}})}{\sum_{n=1}^{N} \sum_{m=1}^{M} \sigma_m^2}.$$
parameters to be estimated are assumed to be placed on the grid specified in (19), and the indices can be obtained through the positions of the non-zero rows in $\zeta$. The number of grid points in the angle and delay domains are both set to 25. The row-sparse matrix $\zeta$ to be recovered has size $625 \times 20$, and contains 5 non-zero rows, i.e., there are $K = 5$ devices in the service region. DQPSK is used for the modulation of the data emitted by the RISs. It is observed that the proposed Algorithm 1 can approach the BCRB for $\zeta_{\text{real}}$ at high SINR. Moreover, compared with other algorithms listed in Table I, Algorithm 1 has the best performance in terms of the NMSE of $\zeta_{\text{real}}$ and the BER of the passive information transferred by RISs. While simulating the BG-GAMP, OMP, Turbo-CS and SBL algorithms, each column of $\zeta$ is independently estimated until all the columns are obtained.

Moreover, to analyze the impact of various parameters on the system performance Fig. 7 shows the NMSE of $\zeta_{\text{real}}$ against $L$, $N_t$, $K$ and $N_r$ under the scenario with fixed noise power $\sigma^2 = 0.005$. The upper figures show that the NMSE improves with increasing the reflecting elements $L$ and the transmit antennas $N_t$. This is attributed to the fact that the power of the reflected echo is stronger when $L$ and $N_t$ increase. In particular, the red line illustrates the power consumption of RIS, $P_{\text{RIS}}$, which is calculated as [4]

$$P_{\text{RIS}} = LP_l,$$

with $P_l$ being the power consumption of each phase shifter. We assume $P_l = 7.8\text{mW}$ [4]. Furthermore, the lower figures show the NMSE against the number of devices $K$ in the service region and the number of receive antennas $N_r$ at the
BS. We see that the proposed algorithm outperforms the other algorithms.

We now consider the possible mismatch between the true localization parameters and the parametric model, namely, the off-grid scenario, where $\theta_{\text{real}}$ and $\tau_{\text{real}}$ are randomly selected within their corresponding ranges, and $\theta_\Delta$ and $\tau_\Delta(1)$ are initialized to zero. The maximum number of EM-learning iterations is 70. In Fig. 8, the decreasing NMSE curves of Algorithm 2 show that the model mismatch problem improves gradually when increasing the number of EM-learning iterations, which verifies the effectiveness of the proposed EM-based parameter learning method. Moreover, Fig. 9 shows the performance of the EM-based parameters learning algorithm against the SINR in the off-grid system. It can be seen that, with the proposed learning algorithm, both the NMSE and the BER significantly outperform those of the algorithms that do not rely on learning methods, especially at high SINR. Note that the proposed EM-based parameters learning method in Algorithm 2 can be utilized with other CS algorithms as long as the posterior mean and variance of $\zeta$ can be obtained. Specifically, step 4 in Algorithm 2 can be replaced by various CS algorithms. Based on that, we also present the performance of BG-GAMP and SBL with the proposed EM-learning method in Fig. 9, which shows the effectiveness of the proposed learning method as well.

VII. Conclusion

In this work, we have proposed a new RIS-aided integrated sensing and communication scenario, where a BS communicates with multiple devices in full-duplex, and senses the positions of these devices simultaneously. A grid based parametric model was constructed, and the joint estimation problem was formulated as a CS problem. A novel message-passing algorithm was used to solve the problem, and the progressive approximation method was proposed to reduce the computational complexity during the message passing. To tackle the issue of model mismatch, the EM algorithm was utilized for parameters learning. Simulation results have substantiated the good performance of the proposed method.

Appendix

In probability theory, the Kullback-Leibler (KL) divergence [39] is popularly used to measure the distance between two probability distributions. Specifically, the KL divergence of $f(x)$ from $\tilde{f}(x)$, denoted $D_{\text{KL}}(f(x)||\tilde{f}(x))$, is a measure of the information loss when $\tilde{f}(x)$ is used to approximate $f(x)$. The smaller $D_{\text{KL}}(f(x)||\tilde{f}(x))$, the closer $f(x)$ and $\tilde{f}(x)$. Letting $f(\nu_i)$ be the exact Gaussian mixture in (34) and $\tilde{f}(\nu_i)$ be the approximation in (37), $D_{\text{KL}}(f(\nu_i)||\tilde{f}(\nu_i))$ is calculated as

$$D_{\text{KL}}(f(\nu_i)||\tilde{f}(\nu_i)) = \int_{\nu_i} f(\nu_i) \ln \frac{f(\nu_i)}{\tilde{f}(\nu_i)}. \quad (75)$$

However, it is difficult to evaluate $D_{\text{KL}}(f(\nu_i)||\tilde{f}(\nu_i))$ in closed form for Gaussian mixtures. As such, we employ numerical integrations considering several examples, as shown in Fig. 10. It can be seen that $D_{\text{KL}}(f(\nu_i)||\tilde{f}(\nu_i))$ for any $i$ approaches zero, which demonstrates the effectiveness of the proposed approximation method. In particular, $D_{\text{KL}}(f(x)||\tilde{f}(x)) = 0$ when $f(x) = \tilde{f}(x)$.

