

Pilot Scheduling on Demand: Adaptive Channel Estimation Scheme for TDD Massive MIMO With Diverse Coherence Times

Zhouyi Qian[✉] and Shaowei Wang[✉], *Senior Member, IEEE*

Abstract—In time-division duplex (TDD) massive multiple-input multiple-output (MIMO) systems, users send pilot sequences for channel state estimation in a fixed time interval, which leads to pilot redundancy if users undergo longer coherence time than the prescribed channel estimation interval, and consequently reduces the net sum spectral efficiency of the systems. In this paper, we propose an adaptive channel estimation scheme for TDD massive MIMO systems, where base stations can reuse the aged channel state information (CSI) by exploiting the temporal correlation inherent in channel aging. We introduce a Fisher transformation method for temporal correlation estimation to determine the CSI estimation interval. We also derive the closed-form expressions of the spectral efficiency with channel aging effect and design a threshold to control the channel aging error introduced by CSI reuse. Numerical results demonstrate that our proposed adaptive scheme yields significant performance gains in spectral efficiency across various communication scenarios.

Index Terms—Channel estimation, coherence time, massive MIMO, temporal correlation.

I. INTRODUCTION

MASSIVE multiple-input multiple-output (MIMO) system is deemed as a critical component in 5G and beyond mobile networks to offer significant spectral and energy efficiency [1]. By deploying hundreds or thousands of antennas at base stations (BSs), massive MIMO systems can employ simple linear precoding techniques to serve a large number of users simultaneously [2]. Besides, massive MIMO can mitigate small-scale fading and frequency dependence owing to channel hardening effect [3].

Massive MIMO relies highly on accurate channel state information (CSI) to realize the aforementioned potential gains [4]. In frequency-division-duplex (FDD) massive MIMO systems, the uplink and downlink channels transmit data over different frequency bands, requiring both the BS and users to transmit pilots for channel estimation. In particular, the overhead of downlink channel estimation scales linearly with

the number of antennas, leading to substantial estimation overhead. In contrast, time-division-duplex (TDD) massive MIMO systems solely need to estimate the uplink CSI due to channel reciprocity, and the estimation overhead is proportional to the number of users [5], [6]. Therefore, massive MIMO systems are generally recommended to operate in TDD mode. Although TDD systems significantly reduce estimation overhead as compared to FDD systems, it still suffers an enormous amount of CSI overhead in scenarios with high mobile device density, such as concerts and train station waiting areas. To reduce channel estimation overhead, many prior efforts focus on leveraging the spatial channel correlation to design pilot allocation strategies. In [7], [8], [9], and [10], the BS divides users into different groups according to the channel covariance matrix and allocates a distinct pilot sequence to each group for pilot reuse. Machine learning and compressive sensing are also utilized to design uplink CSI feedback schemes, as discussed in [11] and [12]. While these schemes effectively mitigate channel estimation overhead, the reliance on the assumption of spatial channel correlation, which forms the core of these proposals, requires further validation. Moreover, these proposed schemes do not explore the potential of temporal channel correlation in the downlink transmission.

In practical scenarios, the channels exhibit temporal correlation due to the nearly stationary scattering environment and user mobility [13], which is known as channel aging and has been widely investigated in prior studies. The channel aging process is typically modeled as an autoregressive (AR) process [14], [15], [16], and the aged CSI sequences can be leveraged to refine the noisy channel estimates by employing techniques such as linear extrapolation, Kalman filtering, and machine learning [17], [18], [19], [20]. In [21], the authors analyze the asymptotic achievable rates of TDD massive MIMO systems with channel aging, where the maximum ratio combining receiver is used for uplink detection and the matched filtering precoder is applied for the downlink transmission. The asymptotic bound is extended to scenarios using minimum mean-square error (MMSE) detectors and regularized zero-forcing precoders in [22], and the power scaling law with the aged CSI is also studied in [23]. More recently, the impact of channel aging in cell-free massive MIMO systems with pilot contamination has also been investigated [24]. These studies indicate that the aged CSI remains exploitable, with its utility increasing as temporal correlation strengthens. This

Received 16 September 2025; revised 16 January 2026 and 24 April 2026; accepted 11 July 2026. Date of publication 27 July 2026; date of current version 29 July 2026. An earlier version of this paper was presented in part at the 2025 IEEE International Conference on Communications, Montreal, Canada, June 2025 [DOI: 10.1109/ICC52391.2025.11161419]. The associate editor coordinating the review of this article and approving it for publication was T. Tsiftsis. (Corresponding author: Shaowei Wang.)

The authors are with the School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China (e-mail: zhouyiqian@smail.nju.edu.cn; wangsw@nju.edu.cn).

Digital Object Identifier 10.1109/TWC.2026.3715022

1536-1276 © 2026 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

Authorized licensed use limited to: Nanjing University. Downloaded on October 29, 2026 at 07:16:53 UTC from IEEE Xplore. Restrictions apply.

further suggests that reusing aged CSI leads to limited channel aging effects and reduces CSI acquisition overhead in slowly varying channels. In TDD massive MIMO systems, all users estimate CSI at the same short interval to accommodate extreme communication scenarios with high Doppler shifts [25]. Users in diverse propagation environments, however, have varying channel coherence times, many of which exceed the predefined CSI estimation interval. The mismatch between channel coherence time and the fixed CSI estimation interval leads to significant estimation redundancy in conventional CSI estimation schemes. Therefore, it is promising to develop more efficient CSI estimation schemes to exploit the pilot redundancy inherent in diverse user coherence times by reusing the aged CSI.

A few studies have explored the temporal channel correlation underlying pilot redundancy. In [26], a multiuser massive MIMO model is proposed where users are grouped according to their velocities, and the grouping threshold is optimized to maximize system capacity. The idea is extended to multi-cell scenarios in [27], where users are similarly grouped by velocity and distinct channel estimation intervals are assigned to each group to improve the achievable sum rate. While grouping strategies reduce the complexity of pilot reuse scheduling, a certain performance loss is inevitably incurred. In [28], an upper bound on the spectral efficiency of multiuser MIMO systems is derived as a function of the channel estimation interval, and a heuristic search algorithm is proposed to determine the optimal interval for each user. In [29], the authors introduce intermittent channel estimation schemes in which the BS schedules a fixed number of users to transmit pilot sequences in each slot to maximize the achievable sum rate. Channel prediction is further incorporated in [30], where a convolutional neural network is trained on a channel dataset to extract channel aging information from historical CSI. A pretrained recurrent neural network then predicts future CSI based on historical CSI and the extracted aging information. Though the aforementioned channel estimation schemes achieve substantial performance improvements, they rely on Jakes model [31] and assume that user velocity is known a priori to characterize temporal channel correlation, which limits their applicability in non-uniform scattering environments. Additionally, the real-world channel coherence time depends not only on user mobility but also on the dynamic propagation environment. Therefore, assessing the strength of temporal channel correlation solely according to user mobility is unreasonable, leading to the aforementioned methods not fully exploiting the diversity and redundancy among user coherence times. We try to overcome these shortcomings by estimating the temporal channel correlation in real time. In contrast to [26], [27], and [32], we focus on determining the channel estimation interval for each user rather than scheduling users to estimate channel in each time slot, and adopt a channel model consistent with those considered in [26], [27], [29], and [33]. Another concern is that parameter estimation algorithms relying on point estimates may yield overconfident temporal correlation assessments due to imbalanced data distributions arising from the limited amount of training data available in real-time estimation

scenarios. The dynamic propagation environment and unpredictable user mobility also intensify uncertainty in the data distribution.

