Received: 29 June 2018 | Revised: 9 September 2018 | Accepted: 22 October 2018

DOI: 10.4218/etrij.2018-0331

ORIGINAL ARTICLE

WILEY ETRI Journal

Blind downlink channel estimation for TDD-based multiuser massive MIMO in the presence of nonlinear HPA

Parisa Pasangi[1] | **Mahmoud Atashbar**[1] | **Mahmood Mohassel Feghhi**[2]

[1]Department of Electrical Engineering, Azarbaijan Shahid Madani University, Tabriz, Iran

[2]Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran

Correspondence
Mahmoud Atashbar, Department of Electrical Engineering, Azarbaijan Shahid Madani University, Tabriz, Iran.
Email: atashbar@azaruniv.ac.ir

In time division duplex (TDD)-based multiuser massive multiple input multiple output (MIMO) systems, the uplink channel is estimated and the results are used in downlink for signal detection. Owing to noisy uplink channel estimation, the downlink channel should also be estimated for accurate signal detection. Therefore, recently, a blind method was developed, which assumes the use of a linear high-power amplifier (HPA) in the base station (BS). In this study, we extend this method to a scenario with a nonlinear HPA in the BS, where the Bussgang decomposition is used for HPA modeling. In the proposed method, the average power of the received signal for each user is a function of channel gain, large-scale fading, and nonlinear distortion variance. Therefore, the channel gain is estimated, which is required for signal detection. The performance of the proposed method is analyzed theoretically. The simulation results show superior performance of the proposed method compared to that of the other methods in the literature.

KEYWORDS
Blind channel estimation, channel gain, flat fading, massive MIMO

1 | INTRODUCTION

The massive multiple input multiple output (MIMO) system is one of the most important technologies for meeting various demands in fifth-generation communication networks [1–4]. In a massive MIMO system, the base station (BS) is equipped with an antenna array with a few hundred antennas, simultaneously serving many tens of mobile stations (MSs), which have one or a few antennas. The massive MIMO offers huge advantages over conventional MIMO, such as improved spectral efficiency, improved energy efficiency [1], enhanced reliability [5], and reduced interference [5]. In massive MIMO systems, accurate channel estimation is essential for functions such as signal detection, beamforming, and resource allocation. Owing to the large number of antennas at the BS, the channel estimation of the massive MIMO is distinct from that of the conventional MIMO, and it has many problems, such as pilot contamination, pilot overhead, and computational complexity [6]. Thus, channel estimation is a major challenge in massive MIMO systems.

Massive MIMO systems can operate in the time division duplex (TDD) or frequency division duplex (FDD) mode, where the FDD mode has several advantages with respect to the TDD mode. In FDD systems, the uplink and downlink use different frequency bands, which lead to different channels for the uplink and downlink. However, in the TDD system, owing to the use of the same frequency band, the uplink and downlink channels are the same. Thus, uplink channel estimation can be used in downlink and the linear pre-coding can be applied to focus each signal at its desired user. On the other hand, in the TDD system, the uplink estimation overhead is proportional to the number of active users, whereas in FDD, this overhead is proportional to the BS antenna number, which is larger than the number of active users in

This is an Open Access article distributed under the term of Korea Open Government License (KOGL) Type 4: Source Indication + Commercial Use Prohibition + Change Prohibition (http://www.kogl.or.kr/info/licenseTypeEn.do).
1225-6463/$ © 2019 ETRI

massive MIMO. Finally, in contrast to FDD, in TDD, the estimation quality (per antenna) cannot be reduced by adding more antennas at the BS [7].

The channel estimation should be implemented on both downlink and uplink. Owing to the existence of several receiver antennas in the uplink, the channel estimation is simple in this direction. However, in downlink, the number of receiver antennas is small, which leads to many problems in channel estimation. To address this issue, the TDD technique is used in massive MIMO. The uplink and downlink channels are approximately the same in TDD. Thus, in TDD-based massive MIMO, the uplink channel is estimated and the result is used in the downlink [8]. However, the uplink channel estimation may be inaccurate or even outdated for the downlink under rapid time-varying channel conditions, which consequently leads to performance deterioration [9,10]. To solve this problem, in TDD-based multiuser massive MIMO systems, along with the uplink channel estimation, the downlink channel estimation should also be performed for accurate signal detection.

Conventional methods for downlink channel estimation of TDD-based massive MIMO are the pilot-based methods [9–11]. In these methods, some training data, which are known in both transmitter and receiver, are sent from the BS to each MS user, and each user estimates channels between itself and all BS antennas. A large pilot overhead and less information transferring rate are some drawbacks of the pilot-based method, which have been solved by blind channel estimation methods.

In the blind methods, statistical properties of the received signals are often used for channel estimation. One of these methods, called independent component analysis (ICA), is used for the blind estimation of the flat fading massive MIMO channel [12]. Given that at ICA, the number of observations should be larger than the unknown parameter number, this method is not applicable to downlink channel estimation in massive MIMO. Another method is the subspace-based method, which is presented for the blind estimation of the frequency-selective MIMO channel [13,14]. In addition, in [15], blind MIMO channel estimation was implemented based on subspace and circulant properties of the received signal. However, none of these methods can be applied for estimating the massive MIMO downlink channel.

For TDD-based massive MIMO, Ngo et al. showed that without sending any pilot, downlink channel estimation, and therefore, signal detection can be performed if the channel has hardening property [16]. For this purpose, in flat fading channels, by choosing the maximum ratio (MR) or zero forcing (ZF) pre-coding [17], the channel gain and large-scale fading coefficients are necessary to perform the signal detection in MS. Consequently, the TDD-based massive MIMO downlink channel estimation is restricted to the estimation of channel gain [18]. In the conventional methods for channel gain estimation, only the mean of channel gain is used, which leads to a large estimate error [19,20]. Ngo et al. [16] showed that using the ZF and MR pre-coding, the mean power of the received signal in MS is a function of channel gain and large-scale fading [16]. Therefore, assuming that large-scale fading is known, they estimated the channel gain from the mean power of the received signal and used it for signal detection. To the best of our knowledge, this is the first method presented for blind estimation of the downlink channel in TDD-based multiuser massive MIMO.

In all abovementioned methods, a linear channel model was assumed for channel estimation. However, in practice, the massive MIMO channel contains nonlinearities, mainly caused by nonlinear high-power amplifiers (HPAs), which amplify the signals before they are applied to the transmitter antennas [21]. Nonlinear HPA can be described using two models [22]: memoryless models with flat frequency responses, and models with memory having frequency-selective responses.

The performance of MIMO systems in the presence of nonlinear HPA has been extensively studied [23]. However, in the context of MIMO channel estimation in the presence of nonlinear HPA, the following studies have been conducted. In [22], the MIMO channel estimation was performed in the presence of a nonlinear amplifier. In this method, a least square (LS) estimation was applied for channel estimation through some pilot transmissions. In another method, nonlinear channel estimation was presented based on support vector machines (SVMs) [24]. These methods could not be applied to estimate the massive MIMO downlink channel, because in this scenario, the number of unknown parameters was larger than that of known data. In addition, the use of pilot is another drawback of these methods, which leads to pilot overhead and less information transferring rate. Therefore, the blind downlink channel estimation of a massive MIMO system in the presence of nonlinear HPA is a significant problem, which has not been studied yet.

This paper proposes a new blind method to estimate the downlink flat fading channel of a TDD-based massive MIMO system in the presence of nonlinear HPA. In the proposed method, we use the Bussgang decomposition for converting the nonlinear effect of HPA into a linear problem. We apply ZF and MR pre-coding in the BS based on uplink channel estimates. Then, we show that the average power of the received signal at each user is a function of channel gain, large-scale fading, and nonlinear distortion variance. Therefore, knowing the other two quantities, the channel gain is estimated according to the average power of the received signal, and finally, used for signal detection. The simulation results show better performance of the proposed method compared to that proposed in [16].

