Efficient Downlink Channel Reconstruction for FDD Multi-Antenna Systems

Yu Han, *Student Member, IEEE,* Tien-Hao Hsu, Chao-Kai Wen, *Member, IEEE,*
Kai-Kit Wong, *Fellow, IEEE,* and Shi Jin, *Member, IEEE*

***Abstract*— In this paper, we propose a novel scheme to reconstruct the downlink channel of a frequency-division-duplex (FDD) multi-antenna system utilizing uplink channel state information (CSI) combined with limited feedback. Our finding is that spatial reciprocity holds among frequency-independent parameters, including the gain, delay and angle of each propagation path in a wireless channel. Based on this, we first introduce the Newtonized orthogonal matching pursuit (NOMP) algorithm to estimate these frequency-independent parameters during uplink sounding. The gains are then refined through downlink training and sent back to the base station (BS). With only a limited amount of overhead, the refinement is able to improve the accuracy of downlink channel reconstruction substantially. Utilizing the uplink-estimated delays and angles and the downlink-refined gains, the BS can reconstruct the downlink channel from the uplink estimation. We carry out both simulations and *over-the-air* tests to assess the performance of the proposed downlink channel reconstruction scheme. Results demonstrate that the proposed scheme is promising.**

***Index Terms*— MIMO antenna system, frequency division duplex, downlink channel reconstruction, over-the-air test.**

I. Introduction

Frequency division duplex (FDD) is one of the most widely used duplexing modes for mobile communications systems, as both directions of communication can take place simultaneously without interference. FDD mode working with multiple-input multiple-output (MIMO) antenna systems have achieved great success in commercial mobile communication networks during the third and fourth generations. In recent years, large-scale MIMO or massive MIMO, capable of using the spatial dimensions for extraordinary spectral efficiency, is regarded as a key enabler for fifth generation (5G) networks.

However, the use of large-scale antenna arrays in 5G [1]–[5] and future networks brings huge challenges to the acquisition of downlink channel state information (CSI) at the base station (BS) for FDD-MIMO systems, which is critical to the system performance especially in the downlink. The fact that there is lack of reciprocity between uplink and downlink channels on different frequency bands makes the downlink CSI acquisition process difficult to achieve. For this reason, downlink CSI is usually acquired through downlink training and feedback.

Before 5G, only a few antennas are adopted at the BS and time-and-frequency resources are abundant to form orthogonal downlink pilots and the feedback amount is relatively small. However, in 5G and future networks, the use of hundreds or even thousands of antenna ports makes it impossible to design completely orthogonal pilot patterns. Pilot reuse is therefore inevitable [6], causing pilot contamination and impacting the accuracy of the downlink CSI estimation. Moreover, the huge amount of feedback for a high-dimensional complex channel matrix is impractical. Thus, downlink CSI acquisition remains a key problem in FDD massive MIMO systems.

To address this, much have been done. The studies can be classified into two approaches. In the first approach, downlink CSI is still solely obtained from downlink training and feedback, but does not require the orthogonality amongst the pilots transmitted from different antennas. For example, codebooks are used to quantize the space and only the codebook indices are sent back to the BS. In [7] and [8], a trellis-based codebook and an angle-of-departure (AoD)-adaptive subspace codebook were proposed, respectively, to quantize the channel of FDD massive MIMO systems. These methods require training and have feedback overhead. Some other methods tried to exploit the slowly-varying nature of the space. In [9], the channels were assumed to be correlated in both time and space and the authors proposed an open- and a closed-loop training with CSI memory which could be derived from previous time instances. If channel is sparse, then compressed sensing will be effective to reduce the training and feedback overhead [10], [11].

The second approach for downlink CSI acquisition in FDD massive MIMO systems is to use spatial reciprocity that exists between channels on two close frequency bands [12]. Spatial congruence has been validated through measurements in [13] and it was also demonstrated that the deviation of the dominant directions of arrival at frequencies 1935MHz and 2125MHz is small. Recently, [14] identified a frequency-invariant transform for mapping channels to paths and introduced a system to infer RF channels on one band by observing them on another band. Also, [15] reconstructed the downlink channel with the aid of the downlink channel covariance matrix which was inferred from the uplink channel covariance matrix. Furthermore, [16] and [17] proposed to exploit channel sparsity for estimating the propagation directions via uplink training, and the direction estimates were used in the downlink training process to reduce the feedback overhead. Moreover, [18] proposed a compressed downlink CSI acquisition method using partial support information obtained from the uplink that was shown to achieve a significant training overhead reduction. Despite the positive results, however, the above-mentioned methods were designed for narrowband systems only and will not be applicable for wideband implementation such as required in 5G.

Motivated by this, this paper considers a FDD multi-antenna orthogonal frequency-division multiplexing (OFDM) system[1]

[1]The proposed scheme can be straightforwardly applied to FDD massive MIMO systems.

and we propose an efficient downlink channel reconstruction scheme by utilizing the frequency-independent features of the gains, delays, and angles of the multipath channel. Our major contributions are summarized as follows.

- *Extension of the Newtonized orthogonal matching pursuit (NOMP) algorithm*: NOMP was designed to extract two parameters originally, i.e., gains and frequencies, for a noisy mixture of sinusoids [19]. In this paper, we extend and adapt it to the trivariate case, where the gain, delay and angle of each path are estimated. During each NOMP iteration, a two-dimensional dictionary is utilized and the Newton step refines delay and angle simultaneously. After updating the stopping criteria, we evaluate the accuracy of the estimations through deriving the lower bounds of the estimation errors and observe that the extracted delays and angles are strongly close to the real values.
- *Downlink channel reconstruction*: We develop an efficient downlink channel reconstruction scheme. The frequency-independent gains, delays and angles are first estimated by the extended NOMP algorithm in the uplink training process. Then, the gains are refined using least squares (LS) algorithm in the downlink with a small amount of pilots and feedback. With the uplink-estimated delays and angles and the downlink-refined gains, the downlink channel can be reconstructed at the BS side. Necessity of downlink refinement is proved through both theoretical analysis and simulations. Numerical results demonstrate that we can reconstruct highly accurate downlink channel through the proposed NOMP-LS based scheme.
- *Over-the-air (OTA) test results*: We set up an OTA testbed to assess the system performance of the proposed downlink reconstruction scheme in practical wireless communication scenarios. We observe that the reconstructed channel using the proposed NOMP-LS based scheme is near the minimum mean square error (MMSE)-estimated channel. Moreover, the power ratio of each propagation path is revealed. The results demonstrate that with more antennas, the receiver can see more propagation paths.

The remainder of this paper is organized as follows. Section II first introduces the multipath channel between BS and a user and studies the frequency-independent spatial parameters over different frequency channels. Section III presents the extended NOMP algorithm to estimate the frequency-independent spatial parameters through uplink training process and analyzes the estimation accuracy. Afterwards, Section IV then provides an analytical justification on the importance of refining the gains in the downlink and proposes a channel reconstruction scheme based on the uplink estimated frequency-independent parameters. Both simulation results and OTA test results will be discussed in Section V. Section VI concludes the paper.

Notations—In this paper, matrices and vectors are denoted by uppercase and lowercase boldface letters, respectively. The superscripts $(\cdot)^\dagger$, $(\cdot)^H$, and $(\cdot)^T$ denote the pseudo-inverse, conjugate-transpose, and transpose, respectively. In addition, $\mathcal{R}\{\cdot\}$ takes the real component of a complex number, and $\mathbb{E}\{\cdot\}$ represents the expectation with respect to all random variables inside the brackets. We also use $|\cdot|$ and $\|\cdot\|$ to denote taking the absolute value and modulus operations, respectively, and the notations $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ represent rounding a decimal number to its nearest lower and higher integers, respectively.

II. Channel Model

In this section, we describe the wireless channel between the BS and its serving user by tracing the propagation paths of the signal. A single cell of a mobile communication system is considered to operate in the FDD mode by employing OFDM. We denote the difference between the uplink and the downlink carrier frequencies as $\triangle F$, and assume that each of the uplink and downlink frequency bands has N sub-carriers with spacing $\triangle f$. The BS is equipped with a uniform linear array (ULA) with M antenna elements and the user has one antenna.

The uplink multipath channel between the user and the BS antenna element m on subcarrier n can be modeled as

$$h_m^{\mathrm{UL}}(n) = \sum_{l=0}^{L-1} g_l e^{-j2\pi n \triangle f \tau_l - j2\pi\Theta_{l,m}}, \tag{1}$$

in which $m = -\lfloor M/2\rfloor, \ldots, -1, 0, 1, \ldots, \lceil M/2\rceil - 1$, $n = -\lfloor N/2\rfloor, \ldots, -1, 0, 1, \ldots, \lceil N/2\rceil - 1$, L is the number of propagation paths, τ_l represents the delay of the lth path when it arrives at antenna element 0, satisfying $0 < \tau_l < 1/\triangle f$, and $\Theta_{l,m}$ is the phase difference between antenna element m and 0 caused by the time difference of arrival of the lth path. For simplification, we treat $\triangle f\tau$ as a whole which is simplified by $\tau \in [0, 1)$ in the subsequent part of this paper.

