# Message Passing-Based Sparse Channel Estimation From Partially Coherent Measurement Vectors

Masoumi, Hamed; Myers, Nitin Jonathan

**DOI**
10.1109/EuCNC/6GSummit63408.2025.11037155

**Publication date**
2025

**Document Version**
Final published version

**Published in**
Proceedings of the Joint European Conference on Networks and Communications and 6G Summit, EuCNC/6G Summit 2025

2025 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit) | 979-8-3503-9180-0/25/$31.00 ©2025 IEEE | DOI: 10.1109/EUCNC/6GSUMMIT63408.2025.11037155

# Message Passing-Based Sparse Channel Estimation from Partially Coherent Measurement Vectors

Hamed Masoumi, Nitin Jonathan Myers
*Delft Center for Systems and Control, Delft University of Technology, The Netherlands*
Email: {H.Masoumi, N.J.Myers}@tudelft.nl

***Abstract*—Phase jitter at oscillators in high-frequency wireless systems perturbs the phase of the acquired channel measurements. As a result, standard sparse channel estimation algorithms that ignore phase errors fail. In this paper, we consider a frame structure in which channel measurements are acquired over two packets. Our model assumes that the phase errors are nearly constant over a packet and they change considerably across the packets, leading to partially coherent channel measurements. In this paper, we develop a message-passing-based technique that leverages the partially coherent structure in the measurements for sparse channel estimation robust to unknown phase offset across the packets. Simulation results show that our approach achieves a lower mean-squared error in the reconstructed channel than comparable benchmarks.**

***Keywords*—Message passing, phase error, mmWave, terahertz, channel estimation, compressed sensing.**

## I. Introduction

The high scattering at mmWave and THz frequencies leads to channels that are approximately sparse in the angle domain [1]. Compressive sensing (CS) leverages this sparsity to estimate channels from sub-Nyquist measurements. In practice, the measurements are corrupted by jitter at the oscillators [2]. Standard CS-based channel estimation techniques [3], [4], which assume no phase errors, fail due to a model mismatch. The phase errors are more pronounced in mmWave and THz frequency bands [5] than at lower frequencies [6].

In IEEE 802.11ad/ay systems, phase errors in channel measurements are similar within a packet but can vary considerably between packets [7]. Measurements acquired under such phase errors are termed partially coherent CS measurements [8]. Prior work on CS-based channel estimation [9]–[11] addressed this issue by relying only on the magnitude of the CS measurements for sparse recovery. These methods, however, do not leverage the partially coherent phase error structure in recovery. The work in [12] uses an extended Kalman filter within CS to track phase errors and obtains a one-sparse estimate of the channel. Such an estimate is usually insufficient to configure hybrid beamformers. Lastly, the lifting technique [13] was discussed in [7] to estimate the unknown phase errors and the sparse signal. Lifting-based channel estimation is computationally demanding, as it solves for the lifted signal whose dimension is higher than the original sparse signal [7].

Several sparse recovery algorithms that leverage partial phase coherence in CS measurements were developed in [7], [8], [14]. In [8], [14], the partial coherence structure is employed to first compute a coarse estimate of the sparse angle-domain channel, which is then refined through iterative updates. The work in [7] combines matching pursuit with alternating optimization to jointly recover the sparse angle-domain channel vector and the associated phase errors. The approaches in [7], [8], [14], however, assume prior knowledge of the number of non-zero elements in the sparse channel—an assumption that is impractical in real-world scenarios. We do not make such an assumption to derive the proposed method.

In this paper, we develop a sparse channel estimation method that exploits the partially coherent structure in the phase perturbed measurements. In our approach, the phase errors are absorbed into the sparse channel resulting in a pair of phase mismatched sparse vectors. These vectors have the same support and magnitude as the channel. We develop a message passing-based algorithm that exploits the shared support and magnitude structure across the phase mismatched sparse vectors to reconstruct the channel. Although we consider channel estimation using measurements from two packets, our technique can also be extended to the multi-packet case by considering two successive packets at a time. This extension, however, is beyond the scope of this paper.

**Notation**: We use $a$, $\mathbf{a}$ and $\mathbf{A}$ to denote a scalar, vector, and a matrix. We use $(\cdot)^{\mathrm{T}}$and $(\cdot)^*$ to denote the transpose and conjugate-transpose operators. The $i^{\text{th}}$ entry of $\mathbf{a}$ is indicated by $a[i]$. We denote the $(i,j)^{\text{th}}$ entry of $\mathbf{A}$ by $A(i,j)$. We use $\mathsf{x}$ for a random variable and $x$ for its realization, with $p(x)$ as the probability density function (PDF) of $\mathsf{x}$ evaluated at $x$. Lastly, $\mathbb{R}_+^{N\times 1}$ is the set of $N\times 1$ vectors with non-negative real entries, and $\mathrm{j}=\sqrt{-1}$.

## II. System model

We consider an $N\times 1$ half-wavelength spaced uniform linear phased array (ULA) at the transmitter (TX) and a single antenna receiver (RX). We use $\mathbf{h}$ to denote the $N$ dimensional spatial channel between the TX and RX. Let $L$ denote the number of propagation rays in the environment with the $\ell^{\text{th}}$ ray having a complex gain of $\beta_\ell$, an angle-of-departure (AoD) $\theta_\ell$. By defining the beamspace angle as $\omega_\ell = \pi\sin\theta_\ell$ [15] and the $N\times 1$ Vandermonde vector $\mathbf{a}_N(\omega)$ as $\mathbf{a}_N(\omega) = \left[1, e^{\mathrm{j}\omega}, e^{\mathrm{j}2\omega}, \cdots, e^{\mathrm{j}(N-1)\omega}\right]^{\mathrm{T}}$, the baseband channel matrix $\mathbf{h}$ is given by

$$\mathbf{h} = \sum_{\ell=1}^{L} \beta_\ell \mathbf{a}_N(\omega_\ell). \tag{1}$$

The channel dimension $N$ in typical mmWave or THz phased arrays can be in the order of hundreds to thousands.