The rest of this paper is organized as follows. In the next section, the nonlinear HPA models are briefly reviewed. In

TABLE 1 Notations used in this paper

$A[\rho(t)]$	Amplitude nonlinearity
$\Phi[\rho(t)]$	Phase nonlinearity
α	A constant coefficient for nonlinear HPA
$d(t)$	Nonlinear distortion of HPA
σ_d^2	Variance for nonlinear distortion
$E\{.\}$	Expectation operator
M	Number of antennas in the base station
K	Number of mobile receivers
β_k	Large-scale fading
$\mathbf{h}_k$	Small-scale fading coefficients vector
$\tau_{u,p}$	Pilot symbols number of users
ρ_u	Transmit signal-to-noise ratio of each pilot symbol
v_k	Uplink channel estimation error
γ_k	Variance of uplink channel estimation
$\boldsymbol{\alpha}_k$	Pre-coding vector
$[.]_k$	The kth column of a matrix
$\|.\|$	Euclidean norm
$(.)^H$	Hermitian
$w_k(n)$	Complex Gaussian noise
$CN(0,1)$	The complex Gaussian distribution with zero mean and unit variance
ξ_k	The average sample power of the received signal in kth user
$\lvert . \rvert$	Absolute value
$\xrightarrow{p}$	Convergence in probability
$[.]_{k,k'}$	The k,k'sth element of the matrix
τ_c	Length of the coherence interval in symbols
τ_d	The number of downlink transmitted symbols in the coherence interval

Section 3, we present the system model, and then, the proposed blind channel estimation method is introduced in Section 4. Comparisons are presented in Section 5, followed by the conclusions presented in Section 6.

We use bold-face large and small case letters to represent the matrices and vectors, respectively. The notations used in the paper are as follows (Table 1).

2 | BRIEF REVIEW OF NONLINEAR HPA MODELS

In the signal transmission process, the modulated signal is amplified by an HPA and then sent via an antenna. In fact, the HPA may operate in a nonlinear region, which leads to distortion in amplitude or phase of the signal. The input signal of the HPA is represented as follows:

$$x(t)=\rho(t)e^{j\varphi(t)}, \tag{1}$$

where $\rho(t)$ and $\varphi(t)$ are the amplitude and phase of the signal, respectively. The HPA output signal can be expressed as [25]:

$$z(t)=f(x(t))=A[\rho(t)]e^{j\{\varphi(t)+\Phi[\rho(t)]\}}, \tag{2}$$

where $A[\rho(t)]$ is the amplitude nonlinearity, called AM/AM conversion, and $\Phi[\rho(t)]$ is the phase nonlinearity, called AM/PM conversion.

Three examples of flat frequency models include a traveling wave tube amplifier (TWTA), a solid-state power amplifier (SSPA), and a soft-envelope limiter (SEL). In the TWTA model, according to the Saleh memory model [26], the AM/AM and AM/PM characteristics are as follows:

$$A[\rho(t)]=A_{\text{sat}}^2\frac{\rho(t)}{\rho(t)^2+A_{\text{sat}}^2}, \tag{3}$$

$$\Phi[\rho(t)]=\frac{\pi}{3}\frac{\rho(t)^2}{\rho(t)^2+A_{\text{sat}}^2}, \tag{4}$$

where A_{sat} is the amplifier input saturation voltage. In the SSPA model, the functions $A[\rho(t)]$ and $\Phi[\rho(t)]$ are given as [25–27]:

$$A[\rho(t)]=\frac{\rho(t)}{\left[1+\left(\frac{\rho(t)}{A_0}\right)^{2p}\right]^{\frac{1}{\{2p\}}}}, \tag{5}$$

$$\Phi\,[\rho\,(t)]=0, \tag{6}$$

where A_0 is the maximum output amplitude and p is a control parameter of the transition from a linear region to a saturation region [28]. Finally, the AM/AM and AM/PM characteristics based on the SEL model [29] are

$$A\,[\rho\,(t)]=\begin{cases}\rho\,(t) & \rho\,(t)\le A_{\text{sat}}\\ A_{\text{sat}} & \rho\,(t)>A_{\text{sat}}\end{cases}, \tag{7}$$

$$\Phi[\rho(t)]=0. \tag{8}$$

Besides, based on the Bussgang's theorem, the nonlinear effect of the model (2) can be expressed as a nonlinear disturbance with Gaussian distribution [29]. As a result, the output of the nonlinear HPA decomposes into the sum of two uncorrelated parts as [23]:

$$z(t)=f(x(t))=\alpha x(t)+d(t), \tag{9}$$

where α is a constant coefficient, which is dependent on the amplifier's characteristics and is expressed as follows:

$$\alpha=\frac{\mathbb{E}\,[x^*\,(t)\,z\,(t)]}{\mathbb{E}\,\left[\lvert x\,(t)\rvert^2\right]}. \tag{10}$$

In addition, $d(t)$ denotes the nonlinear distortion, which is a random variable with Gaussian distribution with zero mean and variance σ_d^2 [22]:

$$\sigma_d^2 = \mathbb{E}[z^*(t)z(t)] - \alpha^2 \mathbb{E}[x^*(t)x(t)]. \tag{11}$$

3 | SYSTEM MODELING IN THE PRESENCE OF NONLINEAR HPA

We consider a single-cell massive MIMO system with a single BS equipped with M antennas and K single-antenna users, where $M > K$. The Rayleigh channel between the BS antennas and the kth user is a $M \times 1$ vector channel, which is modeled as [16]:

$$\boldsymbol{g}_k = \sqrt{\beta_k}\boldsymbol{h}_k. \tag{12}$$

In this model, β_k represents large-scale fading, which is constant in the coherence interval, and $\boldsymbol{h}_k$ is an $M \times 1$ vector, which contains small-scale fading coefficients. We assume that the elements of $\boldsymbol{h}_k$ are uncorrelated random variables with zero mean and unit variance. Moreover, $\boldsymbol{h}_k$ and $\boldsymbol{h}_{k'}$ are assumed to be independent for $k \neq k'$. The mth elements of $\boldsymbol{g}_k$ and $\boldsymbol{h}_k$ are represented as g_k^m and h_k^m, respectively.

We focus on the channel estimation of the downlink in a TDD-based multiuser massive MIMO. The system's block diagram is shown in Figure 1. In the first step, the user's signals are pre-coded using the uplink channel estimates, and then, sent to HPA. Thus, the channel vectors should be estimated in the uplink direction. For this purpose, we assume that for each coherence interval, the uplink channel is estimated by simultaneously sending $\tau_{u,p}$ orthogonal pilot symbols by all users, where $\tau_{u,p} \geq K$. When a linear HPA is used in MS during uplink training, the $M \times \tau_{u,p}$ received signals at the BS are given by [30]:

$$\boldsymbol{Y}_u = \sqrt{\rho_u}\boldsymbol{G}\boldsymbol{X}_u + \boldsymbol{N}_u, \tag{13}$$

where $\boldsymbol{X}_u$ is an $K \times \tau_{u,p}$ vector with the k, nth element being the pilot of the kth user at the nth time index, $\boldsymbol{G}$ is the uplink channel matrix as $\boldsymbol{G} \triangleq [\boldsymbol{g}_1, \ldots, \boldsymbol{g}_K]$, $\boldsymbol{N}_u$ is uplink noise, which is an $M \times \tau_{u,p}$ matrix with i.i.d $CN(0,1)$ elements, and ρ_u is the uplink signal-to-noise ratio (SNR) of each pilot symbol. In this case, the linear MMSE estimate of $\boldsymbol{g}_k$ is given by [30]:

$$\hat{\boldsymbol{g}}_k = \frac{\tau_{u,p}\rho_u\beta_k}{\tau_{u,p}\rho_u\beta_k + 1}\boldsymbol{g}_k + \frac{\sqrt{\tau_{u,p}\rho_u}\beta_k}{\tau_{u,p}\rho_u\beta_k + 1}\boldsymbol{v}_k, \tag{14}$$

where $\boldsymbol{v}_k$ is an $M \times 1$ uplink channel estimation error vector, which is independent of $\boldsymbol{g}_k$. The variance of the mth element of $\hat{\boldsymbol{g}}_k$ is given by:

$$Var\{\hat{g}_k^m\} = E\{|\hat{g}_k^m|^2\} = \frac{\tau_{u,p}\rho_u\beta_k^2}{\tau_{u,p}\rho_u\beta_k + 1} \triangleq \gamma_k. \tag{15}$$

If there is a nonlinear HPA in the MS, by assuming that all users have used a similar nonlinear HPA, by using Bussgang's theorem and combining (9) and (13), the received signals at the BS are given by:

$$\boldsymbol{Y}_u = \sqrt{\rho_u}\boldsymbol{G}(\alpha\boldsymbol{X}_u + \boldsymbol{D}_u)^T + \boldsymbol{N}_u, \tag{16}$$

where $\boldsymbol{D}_u$ is the $K \times \tau_{u,p}$ matrix, where the k, nth element is the nonlinear HPA distortion of the kth user at the nth time index. In this case, (16) can be rewritten as:

$$\boldsymbol{Y}_u = \sqrt{\rho_u}\boldsymbol{G}\alpha\boldsymbol{x}_u^T + \boldsymbol{N}'_u, \tag{17}$$

where $\boldsymbol{N}'_u = \sqrt{\rho_u}\boldsymbol{G}\boldsymbol{D}_u^T + \boldsymbol{N}_u$. Given that $\boldsymbol{D}_u$ and $\boldsymbol{N}_u$ have Gaussian distribution, $\boldsymbol{N}'_u$ also has Gaussian distribution. In this case, the linear MMSE estimate of $\boldsymbol{g}_k$ is given by:

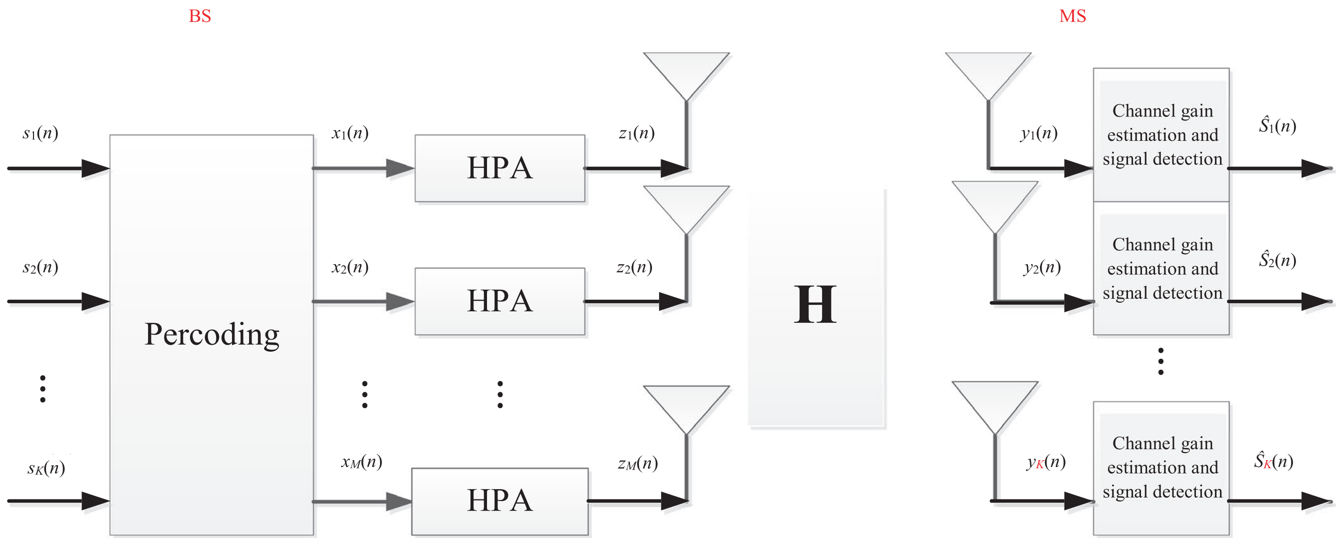

FIGURE 1 Block diagram of the considered massive MIMO system in the presence of nonlinear HPA

$$\hat{\boldsymbol{g}}_k = \frac{\alpha^2 \tau_{\mathrm{u,p}} \rho_{\mathrm{u}} \beta_k}{\alpha^2 \tau_{\mathrm{u,p}} \rho_{\mathrm{u}} \beta_k + 1} \boldsymbol{g}_k + \frac{\sqrt{\alpha^2 \tau_{\mathrm{u,p}} \rho_{\mathrm{u}} \beta_k}}{\alpha^2 \tau_{\mathrm{u,p}} \rho_{\mathrm{u}} \beta_k + 1} \boldsymbol{v}'_k. \tag{18}$$

After estimating the uplink channel, we use it in downlink transmission. For this purpose, suppose that $s_k(n)$ is the nth symbol of the signal, which must be sent to the kth user. In addition, $\boldsymbol{s}(n) \triangleq [s_1(n) \ \dots \ s_K(n)]^T$ displays a $K \times 1$ vector containing all signals sent to the K user, and $\{s(n)s^H(n)\} = \boldsymbol{I}_K$, where $\boldsymbol{I}_K$ is $K \times K$ identity matrix. Let $\boldsymbol{x}(n) \triangleq [x_1(n) \dots x_M(n)]^T$ be $M \times 1$ vector of the pre-coder output, in which $x_m(n)$ is the nth symbol of signal sent to the mth antenna. For linear pre-coding, the $\boldsymbol{x}(n)$ is considered as follows [[16]]:

$$\boldsymbol{x}(n) = \sqrt{\rho_d} \sum_{k=1}^{K} \sqrt{\eta_k} \boldsymbol{a}_k s_k(n), \quad 1 \le n \le T_d, \tag{19}$$

where $\boldsymbol{a}_k$ is the $M \times 1$ vector determined using uplink channel estimation values $\hat{\boldsymbol{G}}$, ρ_d is the transfer mean power, and η_k is the power factor associated with the kth user. In this paper, MR and ZF pre-coding are used. For MR pre-coding, $\boldsymbol{a}_k$ is chosen as [17]:

$$\frac{\boldsymbol{a}_k = \hat{\boldsymbol{g}}_k}{\|\hat{\boldsymbol{g}}_k\|, \quad k = 1, \dots, K,} \tag{20}$$

and in the ZF pre-coding $\boldsymbol{a}_k$ is chosen as [[17]]:

$$\boldsymbol{a}_k = \frac{1}{\|[\hat{\boldsymbol{G}}(\hat{\boldsymbol{G}}^H \hat{\boldsymbol{G}})^{-1}]_k\|} [\hat{\boldsymbol{G}}(\hat{\boldsymbol{G}}^H \hat{\boldsymbol{G}})^{-1}]_k. \tag{21}$$

According to Figure 1, after pre-coding, the signal passes through a nonlinear HPA. In this paper, we use the Bussgang's theorem for modeling the nonlinear effect of HPA. Thus, by using (9), the output of nonlinear HPA is as follows:

$$\begin{aligned} \boldsymbol{z}(n) &= \alpha \boldsymbol{x}(n) + \boldsymbol{d}(n) \\ &= \alpha \sqrt{\rho_d} \sum_{k=1}^{K} \sqrt{\eta_k} \boldsymbol{a}_k s_k(n) + \boldsymbol{d}(n), \end{aligned} \tag{22}$$

where $\boldsymbol{z}(n) \triangleq [z_1(n) \ \dots \ z_M(n)]^T$ and $\boldsymbol{d}(n) \triangleq [d_1(n) \ \dots \ d_M(n)]^T$. In this model, we assume that the nonlinear HPAs of all branches have similar characteristics, which leads to constant α in all branches.

On the other hand, in MS, the nth sample of the signal received for the kth user is given by:

$$y_k(n) = \boldsymbol{g}_k^H \boldsymbol{z}(n) + w_k(n), \tag{23}$$

where $w_k(n) \sim CN(0,1)$ is additive complex Gaussian noise. By replacing $\boldsymbol{z}(n)$ from (22), we have:

$$\begin{aligned} y_k(n) =& \alpha \sqrt{\rho_d \eta_k} a_{kk} s_k(n) \\ &+ \alpha \sum_{k' \neq k}^{K} \sqrt{\rho_d \eta_{k'}} a_{kk'} s_{k'}(n) + \boldsymbol{d}(n) \boldsymbol{g}_k^H + w_k(n), \end{aligned} \tag{24}$$

where $a_{kk'}$ is defined as follows:

$$a_{kk'} \triangleq \boldsymbol{g}_k^H \boldsymbol{a}_{k'} \tag{25}$$

and a_{kk} is named as the kth channel gain.

In (24), the second term on the right side indicates the interference from other users and the third term shows the effect of nonlinear HPA. In this model, to detect the $s_k(n)$ at the kth user, we need to determine the value of channel gain a_{kk}. Therefore, the downlink channel estimation of the TDD-based multiuser massive MIMO system is reduced to the estimation of the channel gain a_{kk}.