The wireless signal travels different distances when arriving at different BS antenna elements, as illustrated in Fig. 1. We denote the angle of the propagation path as θ, which satisfies $0 < \theta < 2\pi$. Signal that comes from direction θ travels a longer distance of $d^{(m)} = md\sin\theta$ when arriving at element m when compared with element 0. Then, the phase difference between element m and 0 for the lth path is calculated as

$$\Theta_{l,m} = \frac{fd^{(m)}}{c} = m\frac{d}{\lambda}sin\theta, \tag{2}$$

where f denotes the carrier frequency, c is the speed-of-light, d is the distance between two adjacent antenna elements which equals $\lambda/2$, with λ being the carrier wavelength. After denoting $\frac{d}{\lambda}\sin\theta$ by $\theta \in [0, 1)$, we rewrite (1) as

$$h_m^{\mathrm{UL}}(n) = \sum_{l=0}^{L-1} g_l e^{-j2\pi n\tau_l - j2\pi m\theta_l}. \tag{3}$$

By stacking the channels on all subcarriers into a vector, we get the multicarrier channel between the user and the BS antenna element m which is expressed as

$$\mathbf{h}_m^{\mathrm{UL}} = \sum_{l=0}^{L-1} g_l \mathbf{p}(\tau_l) e^{-j2\pi m\theta_l}, \tag{4}$$

where

$$\mathbf{p}(\tau) = \left[e^{-j2\pi\lfloor\frac{N_p}{2}\rfloor\tau}, \ldots, e^{j2\pi(\lceil\frac{N_p}{2}\rceil-1)\tau}\right]^H. \tag{5}$$

Similarly, the multi-antenna channel between the user and the

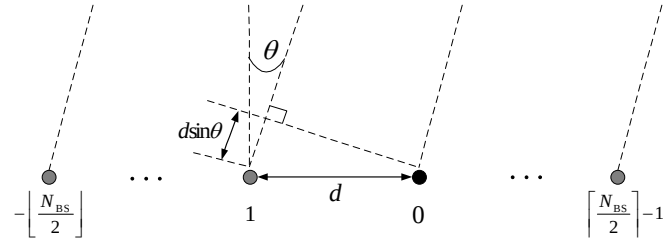

Fig. 1. Difference of the propagation distances when the wireless signal arrives at two adjacent ULA elements. The black circle represents the reference antenna element at the BS.

BS antenna array on subcarrier n is

$$\mathbf{h}^{\mathrm{UL}}(n) = \sum_{l=0}^{L-1} g_l \mathbf{a}(\theta_l) e^{-j2\pi n \tau_l}, \tag{6}$$

where

$$\mathbf{a}(\theta) = \left[e^{-j2\pi \lfloor \frac{M}{2} \rfloor \theta}, \ldots, e^{j2\pi (\lceil \frac{M}{2} \rceil - 1)\theta} \right]^H \tag{7}$$

is the steering vector of the ULA.

An obvious special case is that the BS has only one antenna and under this condition, the angles of propagation paths are not modeled in this channel, but the uplink channel vector on all the subcarriers can be written as

$$\mathbf{h}_{\mathrm{sa}}^{\mathrm{UL}} = \sum_{l=0}^{L-1} g_l \mathbf{p}(\tau_l). \tag{8}$$

In FDD systems, reciprocity does not normally apply because of the different operating frequencies in the uplink and downlink. Nonetheless, the uplink and downlink channels clearly share a common propagation space between the BS and the user and some partial reciprocity is expected if the frequency bands are within a certain coherent bandwidth.

The uplink and downlink signals propagate along common paths and are reflected by the same scatterers. The amplitude, delay, and direction of each path are frequency-independent, suggesting that the parameters $\{g_l, \tau_l, \theta_l\}_{l=0,\ldots,L-1}$ are the same for uplink and downlink [14], and these frequency-independent parameters reveal spatial reciprocity in both links. Regarding the uplink carrier frequency as the reference, i.e., 0Hz, we denote the downlink carrier frequency as $\triangle F$ and model the downlink multi-antenna channel between BS antenna array and the user on subcarrier n as

$$\mathbf{h}^{\mathrm{DL}}(n) = \sum_{l=0}^{L-1} g_l \mathbf{a}(\theta_l) e^{-j2\pi (\frac{\triangle F}{\triangle f} + n)\tau_l}. \tag{9}$$

This spatial reciprocity inspires us to reconstruct the FDD downlink channels through the frequency-independent parameters (g, τ, θ) from estimation in the uplink, instead of estimating downlink CSI through downlink training and feedback.

III. Frequency-Independent Parameters Extraction in Uplink

In this section, we investigate how to obtain the three-tuple (g, τ, θ) from the uplink. During the uplink sounding process, the BS receives pilots sent from the user, and meanwhile acquires the opportunity to estimate these frequency-independent parameters with the aid of pilot transmission.

A. NOMP Signal Model

The pilots received by the ULA on subcarrier n can be expressed as

$$\begin{aligned} \mathbf{y}^{\mathrm{UL}}(n) &= \mathbf{h}^{\mathrm{UL}}(n) s(n) + \mathbf{z}(n) \\ &= \sum_{l=0}^{L-1} g_l \mathbf{a}(\theta_l) e^{-j2\pi n \tau_l} s(n) + \mathbf{z}(n), \end{aligned} \tag{10}$$

where $s(n)$ is the pilot on subcarrier n, $\mathbf{z}(n)$ is the complex Gaussian noise vector on subcarrier n with independently and identically distributed (i.i.d.) elements, and each element of $\mathbf{z}(n)$ has zero mean and unit variance.

If we denote $a_l = g_l e^{-j2\pi n \tau_l} s(n)$, the derived expression

$$\mathbf{y}^{\mathrm{UL}}(n) = \sum_{l=0}^{L-1} a_l \mathbf{a}(\theta_l) + \mathbf{z}(n) \tag{11}$$

satisfies the NOMP signal model in [19]. NOMP behaves well in extracting the features (a, w), i.e., (a, θ) in this paper, from the signals formulated by this model. The difference is that the NOMP algorithm introduced in [19] only estimates two parameters and does not satisfy our requirement. Besides θ, we need to further extract τ and g from a. Therefore, there are totally three parameters to be estimated, and we must extend the original NOMP algorithm to fit the trivariate condition.

In this multi-antenna multi-carrier system, the BS is able to receive pilots on each occupied subcarrier and on each antenna element. It is assumed that there are N_p continuum subcarriers occupied by the pilots and the central subcarrier is DC. Stacking the received pilots on all subcarriers and on all antennas together into a vector, we have

$$\mathbf{y}^{\mathrm{UL}} = \mathbf{S} \sum_{l=0}^{L-1} g_l \mathbf{p}(\tau_l) \otimes \mathbf{a}(\theta_l) + \mathbf{z}, \tag{12}$$

where

$$\mathbf{S} = \mathrm{diag}\left\{ s\left(-\lfloor \frac{N_p}{2} \rfloor \right), \ldots, s\left(\lceil \frac{N_p}{2} \rceil - 1 \right) \right\} \otimes \mathbf{I}_M \tag{13}$$

is the pilot matrix, $\otimes$ represents the Kronecker product, and

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}(-\lfloor \frac{N_p}{2} \rfloor) \\ \vdots \\ \mathbf{z}(\lceil \frac{N_p}{2} \rceil - 1) \end{bmatrix} \tag{14}$$

is the stacked noise vector. Here, we set $s(-\lfloor \frac{N_p}{2} \rfloor) = \cdots = s(\lceil \frac{N_p}{2} \rceil - 1) = s$ and $\mathbf{S} = s\mathbf{I}_{M \times N_p}$. Then, we denote $g_l' = g_l s$. For simplification, we use symbols g_l and $\mathbf{y}$ to represent g_l' and $\mathbf{y}^{\mathrm{UL}}$ in the rest of this section. We further denote

$$\mathbf{u}(\tau, \theta) = \mathbf{p}(\tau) \otimes \mathbf{a}(\theta) \tag{15}$$

which can be viewed as a new $\mathbf{x}(\mathbf{w})$ with two parameters, i.e.,

$\mathbf{w} = (\tau, \theta)$. Then (12) can be rewritten as

$$\mathbf{y} = \sum_{l=0}^{L-1} g_l \mathbf{u}(\tau_l, \theta_l) + \mathbf{z}, \tag{16}$$

which is used as a new signal model for the extended trivariate NOMP algorithm.

B. Trivariate NOMP Algorithm

NOMP is iteration-based. In our extended version, a three-tuple of (g, τ, θ) is estimated in each iteration. The component made by this three-tuple is then removed from the observed pilot. At the end of the ith iteration, the residual is

$$\mathbf{y}_\mathrm{r} = \mathbf{y} - \sum_{l=0}^{i-1} \hat{g}_l \mathbf{u}(\hat{\tau}_l, \hat{\theta}_l), \tag{17}$$

where $(\hat{g}_l, \hat{\tau}_l, \hat{\theta}_l), l = 0, \ldots, i-1$ are the estimated three-tuples in the previous i iterations. Next, in the $(i+1)$th iteration, the goal is to estimate a new three-tuple by minimizing the new residual power $\|\mathbf{y}_\mathrm{r} - g\mathbf{u}(\tau, \theta)\|^2$, which is further translated to maximize the following function

$$S(g, \tau, \theta) = 2\mathcal{R}\left\{\mathbf{y}_\mathrm{r}^H g \mathbf{u}(\tau, \theta)\right\} - |g|^2 \|\mathbf{u}(\tau, \theta)\|^2. \tag{18}$$

The working steps in the $(i+1)$th iteration of the extended trivariate NOMP algorithm are similar to the original algorithm in [19]. We first briefly introduce the steps of the extended algorithm, which are listed below.

- *Step 1 – Newly Detection:* Select the coarse estimates $\hat{\theta}_i$ and $\hat{\tau}_i$ from a two-dimensional over-sampled angle-and-delay grid, and then calculate $\hat{g}_i$ from $\hat{\theta}_i$ and $\hat{\tau}_i$.
- *Step 2 – Single Refinement:* Solely refine the coarsely estimated three-tuple $(\hat{g}_i, \hat{\theta}_i, \hat{\tau}_i)$ through the Newton refinement steps and add the obtained $(\hat{g}_i{}', \hat{\theta}_i{}', \hat{\tau}_i{}')$ into the set of the estimated three-tuples.
- *Step 3 – Cyclic Refinement:* Cyclically refine the set of the estimated three-tuples through the Newton refinement steps and obtain $(\hat{g}_l{}'', \hat{\theta}_l{}'', \hat{\tau}_l{}''), l = 0, 1, \ldots, i$.
- *Step 4 – Gains Update:* Retain the estimated delays and angles and update all the amplitudes through LS estimation $[\hat{\hat{g}}_0, \hat{\hat{g}}_1, \ldots, \hat{\hat{g}}_i]^T = \mathbf{U}^\dagger \mathbf{y}$, where $\mathbf{U} = \left[\mathbf{u}(\hat{\tau}_0{}'', \hat{\theta}_0{}''), \mathbf{u}(\hat{\tau}_1{}'', \hat{\theta}_1{}''), \ldots, \mathbf{u}(\hat{\tau}_i{}'', \hat{\theta}_i{}'')\right]$.