The high scattering at mmWave and THz frequencies results in an approximately sparse channel in the angle domain

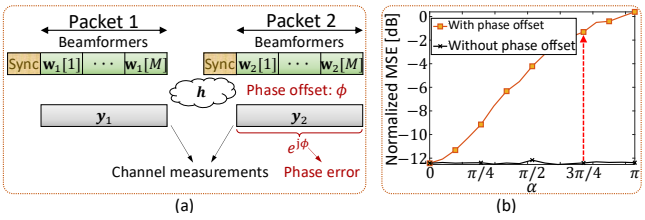


Fig. 1. Channel measurements acquired over different packets in IEEE 802.11ad/ay are partially coherent. Here, $\phi$ in (a) denotes the relative phase offset introduced in the measurements of the second packet. We notice from (b) that CS-based channel estimation with [17] performs poor when the relative phase offset is ignored. Here, $M = 40$ measurements of a $N = 256$ dimension channel are acquired in each packet and $\phi$ is uniformly distributed over $[-\alpha, \alpha)$.

[16]. To exploit this property for channel reconstruction, the channel $\mathbf{h}$ is transformed into the angle domain. Since we assume a ULA at the TX, the DFT of $\mathbf{h}$ is used for its angle-domain representation. Let $\mathbf{U}_N$ denote the standard $N \times N$ unitary DFT matrix and $\mathbf{x}$ denote the angle-domain representation of $\mathbf{h}$. Then, $\mathbf{x}$ and $\mathbf{h}$ are related as

$$\mathbf{h} = \mathbf{U}_N \mathbf{x}, \tag{2}$$

where $\mathbf{x}$ is the unknown sparse vector to be estimated.

We use the frame structure shown in Fig. 1, inspired by the signaling structure in IEEE 802.11ad/ay [18]. We consider only 2 packets, and extending our method to measurements acquired over multiple packets is beyond the scope of this paper. We use $M$ to denote the number of spatial channel measurements acquired in each packet. The RX acquires one spatial measurement of the channel when the TX applies a beam training vector to its array. This spatial measurement is corrupted by phase error due to hardware impairment, such as the jitter at the oscillators [2]. The phase error depends on the time interval between the measurements [2]. Since the time interval between packets is typically large (e.g., 44 $\mu$s) compared to the time needed to acquire each measurement (e.g., 128 ns) [7], [14], the phase variations within each packet can be ignored. The phase offsets across different packets, however, are significant due to the large time lapse and cannot be ignored.

We denote the relative phase error induced in the second packet measurements as $\phi$. Here, $\phi$ is modeled as a uniformly distributed random variable $\mathcal{U}([-\pi, \pi))$. We use $\mathbf{w}_\ell[m] \in \mathbb{C}^{N\times 1}$ to denote the beam training vector applied at the TX in the $\ell^{\text{th}}$ packet, for the RX to obtain the $m^{\text{th}}$ spatial channel measurement denoted by $y_\ell[m]$. We use the random variable $\mathsf{q}_\ell[m] \sim \mathcal{CN}(0, \sigma^2)$ to model the measurement noise and $q_\ell[m]$ to denote a realization of $\mathsf{q}_\ell[m]$. The $m^{\text{th}}$ spatial channel measurement acquired in packet 1 and 2 is

$$y_1[m] = \mathbf{w}_1^*[m]\mathbf{h} + q_1[m], \quad y_2[m] = e^{j\phi}\mathbf{w}_2^*[m]\mathbf{h} + q_2[m]. \tag{3}$$

Our goal is to estimate $\mathbf{h}$ from $2M$ phase-perturbed measurements in (3) by exploiting its angle-domain sparsity. To this end, we define $\mathbf{A}_\ell \in \mathbb{C}^{M\times N}$ as the CS matrix associated with the $M$ measurements acquired with (3) for the $\ell^{\text{th}}$ packet. The $m^{\text{th}}$ row of $\mathbf{A}_\ell$ is $\mathbf{w}_\ell^*[m]\mathbf{U}_N$ [19]. Let $\mathbf{y}_\ell$ denote the vector of $M$ measurements $\{y_\ell[m]\}_{m=1}^M$ in (3) and $\mathbf{q}_\ell$ contain the corresponding measurement noise $\{q_\ell[m]\}_{m=1}^M$. Now, we can write the measurement vector $\mathbf{y}_\ell = [y_\ell[1], \cdots y_\ell[M]]^{\mathrm{T}}$ as

$$\mathbf{y}_1 = \mathbf{A}_1\mathbf{x} + \mathbf{q}_1, \qquad \mathbf{y}_2 = e^{j\phi}\mathbf{A}_2\mathbf{x} + \mathbf{q}_2. \tag{4}$$

Here, $\{\mathbf{y}_1, \mathbf{y}_2\}$ are acquired at the RX, and the CS matrices $\{\mathbf{A}_1, \mathbf{A}_2\}$ are known from the beams applied at the TX. Further, $\mathbf{x}$ is the unknown to be estimated from the measurements acquired under an unknown phase offset $\phi$. As we observe from Fig. 1(b), standard CS-based channel estimation with [17], that is agnostic to the random phase offset $\phi$, fails when the standard deviation of the offset is large.

## III. Proposed Channel Estimation Method

To solve (4) under a sparse prior on $\mathbf{x}$, we absorb the unknown phase offset into $\mathbf{x}$ and define a new unknown vector

$$\mathbf{z} = e^{j\phi}\mathbf{x}. \tag{5}$$

Since $\mathbf{z}$ is just a non-zero scalar multiple of $\mathbf{x}$, a sparse $\mathbf{x}$ leads to a sparse $\mathbf{z}$. Let $\mathcal{S}_\mathbf{x}$ denote the support of $\mathbf{x}$, i.e., $\mathcal{S}_\mathbf{x} \triangleq \{n : x[n] \neq 0\}$, and $\mathcal{S}_\mathbf{z}$ denote the support of $\mathbf{z}$. From (5), we observe that $\mathbf{z}$ has the same support as $\mathbf{x}$, i.e., $\mathcal{S}_\mathbf{z} = \mathcal{S}_\mathbf{x}$. In addition, the entry-wise magnitudes of $\mathbf{x}$ and $\mathbf{z}$ are equal, i.e., $|\mathbf{z}| = |\mathbf{x}|$. We use (5) to rewrite $\mathbf{y}_2$ in (4) to obtain

$$\mathbf{y}_1 = \mathbf{A}_1\mathbf{x} + \mathbf{q}_1, \qquad \mathbf{y}_2 = \mathbf{A}_2\mathbf{z} + \mathbf{q}_2. \tag{6}$$

By absorbing the phase error into the sparse vector, the measurement model for $\mathbf{y}_2$ in (6) is a linear measurement model in the unknown vector $\mathbf{z}$. A naive approach to reconstruct $\mathbf{x}$ using (6) is using standard CS over $\mathbf{y}_1$. This approach, referred to as *Packet-1-only* baseline, is suboptimal as it could have used $\mathbf{y}_2$ while exploiting the correlation between $\mathbf{x}$ and $\mathbf{z}$.