In this paper, we aim to design an adaptive channel estimation scheme that exploits the temporal correlation of user channels to reduce CSI estimation overhead for TDD massive MIMO systems. The proposed scheme comprises two parts: temporal channel estimation and estimation interval allocation. For the former, we employ the Fisher transformation to mitigate data uncertainties in real-time channel correlation estimation [34]. For the estimation interval allocation, a novel method is proposed to determine the CSI estimation interval according to the diverse temporal correlations of users and a predefined threshold, achieving a balance between estimation overhead and channel aging. Specifically, our main contributions are summarized as follows:

- We propose a Fisher transformation temporal correlation estimation method so as to mitigate the uncertainties of data distribution, which can effectively control the reduction of system performance caused by the overestimation of temporal correlation as compared to common machine learning based estimation algorithms.
- We theoretically analyze the achievable rate of our proposed scheme with zero-forcing (ZF) precoding in the downlink transmission, where the error introduced by channel aging is independent of the channel estimation error. The achievable rate serves as a guiding principle to set a threshold for controlling the channel aging error caused by CSI reuse.
- We conduct extensive experiments to evaluate the performance of the proposed adaptive CSI estimation scheme, whose results indicate that the adaptive channel estimation scheme significantly improves the net spectral efficiency as compared to the conventional scheme and achieves a good balance between the channel aging error and CSI overhead reduction.

The remainder of this paper is structured as follows. Section II describes the system model under consideration. Section III elaborates on the advantages of the adaptive channel estimation scheme and derives the lower bound for the net sum spectral efficiency of the adaptive scheme. In Section IV, we illustrate the main steps of our scheme. Finally, Sections V and VI present the numerical results and conclusions, respectively.

Notations: Throughout the paper, we use boldface uppercase letters, boldface lowercase letters and lowercase letters to denote matrices, column vectors and scalars, respectively. $\mathbf{X}^T$, $\mathbf{X}^*$, $\mathbf{X}^H$, $\mathbf{X}^{-1}$, $\text{tr}(\mathbf{X})$, $|\mathbf{X}|$, and $\|\mathbf{X}\|$ correspond to the transpose, complex conjugate, complex conjugate transpose, inverse, trace, modulus, two-norm of $\mathbf{X}$, respectively. The notation $\mathbb{E}\{\cdot\}$ denotes expectation operation.

II. SYSTEM MODEL

A. Channel Model

Consider a single-cell TDD massive MIMO system where a BS equipped with M antennas serves K single-antenna users. The transmission process is organized into discrete time slots,

each comprising N_c symbols, with an uplink channel training phase of τ_{ul} symbols and a downlink data transmission phase of $N_c - \tau_{ul}$ symbols. The channel vector $\mathbf{g}_k[n] \in \mathbb{C}^{M \times 1}$ between the BS and the k th user in the n th time slot is modelled as

$$\mathbf{g}_k[n] = \sqrt{\beta_k} \mathbf{h}_k[n], \quad (1)$$

where β_k represents the large-scale fading between the BS and the k th user, and $\mathbf{h}_k[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ is a small-scale fading vector.¹ To consider the temporal correlation of the channel, we adopt a quasi-static block fading channel model where $\mathbf{g}_k[n]$ stays constant in a slot but varies across slots [21]. Given the l -advanced channel vector $\mathbf{g}_k[n-l]$, $\mathbf{g}_k[n]$ can be expressed as

$$\mathbf{g}_k[n] = \rho_{k,l} \mathbf{g}_k[n-l] + \boldsymbol{\varepsilon}_{k,l}[n], \quad (2)$$

where $\boldsymbol{\varepsilon}_{k,l}[n] \sim \mathcal{CN}(\mathbf{0}, \beta_k(1 - \rho_{k,l}^2) \mathbf{I}_M)$ represents channel aging noise independent of $\mathbf{g}_k[n-l]$, and $\rho_{k,l}$ is the temporal correlation coefficient of the channel. β_k and $\rho_{k,l}$ are invariant in a stationary block, whose length is N_s time slots.

B. Uplink Training

In the uplink training phase, all users synchronously send pilot sequences $\boldsymbol{\psi}_k \in \mathbb{C}^{\tau_{ul} \times 1}$ to the BS, and the pilot matrix can be denoted by $\boldsymbol{\Psi} \triangleq [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \dots, \boldsymbol{\psi}_K, \dots, \boldsymbol{\psi}_K]$. Since the pilot sequences of K users are pairwise orthogonal, i.e., $\boldsymbol{\Psi}^H \boldsymbol{\Psi} = \mathbf{I}_K$, $\tau_{ul} \geq K$ must hold. Then the received signal matrix $\mathbf{Y}_p[n]$ at the BS in the n th slot can be expressed as

$$\mathbf{Y}_p[n] = \sqrt{\tau_{ul} p_{ul}} \mathbf{G}[n] \boldsymbol{\Psi}^H + \mathbf{W}_p[n], \quad (3)$$

where $\mathbf{G}[n] \triangleq [\mathbf{g}_1[n], \dots, \mathbf{g}_K[n], \dots, \mathbf{g}_K[n]]$, $\tau_{ul} p_{ul}$ denotes the uplink average transmit power and $\mathbf{W}_p[n] \in \mathbb{C}^{M \times \tau_{ul}}$ is the additive noise matrix whose elements are i.i.d. $\mathcal{CN}(0, 1)$. After receiving the signal $\mathbf{Y}_p[n]$, the BS can acquire the linear MMSE estimate of $\mathbf{g}_k[n]$ as

$$\hat{\mathbf{g}}_k[n] = \frac{\tau_{ul} p_{ul} \beta_k}{\tau_{ul} p_{ul} \beta_k + 1} \mathbf{g}_k[n] + \frac{\sqrt{\tau_{ul} p_{ul}} \beta_k}{\tau_{ul} p_{ul} \beta_k + 1} \mathbf{w}_{p,k}[n], \quad (4)$$

where $\mathbf{w}_{p,k}[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ is the noise vector independent of $\mathbf{g}_k[n]$. The average power of the m th component of the channel estimation vector $\hat{\mathbf{g}}_k[n]$ is

$$\gamma_k \triangleq \mathbb{E}\{|\hat{g}_k^m[n]|^2\} = \frac{\tau_{ul} p_{ul} \beta_k^2}{1 + \tau_{ul} p_{ul} \beta_k}. \quad (5)$$

Let $\mathbf{e}_k[n] = \hat{\mathbf{g}}_k[n] - \mathbf{g}_k[n]$ denote the channel estimation error, with $e_k^m[n]$ representing the m th component of the vector $\mathbf{e}_k[n]$. Due to the orthogonality of the MMSE estimator, $e_k^m[n]$ and $\hat{g}_k^m[n]$ are uncorrelated, and the variance of the channel estimation error can be calculated as

$$\mathbb{E}\{|e_k^m[n]|^2\} = \frac{\beta_k}{1 + \tau_{ul} p_{ul} \beta_k} = \beta_k - \gamma_k \geq 0. \quad (6)$$

¹ Although practical massive MIMO channels exhibit spatial correlation due to the limited number of scatterers, we adopt the spatially i.i.d. Rayleigh fading model, as utilized in [29]. This modeling choice isolates the channel aging effect from the coupling effects introduced by spatial correlation, enabling a clear assessment of its impact on system performance.