Then, we provide details about the extensions of this work. Since the delay and the angle jointly determine the phase of the channel and can be represented using a common vector $\mathbf{u}(\tau, \theta)$, these two parameters are estimated and refined together in our design. This combination results in the two-dimensional grid and the extended Newton step.

1) Two-Dimensional Grid: The coarse estimates $(\hat{\tau}_i, \hat{\theta}_i)$ in Step 1 are chosen from a two-dimensional angle-and-delay grid Ω, which consists of $\gamma_1 N_p \times \gamma_2 M$ over-sampled grid points

$$\Omega = \left\{\left(\frac{k_1}{\gamma_1 N_p}, \frac{k_2}{\gamma_2 M}\right) : k_1 = 0, 1, \ldots, \gamma_1 N_p - 1; \; k_2 = 0, 1, \ldots, \gamma_2 M - 1\right\}, \tag{19}$$

where γ_1 and γ_2 are the over-sampling rates for the delay grid and the angle grid, respectively. Each point in the grid forms a vector $\mathbf{u}(k_1/(\gamma_1 N_p), k_2/(\gamma_2 M))$. The coarsely estimated delay and angle are obtained by exhaustive searching the grid points as follows

$$(\hat{\tau}_i, \hat{\theta}_i) = \arg \max_{(\tau,\theta)\in\Omega} \frac{|\mathbf{u}^H(\tau, \theta)\mathbf{y}_\mathrm{r}|^2}{\|\mathbf{u}(\tau, \theta)\|^2}. \tag{20}$$

Then the gain is calculated as

$$\hat{g}_i = \frac{\mathbf{u}^H(\hat{\tau}_i, \hat{\theta}_i)\mathbf{y}_\mathrm{r}}{\|\mathbf{u}(\hat{\tau}_i, \hat{\theta}_i)\|^2}. \tag{21}$$

2) Extended Newton Step: The extended Newton step in Step 2 and Step 3 is able to refine the delay and the angle simultaneously, with one more parameter than the original Newton step. In this bivariate problem, the coarsely estimated $(\hat{\tau}_i, \hat{\theta}_i)$ are refined through

$$\begin{bmatrix}\hat{\theta}_i' \\ \hat{\tau}_i'\end{bmatrix} = \begin{bmatrix}\hat{\theta}_i \\ \hat{\tau}_i\end{bmatrix} - \ddot{\mathbf{S}}\left(\hat{g}_i, \hat{\theta}_i, \hat{\tau}_i\right)^{-1} \dot{\mathbf{S}}\left(\hat{g}_i, \hat{\theta}_i, \hat{\tau}_i\right), \tag{22}$$

where

$$\dot{\mathbf{S}}(g, \theta, \tau) = \begin{bmatrix}\frac{\partial S}{\partial \theta} \\ \frac{\partial S}{\partial \tau}\end{bmatrix} \tag{23}$$

is the first-order partial derivative vector, and

$$\ddot{\mathbf{S}}(g, \theta, \tau) = \begin{bmatrix}\frac{\partial^2 S}{\partial \theta^2} & \frac{\partial^2 S}{\partial \tau \partial \theta} \\ \frac{\partial^2 S}{\partial \theta \partial \tau} & \frac{\partial^2 S}{\partial \tau^2}\end{bmatrix} \tag{24}$$

is the second-order partial derivative matrix. According to (18), we can write the first-order partial derivatives of $S(g, \tau, \theta)$ as

$$\frac{\partial S}{\partial x} = 2\mathcal{R}\left\{g\left(\mathbf{y}_\mathrm{r} - g\mathbf{u}\right)^H \frac{\partial \mathbf{u}}{\partial x}\right\}, \tag{25}$$

where x can be τ or θ. The second-order partial derivative of $S(g, \tau, \theta)$ is calculated as

$$\frac{\partial^2 S}{\partial x_1 \partial x_2} = 2\mathcal{R}\left\{g\left(\mathbf{y}_\mathrm{r} - g\mathbf{u}\right)^H \frac{\partial^2 \mathbf{u}}{\partial x_1 \partial x_2} - |g|^2 \frac{\partial \mathbf{u}^H}{\partial x_2} \frac{\partial \mathbf{u}}{\partial x_1}\right\}, \tag{26}$$

where x_1 and x_2 can be τ or θ, respectively. Since we are pursuing the maximum value of $S(g, \theta, \tau)$, one requirement is that $S(g, \theta, \tau)$ is locally convex in the neighborhood of $(\hat{\tau}, \hat{\theta})$. Therefore, if and only if $\det\left(\ddot{\mathbf{S}}\left(\hat{g}, \hat{\theta}, \hat{\tau}\right)\right) > 0$ and the first element of $\ddot{\mathbf{S}}(g, \theta, \tau)$ is lower than 0, the Newton refinement (22) will be carried out. At the end of each Newton step, the gain is also updated using (21).

It should be emphasized that the two-dimensional grid and the extended Newton step are the required major extensions to the original NOMP algorithm. Other minor modifications to fit the trivariate condition are trivial and omitted here.

C. Stopping Criterion

One major challenge is that BS does not know the number of propagation paths in the real channel, which directly determines when to terminate the iteration process. If the estimated three-tuples are precise enough, all the paths will be accurately identified and the residual will be reduced to the noise at the

end of the NOMP algorithm, i.e., $\mathbf{y}_{\mathrm{r}} \approx \mathbf{z}$. In this paper, we utilize this assumption to design the stopping criterion.

1) Power Based Criterion: One choice is to terminate the NOMP iterations when the residual power is less than the total noise power. As is mentioned before, the noise power is normalized to 1. That is to say, if

$$\|\mathbf{y}_{\mathrm{r}}\|^2 < \kappa, \tag{27}$$

where

$$\kappa = \mathbb{E}\{\|\mathbf{z}\|^2\} = MN_p, \tag{28}$$

the trivariate NOMP algorithm is stopped.

2) False Alarm Rate Based Criterion: Alternatively, we can design the stopping criterion based on the false-alarm rate. If we "detect" a fake path that does not exist, we say that a fault appears. It happens when all the paths have been detected but the algorithm is still not working. The following theorem introduces the false alarm rate based stopping criterion.

Theorem 1: If the trivariate NOMP algorithm terminates when

$$\|\mathbf{u}(\tau,\theta)^H \mathbf{y}_{\mathrm{r}}\|^2 < \kappa' \tag{29}$$

holds for all grid points

$$(\tau,\theta) \in \left\{ \left(\frac{{k_1}'}{N_p}, \frac{{k_2}'}{M} \right) : {k_1}' = 0,1,\ldots,N_p-1; \right. \\ \left. {k_2}' = 0,1,\ldots,M-1 \right\}, \tag{30}$$

where

$$\kappa' = \ln(MN_p) - \ln(-\ln(1-P_{\mathrm{fa}})), \tag{31}$$

then the false alarm rate can be approximated by P_{fa}.

Proof: Since the grid points listed in (30) are non-over-sampled points, the corresponding values of $\mathbf{u}(\tau,\theta)^H \mathbf{y}_{\mathrm{r}}$ can be viewed as the Fourier transformed values of $\mathbf{y}_{\mathrm{r}}$ and remain the same statistic property of $\mathbf{y}_{\mathrm{r}}$. As is mentioned before, $\mathbf{y}_{\mathrm{r}} \approx \mathbf{z}$ when all the paths are precisely detected. Then the condition (29) can be translated to

$$\|\mathbf{z}\|_\infty^2 < \kappa'. \tag{32}$$

From [20], we know that

$$\mathbb{E}\left\{\|\mathbf{z}\|_\infty^2\right\} \approx \ln(MN_p) \tag{33}$$

when MN_p grows large. Denoting $x = \|\mathbf{z}\|_\infty^2 - \ln(MN_p)$, we can derive that

$$\mathcal{P}\{x \le X\} = \mathcal{P}\left\{\|\mathbf{z}\|_\infty^2 < X + \ln(MN_p)\right\}. \tag{34}$$

Since each element of $\mathbf{z}$ is i.i.d., it holds that

$$\begin{aligned} &\mathcal{P}\left\{\|\mathbf{z}\|_\infty^2 < X + \ln(MN_p)\right\} \\ =&\mathcal{P}\left\{|z_1|^2 \le X + \ln(MN_p)\right\}^{MN_p} \\ =&\left(1 - \frac{1}{MN_p} e^{-X}\right)^{MN_p}, \end{aligned} \tag{35}$$

where z_1 is the first element of $\mathbf{z}$ and is a Gaussian variable with zero mean and unit variance. Because

$$\exp\{a\} = \lim_{n\to\infty} \left(1 + \frac{a}{n}\right)^n, \tag{36}$$

we have

$$\left(1 - \frac{1}{MN_p} e^{-X}\right)^{MN_p} \approx \exp\{-\exp\{-X\}\} \tag{37}$$

when MN_p grows without limit. Applying (35)–(37) into (34) and denoting $\kappa' = X + \ln(MN_p)$, we can obtain

$$\mathcal{P}\left\{\|\mathbf{z}\|_\infty^2 \le \kappa'\right\} \approx \exp\{-\exp\{\ln(MN_p) - \kappa'\}\}. \tag{38}$$

If we further apply (31), then it holds that

$$\mathcal{P}\left\{\|\mathbf{z}\|_\infty^2 > \kappa'\right\} = 1 - \mathcal{P}\left\{\|\mathbf{z}\|_\infty^2 \le \kappa'\right\} \approx P_{\mathrm{fa}}, \tag{39}$$

which means that the false alarm rate approximates P_{fa}. ■

The false alarm rate based stopping criterion is adopted in the trivariate NOMP algorithm. We denote the final estimation results of the trivariate NOMP algorithm as $(\hat{g}_l, \hat{\theta}_l, \hat{\tau}_l), l = 0,\ldots,\hat{L}-1$, where $\hat{g}_l$ has been normalized by s.

D. Estimation Accuracy

To evaluate the estimation accuracy of the delay and angle, we calculate the normalized mean square errors (MSEs) by

$$\varepsilon_\tau = \mathbb{E}\left\{\frac{|\hat{\tau}-\tau|^2}{1/N_p^2}\right\} \tag{40}$$

and

$$\varepsilon_\theta = \mathbb{E}\left\{\frac{|\hat{\theta}-\theta|^2}{1/M^2}\right\}. \tag{41}$$

We have the following theorem to study the estimation accuracy of the extended trivariate NOMP algorithm.