References

[1] L. Subrt, D. Grace, and P. Pechac, “Controlling the short-range propagation environment using active frequency selective surfaces,” Radioengineering, vol. 19, no. 4, pp. 610–615, Dec. 2010.
[2] M. Di Renzo et al., “Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead,” IEEE J. Sel. Areas Commun., vol. 38, no. 11, pp. 2450–2525, Nov. 2020.
[3] Y. Liu et al., “Reconfigurable intelligent surfaces: Principles and opportunities,” IEEE Commun. Surveys Tuts., vol. 23, no. 3, pp. 1546–1577, 3rd Quart., 2021.
[4] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, “Reconfigurable intelligent surfaces for energy efficiency in wireless communication,” IEEE Trans. Wireless Commun., vol. 18, no. 8, pp. 4157–4170, Aug. 2019.
[5] Q. Wu and R. Zhang, “Weighted sum power maximization for intelligent reflecting surface aided SWIPT,” IEEE Wireless Commun. Lett., vol. 9, no. 5, pp. 586–590, May 2020.
[6] Q. Wu and R. Zhang, “Joint active and passive beamforming optimization for intelligent reflecting surface assisted SWIPT under QoS constraints,” IEEE J. Sel. Areas Commun., vol. 38, no. 8, pp. 1735–1748, Aug. 2020.
[7] S. Li, B. Dua, X. Yuan, Y.-C. Liang, and M. Di Renzo, “Reconfigurable intelligent surface assisted UAV communication: Joint trajectory design and passive beamforming,” IEEE Wireless Commun. Lett., vol. 9, no. 5, pp. 716–720, May 2020.
[8] A. U. Makarfi, K. M. Rabie, O. Kaiwartya, O. S. Badarneh, X. Li, and R. Kharel, “Reconfigurable intelligent surface enabled IoT networks in generalized fading channels,” in Proc. IEEE Int. Conif. Commun. (ICC), Jun. 2020, pp. 1–6.
[9] F. Liu et al., “Integrated sensing and communications; Toward dual-functional wireless networks for 6G and beyond,” IEEE J. Sel. Areas Commun., vol. 40, no. 6, pp. 1728–1767, Jun. 2022.
[10] D. K. Pin Tan et al., “Integrated sensing and communication in 6G: Motivations, use cases, requirements, challenges and future directions,” in Proc. 1st IEEE Int. Online Symp. Joint Commun. Sens. (JC&S), Feb. 2021, pp. 1–6.

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
[11] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, “Joint radar and communication design: Applications, state-of-the-art, and the road ahead,” IEEE Trans. Commun., vol. 68, no. 6, pp. 3834–3862, Jun. 2020.

[12] Z.-M. Jiang et al., “Intelligent reflecting surface aided dual-function radar and communication system,” IEEE Syst. J., vol. 16, no. 1, pp. 375–387, Mar. 2022.

[13] Q. Li, P. Sarkar, P. U. Deepak, and S. Chepuri, “Joint communication and radar sensing with reconfigurable intelligent surfaces,” 2021, arXiv:2105.01966.

[14] A. Zhang, M. L. Rahman, X. Huang, J. Guo, S. Chen, and R. W. Heath, “Perceptive mobile networks: Cellular networks with radio vision via joint communication and radar sensing,” IEEE Veh. Technol. Mag., vol. 16, no. 2, pp. 20–30, Jun. 2021.

[15] Z. Ni, J. A. Zhang, X. Huang, K. Yang, and J. Yuan, “Uplink sensing in perceptive mobile networks with asynchronous transceivers,” IEEE Trans. Signal Process., vol. 69, pp. 1287–1300, 2021.

[16] W. Yan, X. Yuan, and X. Kuai, “Passive beamforming and information transfer via large intelligent surface,” IEEE Wireless Commun. Lett., vol. 9, no. 4, pp. 533–537, Apr. 2020.