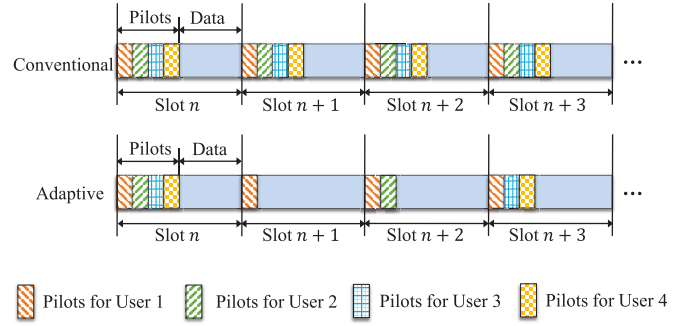


Fig. 1. An example of our proposed channel estimation scheme.

C. Downlink Data Transmission

Let s_k be the symbol the BS transmits to user k during the downlink transmission phase. The transmitted symbol vector is denoted as $\mathbf{s} \triangleq [s_1, \dots, s_K]^T$, which satisfies $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}_K$. The transmitted signal after linear precoding is given by

$$\mathbf{x}[n] = \mathbf{A}[n] \mathbf{D}_\eta^{1/2} \mathbf{s}, \quad (7)$$

where $\mathbf{A}[n] \in \mathbb{C}^{M \times K}$ is the precoding matrix [25], which is a function of CSI matrix $\hat{\mathbf{G}}[n] \triangleq [\hat{\mathbf{g}}_1[n], \dots, \hat{\mathbf{g}}_K[n]]$. $\mathbf{D}_\eta \triangleq \text{diag}\{\eta_1, \eta_2, \dots, \eta_K\}$ is a power scaling matrix satisfying the power normalization $\mathbb{E}\{\|\mathbf{x}\|^2\} = 1$, where η_k is the power scaling parameter corresponding to the signal for user k . We employ the ZF precoding to cancel the inter-user interference and the precoding matrix is given by

$$\mathbf{A}[n] = \sqrt{M - K} \hat{\mathbf{G}}^*[n] (\hat{\mathbf{G}}^T[n] \hat{\mathbf{G}}^*[n])^{-1} \mathbf{D}_\gamma^{1/2}, \quad (8)$$

where $\mathbf{D}_\gamma \triangleq \text{diag}\{\gamma_1, \gamma_2, \dots, \gamma_K\}$. After the BS transmits the precoded signal $\mathbf{x}$, the received signal vector which denoted as $\mathbf{y}$ is given by

$$\mathbf{y}[n] = \sqrt{p_{dl}} \mathbf{G}^T[n] \mathbf{x}[n] + \mathbf{w}[n], \quad (9)$$

where $\mathbf{y}[n] \triangleq [y_1[n], \dots, y_K[n]]^T$, p_{dl} is the average downlink transmit power, and $\mathbf{w}[n]$ is the additive noise vector, with each element $w_k[n] \sim \mathcal{CN}(0, 1)$ representing the noise at the k th user. Let $\mathbf{a}_k[n]$ be the k th column of $\mathbf{A}[n]$. Accordingly, the k th user observes

$$y_k[n] = \sqrt{p_{dl} \eta_k} \mathbf{g}_k^T[n] \mathbf{a}_k[n] s_k + w_k[n] + \sum_{j \neq k} \sqrt{p_{dl} \eta_j} \mathbf{g}_k^T[n] \mathbf{a}_j[n] s_j. \quad (10)$$

III. ADAPTIVE CHANNEL ESTIMATION SCHEME FOR TDD MASSIVE MIMO SYSTEMS

For the uplink training in the conventional TDD massive MIMO systems, a fixed CSI estimation interval T_s is applied to all users, where T_s represents the duration of a time slot. Each user needs to send an orthogonal pilot sequence of length τ_{ul} in each time slot, which results in pilot redundancy since the channel coherence time of many users exceeds T_s . Due to the impact of channel aging, the BS can reuse the aged CSI to precode or decode data in the current time slot when the channel variation is minimal, and the resources freed by reusing the

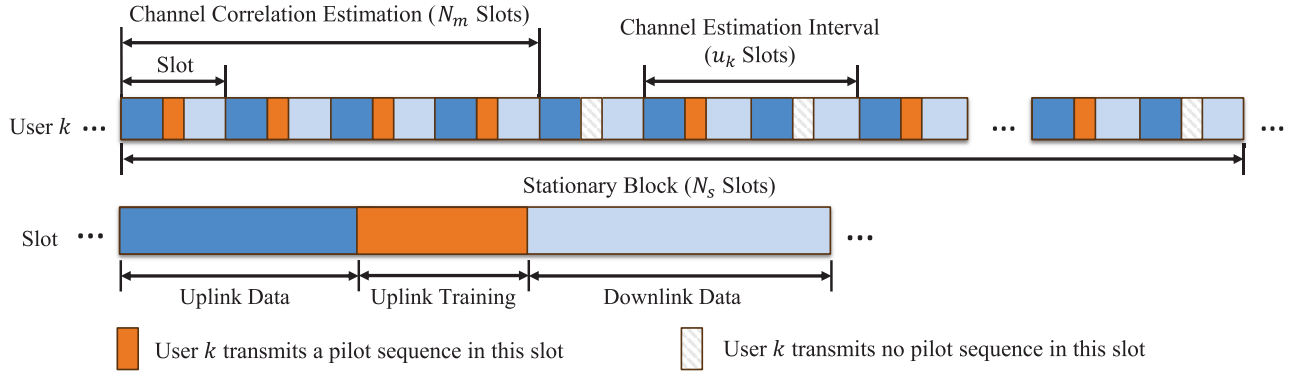


Fig. 2. The proposed adaptive channel estimation scheme.

aged CSI can be allocated to data transmission. Building on this idea, we propose an adaptive channel estimation scheme and derive the formula for the net sum spectral efficiency of the proposed scheme. Recall that the temporal channel correlation for users remains constant in each stationary block. To facilitate scheme design, the duration of a stationary block is used as the update period for channel estimation intervals. The channel estimation intervals for each user in a stationary block are represented as $u_1 T_s, \dots, u_k T_s, \dots, u_K T_s$, where $u_k \in \mathbb{N}^+$. The BS estimates the channel in a conventional scheme when $u_k = 1$. For users with $u_k > 1$, the BS reuses the latest aged CSI for precoding and data transmission. We present an example in Fig. 1 to illustrate the proposed scheme, where the BS serves four users. In the conventional channel estimation scheme, $\{u_k\}_{1 \leq k \leq 4}$ are set to 1 and users estimate CSI in each time slot, while in the proposed adaptive channel estimation scheme, $\{u_k\}_{1 \leq k \leq 4}$ are assigned as 1, 2, 3, 3 in the stationary block, respectively. It is observed from this example that allowing users to send pilot signals at their respective frequencies saves resources for channel estimation and further improves the payload transmission rates as compared to the conventional scheme.