4 | PROPOSED BLIND CHANNEL ESTIMATION METHOD

To estimate the channel gain in the proposed method, we calculate the average power of the receiving signal for each user in the coherence interval as:

$$\xi_k \triangleq \frac{|y_k(1)|^2 + |y_k(2)|^2 + \dots + |y_k(\tau_{\mathrm{d}})|^2}{\tau_{\mathrm{d}}}, \tag{26}$$

where ξ_k is the average sample power of the received signal for the kth user and τ_{d} is the number of downlink transmitted symbols in the coherence interval. According to the law of large numbers, for $\tau_{\mathrm{d}} \to \infty$ we have:

$$\xi_k - E\left\{|y_k(n)|^2\right\} \xrightarrow{p} 0. \tag{27}$$

According to (24) we have:

$$\begin{aligned} E\left\{|y_k(n)|^2\right\} =& E\Big\{\rho_d \eta_k |a_{kk}|^2 \alpha^2 s_k^2(n) \\ &+ \sum_{k' \neq k}^{K} \sum_{k'' \neq k}^{K} \alpha^2 \rho_d \sqrt{\eta_{k'} \eta_{k''}} a_{kk'} a_{kk'} s_{k'}(n) s_{k''}(n) \\ &+ \boldsymbol{g}_k^H \boldsymbol{d}(n) \boldsymbol{d}^H(n) \boldsymbol{g}_k + w_k^2(n) \\ &+ 2\sqrt{\rho_d \eta_k} |a_{kk}| \alpha^2 s_k(n) \sum_{k' \neq k}^{k} \sqrt{\rho_d \eta_{k'}} |a_{kk'}| s_{k'}(n) \\ &+ 2\sqrt{\rho_d \eta_k} |a_{kk}| \alpha s_k(n) \boldsymbol{g}_k^H \boldsymbol{d}(n) \\ &+ 2\sqrt{\rho_d \eta_k} |a_{kk}| \alpha s_k(n) w_k(n) \\ &+ 2\boldsymbol{g}_k^H \boldsymbol{d}(n) \alpha \sum_{k' \neq k}^{k} \sqrt{\rho_d \eta_{k'}} |a_{kk'}| s_{k'}(n) \\ &+ 2w_k(n) \alpha \sum_{k' \neq k}^{k} \sqrt{\rho_d \eta_{k'}} |a_{kk'}| s_{k'}(n) \\ &+ 2\boldsymbol{g}_k^H \boldsymbol{d}(n) w_k(n)\Big\}. \end{aligned} \tag{28}$$

In (28), as the users' signals have zero mean and unit average energy, and are independent for $k \neq k'$, we have:

$$E\left\{\rho_d \eta_k |a_{kk}|^2 \alpha^2 s_k^2(n)\right\} = \rho_d \eta_k |a_{kk}|^2 \alpha^2, \tag{29}$$

$$\begin{aligned} E&\left(\sum_{k'' \neq k}^{K} \alpha^2 \rho_d \sqrt{\eta_{k'} \eta_{k''}} a_{kk'} a_{kk'} s_{k'}(n) s_{k''}(n)\right) \\ &= E\left(\alpha^2 \rho_d \eta_{k'} |a_{kk'}|^2 s_{k'}^2(n)\right) \\ &= \alpha^2 \rho_d \eta_{k'} |a_{kk'}|^2, \end{aligned} \tag{30}$$

$$E\left\{2\sqrt{\rho_d\eta_k}|a_{kk}|\alpha^2 s_k(n)\sum_{k'\neq k}^{k}\sqrt{\rho_d\eta_{k'}}|a_{kk'}|s_{k'}(n)\right\}=0. \quad (31)$$

For the third sentence in (28), we have:

$$E\{\boldsymbol{g}_k^H\boldsymbol{d}(n)\boldsymbol{d}^H(n)\boldsymbol{g}_k\}=\boldsymbol{g}_k^H\mathrm{E}\{\boldsymbol{d}(n)\boldsymbol{d}^H(n)\}\boldsymbol{g}_k = \delta_{\mathrm{d}}^2\boldsymbol{g}_k^H\boldsymbol{g}_k = M\beta_k\delta_{\mathrm{d}}^2. \quad (32)$$

In addition, the nonlinear disturbance $d_m(n)$ is independent of the users' signals. Thus,

$$E\{2\sqrt{\rho_d\eta_k}|a_{kk}|\alpha s_k(n)\boldsymbol{g}_k^H\boldsymbol{d}(n)\}=0, \quad (33)$$

$$E\left\{2\boldsymbol{g}_k^H\boldsymbol{d}(n)\,\alpha\sum_{k'\neq k}^{k}\sqrt{\rho_d\eta_{k'}}|a_{kk'}|s_{k'}(n)\right\}=0. \quad (34)$$

Finally, assuming that the signal and noise are uncorrelated, we have:

$$E\left\{2\sqrt{\rho_d\eta_k}|a_{kk}|\alpha s_k(n)\,w_k(n)\right\}=0, \quad (35)$$

$$E\left\{2w_k(n)\,\alpha\sum_{k'\neq k}^{k}\sqrt{\rho_d\eta_{k'}}|a_{kk'}|s_{k'}(n)\right\}=0. \quad (36)$$

As a result, (28) is simplified as follows:

$$E\left\{|y_k(n)|^2\right\}=\rho_d\eta_k|a_{kk}|^2\alpha^2 + \sum_{k'\neq k}^{k}\rho_d\eta_{k'}|a_{kk'}|^2\alpha^2 + M\beta_k\delta_{\mathrm{d}}^2+1. \quad (37)$$

Given that $\sum_{k'\neq k}^{K}\rho_d\eta_{k'}|a_{kk'}|^2\alpha^2\sim$ is the sum of several sentences, according to the law of large numbers, it can be estimated as $E\left(\alpha^2\rho_d\sum_{k'\neq k}^{K}\eta_{k'}|a_{kk'}|^2\right)$ As a result, when K and τ_{d} are large, ξ_k can be approximated as follows:

$$\xi_k\approx\alpha^2\rho_d\eta_k|a_{kk}|^2 + E\left(\alpha^2\rho_d\sum_{k'\neq k}^{K}\eta_{k'}|a_{kk'}|^2\right)+M\beta_k\delta_{\mathrm{d}}^2+1. \quad (38)$$

The approximation is acceptable even for small values of K, because when K is small, the term $\sum_{k'\neq k}^{K}\eta_{k'}|a_{kk'}|^2$ is much less than $\eta_k|a_{kk}|^2$, (in fact $|a_{kk'}|^2|a_{kk}|^2$ [[16]].

On the other hand, for MR pre-coding in the Rayleigh channel, given that for $k'\neq k$ the vectors $\boldsymbol{a}_{k'}$ and $\boldsymbol{g}_k$ are approximately independent, we have:

$$\begin{aligned}E\{|a_{kk'}|^2\}&=E\{a_{kk'}^H a_{kk'}\}\\&=E\left\{\boldsymbol{a}_{k'}^H\boldsymbol{g}_k\boldsymbol{g}_k^H\boldsymbol{a}_{k'}\right\}\\&=\beta_k E\left\{\|\boldsymbol{a}^2\|\right\}\\&=\beta_k.\end{aligned} \quad (39)$$

This leads to:

$$E\left\{\sum_{k'\neq k}^{K}\eta_{k'}|a_{kk'}|^2\right\}=\sum_{k'\neq k}^{K}\eta_{k'}\beta_k. \quad (40)$$

Combining (40) with (38), for MR pre-coding, we have:

$$\xi_k\approx\alpha^2\rho_d\eta_k|a_{kk}|^2+\alpha^2\rho_d\sum_{k'\neq k}^{K}\eta_{k'}\beta_k+M\beta_k\delta_{\mathrm{d}}^2+1. \quad (41)$$

Solving (41), the channel gain estimation is as follows:

$$|\hat{a}_{kk}|=\begin{cases}\sqrt{\dfrac{\xi_k-1-M\beta_k\delta_{\mathrm{d}}^2-\alpha^2\rho_d\sum_{k'\neq k}^{K}\eta_{k'}\beta_k}{\rho_d\eta_k\alpha^2}} \\ if\,\xi_k>1+M\beta_2\sigma_{\mathrm{d}}^2+\alpha^2\rho_d\sum_{k'\neq k}^{K}\eta_{k'}\beta_k \\ E\left\{|a_{kk}|\right\} \qquad \text{otherwise.}\end{cases} \quad (42)$$

In contrast to MR pre-coding, in ZF pre-coding, $\boldsymbol{a}_{k'}$ and $\boldsymbol{g}_k$ vectors are not independent. Therefore, (40) is not valid. Instead, the following two facts are used: (1) for the channel vector, the estimated value ($\hat{\boldsymbol{g}}_k$) and estimation error ($\boldsymbol{g}_k-\hat{\boldsymbol{g}}_k$) are independent, which leads to the independence of $\boldsymbol{a}_k$ from $\boldsymbol{g}_k-\hat{\boldsymbol{g}}_k$; and (2) the inner product of $\hat{\boldsymbol{g}}_k$ and $\boldsymbol{a}_{k'}$ is zero, that is, $\hat{\boldsymbol{g}}_k^H\boldsymbol{a}_{k'}=0$. Thus, for ZF pre-coding, we have:

$$\begin{aligned}E\left\{|a_{kk'}|^2\right\}&=E\left\{|\boldsymbol{g}_k^H\boldsymbol{a}_{k'}|^2\right\}\\&=E\left\{|(\boldsymbol{g}_k^H-\hat{\boldsymbol{g}}_k^H)\boldsymbol{a}_{k'}+\hat{\boldsymbol{g}}_k^H\boldsymbol{a}_{k'}|^2\right\}\\&=E\left\{|(\boldsymbol{g}_k^H-\hat{\boldsymbol{g}}_k^H)\boldsymbol{a}_{k'}|^2\right\}\\&=\left(\beta_k-\gamma_k\right)E\left\{\|\boldsymbol{a}_{k'}\|^2\right\}\\&=\beta_k-\gamma_k.\end{aligned} \quad (43)$$