Now, we propose two methods that exploit different levels of correlation between $\mathbf{x}$ and $\mathbf{z}$ to reconstruct $\mathbf{x}$ from (6). The first method exploits sparsity and the common support in $\mathbf{x}$ and $\mathbf{z}$, ignoring the shared magnitude $|\mathbf{x}| = |\mathbf{z}|$. We refer to this method as *Proposed SuppOnly*, which is solved using the multiple measurement vector approximate message passing (MMV-AMP) algorithm from [20]. The focus of this paper is on our second method, which not only exploits sparsity and the shared support in $\mathbf{z}$ and $\mathbf{x}$ but also leverages their common magnitude structure. Hence, we refer to the second method as *Proposed MagSupp*.

### A. Proposed SuppOnly method

This method reconstructs $\mathbf{x}$ by exploiting sparsity and the common support structure using the MMV-AMP algorithm [20]. To explain how MMV-AMP is used, we model $\mathbf{z}$ as a realization of a random vector $\mathbf{z}$ and $\mathbf{x}$ as a realization of a random vector $\mathbf{x}$. We define $\mathbf{s} \in \{0, 1\}^N$ as the $N$-dimensional random binary vector, with independent and identically distributed (IID) entries, to model the support of $\mathbf{z}$ and $\mathbf{x}$. We use $\lambda$ to denote the probability that $\mathsf{x}[n]$, which is equivalent to the probability that $\mathsf{z}[n]$ is non-zero. Then,

$$p(s[n]) = \lambda^{s[n]}\,(1-\lambda)^{1-s[n]}, \quad s[n] \in \{0, 1\}. \tag{7}$$

We use $\mathbf{u}$ to model the amplitude of the random vector $\mathbf{x}$ whose entries are independently and identically distributed (IID) as a zero mean Gaussian distribution with variance $\gamma$, i.e.,

$$p(u[n]) = \mathcal{CN}\left(u[n]; 0, \gamma\right). \tag{8}$$

Let $\mathbf{v}$ model the amplitude of the random vector $\mathbf{z}$. Since the entries of the amplitude of $\mathbf{x}$ are IID circularly symmetric Gaussian distributed and the entries of $\mathbf{z}$ are just phase perturbed versions of the entries of $\mathbf{x}$ from (5), the entries of the random vector $\mathbf{v}$ are also IID as a zero mean Gaussian distribution with variance $\gamma$, i.e., $p(v[n]) = \mathcal{CN}(v[n]; 0, \gamma)$.

Now, $\mathsf{x}[n]$ and $\mathsf{z}[n]$ can be decomposed as the product of their corresponding amplitude variable and the support variable as $\mathsf{x}[n] = \mathsf{s}[n]\mathsf{u}[n]$ and $\mathsf{z}[n] = \mathsf{s}[n]\mathsf{v}[n]$, equivalently

$$p\Big(x[n]\Big|s[n], u[n]\Big) = \delta(x[n] - s[n]u[n]), \tag{9}$$

$$p\Big(z[n]\Big|s[n], v[n]\Big) = \delta(z[n] - s[n]v[n]). \tag{10}$$

From (9) and (10) we observe that $x[n] = 0$ and $z[n] = 0$ when $s[n] = 0$, thereby aiding sparse priors. The shared support structure between $\mathbf{x}$ and $\mathbf{z}$ is inherently incorporated as their supports are modeled by a single random vector $\mathbf{s}$.

To estimate $\mathbf{x}$ while leveraging its shared support with $\mathbf{z}$, the MMV-AMP algorithm from [20] can be employed. The MMV-AMP uses the Bayes' rule to factorize the joint distribution of $\mathbf{z}$, $\mathbf{x}$, $\mathbf{s}$, $\mathbf{u}$, and $\mathbf{v}$, in terms of the probabilities in (7)–(9), $\{p(v[n])\}_{n=1}^{N}$ and the likelihoods

$$p\Big(y_1[m]\Big|\mathbf{x}\Big) = \mathcal{CN}\big(y_1[m]; \mathbf{A}_1(m,:)\mathbf{x}, \sigma^2\big), \tag{11}$$

$$p\Big(y_2[m]\Big|\mathbf{z}\Big) = \mathcal{CN}\big(y_2[m]; \mathbf{A}_2(m,:)\mathbf{z}, \sigma^2\big), \tag{12}$$

where $\mathbf{A}_1(m,:)$ and $\mathbf{A}_2(m,:)$ denote the $m^{\text{th}}$ rows of $\mathbf{A}_1$ and $\mathbf{A}_2$. Then, multiple message exchanges in the factor graph corresponding to the factorization are performed to determine the marginal posterior distribution of $\{\mathsf{x}[n]\}_{n=1}^{N}$. The marginal posterior is finally used to compute the minimum mean squared error (MMSE) estimate of $\mathbf{x}$.

### *B. Proposed MagSupp method*

Since MMV-AMP cannot leverage the shared magnitude structure between $\mathbf{x}$ and $\mathbf{z}$, we propose a new message-passing method to address this limitation. In our method, the amplitude random vectors $\mathbf{u}$ and $\mathbf{v}$ are decomposed into their magnitude and phase to exploit the shared magnitude structure. To this end, we use $\mathsf{r}[n]$ to model the magnitude of $\mathsf{x}[n]$ and $\mathsf{z}[n]$. This means that $r[n] \in \mathbb{R}_+$. Furthermore, we use $\Theta_1[n]$ to model the phase of the random variable $\mathsf{x}[n]$ and $\Theta_2[n]$ to model the phase of the random variable $\mathsf{z}[n]$. Since $\mathsf{u}[n]$ and $\mathsf{v}[n]$ have a complex Gaussian PDF with mean zero and variance $\gamma$, their magnitude $\mathsf{r}[n]$ has the Rayleigh PDF

$$p(r[n]) = 2\frac{r[n]}{\gamma}e^{-r^2[n]/\gamma}, \tag{13}$$

and their phase values are uniform IID over $[-\pi, \pi)$, i.e.,

$$p(\theta_1[n]) = p(\theta_2[n]) = \frac{1}{2\pi}, \quad \{\theta_1[n], \theta_2[n]\} \in [-\pi, \pi). \tag{14}$$