In contrast to conventional schemes where the BS estimates the channel in each time slot, the BS reuses the latest channel estimation vector $\hat{\mathbf{g}}_k[n - u_k]$ to precode data for k th user in adaptive channel estimation scheme. After the BS sends the precoded signal $\mathbf{x}[n]$, the received signal at the k th user is

$$y_k[n] = \sqrt{p_{\text{dl}} \eta_k} \rho_{k, u_k}[n] \hat{\mathbf{g}}_k[n - u_k]^T \mathbf{a}_k[n] s_k + w_k[n] - \sum_{j=1}^K \sqrt{p_{\text{dl}} \eta_j} (\rho_{k, u_k}[n] \hat{\mathbf{g}}_k[n - u_k] - \mathbf{g}_k[n])^T \mathbf{a}_j[n] s_j. \quad (11)$$

Let $e_{k, u_k}^m[n]$ be the m th component of the vector $\mathbf{e}_{k, u_k}[n]$ where $\mathbf{e}_{k, u_k}[n] \triangleq \rho_{k, u_k}[n] \hat{\mathbf{g}}_k[n - u_k] - \mathbf{g}_k[n]$ represents errors caused by channel estimation and channel aging. The signal received by the k th user can be rewritten as

$$y_k[n] = \sqrt{(M - K) p_{\text{dl}} \gamma_k \eta_k \rho_{k, u_k}[n]} s_k + \underbrace{w_k[n] - \sqrt{p_{\text{dl}}} \mathbf{e}_{k, u_k}^T[n] \mathbf{x}[n]}_{\text{effective noise}}, \quad (12)$$

where the three terms are mutually uncorrelated and we treat the last two terms as effective noise. Therefore, the effective signal-to-interference-plus-noise ratio (SINR) for user k in the adaptive scheme can be expressed as, respectively,

$$\text{SINR}_k[n] = \frac{(M - K) p_{\text{dl}} \gamma_k \eta_k \rho_{k, u_k}[n]^2}{1 + p_{\text{dl}} \mathbb{E} \left\{ |e_{k, u_k}^m[n]|^2 \right\} \sum_{k=1}^K \eta_k}. \quad (13)$$

The achievable rate for user k can be calculated as

$$C_k[n] = \log_2 (1 + \text{SINR}_k[n]). \quad (14)$$

Let $\tau_{\text{ul}}[n]$ be the number of symbols for uplink training in the n th slot of a stationary block, we obtain the net sum spectral efficiency of our proposed scheme as

$$C_{\text{net}}[n] = \frac{N_c - \tau_{\text{ul}}[n]}{N_c} \sum_{k=1}^K C_k[n], \quad (15)$$

and the average net sum spectral efficiency in each stationary block is

$$\bar{C}_{\text{net}} = \frac{1}{N_s} \sum_{n=1}^{N_s} C_{\text{net}}[n] = \frac{1}{N_s} \sum_{n=1}^{N_s} \frac{N_c - \tau_{\text{ul}}[n]}{N_c} \sum_{k=1}^K C_k[n]. \quad (16)$$

Note that the channel coherence time is prolonged with strong temporal channel correlations in each stationary block, which enables a longer u_k to increase the net sum spectral efficiency. Conversely, when the temporal channel correlations are weak, long u_k can lead to more CSI aging errors, thereby decreasing system performance. Hence it is natural to think of setting $\{u_k\}_{1 \leq k \leq K}$ based on the temporal channel correlation in each stationary block.

IV. ADAPTIVE PILOT INTERVAL VIA HIGH-CONFIDENCE CHANNEL ESTIMATION

Due to the absence of closed-form expressions for channel autocorrelation function in outdoor environments or scenarios with non-uniformly distributed scatters, we propose a Fisher transformation method for estimating temporal correlation in real time. We also derive the lower bound on the asymptotic achievable rate in an ideal propagation environment where p_{ul} and p_{dl} go to infinity and set a threshold according to the lower bound to assess the strength of the temporal correlation.

The process of the adaptive channel estimation is illustrated in Fig. 2, where N_s is the length of a stationary block, and the time window used for correlation coefficient estimation is N_m time slots. The BS first collects CSI samples to form a sample set in the initial N_m time slots of each block. We then apply the Fisher transformation to the CSI sample set to obtain high-confidence interval estimates of temporal channel correlation, and adopt the lower bound as a conservative indicator. The BS determines $\{u_k\}_{1 \leq k \leq K}$ for all users according to the correlation indicator and the threshold, and reuses the latest CSI in time slots when the users send no pilots.

A. Fisher Transmission-Based Correlation Estimation

During the N_m time slots at the beginning of each stationary block, the BS can obtain the channel estimate vector $\hat{\mathbf{g}}_k[n] = [\hat{g}_k^1[n], \dots, \hat{g}_k^m[n], \dots, \hat{g}_k^M[n]]$ for the k th user in each time slot, where $\hat{g}_k^m[n]$ is the CSI of the m th antenna. The temporal channel correlation is uniform across antennas and consistent between the in-phase and quadrature components. Consequently, estimating the temporal correlation of the in-phase component is sufficient to characterize the overall channel correlation. We denote the in-phase component vector of $\hat{\mathbf{g}}_k[n]$ as $\hat{\mathbf{g}}_{I,k}[n] = [\hat{g}_{I,k}^1[n], \dots, \hat{g}_{I,k}^m[n], \dots, \hat{g}_{I,k}^M[n]]$, and the BS can collect the CSI sample set $\{\hat{\mathbf{g}}_{I,k}[n]\}_{1 \leq n \leq N_m}$. Due to channel noise, we can only estimate the temporal correlation between CSIs. Let $\hat{\rho}_{k,l}$ be the temporal correlation coefficient between CSIs separated by l time slots. The relationship between the channel estimates $\hat{\mathbf{g}}_k[n]$ and $\hat{\mathbf{g}}_k[n-l]$ separated by l time slots can be expressed as

$$\hat{\mathbf{g}}_k[n] = \hat{\rho}_{k,l} \hat{\mathbf{g}}_k[n-l] + \hat{\mathbf{e}}_{k,l}[n], \quad (17)$$

where $\hat{\mathbf{e}}_{k,l}[n] \sim \mathcal{CN}(\mathbf{0}, \gamma_k(1 - \hat{\rho}_{k,l}^2)\mathbf{I}_M)$ represents noise independent of $\hat{\mathbf{g}}_k[n-l]$, and $\hat{\rho}_{k,l}$ is the temporal correlation coefficient between $\hat{\mathbf{g}}_k[n]$ and $\hat{\mathbf{g}}_k[n-l]$. Once the CSI samples are collected, $\hat{\rho}_{k,l}$ can be estimated using the sample correlation coefficient $r_{k,l}$, which is calculated as