Theorem 2: The normalized MSEs of the delay and angle are lower bounded, respectively, by

$$\varepsilon_\tau \ge \frac{3N_p}{\mathrm{SNR}\cdot 2\pi^2 M(N_p^2-1)} \tag{42}$$

and

$$\varepsilon_\theta \ge \frac{3M}{\mathrm{SNR}\cdot 2\pi^2 N_p(M^2-1)}. \tag{43}$$

Proof: Cramer-Rao bound (CRB) can be interpreted as a lower bound of the variance of the estimator. The CRBs of the single path case, $\mathbf{y} = g\mathbf{u}(\tau,\theta) + \mathbf{z}$, are introduced, where each element of $\mathbf{z}$ is i.i.d Gaussian with zero mean and unit variance. According to [21], the Fisher information matrix is calculated by

$$\mathbf{F}(\tau,\theta) = 2|g|^2 \mathcal{R}\left\{ \begin{bmatrix} \frac{\partial \mathbf{u}^H}{\partial \tau}\frac{\partial \mathbf{u}}{\partial \tau} & \frac{\partial \mathbf{u}^H}{\partial \tau}\frac{\partial \mathbf{u}}{\partial \theta} \\ \frac{\partial \mathbf{u}^H}{\partial \theta}\frac{\partial \mathbf{u}}{\partial \tau} & \frac{\partial \mathbf{u}^H}{\partial \theta}\frac{\partial \mathbf{u}}{\partial \theta} \end{bmatrix} \right\}. \tag{44}$$

Applying (15) into (44), we can get an analytical expression of the Fisher information matrix as

$$\mathbf{F}(\tau,\theta) = 2|g|^2 \begin{bmatrix} \frac{\pi^2 MN_p(M^2-1)}{3} & 0 \\ 0 & \frac{\pi^2 MN_p(N_p^2-1)}{3} \end{bmatrix}. \tag{45}$$

Then, the CRB of the delay is expressed as

$$\mathrm{CRB}_\tau = \mathbf{F}_{1,1}^{-1}(\tau,\theta) = \frac{3}{2|g|^2\pi^2 MN_p(N_p^2-1)}. \tag{46}$$

Similarly, the CRB of the angle is

$$\mathrm{CRB}_\theta = \mathbf{F}_{2,2}^{-1}(\tau,\theta) = \frac{3}{2|g|^2\pi^2 MN_p(M^2-1)}. \tag{47}$$

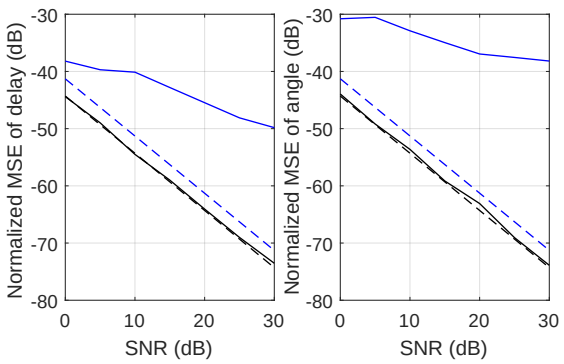

Fig. 2. Comparison of the practical normalized MSEs and the lower bounds for (a) delay and (b) angle.

Since g contains the pilot and the noise power equals 1, the signal-to-noise ratio (SNR) here is measured through $|g|^2$, that is, $|g|^2 = \text{SNR}$. Moreover, $\text{CRB}_\tau \leq \mathbb{E}\left\{|\hat{\tau}-\tau|^2\right\}$, $\text{CRB}_\theta \leq \mathbb{E}\left\{|\hat{\theta}-\theta|^2\right\}$, and we then obtain (42) and (43). ■

When we set $M = 1$, the problem is reduced to the bivariate case that only the gain and the delay are to be estimated. If $\mathbf{p}(\tau) = \left[e^{-j\lfloor N_p/2\rfloor\tau}, \ldots, e^{j(\lceil N_p/2\rceil-1)\tau}\right]^H/\sqrt{N_p}$, then the CRB of delay is written as

$$\text{CRB}_\tau^{(M=1)} = \frac{6}{\text{SNR}(N_p^2-1)}, \tag{48}$$

which is exactly in accordance with the CRB bound given in [19]. It proves the correctness of *Theorem 2*.

Corollary 1: When N_p or M grows large, the lower bounds of the normalized MSEs of the delay and angle coincide, i.e.,

$$\varepsilon_\tau, \varepsilon_\theta \geq \frac{3}{\text{SNR}\cdot 2\pi^2 M N_p}. \tag{49}$$

Proof: It holds that $N_p^2/(N_p^2-1) \approx 1$ when N_p grows large. Then (42) approaches (49). Similarly, (43) approaches (49) when M grows large. ■

Remark: From *Theorem 2*, we can find that the bounds can be further lowered if we increase the number of subcarriers occupied by the pilots or the number of BS antenna elements. This is because with more observed samples, we can see more details about the spatial channel. What should be emphasized are the preconditions of high estimation accuracy, i.e., the angles and delays of different paths are well-separated and that the number of paths is far less than M or N_p. In addition, only if the channel satisfies these preconditions, the algorithm can achieve the lower-bound performances.

Results in Fig. 2 give an intuitive comparison of the normalized MSEs with the derived lower bounds. Solid lines represent the practical MSEs of the trivariate NOMP algorithm, and the dotted lines are the lower bounds. We set $\gamma_1 = \gamma_2 = 2$. There exist 15 equal-power paths in the channel. The minimum separations among the delays and the angles are no less than $1/N_p$ and $1/M$, respectively. Lines colored in black correspond to the case when $M = 32, N_p = 128$. The values of M and N_p satisfy the condition in *Corollary 1* and we find that the lower bounds of the delay and the angle are nearly the same with each other. Besides, with the increase of SNR, the estimation accuracy is enhanced proportionally. The practical MSEs closely coincide with their theoretical lower bounds for both the delay and the angle, which demonstrates the high accuracy of the trivariate NOMP algorithm. Moreover, even though $M \ll N_p$, the practical estimation accuracy of the angle is not inferior to that of the delay. This is due to the well-separated spatial angles of each path, and that the number of paths is far less than M or N_p. Results in Fig. 2 also compare the performances of the algorithm in the cases when $M = 32, N_p = 64$ in the left subfigure and when $M = 16, N_p = 128$ in the right subfigure, and the results are colored in blue. Results demonstrate that the practical estimation accuracy degrades and the MSE lines deviate with the theoretical lower bounds. The lower bounds are accessed when the observations are far more than the propagation paths. Despite this, the MSEs of the delay and angle are below -30 dB and the practical estimation accuracy is still high.

IV. Downlink Channel Reconstruction Using Uplink CSI

Having obtained the gain, delay and angle of each path through the uplink training and the trivariate NOMP estimation, we move on to the next step, that is, reconstructing the downlink channel for the FDD transmission system.

A. The Issue

As mentioned before, the delays, gains and angles are all frequency-independent parameters, and their uplink estimates are still applicable to the downlink frequency band channel reconstruction. Based on these, [14] proposed an R2-F2 system which extracts the information of the propagation paths from channels on band 1 to reconstruct the corresponding channels on band 2. R2-F2 enables the LTE BSs to infer the downlink channels using the uplink derived CSI, and suggests to eliminate CSI feedback. Eliminating the feedback will result in a significant improvement of the time-frequency resource utilization in the uplink. However, the question remains whether the reconstructed downlink channel is accurate enough if one directly applies the uplink estimates into the downlink channel model without using any of downlink CSI.

Actually, estimation error is inevitable for any estimation method, including NOMP as well as the optimization method used in [14]. The error has a harmful impact on the reconstructions on another frequency band. We first use the simplest single-antenna single-path case as an example to clarify this impact. We denote the real gain and delay as g and τ, and the real channel on frequency f_1 is $h_1 = ge^{-j2\pi f_1\tau}$. Assume the estimated gain and delay on frequency f_1 are $\hat{g}$ and $\hat{\tau}$, and the estimation error of delay is $\triangle\tau = \hat{\tau}-\tau$. For frequency f_1, $\hat{g}$ will compensate for the phase error caused by $\triangle\tau$, because the gain is updated using LS estimation at the end of the NOMP algorithm. The reconstructed channel $\hat{h}_1 = \hat{g}e^{-j2\pi f_1\hat{\tau}}$ will be very much the same as the original channel h_1. In this case, we obtain "global accuracy" instead of "local accuracy".

However, when $\hat{g}$ and $\hat{\tau}$ are used directly to reconstruct the channel on frequency f_2, the phase error $2\pi f_2\triangle\tau$ is considerable if either f_2 or $\triangle\tau$ is large, as shown in Fig. 3(c).

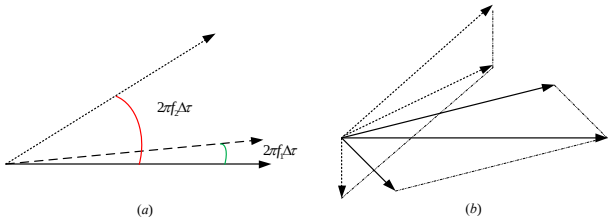

Fig. 3. Impact of phase error: (a) The phase error of a path on target frequency f_2 increases significantly when f_2 is large. The solid line with arrow represents the true phase of the path. The dotted lines with arrows are the phase errors on frequencies f_1 and f_2. (b) The phase errors from different paths result in a false superposition. The solid lines with arrows are the real multi-path components and the real superposition. The dotted lines with arrows are the estimated components and their superposition.

Meanwhile, $\hat{g}$ derived on frequency f_1 is no longer able to compensate this phase error on frequency f_2. The reconstructed SISO channel on frequency f_2 is expressed as

$$\hat{h}_2 = \hat{g}e^{-j2\pi f_2\hat{\tau}} = h_2\frac{\hat{g}}{g}e^{-j2\pi f_2\triangle\tau}, \tag{50}$$

where $\frac{\hat{g}}{g}e^{-j2\pi f_2\triangle\tau}$ is the multiplicative estimation error from the global perspective. We can find that the derived channel on frequency f_2 is far from the real channel. Even though $\hat{g}$ has the same absolute value with g, that is, $|\hat{g}| = |g|$, the phase difference between $\hat{h}_2$ and h_2 would be unacceptably large since the phase information of the wireless channel is of great importance to the transceiver design.