We note that incorporating a common phase error across all the elements of $\mathbf{z}$ is challenging as it leads to numerous loops in the corresponding factor graph. Although $\theta_1[n] - \theta_2[n] = \phi\,\forall n$ in our model due to a scalar phase offset, we assume that $\theta_1[n]$ and $\theta_2[n]$ are independent for each $n$ to develop a tractable message-passing method. Finally, we express $\mathsf{x}[n] = \mathsf{s}[n]\mathsf{r}[n]e^{\mathrm{j}\Theta_1[n]}$ and $\mathsf{z}[n] = \mathsf{s}[n]\mathsf{r}[n]e^{\mathrm{j}\Theta_2[n]}$, equivalently

$$f_1[n] \triangleq p\Big(x[n]\Big|s[n], r[n], \theta_1[n]\Big) = \delta\Big(x[n] - s[n]r[n]e^{\mathrm{j}\theta_1[n]}\Big), \tag{15}$$

$$f_2[n] \triangleq p\Big(z[n]\Big|s[n], r[n], \theta_2[n]\Big) = \delta\Big(z[n] - s[n]r[n]e^{\mathrm{j}\theta_2[n]}\Big). \tag{16}$$

As the supports of $\mathbf{x}$ and $\mathbf{z}$ are modeled by a single random vector $\mathbf{s}$ and their magnitudes are modeled by a single random vector $\mathbf{r}$, our method inherently incorporates the shared support and magnitude structure.

To develop our algorithm to estimate $\mathbf{x}$, we use the Bayes' rule and the dependencies between $\mathbf{z}$, $\mathbf{s}$, $\mathbf{r}$, and $\{\Theta_\ell\}_{\ell=1}^{2}$ to factorize their posterior joint PDF as

$$\begin{aligned} &p\Big(\mathbf{x}, \mathbf{z}, \mathbf{s}, \mathbf{r}, \{\boldsymbol{\theta}_\ell\}_{\ell=1}^{2}\Big|\{\mathbf{y}_\ell\}_{\ell=1}^{2}\Big) \\ &\propto \prod_{m=1}^{M} p\Big(y_1[m]\Big|\mathbf{x}\Big)\, p\Big(y_2[m]\Big|\mathbf{z}\Big) \prod_{n=1}^{N}\Big(f_1[n]f_2[n]p(\theta_1[n]) \\ &\qquad p(\theta_2[n])p\left(s[n]\right)p\left(r[n]\right)\Big), \end{aligned} \tag{17}$$

where $\propto$ denotes equality up to a constant scale factor. The dependencies among different variables in (17) are represented by the factor graph shown in Fig. 2. This factor graph consists of 2 planes, each containing a collection of round nodes indicating the unknown variables and rectangular nodes indicating probability distributions. In plane 1, standard AMP is used to compute marginal posteriors on $\{\mathsf{x}[n]\}_{n=1}^{N}$ where the sparsity is enforced through the messages from $\{f_1[n]\}_{n=1}^{N}$ and faithfulness to measurements through $\{p\Big(y_1[m]\Big|\mathbf{x}\Big)\}_{m=1}^{M}$. A similar procedure is performed in plane 2 to compute marginal posteriors on $\{\mathsf{z}[n]\}_{n=1}^{N}$. We observe that the planes are connected through $\{\mathsf{s}[n]\}_{n=1}^{N}$ and $\{\mathsf{r}[n]\}_{n=1}^{N}$. The messages that are exchanged through these connections across the planes make sure that $\mathbf{x}$ and $\mathbf{z}$ have the same support and magnitude. After a few iterations of message exchanges between the nodes in the factor graph, an estimate of the marginal posteriors of $\{x[n]\}_{n=1}^{N}$, conditioned on the observed measurements, is obtained and its mean is calculated for the MMSE estimate.

We explain the messages exchanged in our factor graph in Fig. 2. Let $\nu_{i\to j}$ denote the message passed from node $i$ to node $j$. For ease of exposition, we focus on the messages between nodes with index $n$. Similar to MMV-AMP [20], the message exchanged between the factors $f_\ell[n]$ and the support variable $\mathsf{s}[n]$ is a Bernoulli probability mass function (PMF), e.g., $\nu_{f_\ell[n]\to\mathsf{s}[n]} = (\vec{\pi}_\ell[n], 1 - \vec{\pi}_1[n])$ where $\vec{\pi}_\ell[n]$ is the belief that $\mathsf{s}[n]$ is 1. In compact form, we denote this Bernoulli PMF as $\nu_{f_\ell[n]\to\mathsf{s}[n]} = \vec{\pi}_\ell[n]$. Similarly, $\nu_{\mathsf{s}[n]\to f_\ell[n]} = \overleftarrow{\pi}_\ell[n]$.

**Initial message passes with plane 1:** Message passing starts with the messages sent from the support $\mathsf{s}[n]$, magnitude $\mathsf{r}[n]$, and phase $\Theta_1[n]$ variable nodes into $f_1[n]$ in plane 1. As no side information is available for plane 1, these messages are merely the priors of each variable. Since $p(s[n])$ is a Bernoulli PMF from (7), $p(r[n])$ is a Rayleigh PDF from (13) and $p\left(\Theta_1[n]\right)$ is a uniform PDF over $[-\pi, \pi)$, it follows from (15) and the sum-product rule that $\nu_{f_1[n]\to\mathsf{z}_1[n]}$ has the Bernoulli-Gaussian PDF

$$\nu_{f_1[n]\to\mathsf{x}[n]} = (1 - \lambda)\delta(x[n]) + \lambda\mathcal{CN}(x[n]; 0, \gamma), \tag{18}$$

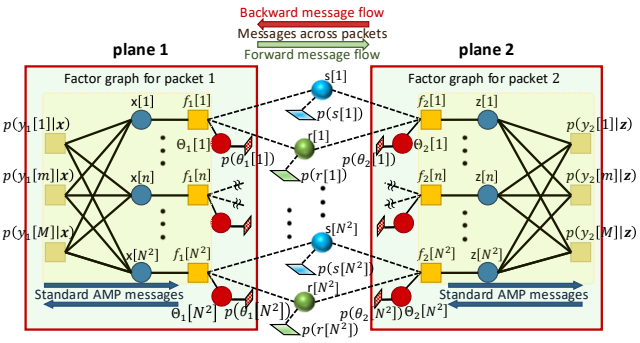


Fig. 2. Our factor graph to estimate $\mathbf{x}$ from the partially coherent measurements $\{\mathbf{y}_1, \mathbf{y}_2\}$ in (4). Standard AMP is performed in each plane corresponding to a measurement vector. Our key idea is to use posteriors from one plane as side-information for the other plane.

which is used as a local prior for $\mathsf{x}[n]$. Next, these priors and the likelihoods $p\Big(y_1[m]\Big|\mathbf{x}\Big)$s are used to perform standard AMP. After AMP iterations, the estimated posterior of $\mathsf{x}[n]$ based only on the measurements from packet 1 is obtained. We use $e_1[n]$ to denote the mean and $c_1[n]$ to denote the variance of this posterior. At the end of AMP iterations, the messages $\{\nu_{\mathsf{x}[n]\to f_1[n]}\}_{n=1}^{N}$ given by

$$\nu_{\mathsf{x}[n]\to f_1[n]} = \mathcal{CN}(x[n]; e_1[n], c_1[n])\,, \tag{19}$$

are passed through the common support and magnitude variable nodes $\mathsf{s}[n]$ and $\mathsf{r}[n]$ to plane 2 as side information.