$$r_{k,l} = \frac{\sum_{m=1}^M (\hat{g}_{I,k}^m[n-l] - \bar{\hat{g}}_{n-l}) (\hat{g}_{I,k}^m[n] - \bar{\hat{g}}_n)}{\sqrt{\sum_{m=1}^M (\hat{g}_{I,k}^m[n-l] - \bar{\hat{g}}_{n-l})^2 \sum_{m=1}^M (\hat{g}_{I,k}^m[n] - \bar{\hat{g}}_n)^2}}, \quad (18)$$

where $\bar{\hat{g}}_{n-l} = \frac{1}{M} \sum_{m=1}^M \hat{g}_{I,k}^m[n-l]$ and $\bar{\hat{g}}_n = \frac{1}{M} \sum_{m=1}^M \hat{g}_{I,k}^m[n]$. Subsequently, we obtain the sample correlation coefficient vector $\mathbf{r}_k = [r_{k,1}, \dots, r_{k,l}, \dots, r_{k,N_m-1}]$.

Sample correlation coefficients cannot be directly used as correlation indicators, as they almost inevitably yield biased estimates of the true temporal channel correlation when the data distribution is imbalanced. In particular, overestimation can lead the BS to allocate a channel estimation interval that exceeds the actual coherence time. The mismatch weakens the temporal correlation between aged and current CSIs, in which case the reuse of the outdated CSI can no longer provide effective spectral efficiency, potentially leading to worse performance than the conventional scheme. To mitigate the overestimation caused by data uncertainties, we employ Fisher transformation to $\mathbf{r}_k$ and use the lower confidence bound of $\mathbf{r}_k$ as the temporal correlation indicator. Let the joint

confidence level for $\hat{\rho}_k \triangleq [\hat{\rho}_{k,1}, \dots, \hat{\rho}_{k,l}, \dots, \hat{\rho}_{k,N_m-1}]$ be $1 - \alpha$, the confidence level for $\hat{\rho}_{k,l}$ is $1 - \zeta = 1 - \frac{\alpha}{N_m-1}$ by the Bonferroni correction. The Fisher transformation of $r_{k,l}$ is

$$z_{k,l} = \tanh^{-1} r_{k,l} = \frac{1}{2} \ln \left(\frac{1 + r_{k,l}}{1 - r_{k,l}} \right), \quad (19)$$

where $z_{k,l} \sim \mathcal{N}\left(\frac{1}{2} \ln \left(\frac{1 + \hat{\rho}_{k,l}}{1 - \hat{\rho}_{k,l}} \right), \frac{1}{M-3}\right)$. Hence the $1 - \zeta$ confidence interval for $\frac{1}{2} \ln \left(\frac{1 + \hat{\rho}_{k,l}}{1 - \hat{\rho}_{k,l}} \right)$ can be expressed as

$$\left(L_{k,l}^z, U_{k,l}^z \right) = \left(z_{k,l} - \frac{u_{\zeta/2}}{\sqrt{M-3}}, z_{k,l} + \frac{u_{\zeta/2}}{\sqrt{M-3}} \right), \quad (20)$$

where $u_{\zeta/2}$ represents the upper $\zeta/2$ quantile of the standard normal distribution. We apply the hyperbolic tangent transformation to $(L_{k,l}^z, U_{k,l}^z)$, and then the $1 - \zeta$ confidence interval for the channel correlation coefficient $\hat{\rho}_{k,l}$ is calculated as

$$\begin{aligned} (L_{k,l}, U_{k,l}) &= (\tanh L_{k,l}^z, \tanh U_{k,l}^z) \\ &= \left(\frac{e^{2L_{k,l}^z} - 1}{e^{2L_{k,l}^z} + 1}, \frac{e^{2U_{k,l}^z} - 1}{e^{2U_{k,l}^z} + 1} \right). \end{aligned} \quad (21)$$

Define the joint confidence interval as $A = \bigcap_{l=1}^{N_m-1} \{L_{k,l} \leq \hat{\rho}_{k,l} \leq U_{k,l}\}$. According to the Bonferroni inequality, we can infer that

$$\begin{aligned} \Pr(A) &= \Pr\left(\bigcap_{l=1}^{N_m-1} L_{k,l} \leq \hat{\rho}_{k,l} \leq U_{k,l}\right) \\ &\geq \sum_{l=1}^{N_m-1} \Pr(L_{k,l} \leq \hat{\rho}_{k,l} \leq U_{k,l}) - (N_m - 2) \\ &= 1 - \frac{\alpha}{N_m - 1} (N_m - 1) = 1 - \alpha. \end{aligned} \quad (22)$$

Therefore, A serves as a joint confidence interval for $\hat{\rho}_k$ with a confidence level of at least $1 - \alpha$.

B. System Performance Analysis

To assess the strength of temporal channel correlation, we use $L_{k,l}$ as the temporal correlation indicator and define b as the threshold. When $L_{k,l} \geq b$, the temporal correlation between $\hat{\mathbf{g}}_k[n+l]$ and $\hat{\mathbf{g}}_k[n]$ is strong, indicating that the coherence time is longer than l slots. Conversely, the temporal channel correlation is weak and the coherence time is less than l slots between channels when $L_{k,l} \leq b$. So the coherence time can be estimated by sequentially comparing $L_{k,l}$ against b and we have

$$u_k = \max\{l \mid L_{k,l} \geq b \geq L_{k,l+1}, 0 \leq l \leq N_m - 1, l \in \mathbb{Z}\}. \quad (23)$$

We know that reusing CSI inevitably introduces channel aging errors, with the extent of aging determined by threshold b . To ensure the channel aging error does not outweigh the system performance gains from CSI reuse, we analyze the system performance and set the threshold accordingly. Thus, we analyze the lower bound on the asymptotic achievable rate as p_{ul} and p_{dl} approach infinity. Before derivation, we suggest the following lemmas.

Lemma 1: The variance of the m th component of the vector $\mathbf{e}_{k,u_k}[n]$ is given by

$$\mathbb{E}\{|e_{k,u_k}^m[n]|^2\} = \beta_k(1 - \rho_{k,u_k}[n]^2) + \rho_{k,u_k}[n]^2 \mathbb{E}\{|e_k^m[n]|^2\}. \quad (24)$$

Proof: The proof is presented in Appendix A. ■

Lemma 2: Let $\hat{\rho}_{k,u_k}[n]$ be the temporal correlation between $\hat{\mathbf{g}}_k[n]$ and $\hat{\mathbf{g}}_k[n - u_k]$, then we have

$$\rho_{k,u_k}[n] = \frac{\hat{\rho}_{k,u_k}[n]\beta_k}{\gamma_k} \geq \hat{\rho}_{k,u_k}[n]. \quad (25)$$