The phase error will exert severe impact on the multi-path channel reconstructions. As shown in Fig. 3(d), the original channel is composed of two paths with different amplitudes and delays which are illustrated by solid lines with arrows. Due to the phase error, these multi-path components rotate, and form an incorrect superposition. In more complicated multi-antenna multi-path scenarios, the angular error exists as well, further harming the reconstruction on another band. Therefore, it is ineffective to use only uplink CSI for reconstructing the downlink channel for FDD transmission systems.

B. Re-estimation and Reconstruction Scheme

Due to the inevitable estimation error of the delay and angle, there exists unacceptable amplitude and phase errors in the reconstruction of the uplink channel. To address this, the gains are LS-estimated at the last step of NOMP to compensate for these error. As a consequence, we follow a similar approach and re-estimate using LS the gains to compensate for the errors in reconstructing the downlink channel. This approach needs extra downlink overhead. Fortunately, only the gains need to be refined. Both the delays and angles estimated in the uplink are still applicable to the downlink channel reconstruction. Hence, a small amount of overhead is required to refine the gains.

The gains are refined with the aid of pilots transmitted in the downlink frequency band. To retrieve the feature of the whole downlink frequency band, these pilots are designed to be sparsely distributed in the whole band in comb-type. Suppose that K subcarriers are occupied by the pilots and the indices of the subcarriers are $n_0, n_2, \ldots, n_{K-1}$.

Different from the uplink, there are multiple antennas at the transmitter and a single antenna at the receiver. Pilots transmitted by the antenna array will be additively received at a single antenna. To improve the received power, the pilots are beamformed before transmission. Given the angles of the propagation paths $\hat{\theta}_l, l = 0, \ldots, \hat{L}-1$ estimated in the uplink, we target the pilots to these directions and concentrate the transmit power onto the propagation path of the channel. Two different beamforming types are considered here.

1) Pilots in one OFDM symbol are beamformed to target one specific direction. We need $\hat{L}$ OFDM symbols to send the pilots. Different OFDM symbols correspond to different directions. For subcarrier n_i on the jth OFDM symbol, the received pilot can be expressed as

$$y_j^{\mathrm{DL}}(n_i) = \sum_{l=0}^{L-1} g_l e^{-j2\pi n_i\triangle f\tau_l}\mathbf{a}^H(\theta_l)\mathbf{a}(\hat{\theta}_j)p_j(n_i) + z_j(n_i), \tag{51}$$

where $p_j(n_i)$ is the downlink pilot signal on subcarrier n_i and on the jth OFDM symbol, $z_j(n_i)$ is the noise, $i = 0, \ldots, K-1$ and $j = 0, \ldots, \hat{L}-1$.

2) Pilots are frequency-division multiplexed onto different directions. Only one OFDM symbol is needed. Subcarriers $n_1, \ldots, n_{\hat{L}}$ correspond to directions $\hat{\theta}_1, \ldots, \hat{\theta}_{\hat{L}}$, respectively. The received pilot on subcarrier n_i can be expressed as

$$y^{\mathrm{DL}}(n_i) = \sum_{l=0}^{L-1} g_l e^{-j2\pi n_i\triangle f\tau_l}\mathbf{a}^H(\theta_l)\mathbf{a}(\hat{\theta}_{\underline{i}})p(n_i) + z(n_i), \tag{52}$$

where $\underline{i} = i \bmod \hat{L}$, $p(n_i)$ is the downlink pilot signal on subcarrier n_i, $z(n_i)$ is the noise, and $i = 0, \ldots, K-1$.

In order to re-estimate the vector $\mathbf{g}$ at the user side, the BS needs to inform the user with the uplink estimated parameters $(\hat{\tau}_l, \hat{\theta}_l), l = 0, \ldots, \hat{L}-1$, as well as the beamforming type. The user applies these estimates into (51) or (52) according to the beamforming type and rewrites the signal models as

$$y_j^{\mathrm{DL}}(n_i) = \sum_{l=0}^{\hat{L}-1} g_l e^{-j2\pi n_i\triangle f\hat{\tau}_l}\mathbf{a}^H(\hat{\theta}_l)\mathbf{a}(\hat{\theta}_j)p_j(n_i) + z_j(n_i), \tag{53}$$

and

$$y^{\mathrm{DL}}(n_i) = \sum_{l=0}^{\hat{L}-1} g_l e^{-j2\pi n_i\triangle f\hat{\tau}_l}\mathbf{a}^H(\hat{\theta}_l)\mathbf{a}(\hat{\theta}_{\underline{i}})p(n_i) + z(n_i). \tag{54}$$

After stacking the received pilots on all the subcarriers and all the OFDM symbols, we get the following unified signal model for both the above two types as

$$\mathbf{y} = \mathbf{A}\mathbf{g} + \mathbf{n}, \tag{55}$$

where $\mathbf{y}$, $\mathbf{g}$ and $\mathbf{n}$ are the stacked $X \times 1$ dimensional received pilots, gains and noise vectors, respectively, and $\mathbf{A}$ denotes the $X \times \hat{L}$ dimensional coefficient matrix.

For Type 1, $X = K\hat{L}$ and A is comprised of $\hat{L}$ submatrices

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}^{(0)} \\ \vdots \\ \mathbf{A}^{(\hat{L}-1)} \end{bmatrix}^T, \tag{56}$$

where the (i,l)th entry of the submatrix $\mathbf{A}^{(j)}$ equals

$$\mathbf{A}^{(j)}_{i,l} = e^{-j2\pi n_i \triangle f \tau_l} \mathbf{a}^H(\hat{\theta}_l)\mathbf{a}(\hat{\theta}_j)p_j(n_i), \tag{57}$$

where $j = 0, \ldots, \hat{L}-1$, $i = 0, \ldots, K-1$, and $l = 0, \ldots, \hat{L}-1$. For Type 2, $X = K$, and the (i,l)th entry of A equals

$$\mathbf{A}_{i,l} = e^{-j2\pi n_i \triangle f \tau_l} \mathbf{a}^H(\hat{\theta}_l)\mathbf{a}(\hat{\theta}_{\underline{i}})p(n_i), \tag{58}$$

where $i = 0, \ldots, K-1$, and $l = 0, \ldots, \hat{L}-1$.

Since the coefficient matrix $\mathbf{A}$ is also known at the user side, the user is able to refine the gains through LS estimation

$$\hat{\tilde{\mathbf{g}}} = \mathbf{A}^{\dagger}\mathbf{y}^{\mathrm{DL}}, \tag{59}$$

where $\mathbf{A}^{\dagger}$ represents the pseudo-inverse of $\mathbf{A}$ and the dimension of the refined gain vector $\hat{\tilde{\mathbf{g}}}$ is still $\hat{L}$. Then the refined gains are sent back to the BS.

Finally, the BS obtains all the information required in order to reconstruct the downlink channel, which are the parameters $(\hat{\tilde{g}}_l, \hat{\tau}_l, \hat{\theta}_l), l = 0, \ldots, \hat{L}-1$. In particular, the downlink multipath channel on subcarrier n is reconstructed as

$$\mathbf{h}^{\mathrm{DL}}(n) = \sum_{l=0}^{\hat{L}-1} \hat{\tilde{g}}_l \mathbf{a}^T(\hat{\theta}_l) e^{-j2\pi(\frac{\triangle F}{\triangle f}+n)\hat{\tau}_l}. \tag{60}$$

We refer to our proposed reconstruction as the NOMP-LS based downlink channel reconstruction scheme. For clarity, we here provide a brief summary of the procedures discussed in *Sections III-B and IV* and used in the NOMP-LS based downlink channel reconstruction scheme below:

- Uplink NOMP estimation. The user sends uplink pilots to the BS. The BS uses the extended triviriate NOMP algorithm to estimate the gain, delay and angle of each propagation path of the channel.
- Downlink gain refinement and feedback. The BS transmits downlink pilots to the user and informs the user with the beamforming type and the uplink estimated delays and angle-of-arrivals (AoAs). The user re-estimates the gains and feeds the gains back to the BS.
- Downlink channel reconstruction. The BS reconstructs the downlink channel using the uplink estimated delays and angles and the downlink refined gains.

V. Performance Evaluation

In this section, we evaluate the performance of the proposed NOMP-LS based downlink channel reconstruction scheme. We first discuss our computer simulation results and then move on to our hardware OTA tests for validation.

A. Simulation Results

Computer simulations are realized through MATLAB. For the NOMP algorithm, the over-sampling rates of the delay and angle are set to 2 and 4, respectively. One round of a single refinement and three rounds of cyclic refinement are implemented during each NOMP iteration. We set the number of FFT points as 2048 and the subcarrier spacing as 75 kHz.

We first validate the necessity of the gain refinement through comparing the amplitude of the reconstructed channel with that of the real channel. The total bandwidth is 153.6 MHz. The center frequency 45 MHz is regarded as the in-band to estimate the frequency-independent parameters. Then the out-of-band channels on the other bands are inferred or reconstructed using these parameters. We consider a simple example that the BS is equipped with one antenna. Fig. 4(a) reveals the difference between the actual full-band channel and the NOMP-based inferred channel. The latter is obtained by applying the in-band estimated parameters into the channel model as suggested in [14]. It can be observed that within the 45 MHz in-band, the inferred channel matches well with the actual channel, corroborating the precision of the NOMP algorithm. On the other hand, for the out-of-band, an obvious deviation can be seen between the inferred and the actual channels. The large performance degradation indicates that the gains derived from the in-band estimation are insufficiently accurate in inferring the out-of-band channel. Therefore, we refine the gains using LS estimation with the aid of the out-of-band pilots. In the simulations, the out-of-band pilots are inserted in every four subcarriers. The results of the refinement are given in Fig. 4(b). Results demonstrate that the reconstructed out-of-band channel matches closely to the actual channel, thereby validating the necessity and effectiveness of the gain refinement.