This leads to

$$\left\{\sum_{k'\neq k}^{K}\eta_{k'}|a_{kk'}|^2\right\}=\sum_{k'\neq k}^{K}\eta_{k'}\left(\beta_k-\gamma_k\right). \quad (44)$$

Combination (44) with (41) for the ZF pre-coding, we have:

$$\xi_k\approx\alpha^2\rho_d\eta_k|a_{kk}|^2+M\beta_k\delta_{\mathrm{d}}^2+1 +\alpha^2\rho_d\sum_{k'\neq k}^{K}\eta_{k'}\left(\beta_k-\gamma_k\right). \quad (45)$$

So, the estimated value of $|a_{kk}|$ is:

$$|\hat{a}_{kk}|=\begin{cases}\sqrt{\dfrac{\xi_k-1-M\beta_k\delta_{\mathrm{d}}^2-\alpha^2\rho_d\sum\limits_{k'\neq k}^{K}\eta_{k'}\left(\beta_k-\gamma_k\right)}{\rho_d\eta_k\alpha^2}}\\ if \quad \xi_k>1+M\beta_2\sigma_{\mathrm{d}}^2+\alpha^2\rho_d\sum\limits_{k'\neq k}^{K}\eta_{k'}\left(\beta_k-\gamma_k\right)\\ E\left\{|a_{kk}|\right\} \qquad \text{otherwise.}\end{cases} \tag{46}$$

Besides, when M is large, the real part of a_{kk} is much larger than its imaginary part. Therefore, the phase of a_{kk} is very small and can be approximated with zero [16]. Hence, $|\hat{a}_{kk}|=\hat{a}_{kk}$.

Finally, the proposed algorithm for the downlink channel estimation of TDD-based massive MIMO is summarized in Table 2.

5 | PERFORMANCE ANALYSIS

In this section, the convergence of the proposed method for $\tau_{\mathrm{c}}\to\infty$ (lead to $\tau_{\mathrm{d}}\to\infty$) and $M\to\infty$ is investigated for the Rayleigh fading channel. For this purpose, for both MR and ZF pre-coding, we rewrite (42) and (46) as

$$|\hat{a}_{kk}|=\begin{cases}\sqrt{\dfrac{\xi_k-1-M\beta_k\delta_{\mathrm{d}}^2-E\left(\alpha^2\rho_d\sum\limits_{k'\neq k}^{K}\eta_{k'}|a_{kk}|^2\right)}{\rho_d\eta_k\alpha^2}}\\ if \ \xi_k>1+M\beta_2\sigma_{\mathrm{d}}^2+E\left(\alpha^2\rho_d\sum\limits_{k'\neq k}^{K}\eta_{k'}|a_{kk}|^2\right)\\ E\left\{|a_{kk'}|\right\} \qquad \text{otherwise.}\end{cases} \tag{47}$$

When $\tau_{\mathrm{c}}\to\infty$, according to (27) and (37), we have:

$$\begin{aligned}\xi_k-(\rho_d\eta_k|a_{kk}|^2\alpha^2+\sum_{k'\neq k}^{K}\rho_d\eta_{k'}|a_{kk'}|^2\alpha^2\\ +M\beta_2\sigma_{\mathrm{d}}^2+1)=0.\end{aligned} \tag{48}$$

Calculating ξ_k in (48) and substituting it in (47), we have:

$$|\hat{a}_{kk}|=\begin{cases}\sqrt{|a_{kk}|^2+\sum\limits_{k'\neq k}^{K}\frac{\eta_{k'}}{\eta_k}(|a_{kk'}|^2-E\{|a_{kk'}|^2\})}\\ if \ \eta_k|a_{kk}|^2+\sum\limits_{k'\neq k}^{K}\eta_{k'}|a_{kk'}|^2>\sum\limits_{k'\neq k}^{K}\eta_{k'}E\left\{|a_{kk'}|^2\right\}\\ E\left\{|a_{kk}|\right\} \qquad \text{otherwise.}\end{cases} \tag{49}$$

TABLE 2 Proposed algorithm

In BS	
	1. Uplink channel estimation using a conventional method
	2. Pre-coding downlink signals using the MR or ZF method
In MS (kth user)	
	1. Calculated ξ_k using (26)
	2. Calculated a_{kk} using (42) or (46)

Since $\tau_{\mathrm{c}}\to\infty$, we can assume that the channel estimation for uplink is perfect, which leads to $\hat{\boldsymbol{G}}=\boldsymbol{G}$. In this case, for MR pre-coding, we have:

$$a_{kk}=\boldsymbol{g}_k^H\boldsymbol{a}_k=\boldsymbol{g}_k^H\frac{\hat{\boldsymbol{g}}_k}{\|\hat{\boldsymbol{g}}_k\|}=\boldsymbol{g}_k^H\frac{\boldsymbol{g}_k}{\|\boldsymbol{g}_k\|}=\|\boldsymbol{g}_k\|>0. \tag{50}$$

By dividing both sides of (49) by a_{kk}, we have:

$$\frac{\hat{a}_{kk}}{\hat{a}_{kk}}=\begin{cases}\sqrt{1+\sum\limits_{k'\neq k}^{k}\frac{\eta'_k}{\eta_k a_{kk}^2}(|a_{kk'}|^2-E\{|a_{kk'}|^2\})}\\ if \ \eta_k|a_{kk}|^2+\sum\limits_{k'\neq k}^{k}\eta_{k'}(|a_{kk'}|^2-E\{|a_{kk'}|^2\})>0\\ \frac{E\{|a_{kk}|\}}{a_{kk}} \qquad \text{otherwise.}\end{cases} \tag{51}$$

Using (39) and (50), the above equation is simplified as follows:

$$\frac{\hat{a}_{kk}}{\hat{a}_{kk}}=\begin{cases}\sqrt{1+\sum\limits_{k'\neq k}^{K}\frac{\eta_{k'}}{\eta_k}\frac{\left|\frac{\mathbf{g}_k^H\mathbf{g}_{k'}}{\|\mathbf{g}_{k'}\|}\right|^2-\beta_k}{\|\mathbf{g}_k\|^2}}\\ \text{if } \eta_k|a_{kk}|^2+\sum\limits_{k'\neq k}^{K}\eta_{k'}\left(|a_{kk'}|^2-\beta_k\right)>0\\ \frac{E\{|a_{kk}|\}}{a_{kk}} \qquad \text{otherwise.}\end{cases} \tag{52}$$

On the other hand, according to the Cauchy–Schwarz inequality, the maximum value of $|\frac{\mathbf{g}_k^H\mathbf{g}_{k'}}{\|\mathbf{g}_{k'}\|}|^2$ occurs when $k=k'$. Therefore,

$$\frac{\left|\frac{\mathbf{g}_k^H\mathbf{g}_{k'}}{\|\mathbf{g}_{k'}\|}\right|^2-\beta_k}{\|\boldsymbol{g}_k\|^2}\leq\frac{\left|\frac{\mathbf{g}_k^H\mathbf{g}_{k}}{\|\mathbf{g}_{k}\|}\right|^2-\beta_k}{\|\boldsymbol{g}_k\|^2}. \tag{53}$$

In addition,

$$\frac{\left|\frac{\mathbf{g}_k^H\mathbf{g}_{k}}{\|\mathbf{g}_{k}\|}\right|^2-\beta_k}{\|\boldsymbol{g}_k\|^2}=\frac{\|\boldsymbol{g}_k\|-\beta_k}{\|\boldsymbol{g}_k\|^2}=\frac{1}{\|\boldsymbol{g}_k\|}-\frac{\beta_k}{\|\boldsymbol{g}_k\|^2}. \tag{54}$$