[17] W. Yan, X. Yuan, Z.-Q. He, and X. Kuai, “Passive beamforming and information transfer design for reconfigurable intelligent surfaces aided multiuser MIMO systems,” IEEE J. Sel. Areas Commun., vol. 38, no. 8, pp. 1793–1808, Aug. 2020.

[18] Y. L. Sit, B. Nuss, and T. Zwick, “On mutual interference cancellation in a MIMO OFDM multiuser radar-communication network,” IEEE Trans. Veh. Technol., vol. 67, no. 4, pp. 3339–3348, Apr. 2018.

[19] L. Zheng and X. Wang, “Super-resolution delay-Doppler estimation for OFDM passive radar,” IEEE Trans. Signal Process., vol. 65, no. 9, pp. 2197–2210, May 2017.

[20] J. F. Gu, J. Moghaddasi, and K. Wu, “Delay and Doppler shift estimation for OFDM-based radar-radio (RadCom) system,” in Proc. IEEE Int. Wireless Symp. (IWS), Mar. 2015, pp. 1–4.

[21] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, “In-band full-duplex wireless: Challenges and opportunities,” IEEE J. Sel. Areas Commun., vol. 32, no. 9, pp. 1657–1672, Sep. 2014.

[22] T. N. Nguyen, N. N. Thang, B. C. Nguyen, T. M. Hoang, and P. T. Tran, “Intelligent-reflecting-surface-aided bidirectional full-duplex communication system with imperfect self-interference cancellation and hardware impairments,” IEEE Syst. J., vol. 17, no. 1, pp. 1352–1362, Mar. 2023.

[23] N. Yu et al., “Light propagation with phase discontinuities: Generalized laws of reflection and refraction,” Science, vol. 334, no. 6054, pp. 333–337, Oct. 2011.

[24] C. Cai, X. Yuan, W. Yan, Z. Huang, Y.-C. Liang, and W. Zhang, “Hierarchical passive beamforming for reconfigurable intelligent surface aided communications,” IEEE Wireless Commun. Lett., vol. 10, no. 9, pp. 1909–1913, Sep. 2021.

[25] A. Albanese, F. Devoti, V. Sciancalepore, M. Di Renzo, and X. Costa-Pérez, “MARISA: A self-configuring metasurfaces absorption and reflection solution towards 6G,” 2021, arXiv:2112.01949.

[26] M. Di Renzo et al., “Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come,” EURASIP J. Wireless Commun. Netw., vol. 2019, no. 1, pp. 1–20, May 2019.

[27] X. Xu, X. Wei, and Z. Ye, “DOA estimation based on sparse signal recovery utilizing weighted 1-norm penalty,” IEEE Signal Process. Lett., vol. 19, no. 3, pp. 155–158, Mar. 2012.

[28] N. Hu, Z. Ye, X. Xu, and M. Bao, “DOA estimation for sparse array via sparse signal reconstruction,” IEEE Trans. Aerosp. Electron. Syst., vol. 49, no. 2, pp. 760–773, Apr. 2013.

[29] J. P. Vila and P. Schniter, “Expectation-maximization Gaussian-mixture approximate message passing,” IEEE Trans. Signal Process., vol. 61, no. 19, pp. 4658–4672, Oct. 2013.

[30] S. Lloyd, “Least squares quantization in PCM,” IEEE Trans. Inf. Theory, vol. IT-28, no. 2, pp. 129–137, Mar. 1982.

[31] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” J. Roy. Stat. Soc., Ser. B Methodolog., vol. 39, no. 1, pp. 1–22, Sep. 1977.

[32] U. S. Kamilov, S. Rangan, A. K. Fletcher, and M. Unser, “Approximate message passing with consistent parameter estimation and applications to sparse learning,” IEEE Trans. Inf. Theory, vol. 60, no. 5, pp. 2969–2985, May 2014.

[33] S. K. Sahoo and A. Makur, “Signal recovery from random measurements via extended orthogonal matching pursuit,” IEEE Trans. Signal Process., vol. 63, no. 10, pp. 2572–2581, May 2015.

[34] J. Ma, X. Yuan, and L. Ping, “Turbo compressed sensing with partial DFT sensing matrix,” IEEE Signal Process. Lett., vol. 22, no. 2, pp. 158–161, Feb. 2015.

[35] D. P. Wipf and B. D. Rao, “An empirical Bayesian strategy for solving the simultaneous sparse approximation problem,” IEEE Trans. Signal Process., vol. 55, no. 7, pp. 3704–3716, Jul. 2007.