Proof: Recall that $\hat{\mathbf{g}}_k^m[n]$ and $\hat{\mathbf{g}}_k^m[n - u_k]$ are Gaussian random variables with zero mean and variance γ_k , we can calculate $\hat{\rho}_{k,u_k}[n]$ as

$$\begin{aligned} & \hat{\rho}_{k,u_k}[n] \\ & \stackrel{(a)}{=} \frac{\mathbb{E}\{\hat{\mathbf{g}}_k^m[n]\} \mathbb{E}\{\hat{\mathbf{g}}_k^m[n - u_k]\}^H}{(\mathbb{E}\{|\hat{\mathbf{g}}_k^m[n]|^2\} \mathbb{E}\{|\hat{\mathbf{g}}_k^m[n - u_k]|^2\})^{\frac{1}{2}}} \\ & \stackrel{(b)}{=} \frac{\rho_{k,u_k}[n] \left(\frac{\tau_{ul} p_{ul} \beta_k}{\tau_{ul} p_{ul} \beta_k + 1} \right)^2 \mathbb{E}\{|g_k^m[n]|^2\}}{(\mathbb{E}\{|\hat{\mathbf{g}}_k^m[n]|^2\} \mathbb{E}\{|\hat{\mathbf{g}}_k^m[n - u_k]|^2\})^{\frac{1}{2}}} \\ & \stackrel{(c)}{=} \rho_{k,u_k}[n] \left(\frac{\tau_{ul} p_{ul} \beta_k}{\tau_{ul} p_{ul} \beta_k + 1} \right)^2 \frac{1 + \tau_{ul} p_{ul} \beta_k}{\tau_{ul} p_{ul} \beta_k^2} \beta_k = \frac{\hat{\rho}_{k,u_k}[n] \gamma_k}{\beta_k}. \end{aligned} \quad (26)$$

In (a), we apply the definition of the correlation coefficient [34]. In (b) we utilize the relation $\mathbf{g}_k[n] = \rho_{k,u_k}[n]\mathbf{g}_k[n - u_k] + \mathbf{e}_{k,u_k}[n]$, and in (c) we substitute (5) and $\mathbb{E}\{|g_k^m[n]|^2\} = \gamma_k$ into (b). The proof is completed by combining (26) with (6). ■

Then we have the following theorem which gives the lower bound on the effective SINR of our proposed scheme.

Theorem 1: The lower bound on effective SINR for user k in the adaptive scheme at a confidence level of $1 - \alpha$ is given by

$$\text{SINR}_k \geq \frac{(M - K)\gamma_k \eta_k}{\frac{1}{b^2 p_{dl}} + (\beta_k (\frac{1}{b^2} - 1) + \beta_k - \gamma_k) \sum_{k=1}^K \eta_k}, \quad (27)$$

where $\frac{1}{b^2 p_{dl}}$ is the error caused by downlink transmission noise, $\beta_k (\frac{1}{b^2} - 1)$ represents the channel aging error, and $\beta_k - \gamma_k$ is the error introduced by channel estimation.

Proof: By Substituting (6), Lemma 1 into (13), the effective SINR for user k can be expressed as

$$\text{SINR}_k = \frac{(M - K)\gamma_k \eta_k}{\frac{1}{\rho_{k,u_k}[n]^2 p_{dl}} + \left(\beta_k \left(\frac{1}{\rho_{k,u_k}[n]^2} - 1 \right) + \beta_k - \gamma_k \right) \sum_{k=1}^K \eta_k}. \quad (28)$$

According to (23) and Lemma 2, we have

$$\rho_{k,u_k}[n] = \frac{\hat{\rho}_{k,u_k}[n]\beta_k}{\gamma_k} \geq \hat{\rho}_{k,u_k}[n] \geq b. \quad (29)$$

The effective SINR increases with respect to $\rho_{k,u_k}[n]$, thus we can derive Theorem 1 by substituting (29) into (28). ■

Theorem 1 indicates that the channel aging error, determined by b , is independent of both uplink training signal-to-noise ratio (SNR) and downlink transmission SNR. Thus, channel aging error persists in an ideal propagation environment, requiring b to be appropriately set to control the channel aging effect. To figure out the effect of channel aging only, we derive

the achievable rate with ideal channel estimation and noise-free downlink transmission.

Corollary 1: We assume ideal uplink training and downlink transmission, i.e., $p_{ul} \rightarrow \infty$, $p_{dl} \rightarrow \infty$. The lower bound on the effective SINR for user k with equal power allocation at a confidence level of $1 - \alpha$ is expressed as

$$\lim_{p_{ul}, p_{dl} \rightarrow \infty} \text{SINR}_k \geq \frac{\frac{M}{K} - 1}{\frac{1}{b^2} - 1}, \quad (30)$$

and the downlink achievable rate is

$$\lim_{p_{ul}, p_{dl} \rightarrow \infty} C_k \geq \log_2 \left(1 + \frac{\frac{M}{K} - 1}{\frac{1}{b^2} - 1} \right). \quad (31)$$

Proof: We can directly derive the Corollary 1 from Theorem 1 when $p_{ul} \rightarrow \infty$ and $p_{dl} \rightarrow \infty$. ■

Then we can apply the asymptotic downlink achievable rate as the metric for threshold design, and the threshold is set to ensure that the channel aging error is negligible as compared to the performance gains from reducing redundant estimation. Specifically, we assume that all users can reuse the aged CSI, i.e., $\tau_{ul}[n] = 0$. From (15), the asymptotic net sum spectral efficiency is the sum of the asymptotic achievable rate in this case. We then set the asymptotic net sum spectral efficiency, which accounts only for aging errors, to be equivalent to that of the conventional channel estimation scheme in a propagation environment with target uplink and downlink SNRs, so that the impact of channel aging errors in the current propagation environment can be negligible.

Corollary 2: Consider an equal power allocation scheme, where $\zeta_{ul} \triangleq K p_{ul} \beta_k$ represents the SNR during the uplink training phase, and $\zeta_{dl} \triangleq \frac{p_{dl} \beta_k}{K}$ denotes the SNR for downlink transmission. Let ϵ_{ul} and ϵ_{dl} represent target uplink and downlink SNRs, and we assume that all training resources can be used for data transmission, the threshold b can be calculated as

$$b = \left(\frac{\frac{M}{K} - 1}{\left(1 + \frac{(M-K)\epsilon_{ul}\epsilon_{dl}}{1 + \epsilon_{ul} + K\epsilon_{dl}} \right)^{1 - \frac{K}{N_s}} - 1} + 1 \right)^{-\frac{1}{2}}, \quad (32)$$

such that the asymptotic net sum spectral efficiency of our proposed adaptive scheme is equivalent to that of the conventional channel estimation scheme in the target propagation environment.

Proof: The proof is presented in Appendix B. ■

From Corollary 2 we can see that b depends on the number of antennas and users. Therefore, the threshold will be adjusted in real time according to the number of users in the system, effectively balancing the gain from reusing CSI and the channel aging error. We summarize the main steps for the adaptive channel estimation scheme in each stationary block in Algorithm 1. Specifically, we first estimate the temporal channel correlation and compute the confidence lower bound according to the Fisher transformation. The channel estimation interval is then determined based on the derived threshold and the confidence lower bound on temporal channel correlation for the current stationary block, followed by the adaptive channel estimation scheme described in Section III. Considering the

Algorithm 1 Algorithm for Setting Adaptive Channel Estimation Intervals in a Stationary Block

Input: The number of users K , the number of antennas M , the target uplink training SNR ϵ_{ul} , the target downlink transmission SNR ϵ_{dl} , stationary block length N_s and time slots for temporal correlation estimation N_m .