If $M\to\infty$, then $\|\boldsymbol{g}_k\|\to\infty$, and hence, (54) will be zero. Thus, (52) becomes:

$$\frac{\hat{a}_{kk}}{a_{kk}}=\begin{cases}1\\ if \ \eta_k|a_{kk}|^2+\sum\limits_{k'\neq k}^{K}\eta_{k'}\left(|a_{kk'}|^2-\beta_k\right)>0\\ \frac{E\{|a_{kk}|\}}{a_{kk}} \qquad \text{otherwise.}\end{cases} \tag{55}$$

Lemma *If* $\tau_d \to \infty$ *and* $\mathrm{M} \to \infty$, *for MR pre-coding, we have:*

$$\Pr\left\{\eta_k|a_{kk}|^2+\sum_{k'\neq k}^{K}\eta_{k'}\left(|a_{kk'}|^2-\beta_k\right)>0\right\}=1.$$

Proof: Note that:

$$\begin{aligned}\Pr&\left\{\eta_k|a_{kk}|^2+\sum_{k'\neq k}^{K}\eta_{k'}\left(|a_{kk'}|^2-\beta_k\right)>\sim 0\right\}\\ &\geq \Pr\left\{\eta_k|a_{kk}|^2>\sum_{k'\neq k}^{K}\eta_{k'}\beta_k\right\}\\ &=\Pr\left\{\|\boldsymbol{g}_k\|^2>\sim\sum_{k'\neq k}^{K}\frac{\eta_{k'}}{\eta_k}\beta_k\right\},\end{aligned} \tag{56}$$

where the inequality is due to ignoring the positive term $\sum_{k'\neq k}^{K}\eta_{k'}|a_{kk'}|^2$ and the third equality is due to replacing a_{kk} with (50). In this case, when $M\to\infty$, the value of $\|g_k\|^2$ is infinite; hence, the above probability equals to one. ■

Consequently, for $M \to \infty$, with probability equal to one, the first condition in (55) is established, which is equivalent to:

$$\frac{\hat{a}_{kk}}{a_{kk}}\to 1,\quad \text{as } M\to\infty. \tag{57}$$

We now prove the convergence in the ZF pre-coding case. For this purpose, we first present the following lemma.

Lemma *If* $\tau_c\to\infty$, *for ZF pre-coding, we have:*

$$a_{kk'}=\begin{cases}\frac{1}{\|[\boldsymbol{G}(\boldsymbol{G}^H\boldsymbol{G})^{-1}]_k\|} & k=k'\\ 0 & k\neq k'\end{cases}. \tag{58}$$

***Proof*:** Given the relations (21) and (26), we have $a_{kk'}=\frac{[G^H\hat{G}(\hat{G}^H\hat{G})^{-1}]_{k,k'}}{\|[\hat{G}(\hat{G}^H\hat{G})^{-1}]\|_{k'}}$**.** For $\tau_c\to\infty$**,** given that $\hat{\boldsymbol{G}}=\boldsymbol{G}$**,** we have:

$$a_{kk'}=\frac{[\boldsymbol{G}^H\boldsymbol{G}(\boldsymbol{G}^H\boldsymbol{G})^{-1}]_{k,k'}}{\|[\boldsymbol{G}(\boldsymbol{G}^H\boldsymbol{G})^{-1}]_{k'}\|}=\frac{[I_K]_{k,k'}}{\|[\boldsymbol{G}(\boldsymbol{G}^H\boldsymbol{G})^{-1}]_{k'}\|}. \tag{59}$$

which is equivalent to (58). ■

Then, by replacing (58) in (49), we have:

$$\frac{\hat{a}_{kk}}{a_{kk}}=\begin{cases}1 & \text{if } \eta_k|a_{kk}|^2>0\\ \frac{E\{|a_{kk}|\}}{a_{kk}} & \text{otherwise}\end{cases}. \tag{60}$$

As a result, similar to MR, in ZF pre-coding $\frac{\hat{a}_{kk}}{a_{kk}}\to 1$ Therefore, convergence is obtained with the proposed method. Now, the proof is completed.

6 | SIMULATION RESULTS

In this section, the simulation results are presented to evaluate the performance of the proposed method. In all experiments, the results of channel estimation for three methods are compared: (a) the estimation method using $E\{a_{kk}\}$, named as "*mean method*," (b) method proposed in [[16]], named as "*Ngo method,*" and (c) our proposed method. For each user, we will consider the normalized mean square error (MSE) as an evaluation metric, which is defined as:

$$\mathrm{MSE}_k\triangleq\frac{E\left\{|\hat{a}_{kk}-a_{kk}|^2\right\}}{\left\{|a_{kk}|^2\right\}}, \tag{61}$$

where $\hat{a}_{kk}$ represents the estimated value of a_{kk}. The results of each simulation are the average of 1,000 independent Monte Carlo simulation running. The downlink signal-to-noise ratio ($\mathrm{SNR_d}$) and the uplink signal-to-noise ratio ($\mathrm{SNR_u}$), respectively, are defined as:

$$\mathrm{SNR_d}=\rho_\mathrm{d}\times\mathrm{MED} \tag{62}$$

and

$$\mathrm{SNR_u}=\rho_\mathrm{u}\times\mathrm{MED}, \tag{63}$$

where MED is the middle amount of large-scale fading for the user located in the middle of the cell.

We consider a circular cell with a radius of $R_{\max}$ m and a BS located in its center. Similar to [16], we assume $K+1$ users located randomly (with uniform distribution) in the cell. In addition, we assume that the users' distances from the BS are larger than $R_{\min}$ m, and the user with the smallest value of β_k is dropped; Hence, K users remain. The large-scale fading with log-normal distribution is assumed as follows:

$$\beta_k=\mathrm{PL}_0\left(\frac{d_k}{R_{\min}}\right)^{v}\times 10^{\frac{\sigma_{\mathrm{sh}}N(0,1)}{10}}, \tag{64}$$

where v is the path loss exponent, d_k is the distance of the kth user from the BS, PL_0 is a path loss constant, and σ_{sh} is the standard deviation of shadow fading. In the simulation, we use $K=10$, $R_{\min}=150$, $R_{\max}=1{,}050$, $v=3.8$, $\sigma_{\mathrm{sh}}=5$, $\mathrm{PL}_0=1$, $\tau_{\mathrm{u,p}}=K$, and $\mathrm{SNR_u}=0\,\mathrm{dB}$.

In addition, based on the max-min power control algorithm, we select the power control coefficients as [17]:

$$\eta_k=\frac{1+\rho_d\beta_k}{\rho_d\gamma_k\left(\frac{1}{\rho_d}\sum_{k'=1}^{K}\gamma_{k'}+\sum_{k'=1}^{K}\frac{\beta_{k'}}{\gamma_{k'}}\right)}. \tag{65}$$

Besides, we consider the SEL nonlinear amplifier model, where parameters α and σ_d^2 are obtained as follows [31]:

$$\alpha=1-e^{-\lambda^2}+\frac{1}{2}\sqrt{\pi}\lambda erfc(\lambda), \tag{66}$$

$$\sigma_\mathrm{d}^2=\sigma^2\left[1-e^{-\lambda^2}-\alpha^2\right], \tag{67}$$

where σ^2 is the nonlinear HPA input signal variance and $\lambda = A_{sat}/\sigma$. In this paper, the nonlinear HPA parameters are $\alpha = 0.769$ and $\sigma_d^2 = 0.000593$, w, which are chosen from table I in [[31]]. For a fair comparison of the proposed method with the other methods, we assume that the HPA of the reference methods in a linear region has the same gain, α.

In the first section of the experiments, the performances of three proposed methods are compared for various SNR_d values, with $M = 300$ and $\tau_d = 300$. The results for both ZF and MR pre-coding are shown in Figure 2. (In all figures of this paper, the vertical axis represents the mean of the normalized MSEs). As can be seen, by increasing SNR_d, the MSE of both the proposed and the reference method are decreased, while that of the mean method is constant. In addition, the MSE of the proposed method is less than that of the other methods. This result is valid for both MR and ZF pre-codings. For example, in $SNR_d = 10$ dB, the MSE of the proposed method for MR pre-coding is −19.79 dB, while that for the reference method is −9.81 dB, and for mean method is −1.79 dB. The corresponding values for ZF pre-coding are −27.97 dB for the proposed method, −11.46 dB for the reference method, and −1.79 dB for the mean method.

In addition, the MSE performance of the three methods vs SNR_d is compared for $M = 500$ and $\tau_d = 300$ in Figure 3 and for $M = 300$ and $\tau_d = 500$ in Figure 4. As can be seen in Figures 3 and 4, similar to Figure 2 for the proposed and reference methods, by increasing SNR_d, MSE decreases and the proposed method outperforms the mean and reference method in accurate channel gain estimation. Especially, the difference in the MSEs of the proposed method with that of the reference method is significant in lower SNRs, where the nonlinear distortion of HPA has a large effect in channel gain estimation in this case. In return, for the higher values of SNR_d, the proposed and reference methods approximately have similar performances.