[36] T. K. Moon, “The expectation-maximization algorithm,” IEEE Signal Process. Mag., vol. 13, no. 6, pp. 47–60, Nov. 1997.

[37] K. L. H. L. Van Trees and Z. Tian, Detection Estimation Modulation Theory, Part I: Detection, Estimation, Filtering Theory, Detection Estimation Modulation Theory. Hoboken, NJ, USA: Wiley, 2013.

[38] W. Tang et al., “Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement,” IEEE Trans. Wireless Commun., vol. 20, no. 1, pp. 421–439, Jan. 2021.

[39] S. Killiback and R. A. Leibler, “On information and sufficiency,” Ann. Math. Statist., vol. 22, no. 1, pp. 79–86, 1951.
Wei Zhang (Fellow, IEEE) received the Ph.D. degree from The Chinese University of Hong Kong in 2005. Currently, he is a Professor with the School of Electrical Engineering and Telecommunications, The University of New South Wales, Sydney, Australia. His current research interests include 5G and beyond. He received six best paper awards from IEEE conferences and ComSoc technical committees. He was an IEEE ComSoc Distinguished Lecturer (2016–2017). He is also the Vice President of IEEE Communications Society. Within the IEEE ComSoc, he has taken many leadership positions, including a Member-at-Large on the Board of Governors (2018–2020), the Chair of the Wireless Communications Technical Committee (2019–2020), the Vice Director of Asia-Pacific Board (2016–2021), the Editor-in-Chief of IEEE WIRELESS COMMUNICATIONS LETTERS (2016–2019), the Technical Program Committee Chair of APCC 2017 and ICC 2019, the Award Committee Chair of Asia-Pacific Board, and the Award Committee Chair of Technical Committee on Cognitive Networks. In addition, he has served as a member for various ComSoc boards/standing committees, including Journals Board, the Technical Committee, the Recertification Committee, the Finance Standing Committee, the Information Technology Committee, the Steering Committee of IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, and the Steering Committee of IEEE NETWORKING LETTERS. He also serves as an Area Editor for IEEE TRANSACTIONS ON COMMUNICATIONS and the Editor-in-Chief for Journal of Communications and Information Networks. Previously, he served as an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, and IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS—Cognitive Radio Series.

Marco Di Renzo (Fellow, IEEE) received the Laurea (cum laude) and Ph.D. degrees in electrical engineering from the University of L’Aquila, Italy, in 2003 and 2007, respectively, and the Habilitation à Diriger des Recherches (Doctor of Science) degree from University Paris-Sud (currently Paris-Saclay University), France, in 2013. Currently, he is a CNRS Research Director (Professor) and the Head of the Intelligent Physical Communications group in the Laboratory of Signals and Systems (L2S) at Paris-Saclay University—CNRS and CentraleSupelec, Paris, France. Also, he is an elected member of the L2S Board Council and a member of the L2S Management Committee, and is a Member of the Admission and Evaluation Committee of the Ph.D. School on Information and Communication Technologies, Paris-Saclay University. He is a Founding Member and the Academic Vice Chair of the Industry Specification Group (ISG) on Reconfigurable Intelligent Surfaces (RIS) within the European Telecommunications Standards Institute (ETSI), where he served as the Rapporteur for the work item on communication models, channel models, and evaluation methodologies. He is a Fellow of the IEEE, IET, and AAIA; an Ordinary Member of the European Academy of Sciences and Arts, an Ordinary Member of the Academia Europaea; and a Highly Cited Researcher. Also, he holds the 2023 France-Nokia Chair of Excellence in ICT, and was a Fulbright Fellow at City University of New York, USA, a Nokia Foundation Visiting Professor, and a Royal Academy of Engineering Distinguished Visiting Fellow. His recent research awards include the 2021 EURASIP Best Paper Award, the 2022 IEEE COMSOC Outstanding Paper Award, the 2022 Michel Monpetit Prize conferred by the French Academy of Sciences, the 2023 EURASIP Best Paper Award, the 2023 IEEE ICC Best Paper Award (wireless), the 2023 IEEE COMSOC Fred W. Ellersick Prize, the 2023 IEEE COMSOC Heinrich Hertz Award, the 2023 IEEE VTS James Evans Avant Garde Award, and the 2023 IEEE COMSOC Technical Recognition Award from the Signal Processing and Computing for Communications Technical Committee. He served as the Editor-in-Chief of IEEE COMMUNICATIONS LETTERS during the period 2019-2023, and he is now serving in the Advisory Board. In 2024 to 2025, he will be serving as a Voting Member of the Fellow Evaluation Standing Committee and as the Director of Journals of the IEEE Communications Society.