Output: Adaptive channel estimation intervals $\{u_k\}_{1 \leq k \leq K}$.

- 1: Initialize the corresponding slot counter $d_k = 0$ for each user k .
- 2: Collect the in-phase component of channel $\hat{\mathbf{g}}_{I,k}[n] = [\hat{g}_{I,k}^1[n], \dots, \hat{g}_{I,k}^M[n]]$ in the initial N_m time slots of a stationary block as sample set $\{\hat{\mathbf{g}}_{I,k}[n]\}_{1 \leq n \leq N_m}$.
- 3: Calculate the sample correlation vector $\mathbf{r}_k$ according to (18) and employ Fisher transformation to $\mathbf{r}_k$ according to (19), (20) and (21) to get $L_k \triangleq [L_{k,1}, \dots, L_{k,N_m-1}]$.
- 4: Compute the threshold b for judging the strength of temporal channel correlation by substituting K and other required parameters into (32) and set $\{u_k\}_{1 \leq k \leq K}$ according to (23).
- 5: **for** $n \in \{N_m + 1, \dots, N_s\}$ **do**
- 6: **for** $k \in \{1, \dots, K\}$ **do**
- 7: $d_k = d_k + 1$.
- 8: **if** $d_k \bmod u_k == 0$ **then**
- 9: Let $d_k = 0$, and user k should send pilot in the current time slot.
- 10: **end if**
- 11: User k reuses the latest CSI in the current time slot.
- 12: **end for**
- 13: **end for**

confidence interval of $\hat{\rho}_k$, we can give the lower bound on the asymptotic average achievable rate for user k with threshold b as

$$\lim_{p_{ul}, p_{dl} \rightarrow \infty} \bar{C}_k \geq (1 - \alpha) \log_2 \left(1 + \frac{(M - K)\epsilon_{ul}\epsilon_{dl}}{1 + \epsilon_{ul} + K\epsilon_{dl}} \right). \quad (33)$$

V. NUMERICAL RESULTS

We consider a TDD massive MIMO system with $M = 128$, and the simulation parameters follow the LTE standard [35]. The OFDM symbol delay spread is $5 \mu s$, so the coherent bandwidth is 100 kHz which contains $N_b = 7$ subcarriers. Each slot in TDD LTE standard lasts $T_s = 1$ ms and consists of $N_t = 14$ symbols, therefore, $N_c = N_t N_b = 98$. The CSI samples are generated using the COST 2100 outdoor channel model, which captures the non-stationary characteristics of radio channels and closely matches realistic massive MIMO channel measurements [36]. The inherent spatial correlation of the COST 2100 model is fully preserved, enabling performance evaluation of the proposed scheme under realistic propagation environments. Using the COST 2100 model, we simulate an outdoor channel environment with a central frequency of 2.6 GHz and a bandwidth of 20 MHz. Mobile users are uniformly located in a circle of radius 500 m with the BS located at the center of the circle. Unless otherwise specified, user speeds are uniformly drawn from [1, 40] km/h,

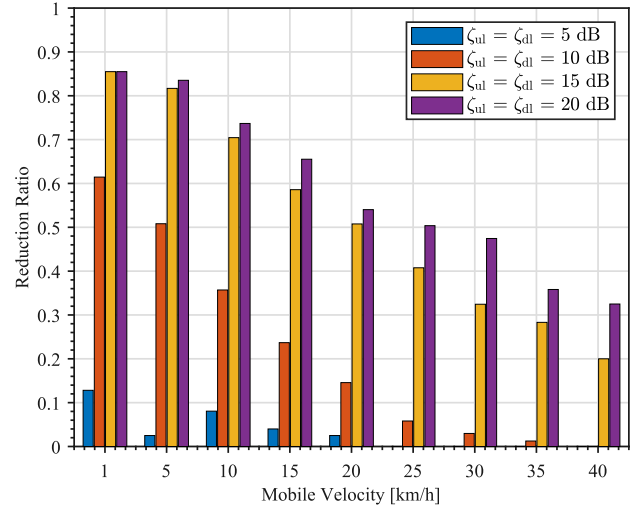


Fig. 3. Reduction ratio of channel estimation versus different user velocities. $K = 60$, $\zeta_{ul} = \zeta_{dl} = 5$ dB, 10 dB, 15 dB, 20 dB, respectively.

and we set $K = 50$, $N_s = 100$, $N_m = 10$, $\zeta_{ul} = \zeta_{dl}$, $\epsilon_{ul} = \epsilon_{dl} = 15$ dB and $\alpha = 0.1$. The number of stationary blocks in the simulation is set to 500. For the performance metrics, we employ the normalized mean square error (NMSE) to evaluate the error caused by channel estimation and channel aging, which is given by

$$\text{NMSE} = \mathbb{E} \left\{ \frac{1}{K} \sum_{k=1}^K \frac{\|\mathbf{g}_k[n] - \hat{\mathbf{g}}_k[n - u_k]\|^2}{\|\mathbf{g}_k[n]\|^2} \right\}, \quad (34)$$

and we use the cumulative distribution function (CDF) of the spectral efficiency across all users to evaluate the effectiveness of the channel estimation scheme. The pilot reduction ratio is also adopted as a metric to evaluate the channel resources freed by the proposed scheme, which is defined as the ratio of reduced pilot overhead to the total pilot overhead of the conventional scheme.

We first compare the pilot reduction ratio using the adaptive channel estimation scheme with respect to the value of ζ_{ul} in Fig. 3. The results are obtained via Monte Carlo simulations averaged over 500 stationary blocks. The mobile user speeds are set to $\{1, 5, 10, 15, 20, 25, 30, 35, 40\}$ km/h, with $K = 60$ and $\epsilon_{ul} = 15$ dB. Fig. 3 clearly illustrates that when $\zeta_{ul} = 5$ dB, our proposed adaptive scheme results in a slight reduction in the number of channel estimations, and the reduction ratio of channel estimation for all users becomes higher as ζ_{ul} increases. The phenomenon is consistent with the theoretical analysis. According to Lemma 2 and Corollary 2, a low uplink training SNR significantly degrades the accuracy of CSI estimation, leading to an underestimation of the temporal channel correlation. In this case, our proposed channel estimation scheme degrades to the conventional channel estimation scheme. Conversely, the temporal channel correlation can be estimated accurately when ζ_{ul} is high, which enables our proposed scheme to set appropriate channel estimation intervals. Moreover, as shown in Fig. 3, we observe that the users with lower mobility exhibit a greater reduction in the number of channel estimations as compared to those with higher mobility,