**Messages sent from plane 1 to plane 2 through the magnitude variables:** At this stage, the beliefs on $\mathsf{r}[n]$s are updated using the posteriors in (19) from plane 1. By the sum-product algorithm, $\nu_{f_1[n]\to\mathsf{r}[n]}$ is derived by first computing the product of $\nu_{\mathsf{s}[n]\to f_1[n]}$, the posterior from (19) and $\nu_{\Theta_1[n]\to f_1[n]} = 1/(2\pi)$. This product is then integrated over $\mathsf{s}[n]$ and $\Theta_1[n]$ to write

$$\begin{aligned}\nu_{f_1[n]\to\mathsf{r}[n]} =& (1-\overleftarrow{\pi}_1[n])\,\mathcal{CN}(0; e_1[n], c_1[n])\\ &+\frac{\overleftarrow{\pi}_1[n]}{2\pi}\int_{\theta_1[n]}\mathcal{CN}\Big(r[n]e^{\mathrm{j}\theta_1[n]}; e_1[n], c_1[n]\Big)\,.\end{aligned} \tag{20}$$

The first term in (20), corresponding to $s[n]=0$, is constant with respect to the variable $r[n]$ and this constant makes $\nu_{f_1[n]\to\mathsf{r}[n]}$ an improper distribution which does not integrate to 1 [20]. To address this issue, we use the idea from [20] that considers $\mathsf{s}[n]\in\{\epsilon, 1\}$ in the limit $\epsilon\to 0$, and define

$$\Omega\left(\overleftarrow{\pi}_1[n]\right) = \frac{\epsilon^2\overleftarrow{\pi}_1[n]}{(1-\overleftarrow{\pi}_1[n])+\epsilon^2\overleftarrow{\pi}_1[n]}. \tag{21}$$

For a small $\epsilon$, the message $\nu_{f_1[n]\to\mathsf{r}[n]}$ in (20) is modified to

$$\begin{aligned}&\bar{\nu}_{f_1[n]\to\mathsf{r}[n]} =\\ &\frac{(1-\Omega(\overleftarrow{\pi}_1[n]))}{2\pi}\int_{\theta_1[n]}\mathcal{CN}\Big(r[n]e^{\mathrm{j}\theta_1[n]}; \frac{1}{\epsilon}e_1[n], \frac{1}{\epsilon^2}c_1[n]\Big)\\ &+\frac{\Omega(\overleftarrow{\pi}_1[n])}{2\pi}\int_{\theta_1[n]}\mathcal{CN}(r[n]e^{\mathrm{j}\theta_1[n]}; e_1[n], c_1[n]),\end{aligned} \tag{22}$$

which can be shown to be a proper distribution. Now, by using the identity $\int_{-\pi}^{\pi}\exp\big(\Re\big(\kappa e^{-\mathrm{j}\theta}\big)\big)\,d\theta = 2\pi\mathrm{I}_0(|\kappa|)$ from [21, 3.5.17] for any complex scalar $\kappa$ with $\Re(\kappa)$ denoting its real part and $\mathrm{I}_k(\cdot)$ denoting the modified Bessel function of the first kind and order $k$, we can simplify $\bar{\nu}_{f_1[n]\to\mathsf{r}[n]}$ as

$$\begin{aligned}&\bar{\nu}_{f_1[n]\to\mathsf{r}[n]} =\\ &\frac{1-\Omega\left(\overleftarrow{\pi}_1[n]\right)}{\pi\epsilon^{-2}c_1[n]}\exp\left(-\frac{r^2[n]+\epsilon^{-2}|e_1[n]|^2}{\epsilon^{-2}c_1[n]}\right)\mathrm{I}_0\left(\frac{2\,|e_1[n]|}{\epsilon^{-1}c_1[n]}r[n]\right)\\ &+\frac{\Omega\left(\overleftarrow{\pi}_1[n]\right)}{\pi c_1[n]}\exp\left(-\frac{r^2[n]+|e_1[n]|^2}{c_1[n]}\right)\mathrm{I}_0\left(\frac{2\,|e_1[n]|}{c_1[n]}r[n]\right).\end{aligned} \tag{23}$$

Now, we can apply the sum-product rule to compute the message $\nu_{\mathsf{r}[n]\to f_2[n]}$ as the product of the prior $p(r[n])$ and the modified message $\bar{\nu}_{f_1[n]\to\mathsf{r}[n]}$ from (23). This message can be simplified to a weighted sum of two Rician PDFs [22]. To this end, we define

$$\tau_1^{\epsilon}[n] = \frac{1}{2}\left(\frac{1}{\gamma}+\frac{1}{\epsilon^{-2}c_1[n]}\right)^{-1},\quad \rho_1^{\epsilon}[n] = \frac{\gamma\epsilon^{-1}|e_1[n]|}{\gamma+\epsilon^{-2}c_1[n]}, \tag{24}$$

$$\tau_1[n] = \frac{1}{2}\left(\frac{1}{\gamma}+\frac{1}{c_1[n]}\right)^{-1},\qquad \rho_1[n] = \frac{\gamma|e_1[n]|}{\gamma+c_1[n]}. \tag{25}$$

The message $\nu_{\mathsf{r}[n]\to f_2[n]}$ is simplified to

$$\begin{aligned}\nu_{\mathsf{r}[n]\to f_2[n]} =& (1-\Omega(\overleftarrow{\pi}_1[n]))\,\mathcal{CN}\big(0;\epsilon^{-1}e_1[n],\gamma+\epsilon^{-2}c_1[n]\big)\\ &\times\mathrm{Rice}(r[n];\rho_1^{\epsilon}[n],\tau_1^{\epsilon}[n])+\Omega\left(\overleftarrow{\pi}_1[n]\right)\\ &\times\mathcal{CN}(0;e_1[n],\gamma+c_1[n])\,\mathrm{Rice}(r[n];\rho_1[n],\tau_1[n]),\end{aligned} \tag{26}$$

where

$$\begin{aligned}\mathrm{Rice}\left(r[n];\rho_1[n],\tau_1[n]\right) =& \frac{r[n]}{\tau_1[n]}\exp\left(-\frac{r^2[n]+\rho_1^2[n]}{2\tau_1[n]}\right)\\ &\times\mathrm{I}_0\left(\frac{\rho_1[n]}{\tau_1[n]}r[n]\right),\end{aligned} \tag{27}$$

denotes a Rice PDF [22] with parameters $\rho_1[n]$ and $\tau_1[n]$.