Besides, comparing the ZF and MR pre-coding results show that the performance of channel estimation by using ZF

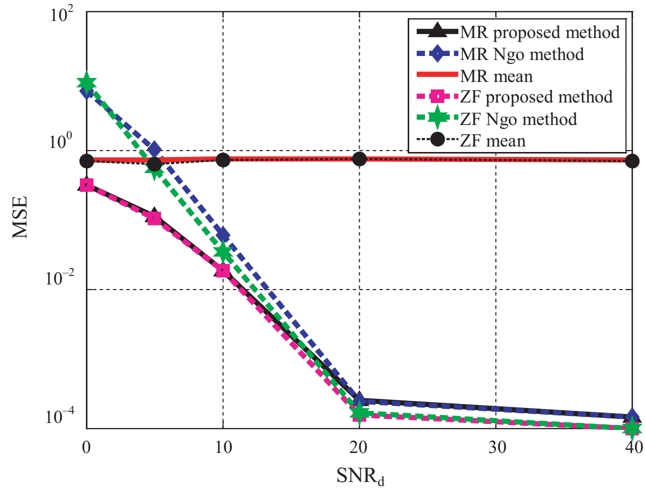

FIGURE 3 MSE vs SNR_d for $M = 500$ and $\tau_d = 300$

pre-coding is better than that of using MR pre-coding. This is because, for the approximation of $E\left\{|a_{kk'}|^2\right\}$ in ZF pre-coding, we have supposed that for $k' \neq k$ the vectors $\boldsymbol{a}_{k'}$ and $\boldsymbol{g}_k$ are independent, which is not accurate. In fact, $\boldsymbol{g}_{k'}$ and $\cap\, g_k$ are approximately (not exactly) independent.

In addition, comparing the results of Figures 2 and 4, show that for $M = 300$, in most SNR_d, such as $SNR_d = 0$ dB, 5 dB, 10 dB, and 20 dB, by increasing τ_d, the MSEs of the proposed method are decreased both pre-coding.

In the fifth simulation of the performance evaluation, three methods are compared for various values of τ_d. For this purpose, we select $M = 100$ and $SNR_d = 10$ dB. The results are shown in Figure 5. In this figure, the horizontal axis represents the values of τ_d and the vertical axis represents the MSE values. As shown in Figure 5, the performance of the proposed method in both pre-coding is better than that of the reference and mean methods. For example, in $\tau_d = 500$, the proposed method has the MSE value of −15.21 dB for MR

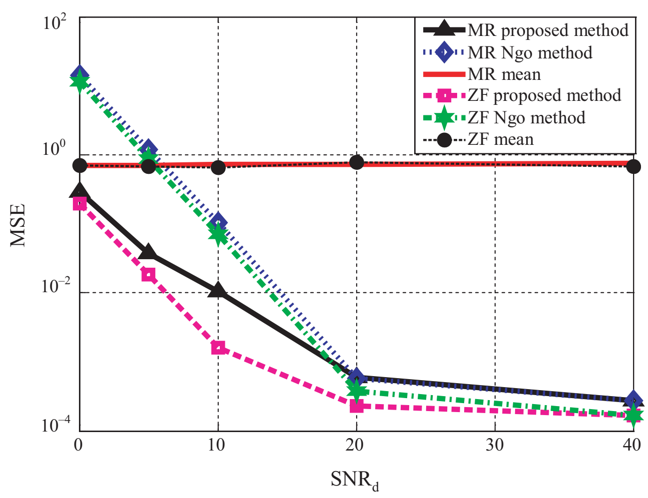

FIGURE 2 MSE vs SNR_d for $M = 300$ and $\tau_d = 300$

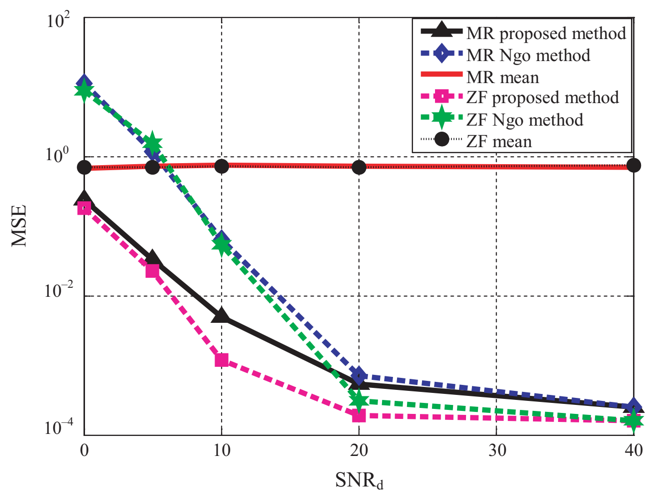

FIGURE 4 MSE vs SNR_d for $M = 300$ and $\tau_d = 500$

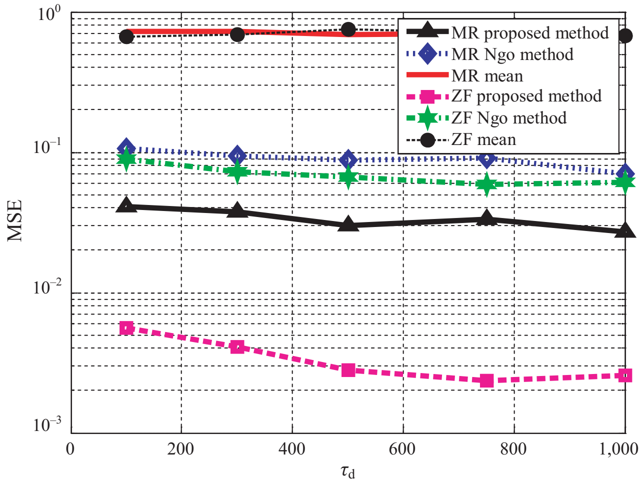

FIGURE 5 MSE vs τ_d for $M = 100$ and $SNR_d = 10$ dB

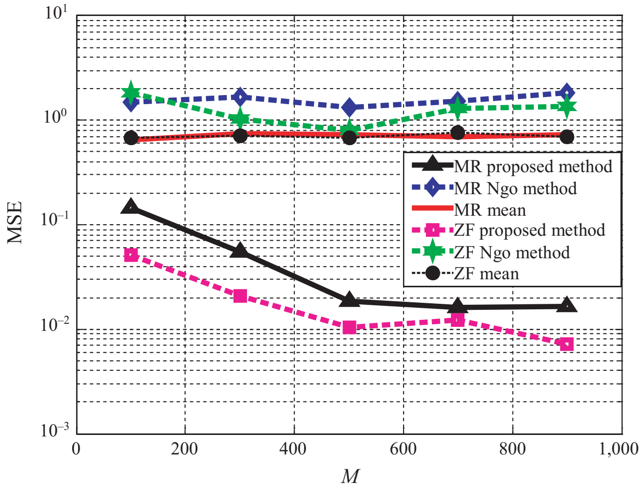

FIGURE 6 MSE vs M for $\tau_d = 1{,}000$ and $SNR_d = 5$

pre-coding case, but for the reference and mean methods, the above values are −10.6 dB and −1.24 dB, respectively. For the ZF pre-coding case, the MSE values in $\tau_d = 500$ are −25.52 dB, −11.76 dB, and −1.24 dB for the proposed, reference, and mean methods, respectively. This result shows that the performance of the proposed method in both pre-coding cases is better than the reference and mean methods.

In the sixth section of the performance test, three methods are compared for different antenna numbers. For this purpose, we select $\tau_d = 1{,}000$, $SNR_d = 5$, and $M = \{100, 300, 500, 700, 900\}$. The results are shown in Figure 6, where the MSEs of the channel gain estimation for both ZF and MR pre-coding are seen. As shown in Figure 6, the performance of the proposed method in both pre-coding cases is better than that of the reference and mean methods. As seen by increasing the value of M, the difference between the proposed method and other methods are increased. The main reason for this result is that in the presence of nonlinear HPA, according to (30) and (41), the average power of the received signal (ξ_k) is a function of antenna number (M), which is not seen in the reference and mean method relations.

Overall, the comparison of all results shows that the proposed method has high accuracy channel gain estimation compared to the reference and mean methods. In addition, a comparison of the results of the proposed method for two types of pre-codings ZF and MR shows that in both pre-coding cases, performance of the proposed method is better than that of the reference methods.