Now, we examine the performance of the proposed NOMP-LS based downlink channel reconstruction scheme in FDD-OFDM systems. In both uplink and downlink OFDM modules, the central 1200 subcarriers around DC compose the transmission band whose bandwidth equals 90 MHz. The separation between the uplink and downlink central subcarriers is 300 MHz. In the uplink, all the 1200 subcarriers are filled with pilots for the NOMP algorithm. Also, in the downlink, pilots are sparsely and uniformly inserted in every K subcarriers. We compare results for $K = 1, 2, 4$ and focus on two propagation scenarios. Scenario (a) is a sparsely scattering scenario, where two distinct paths exist in the channel. The power of the two paths is normalized by $[0.8, 0.2]$ and the angles are i.i.d. and randomly generated in $[0^\circ, 360^\circ)$. Scenario (b) is the clustering channel, where there is one cluster with six close paths. The power of the paths is $[0.6, 0.12, 0.12, 0.08, 0.05, 0.03]$. The angular spread of the cluster is 10°. In addition, it is noted that the noise power is set to 1. Therefore, the SNR measures the pilot power on one antenna and for each subcarrier.

LS and MMSE channel estimation results are introduced as lower and upper benchmarks, respectively. LS is a commonly used estimation method with low complexity, but has a drawback of increasing the noise. Also, MMSE is an improved estimation algorithm which fixes this drawback and achieves much higher accuracy. When conducting LS and MMSE estimation algorithms, we use the pilots on every four subcarriers. Moreover, we compare the NOMP-LS based reconstructed downlink channel and the NOMP-based reconstructed uplink channel with the actual channel by evaluating their MSE

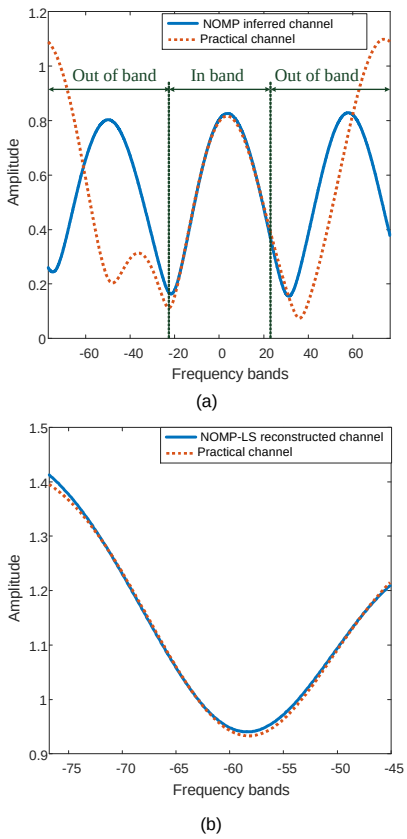

Fig. 4. Full-band reconstruction results, where (a) presents the NOMP-based inferred channel in the full band, and (b) shows the NOMP-LS based reconstructed channel in the out-of-band.

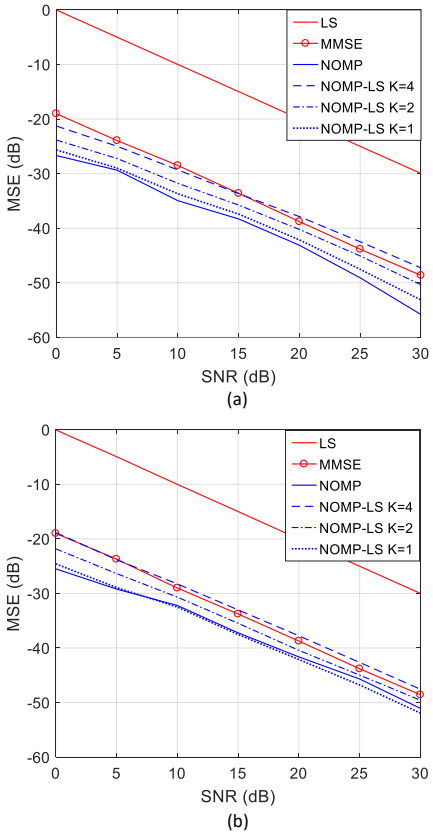

Fig. 5. MSE performance of the reconstructions in Scenario (a) and (b). $M = 4$ and beamforming type 1 is used.

performance, which is calculated as

$$\text{MSE} = \frac{\mathbb{E}\left\{\|\hat{\mathbf{h}} - \mathbf{h}\|^2\right\}}{\sigma_{\mathbf{z}}^2}, \tag{61}$$

where $\hat{\mathbf{h}}$ denotes the reconstructed or inferred channel on one subcarrier, $\mathbf{h}$ is the real channel, and $\sigma_{\mathbf{z}}^2$ is the addition of the noise power on multiple antennas which equals M here.

We first test the case when the BS is equipped with four antennas. Fig. 5(a) demonstrates the MSE performances of the NOMP-based uplink reconstructed channel and the NOMP-LS based downlink reconstructed channel when beamforming type 1 is used in Scenario (a). Results show that the LS estimated channel has the worst MSE, while the NOMP-based uplink reconstructed channel has the best MSE, and its precision is even higher than the MMSE estimated downlink channel. Both of the methods employ the in-band pilots to estimate the in-band CSI and acquire excellent MSE performances. In Scenario (a), the two distinct propagation paths can be easily and exactly detected by the NOMP algorithm. Hence, the uplink reconstructed channel is almost the same as the actual channel. The MSE results of the NOMP-LS based downlink reconstruction scheme when $K = 1, 2, 4$ are provided, which are inferior to the NOMP-based uplink reconstruction scheme. Especially when $K = 4$, there exists a relatively large performance gap between NOMP-LS and NOMP. However, the MSE performance of the downlink reconstruction scheme is still comparable with the MMSE estimation. If we enhance the density of the downlink pilots by decreasing the value of K, the MSE of reconstructed channel drops further. As overly high performance is not necessary and the cost is large, we need to strike a balance between performance and cost.

The numerical results of the four-antenna case when using beamforming type 1 in Scenario (b) are provided in Fig. 5(b). Since the paths are clustered within a small angular-spread area, it becomes hard to extract each path from their spatial superposition. This is particularly true if the angles cannot be accurately estimated, and the number of estimated paths may be more or less than the actual channel has. Hence, the performance of NOMP-based uplink channel reconstruction degrades slightly when compared with Fig. 5(a). Regarding the downlink reconstructions, using beamforming type 1 still achieves excellent MSE performance due to the large amount of beamformed pilots. We see that at high SNR, it performs even better than the uplink channel reconstruction.

We compare the two different beamforming types and show the performance of beamforming type 2 in Fig. 6. The NOMP-LS based downlink reconstructions are greatly impacted by sharply reducing the number of downlink pilots. However, for Scenario (a), the performance degradation is acceptable even when setting $K = 4$, since there are only a small number of estimated paths and each estimated direction can be allocated with enough pilots. While for Scenario (b), it needs 8 dB more SNR for the NOMP-LS-based downlink reconstruction scheme

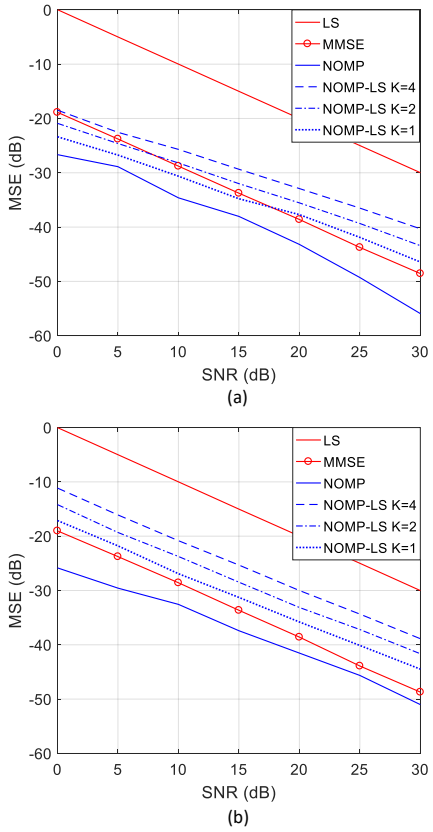

Fig. 6. The MSE performance of the reconstructions in Scenario (a) and (b) where $M = 4$ and beamforming type 2 is used.

to achieve the same MSE performance in Scenario (a).

Now, we scale up the computer simulations by considering the sixteen-antenna configuration when using beamforming type 1 in the more commonly seen Scenario (b). This simulation aims to assess the performance of the proposed reconstruction scheme in massive MIMO environments. The results are shown in Fig. 7. We find that when the number of antenna elements grows large, the reconstructed channel has excellent performance as well. In particular, at low SNR, the MSE performance of the sixteen-antenna case is obviously better than the single-antenna and four-antenna cases due to the help of the multi-antenna gain. Therefore, we can conclude through the numerical results that both the NOMP-based uplink reconstruction and the NOMP-LS-based downlink reconstruction perform well in estimating the actual channel.

B. OTA Test Results

We also set up an OTA testbed which is shown in Fig. 8(a) for validating the results in practical environments. The radio devices are placed along the table for equipment. The yellow circle represents the position of the user and the red squares are the potential positions of the BS. The BS and the user are equipped as shown in Fig. 8(b). The user works as the transmitter and has a single antenna. It is controlled by the SGT100A radio frequency vector signal generator [23]. The BS is the receiver. The received signal at the BS antenna array is first transported to the RTO2000 digital oscilloscopes

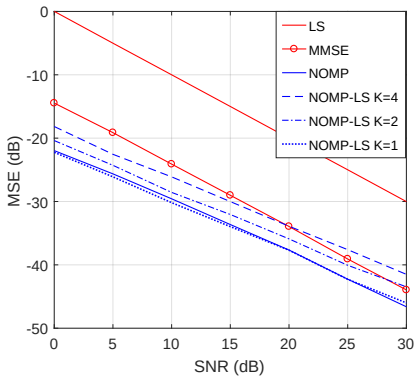

Fig. 7. The MSE performance of the reconstructions in Scenario (b) where $M = 16$ and beamforming type 1 is used.