**Messages sent from plane 1 to plane 2 through the support variables:** Here, we first derive $\nu_{f_1[n]\to\mathsf{s}[n]}$. From $\nu_{\Theta_1[n]\to f_1[n]} = 1/2\pi$ and the sum-product rule,

$$\nu_{f_1[n]\to\mathsf{s}[n]} = \frac{1}{2\pi}\int_{x[n]}\int_{r[n]}\int_{\theta_1[n]} f_1[n]\nu_{\mathsf{x}[n]\to f_1[n]}\nu_{\mathsf{r}[n]\to f_1[n]}. \tag{28}$$

For the Bernoulli PMF $\nu_{f_1[n]\to\mathsf{s}[n]}$, the belief that $\mathsf{s}[n]=1$ is

$$\vec{\pi}_1[n] = \frac{[\nu_{f_1[n]\to\mathsf{s}[n]}]_{\mathsf{s}[n]=1}}{[\nu_{f_1[n]\to\mathsf{s}[n]}]_{\mathsf{s}[n]=1}+[\nu_{f_1[n]\to\mathsf{s}[n]}]_{\mathsf{s}[n]=0}}. \tag{29}$$

To compute $[\nu_{f_1[n]\to\mathsf{s}[n]}]_{\mathsf{s}[n]=1}$ from (28), similar steps in deriving (26) can be employed to calculate the integrals with respect to $x[n]$ and $\theta_1[n]$. Noting that $\nu_{\mathsf{r}[n]\to f_1[n]}$ is equal to the Rayleigh PDF $p(r[n])$ at this stage, we can simplify the result using a Rice PDF as in (26) and then compute the integral with respect to $r[n]$ to obtain

$$[\nu_{f_1[n]\to\mathsf{s}[n]}]_{\mathsf{s}[n]=1} = \mathcal{CN}(0; e_1[n], \gamma+c_1[n]). \tag{30}$$

Similarly $[\nu_{f_1[n]\to\mathsf{s}[n]}]_{\mathsf{s}[n]=0} = \mathcal{CN}(0; e_1[n], c_1[n])$. Now, we can obtain $\vec{\pi}_1[n]$ in (29) as

$$\vec{\pi}_1[n] = \frac{\mathcal{CN}(0; e_1[n], \gamma+c_1[n])}{\mathcal{CN}(0; e_1[n], \gamma+c_1[n])+\mathcal{CN}(0; e_1[n], c_1[n])}. \tag{31}$$

Next, we apply the sum-product rule to compute the probability associated with the Bernoulli PMF $\nu_{\mathsf{s}[n]\to f_2[n]}$ as

$$\overleftarrow{\pi}_2[n] = \frac{\lambda \vec{\pi}_1[n]}{(1-\lambda)\,(1-\vec{\pi}_1[n]) + \lambda \vec{\pi}_1[n]}, \tag{32}$$

where the probability derived in (31) is used for $\vec{\pi}_1[n]$.

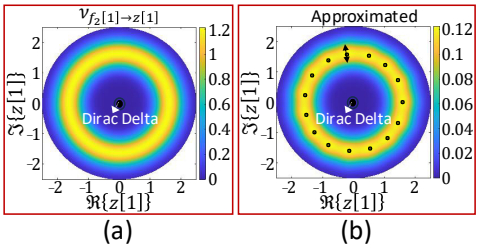


Fig. 3. As $\mathbf{x}$ and $\mathbf{z}$ share the same magnitude, the posterior of $\mathbf{x}$ can be used to obtain a belief on $\mathbf{z}$, and vice versa. This belief ($\nu_{f_2[n]\to z[n]}$) has a donut-like structure in (a) due to the phase offset uncertainty. We approximate this belief as a Bernoulli-Gaussian mixture in (b). The dots on the donut are the means of the Gaussians in this mixture.

**Computing the message $\nu_{f_2[n]\to \mathsf{z}[n]}$:** This message is computed by first evaluating the product $f_2[n]\nu_{\Theta_2[n]\to f_2[n]}\nu_{s[n]\to f_2[n]}\nu_{\mathsf{r}[n]\to f_2[n]}$, where the individual terms are listed in (16), (32) and (26). Then, the product is integrated over $\mathsf{r}[n]$, $\mathsf{s}[n]$, and $\Theta_2[n]$ to arrive at $\nu_{f_2[n]\to \mathsf{z}[n]}$. As this integration is challenging, we approximate it as a Bernoulli-Gaussian Mixture (BGM) distribution. Such an assumption aids computationally tractable messages for the AMP procedure in plane 2. An example of the true $\nu_{f_2[n]\to \mathsf{z}[n]}$ and our BGM approximation is shown in Fig. 3. This belief, originating from the preceding plane, comprises non-zero probabilities concentrated on an annular region, as illustrated in Fig. 3. The radius of this region is governed by the magnitude prior on $\mathsf{z}[1]$, while its angular spread covers the whole $-\pi$ to $\pi$ as $\Theta_1[n]$ uniformly distributed within $[-\pi, \pi)$.