7 | CONCLUSION

In the presence of nonlinear HPA, which is modeled according to the Bussgang's theorem, and by using ZF or MR precoding in the BS, we showed that the signal received to each user in multiuser massive MIMO is a summation of the corresponding transmitted signal multiplied by the channel gain coefficient, other users' interference, nonlinear distortion, and noise. Thus, if the channel gain is known, signal detection can be performed for each user. Therefore, the channel estimation turned into channel gain estimation. By calculating the average power of the signal received for each user, we proposed a new blind method for channel gain estimation in TDD-based multiuser massive MIMO in the presence of nonlinear HPA. In the proposed method, the channel gain was estimated as a function of large-scale fading, nonlinear distortion variance, and noise variance. The simulation results showed that the MSE of the proposed method is less than that of the reference methods in terms of SNR variation, antenna number, and coherence symbol number.

REFERENCES

1. E. G. Larsson et al., *Massive MIMO for next generation wireless systems*, IEEE Commun. Mag. **52** (2014), no. 2, 186–195.
2. Q. Zhang et al., *Power scaling of uplink massive MIMO systems with arbitrary-rank channel means*, IEEE J. Sel. Topics Signal Process. **8** (2014), no. 5, 966–981.
3. A. Liu and V. K. N. Lau, *Phase only RF pre-coding for massive MIMO systems with limited RF chains*, IEEE Trans. Signal Process. **62** (2014), no. 17, 4505–4515.
4. S. Nohn et al., *Pilot beam pattern design for channel estimation in massive MIMO systems*, IEEE J. Sel. Topics Signal Process. **8** (2014), no. 5, 787–801.
5. F. Rusek et al., *Scaling up MIMO: opportunities and challenges with very large arrays*, IEEE Signal Process. Mag. **30** (2013), no. 1, 40–60.
6. H. Xie, F. Gao, and S. Jin, *An overview of low-rank channel estimation for massive MIMO systems*, IEEE Access **4** (2016), 7313–7321.
7. E. Bjornson, E. G. Larsson, and T. L. Mazetta, *Massive MIMO: Ten myths and one critical question*, IEEE Commun. Mag. **54** (2016), no. 2, 114–123.
8. J. Zhang et al., *Pilot contamination elimination for large-scale multiple-antenna aided OFDM systems*, IEEE J. Sel. Topics Signal Process. **8** (2014), no. 5, 759–772.
9. Y. Nan, L. Zhang, and X. Sun, *Efficient downlink channel estimation scheme based on block-structured compressive sensing for*

TDD massive MU-MIMO systems, IEEE Wireless Commun. Letts. **4** (2015), no. 4, 345–348.
10. Y. Nan, L. Zhang, and X. Sun, An efficient downlink channel estimation approach for TDD massive MIMO systems, in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Nanjing, China, May 15–18, 2016, pp. 1–5.
11. J. Zuo et al., *Multi-cell multiuser massive MIMO transmission with downlink training and pilot contamination pre-coding*, IEEE Trans. Veh. Technol. **65** (2016), no. 8, 6301–6314.
12. L. Shen et al., *Blind decoding based on independent component analiysis for a massive MIMO uplink system in microcell Rician/Rayleigh fading channels*, IEEE Trans. Veh. Technol. **56** (2016), no. 10, 8322–8329.
13. C. Shin, R. W. Heath, and E. J. Powers, *Blind channel estimation for MIMO-OFDM systems*, IEEE Trans. Veh. Technol. **56** (2007), no. 2, 670–685.
14. J.-D. Jiang, T.-C. Lin, and S.-M. Phoong, New subspace-based blind channel estimation for orthogonally coded MIMO-OFDM systems, in *Proc. IEEE ICASSP*, Florence, Italy, May 4–9, 2014, pp. 6489–6493.
15. W. Zhang, F. Gao, and Q. Yin, *Blind channel estimation for MIMO-OFDM Systems with low order signal constellation*, IEEE Commun. Lett. **19** (2015), no. 3, 499–502.
16. H. Q. Ngo and E. G. Larsson, *No downlink pilots are needed in TDD massive MIMO*, IEEE Trans. Wireless Commun. **16** (2017), no. 5, 2921–2935.
17. A. Khansefid and H. Minn, *Achievable Downlink Rates of MRC and ZF Precoders in Massive MIMO with Uplink and Downlink Pilot Contamination*, IEEE Trans. Commun. **63** (2015), no. 12, 4849–4864.
18. H. Q. Ngo, *Massive MIMO: Fundamentals and system designs*, Ph.D. Thesis, Linköping University, Sweden, 2015.
19. H. Yang and T. L. Marzetta, *Performance of conjugate and zero-forcing beamforming in large-scale antenna systems*, IEEE J. Sel. Areas Commun. **31** (2013), no. 2, 172–179.
20. J. Hoydis, S. ten Brink, and M. Debbah, *Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?* IEEE J. Sel. Areas Commun. **31** (2013), no. 2, 160–171.
21. I. Yoffe and D. Wulich, *Predistorter for MIMO system with nonlinear power amplifiers*, IEEE Trans. Commun. **65** (2017), no. 8, 3288–3301.
22. J. Qi and S. Aïssa, *Analysis and compensation of power amplifier nonlinearity in MIMO transmit diversity systems*, IEEE Trans. Veh. Technol. **59** (2010), no. 6, 2921–2931.
23. J. Guerreiro et al., *Analytical performance evaluation of pre-coding techniques for nonlinear massive MIMO Systems with channel estimation errors*, IEEE Trans. Commun. **66** (2017), no. 4, 1440–1451.
24. M. Sánchez-Fernández et al., *SVM multiregression for nonlinear channel estimation in multiple-input multiple-output systems*, IEEE Trans. Sig. Process **52** (2004), no. 8, 2298–2307.
25. E. Bogenfeld et al., Influence on nonlinear HPA on trellis-coded OFDM for terrestrial broadcasting of digital HDTV, in *ICC'93*, Houston, TX, USA, Nov. 23–Dec. 2, 1993, pp. 1433–1438.
26. A. A. M. Saleh, *Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers*, IEEE Trans. Commun. **29** (1981), no. 11, 1715–1720.
27. C. Rapp, Effects of HPA-nonlinearity on 4-DPSK-OFDM-signal for a digital sound broadcasting system, in *Proc. Eur. Conf. Satellite Commun*, Liege, Belgium, Oct. 22–24, 1991, pp. 179–184.
28. G. Santella and F. Mazzenga, *A hybrid analytical-simulation procedure for performance evaluation in M-QAM-OFDM schemes in presence of nonlinear distortions*, IEEE Trans. Veh. Technol. **47** (1998), no. 1, 142–151.
29. H. E. Rowe, *Memoryless nonlinearities with Gaussian inputs: elementary results*, Bell Syst. Tech. J. **61** (1982), no. 7, 1519–1525.
30. H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, *Energy and spectral efficiency of very large multiuser MIMO systems*, IEEE Trans. Commun. **61** (2013), no. 4, 1436–1449.
31. D. Dardari, V. Tralli, and A. Vaccari, *A theoretical characterization of nonlinear distortion effects in OFDM systems*, IEEE Trans. Commun. **48** (2000), no. 10, 1755–1764.

AUTHOR BIOGRAPHIES

Parisa Pasangi received her BS degree in electrical engineering from the College of Engineering, Golestan University, Gorgan, Iran, in 2016. Currently, she is working toward MS degree in communication engineering at Azarbaijan Shahid Madani University, Tabriz, Iran. Her research interests include MIMO system and channel estimation.

Mahmoud Atashbar received his BS degree in electrical engineering from the School of Electrical Engineering, Sahand University of Technology, Tabriz, Iran, in 2003, and his MS and PhD degrees in communication engineering from the school of electrical engineering, Iran University of Science and Technology, Tehran, Iran, in 2006 and 2013, respectively. Since 2013, he has been with the Department of Electrical Engineering, Azarbaijan Shahid Madani University, Tabriz, Iran, where he is now an Assistant Professor. His research interests include wireless communication and statistical signal processing.

Mahmood Mohassel Feghhi received his BS and MS degrees (with Honors) in electrical engineering from the Iran University of Science and Technology, Tehran, Iran, in 2006 and 2009, respectively, and the PhD degree in electrical engineering from the College of Engineering, University of Tehran, Iran, in 2015. From 2007 to 2016, he held positions as a senior design engineer in the areas of communication systems design with several Communications Industries and Inc. Since 2016, he has been with the Faculty of Electrical and Computer Engineering, University of Tabriz, Iran, where he is now an Assistant Professor. His current research interests include network information theory, wireless networks, resource allocation, scheduling, and optimization.