TABLE I
OTA CONFIGURATIONS

Parameter	**Value**
Antenna Bandwidth	90 MHz
Carrier Frequency	3.5 GHz
Sampling Rate	153.6 MHz
Number of FFT Points	2048
Subcarrier Spacing	75 kHz
Transmit Power	–20dBm

[24]. After down-converting, synchronizing and sampling, the received signal is imported to the computer and processed through MATLAB. It is noted that the BS antenna array shown in Fig. 8(b) is a four-element ULA, where one column of the array is combined to form a ULA element. When evaluating the single-antenna case, the ULA is replaced by one antenna element like the user antenna.

Configurations of the OTA tests are listed in Table. I. Due to the limitations of the hardware equipments, we use in-band versus out-of-band tests to imitate the uplink versus downlink tests. Considering the bandwidth of the antenna, we select the in-band and out-of-band regions within the central 90 MHz band and distance them to the greatest extent. As shown in Fig. 9, we regard the red region with 45 MHz bandwidth as the in-band to imitate the uplink. The 15 MHz-bandwidth region colored in blue or orange is chosen as the out-of-band to imitate the downlink. The central frequency of the out-of-band region is 60 MHz away from the central frequency of the in-band region. In the gain refinement stage, we set $K = 2$ or 4, which means 100 or 50 subcarriers are allocated for the out-of-band pilots. Due to the high accuracy of MMSE estimation algorithm, we regard the MMSE-estimated channel as the real channel when evaluating the NOMP-LS-based scheme. The out-of-band channel inference method is also evaluated, which represents the method introduced in [14]. Similarly, MSE is used as the metric, which is calculated as

$$\mathrm{MSE} = \frac{\mathbb{E}\left\{\left\|\hat{\mathbf{h}}_{\mathrm{MMSE}} - \hat{\mathbf{h}}\right\|^2\right\}}{\sigma_{\mathbf{z}}^2}, \tag{62}$$

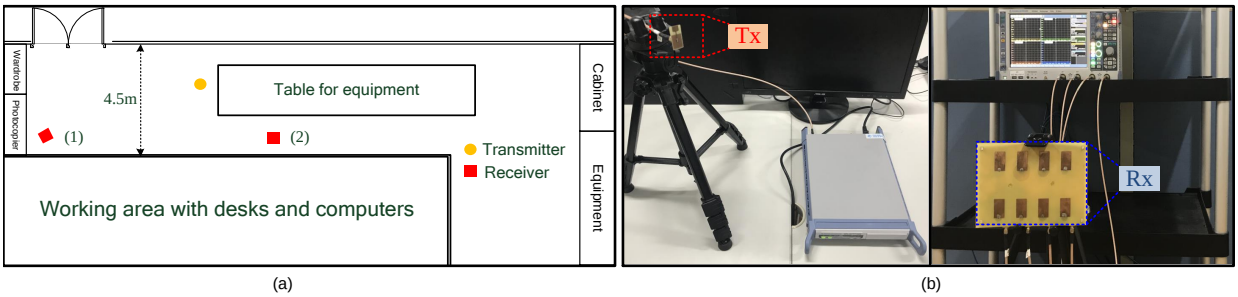

Fig. 8. (a) The OTA test environment in a laboratory and the radio devices are placed along the table for equipment. (b) The left and the right subfigures show the user and the BS, respectively.

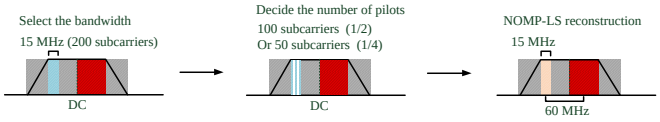

Fig. 9. Illustration of the in-band (red regions) and out-of-band (blue or orange regions) in the OTA tests.

where $\hat{\mathbf{h}}_{\mathrm{MMSE}}$ is the MMSE-estimated channel and is regarded as the real channel. The test results are displayed in the form of a cumulative distribution function (CDF).

We start by reconstructing the simplest channel when the BS is equipped with a single antenna. In the single-antenna tests, we set $K = 4$. Fig. 10(a) provides the CDF of the MSE (in dB) when the BS is located at Position (1). It can be observed that the NOMP-based in-band reconstruction achieves the highest accuracy, with a 90% probability that the MSE is below -9.65 dB. However, the performance of the NOMP-based out-of-band inference is poor, with 90% probability that MSE is lower than 15.36 dB. It demonstrates that the in-band channel can be accurately derived from the in-band NOMP process. While if we infer the out-of-band channel solely from the in-band CSI, then the inferred channel can not accurately depict the actual channel. Fortunately, the accuracy is greatly improved when the gains are refined with the aid of the out-of-band pilots. The NOMP-LS-based out-of-band reconstruction scheme functions well, with a 90% probability that the MSE is below -6.64 dB. These OTA results align with the previous numerical results. We can further investigate the power ratio of the propagation paths through the results in Fig. 10(b). The first detected path occupies 75.2% power of the channel because a strong line-of-sight (LoS) propagation path can be detected at Position (1). The power ratio is increased to 94.3% after adding the second detected paths. Results also indicate that the possibility that more than 4 paths are detected is below 0.036.

Then, we test the performances when the BS is located at Position (2) and the results are given in Fig. 11. We find that the NOMP-based in-band and the NOMP-LS-based out-of-band reconstructions still have excellent performance in the non-LoS (NLoS) propagation scenario. The 90%-probability MSEs of the two reconstructions are -9.00 dB and -6.17 dB, respectively. Although the MSE performance is inferior to the that of the LoS case, this performance degradation is relatively small. As for the power ratio of the propagation paths, the first reconstructed path only occupies 45.2% power of the channel. It has a probability of 0.416 that the number of existing paths is more than 4. These results reveal strongly that the NOMP-LS based reconstruction scheme functions well even when multiple NLoS propagation paths exist in the channel.

Next, we evaluate the performance of the NOMP-LS reconstruction in the single-input multiple-output (SIMO) system where the BS is configured to have four antennas. The CDF of the MSE results when the BS is located at Position (1) is shown in Fig. 12. Different from the single antenna cases, in the multi-antenna case, the BS needs to estimate the AoAs, in addition to the delays and gains. The channel tested from Position (1) is similar to Scenario (b) in computer simulations. A dominant LoS propagation path exists, but it is surrounded by multiple paths. These paths compose a cluster, which makes it more difficult to separate these paths from each other. We see from Fig. 12(b) that the number of estimated paths increases greatly when compared with Fig. 10(b), even these results are derived at the same place. This is because multiple antennas enhance the spatial resolution and explain the spatial channel in a more detailed way. Therefore, the NOMP-based in-band channel reconstruction achieves higher accuracy in multiple-antenna environment than single-antenna environment. When using the NOMP-LS based out-of-band channel reconstruction scheme, we insert pilots in every two or four subcarriers at the gain refinement stage. The reconstruction accuracy for the out-of-band is improved when the density of the pilots is increased. It has a 90% probability that the MSE is below -8.67 dB if we set $K = 2$. These results strongly reveal the spatial reciprocity between the channels in separated bands and validate that the proposed reconstruction scheme is effective.

Fig. 13 presents the results in the setup where the BS is located at Position (2) and the channel has NLoS paths. From the MSE lines of both the NOMP-based in-band reconstruction and the NOMP-LS based out-of-band reconstruction, a slight but negligible performance degradation can be found when compared with Fig. 12. Also, the power ratio of the first two detected paths decreases. These observations are in accordance with what we have seen from Figs. 10 and 11, demonstrating the correctness of the OTA tests about the proposed reconstruction scheme and indicating that the reconstruction scheme is able to work well in cases with more antennas.

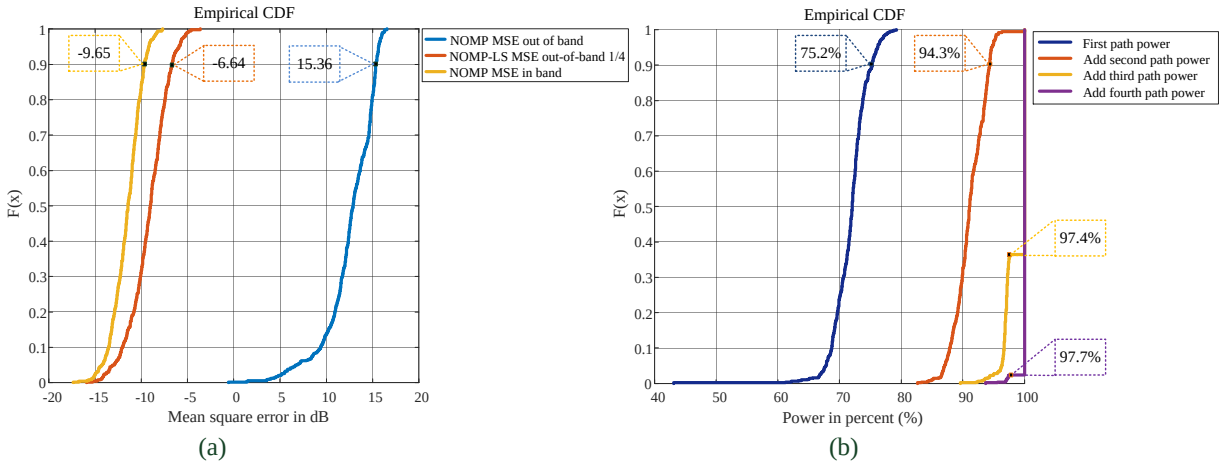

Fig. 10. The MSE performance of the NOMP-LS-based SISO channel reconstruction and the power ratio of reconstructed paths when conducting OTA tests in Position (1).