Now, the BGM approximates of $\nu_{f_2[n]\to \mathsf{z}[n]}$s and the likelihoods $p\Big(y_2[m]\Big|\mathbf{z}\Big)$s are used to perform standard AMP. After AMP iterations, the estimated posterior of $\mathsf{z}[n]$ based on the measurements from packets 1 and 2 is obtained. We use $e_2[n]$ to denote the mean and $c_2[n]$ to denote the variance of this posterior. At the end of AMP iterations, the messages $\{\nu_{\mathsf{z}[n]\to f_2[n]}\}_{n=1}^{N}$ given by

$$\nu_{\mathsf{z}[n]\to f_2[n]} = \mathcal{CN}(z[n]; e_2[n], c_2[n])\,, \tag{33}$$

are passed through the common support and magnitude variable nodes $\mathsf{s}[n]$ and $\mathsf{r}[n]$ to plane 1 as side information to refine the marginal posterior estimation of $x[n]$s. We note that all the messages passed to plane 1 can be computed in the same way as those derived up to this point with the exception of the message $\nu_{f_2[n]\to \mathsf{s}[n]}$. This is because the message $\nu_{\mathsf{r}[n]\to f_2[n]}$ used to compute $\nu_{f_2[n]\to \mathsf{s}[n]}$ is a superposition of two Rice PDFs given in (26) while during the initial message passes from plane 1 to plane 2 the message $\nu_{\mathsf{r}[n]\to f_1[n]}$ used to compute $\nu_{f_1[n]\to \mathsf{s}[n]}$ had the Rayleigh PDF (13).

**Computing the message $\nu_{f_2[n]\to \mathsf{s}[n]}$:** By applying the sum-product rule similar to $\nu_{f_2[n]\to \mathsf{s}[n]}$, and the fact that $\nu_{\mathsf{r}[n]\to f_2[n]}$ is a superposition of two Rice distributions in (26), we derive

$$\begin{aligned}[\nu_{f_2[n]\to \mathsf{s}[n]}]_{\mathsf{s}[n]=1} = &\frac{(1-\Omega(\overleftarrow{\pi}_1[n]))\,\mathcal{CN}\big(0;\epsilon^{-1}e_1[n],\gamma+\epsilon^{-2}c_1[n]\big)}{2\pi\,|e_2[n]|} \\ &\times \text{Rice}\left(|e_2[n]|\,;\ \rho_1^{\epsilon}[n], \frac{2\tau_1^{\epsilon}[n]+c_2[n]}{2}\right) \\ &+ \frac{\Omega(\overleftarrow{\pi}_1[n])\,\mathcal{CN}(0; e_1[n],\gamma+c_1[n])}{2\pi\,|e_2[n]|} \\ &\times \text{Rice}\left(|e_2[n]|\,;\ \rho_1[n], \frac{2\tau_1[n]+c_2[n]}{2}\right).\end{aligned} \tag{34}$$

We used [23, 10.43.28] in computing the integral with respect to $r[n]$ to derive (34). Similarly, $[\nu_{f_b[n]\to \mathsf{s}[n]}]_{s[n]=0}$ is

$$\begin{aligned}\nu_{f_b[n]\to \mathsf{s}[n]}\big|_{s[n]=0} &= \mathcal{CN}(0; e_2[n], c_2[n]) \\ &\times \Big((1-\Omega(\overleftarrow{\pi}_1[n]))\,\mathcal{CN}\big(0;\epsilon^{-1}e_1[n],\gamma+\epsilon^{-2}c_1[n]\big) \\ &\quad + \Omega(\overleftarrow{\pi}_1[n])\,\mathcal{CN}(0; e_1[n],\gamma+c_1[n])\Big).\end{aligned} \tag{35}$$

Using (34) and (35), we compute $\vec{\pi}_2[n]$ similar to $\vec{\pi}_1[n]$.

**Closing the loop for channel estimation:** After repeated exchange of messages from plane 1 to plane 2 (in Fig. 2) and backwards, the marginal posteriors of $\{\mathsf{x}[n]\}_{n=1}^{N}$ converge. The sparse channel estimate $\hat{\mathbf{x}}$ is the mean of this posterior.

## IV. Simulation Results

We consider a uniform linear array of $N = 256$ antennas at the TX. The beam training vectors $\{\mathbf{w}_1[m]\}_{m=1}^{M}$ and $\{\mathbf{w}_2[m]\}_{m=1}^{M}$ are distinct random circular shifts of Zadoff-Chu sequences [19] applied at the ULA. Unlike [7] that assumes a known sparsity level, we only assume known statistics for $\mathbf{x}$. The entries of sparse $\mathbf{x}$ in our simulations are random realizations of a Bernoulli-Gaussian PDF (18) with sparsity level $\lambda = 0.04$ and variance $\gamma = 1.5$. We use $\hat{\mathbf{h}}$ to denote the estimated channel obtained from its angle-domain estimate $\hat{\mathbf{x}}$ by inverse DFT from (2). We define normalized mean squared error (NMSE) as $\mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}\|_2^2/\|\mathbf{h}\|_2^2]$. Signal-to-noise ratio (SNR) is given by $\text{SNR} = \left(\|\mathbf{A}_1\mathbf{x}\|^2 + \|\mathbf{A}_2\mathbf{z}\|^2\right)/(2M\sigma^2)$.

We use *SparseLift* from [13] as one of the benchmark methods. *SparseLift* translates the unknowns $\mathbf{x}$ and $\phi$ into a higher dimension as $\mathbf{\Gamma} = [\mathbf{x}, e^{\mathrm{j}\phi}\mathbf{x}]^{\mathrm{T}}$ [7]. Then, $\mathbf{\Gamma}$ is estimated using the AMP [17]. Finally, the estimate $\hat{\mathbf{\Gamma}}$ is decomposed to find $\hat{\mathbf{x}}$. We consider another benchmark termed as 'NoPhaseError" in which ideal phase error-free measurements are used for sparse recovery. This benchmark uses a linear model obtained by stacking the CS matrices for the two packets and then employs AMP [17] for sparse recovery.

From Fig. 4, we observe that our *proposed MagSupp* technique brings around 2 dB improvement in the NMSE compared to *SparseLift* for small $M$. This is because our technique leverages the structure in partially coherent measurements, such as the common support and magnitude of the phase-perturbed angle-domain channel and the original angle-domain channel in the factor graph in Fig. 2. The gap between our proposed technique and the *SparseLift*, however, is small at large $M$. Next, by comparing the *proposed SuppOnly* with the *proposed MagSupp*, we observe that exploiting the additional

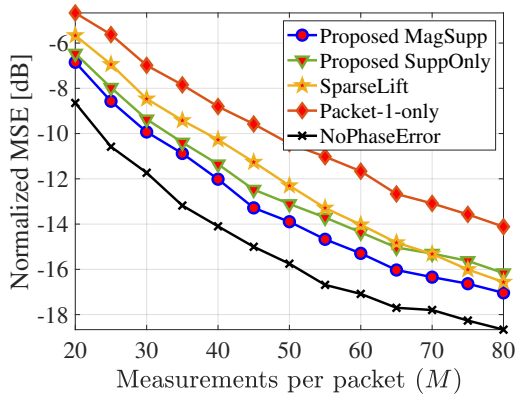


Fig. 4. NMSE with the number of CS measurements per packet for SNR $= 5$ dB. With phase-perturbed CS measurements, the proposed method results in a smaller NMSE than the benchmarks.