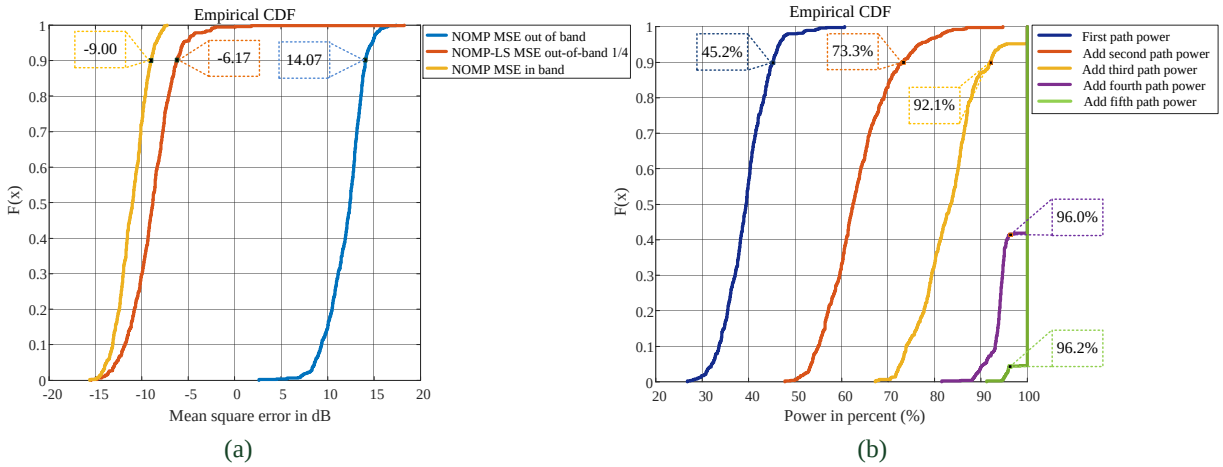

Fig. 11. The MSE performance of the NOMP-LS-based SISO channel reconstruction and the power ratio of reconstructed paths when conducting OTA tests in Position (2).

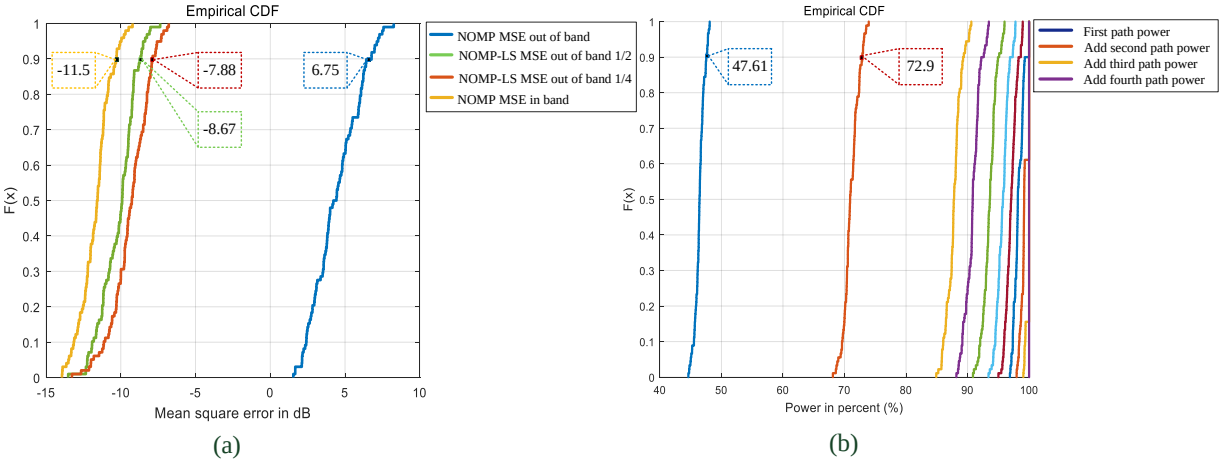

Fig. 12. The MSE performance of the NOMP-LS-based SIMO channel reconstruction and the power ratio of reconstructed paths when conducting OTA tests in Position (1).

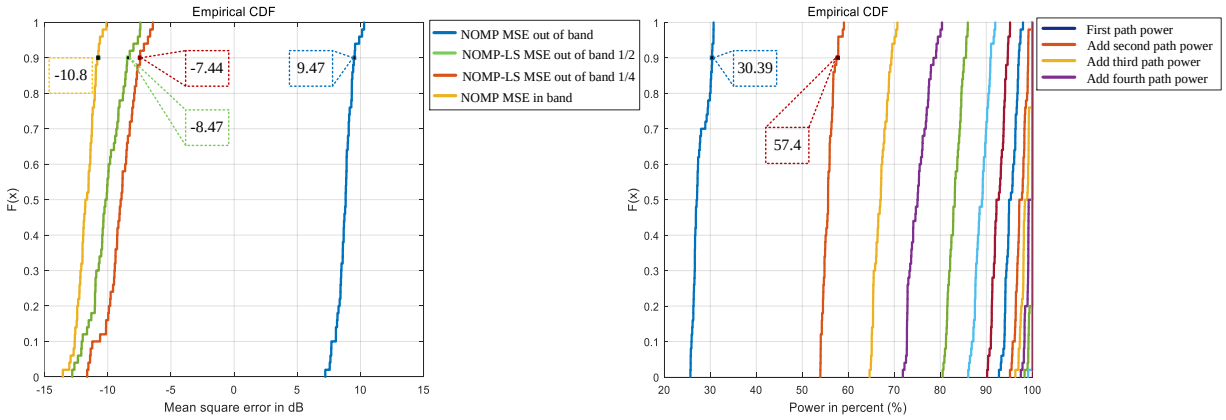

Fig. 13. The MSE performance of the NOMP-LS-based SIMO channel reconstruction and the power ratio of reconstructed paths when conducting OTA tests in Position (2).

VI. Conclusion

In this paper, an efficient downlink channel reconstruction scheme was proposed for FDD multi-antenna systems. It uses the frequency-independent features of the spatial parameters and tackles the problem of downlink CSI acquisition at the BS in the absence of uplink-downlink reciprocity with limited overhead of downlink training and feedback. We have extended the NOMP algorithm to cope with the trivariate condition to extract the frequency-independent spatial parameters. Numerical simulations validated the effectiveness of the gain refinement which causes the downlink training and feedback overhead. Then our OTA tests demonstrated that the downlink reconstruction scheme achieves promising MSE performance. The scheme and the OTA results have directive significance to the design of FDD massive MIMO systems.

References

[1] 3rd Generation Partnership Project, "Technical specification group radio access network; NR; Physical layer; General description (Release 15)," *3GPP TS 38.201 V2.0.0*, Dec. 2017.

[2] 3rd Generation Partnership Project, "Technical Specification Group Radio Access Network; Study on New Radio (NR) Access Technology (Release 14)," *3GPP TR 38.912 V1.0.0*, Mar. 2017.

[3] 3rd Generation Partnership Project, "Technical specification group Radio access network; Study on scenarios and requirements for next generation access technologies (Release 14)," *TR 38.913 V14.3.0*, Jun. 2017.

[4] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065-1082, Jun. 2014.

[5] E. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186-195, Feb. 2014.

[6] J. C. Chen, C. K. Wen, and S. Jin, "A low complexity pilot scheduling algorithm for massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 6, no. 1, pp. 18-21, Feb. 2017.

[7] J. Choi, D. J. Love, and T. Kim, "Trellis-extended codes and successive phase adjustment: A path from LTE-advanced to FDD massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2007-2016, Apr. 2015.

[8] W. Shen, L. Dai, B. Shim, Z. Wang, and R. W. Heath, "Channel feedback based on AoD-adaptive subspace codebook in FDD massive MIMO systems," *arXiv preprint arXiv:1704.00658*, 2017.

[9] J. Choi, D. J. Love, and P. Bidigare, "Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 802-814, Oct. 2014.

[10] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6169-6183, Dec. 2015.

[11] X. Rao, and V. K. N. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3261-3271, Jun. 2014.

[12] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-level localization with a singleWiFi access point," in *Proc. 13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, Mar. 2016.

[13] K. Hugl, K. Kalliola, and J. Laurila, "Spatial reciprocity of uplink and downlink radio channels in FDD systems," COST 273 Technical Document TD(02) 066, 2002.

[14] D. Vasisht, S. Kumar, H. Rahul, and D. Katabi, "Eliminating channel feedback in next-generation cellular networks," in *Proc. ACM SIGCOMM*, Aug. 2016, pp. 398-411.

[15] H. Xie, F. Gao, S. Jin, J. Fang, and Y.-C. Liang, "Channel Estimation for TDD/FDD Massive MIMO Systems with Channel Covariance Computing," *arXiv preprint arXiv:1710.00704*, 2017.

[16] Y. Han, H. Zhang, S. Jin, X. Li, R. Yu and Y. Zhang, "Investigation of transmission schemes for millimeter-wave massive MU-MIMO systems," *IEEE Syst. J.*, vol. 11, no. 1, pp. 72-83, Mar. 2017.

[17] M. B. Khalilsarai, S. Haghighatshoar, and G. Caire, "Efficient Downlink Channel Probing and Uplink Feedback in FDD Massive MIMO Systems," *arXiv:1710.07993v1*, Oct. 2017.

[18] J. Shen, J. Zhang, E. Alsusa, and K. B. Letaief, "Compressed CSI acquisition in FDD massive MIMO: How much training is needed?" *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4145-4156, Jun. 2016.

[19] B. Mamandipoor, D. Ramasamy, and U. Madhow, "Newtonized orthogonal matching pursuit: Frequency estimation over the continuum," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5066-5081, Oct. 2016.

[20] B. Eisenberg, "On the expectation of the maximum of IID geometric random variables," *Statist. Probab. Lett.*, vol. 78, no. 2, pp. 135-143, Feb. 2008.

[21] D. Ramasamy, S. Venkateswaran, and U. Madhow, "Compressive parameter estimation in awgn," *IEEE Trans. Signal Process.*, vol. 62, no. 8, pp. 2012C2027, Oct. 2014.

[22] Cisco, Ericsson, Intel Corp., LG Electronics, Nokia, Qualcomm Technologies Inc., Samsung Electronics, and Verizon, "Physical channels and modulation (Release 1)," `http://www.5gtf.net/V5G_211_v1p7.pdf`.

[23] *R&S SGT100A SGMA Vector RF Source User Manual*, `https://cdn.rohde-schwarz.com/pws/dl_downloads/dl_common_library/dl_manuals/gb_1/s/sgt100a_1/SGT100A_UserManual_en_07.pdf`.

[24] *R&S RTO Digital Oscilloscope User Manual*, `https://cdn.rohde-schwarz.com/pws/dl_downloads/dl_common_library/dl_manuals/gb_1/r/rto_1/rto2000/RTO_UserManual_en_08.pdf`.