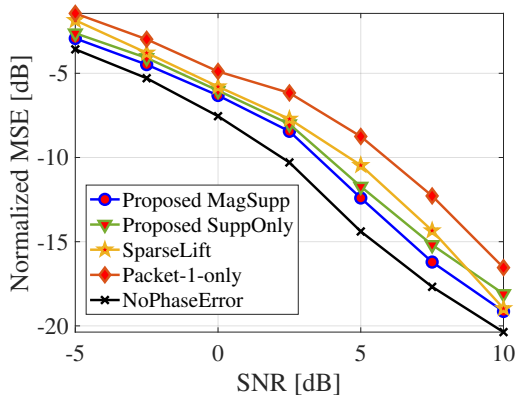


Fig. 5. NMSE with SNR for $M = 40$. For a random phase offset, our method results in a lower NMSE than the benchmarks at all SNRs.

common magnitude structure brings an improvement of up to 1 dB in the NMSE. The *Packet-1-only* method achieves the worst NMSE among the methods as it ignores measurements from the second packet. From Fig. 5, we observe that our *proposed MagSupp* technique achieves a lower NMSE than the *proposed SuppOnly* and *SparseLift*, especially at a low SNR. As SNR increases, NMSE decreases with both the proposed methods and *SparseLift*. The gap between the *proposed MagSupp* and *SparseLift* decreases at high SNR.

## V. Conclusions

In this paper, we developed a message-passing-based technique for sparse channel estimation using partially coherent measurements. Our approach absorbs the phase errors into the channel to define a pair of phase-perturbed sparse vectors. The proposed method exploits the common support and magnitude structure across this pair, in addition to channel sparsity. The central idea is to use the posterior distribution of one phase-perturbed vector to construct a prior for estimating the other vector. After several such iterations, our method obtains the MMSE estimate of the sparse channel.

## References

[1] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, 2016.

[2] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: a unifying theory and numerical methods for characterization," *IEEE Trans. on Circuits and Syst. I: Fundamental Theory and Applications*, vol. 47, no. 5, pp. 655–674, 2000.

[3] A. Alkhateeb, G. Leus, and R. W. Heath, "Compressed sensing based multi-user millimeter wave systems: How many measurements are needed?" in *Proc. IEEE Intl. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2015, pp. 2909–2913.

[4] N. J. Myers, A. Mezghani, and R. W. Heath, "FALP: Fast beam alignment in mmWave systems with low-resolution phase shifters," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8739–8753, 2019.

[5] Z. Sha and Z. Wang, "Channel estimation and equalization for terahertz receiver with RF impairments," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1621–1635, 2021.

[6] R. Zhang, B. Shim, and H. Zhao, "Downlink compressive channel estimation with phase noise in massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5534–5548, 2020.

[7] W. Yi, N. J. Myers, and G. Joseph, "Sparse millimeter wave channel estimation from partially coherent measurements," in *Proc. IEEE Intl. Conf. Commun. (ICC)*. IEEE, 2024, pp. 1897–1902.

[8] C. Hu, X. Wang, L. Dai, and J. Ma, "Partially coherent compressive phase retrieval for millimeter-wave massive MIMO channel estimation," *IEEE Trans. Signal Process.*, vol. 68, pp. 1673–1687, 2020.

[9] M. E. Rasekh and U. Madhow, "Noncoherent compressive channel estimation for mm-wave massive MIMO," in *Proc. Asilomar Conf. Signals Syst. Comput.* IEEE, 2018, pp. 889–894.

[10] X. Li, J. Fang, H. Duan, Z. Chen, and H. Li, "Fast beam alignment for millimeter wave communications: A sparse encoding and phaseless decoding approach," *IEEE Trans. Signal Process.*, vol. 67, no. 17, pp. 4402–4417, 2019.

[11] K.-H. Liu, X. Li, H. Zhao, and G. Fan, "Structured phase retrieval-aided channel estimation for millimeter-wave/sub-terahertz MIMO systems," in *Proc. IEEE 96th Vehicular Tech. Conf. (VTC2022-Fall)*. IEEE, 2022, pp. 1–5.

[12] H. Yan and D. Cabria, "Compressive sensing based initial beamforming training for massive MIMO millimeter-wave systems," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*. IEEE, 2016, pp. 620–624.

[13] S. Ling and T. Strohmer, "Self-calibration and biconvex compressive sensing," *Inverse Problems*, vol. 31, no. 11, p. 115002, 2015.

[14] C. Wei, H. Jiang, J. Dang, L. Wu, and H. Zhang, "Accurate channel estimation for mmwave massive MIMO with partially coherent phase offsets," *IEEE Commun. Lett.*, vol. 26, no. 9, pp. 2170–2174, 2022.

[15] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[16] H. Sarieddeen, M.-S. Alouini, and T. Y. Al-Naffouri, "An overview of signal processing techniques for terahertz communications," *Proc. IEEE*, vol. 109, no. 10, pp. 1628–1665, 2021.

[17] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *IEEE Intl. Symposium on Information Theory Proceedings*. IEEE, 2011, pp. 2168–2172.

[18] Y. Ghasempour, C. R. Da Silva, C. Cordeiro, and E. W. Knightly, "IEEE 802.11ay: Next-generation 60 GHz communication for 100 Gb/s Wi-Fi," *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 186–192, 2017.

[19] N. J. Myers, A. Mezghani, and R. W. Heath, "Spatial Zadoff-Chu modulation for rapid beam alignment in mmwave phased arrays," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, 2018, pp. 1–6.

[20] J. Ziniel and P. Schniter, "Efficient high-dimensional inference in the multiple measurement vector problem," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 340–354, 2012.

[21] K. V. Mardia and P. E. Jupp, *Directional statistics*. John Wiley & Sons, 2009.

[22] S. O. Rice, "Mathematical analysis of random noise," *The Bell System Technical Journal*, vol. 23, no. 3, pp. 282–332, 1944.

[23] F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, "*NIST Digital Library of Mathematical Functions*," https://dlmf.nist.gov/10.43, Release 1.2.1 of 2024-06-15. [Online]. Available: https://dlmf.nist.gov/10.43