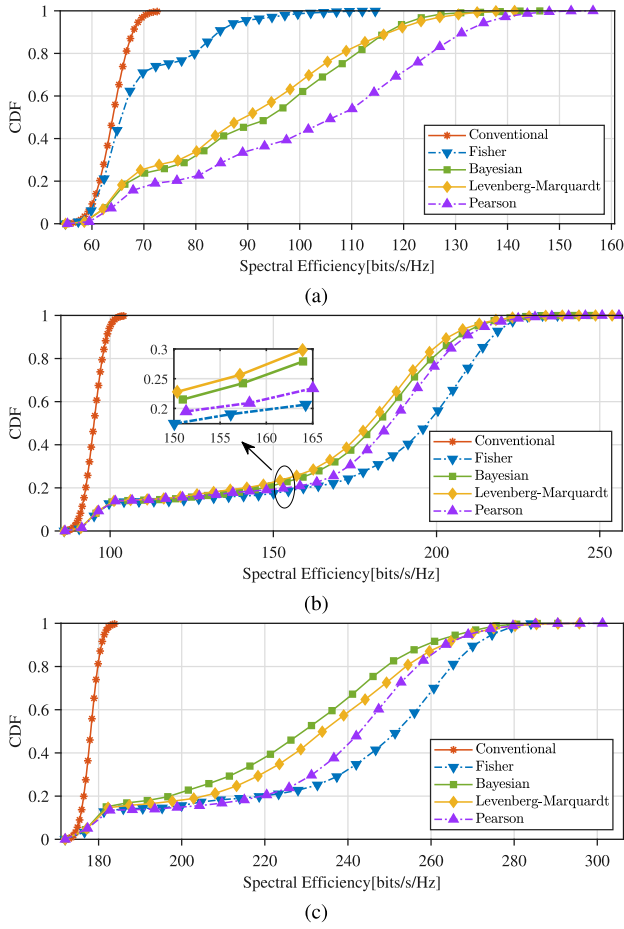


Fig. 4. CDF of net sum spectral efficiency for our proposed scheme employing different temporal correlation estimation algorithms with $K = 50$. a) $\zeta_{ul} = \zeta_{dl} = 10$ dB; b) $\zeta_{ul} = \zeta_{dl} = 15$ dB; c) $\zeta_{ul} = \zeta_{dl} = 20$ dB.

given the same ζ_{dl} . The reason lies in the fact that users of lower velocities usually have stronger temporal channel correlation, in which case longer channel estimation intervals can effectively enhance the net sum spectral efficiency. From Fig. 3, we also see that in the scenario of high uplink training SNR where $\zeta_{ul} \geq 15$ dB, the reduction ratio of the channel estimation frequency is almost the same when the velocities of users are 1 km/h. The reason is that the BS needs to estimate the channel at least once every N_m time slots, so N_m limits the maximum channel estimation interval and thus constrains the maximum reduction ratio of all users. Hence, N_m should be set based on specific scenarios.

Since SNR is the primary factor influencing the frequency of channel estimation, we then compare the CDF of the net sum spectral efficiency for the conventional scheme and the adaptive channel estimation scheme under different SNR, as depicted in Fig. 4. In addition to Fisher transformation, we consider three other algorithms, namely Levenberg-Marquardt algorithm [37], Bayesian estimation [38], and Pearson correlation coefficient for temporal correlation estimation. We set $K = 50$ and ζ_{ul} of Fig. 4a-c to 10 dB, 15 dB and 20 dB, respectively. From Fig. 4, it is evident that the net sum spectral efficiency of Fisher transformation-based adaptive scheme increases with ζ_{ul} . The reason is that Fisher

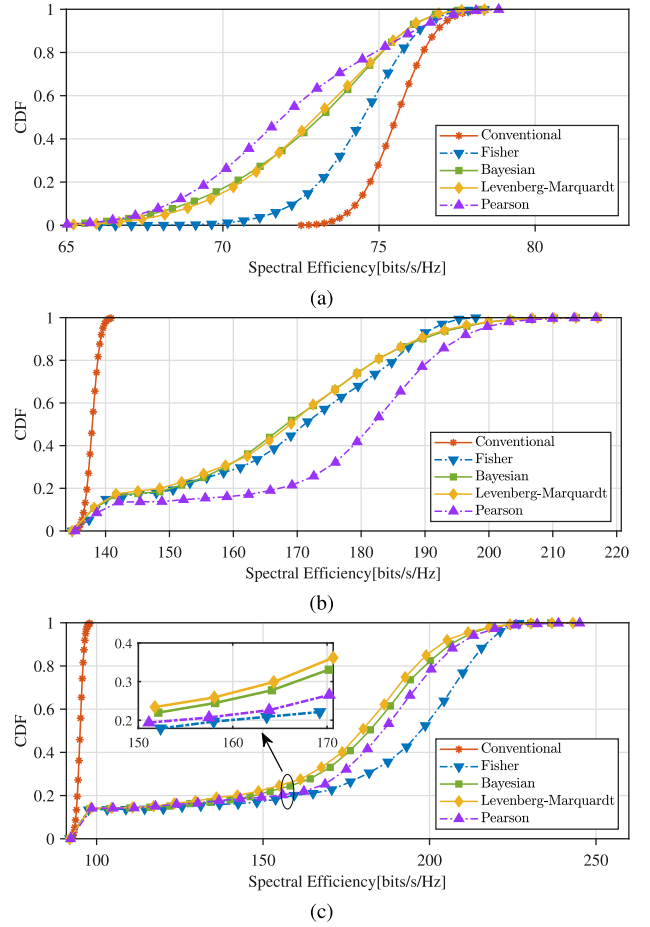


Fig. 5. CDF of net sum spectral efficiency for our proposed scheme employing different temporal correlation estimation algorithms with $\zeta_{ul} = \zeta_{dl} = 15$ dB. a) $K = 10$; b) $K = 40$; c) $K = 70$.

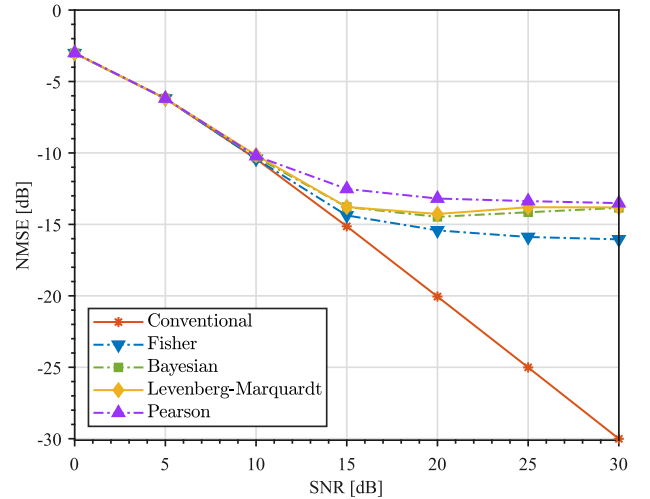


Fig. 6. NMSE of adaptive channel estimation schemes employing different temporal correlation estimation algorithms.

transformation obtains an underestimation of the temporal correlation coefficient due to uplink training noise. Thus, the increase in ζ_{ul} improves the accuracy of CSI samples and enables more precise correlation estimates obtained by Fisher transformation. In this case, the channel estimation interval can be correctly determined, which results in a substantial increase

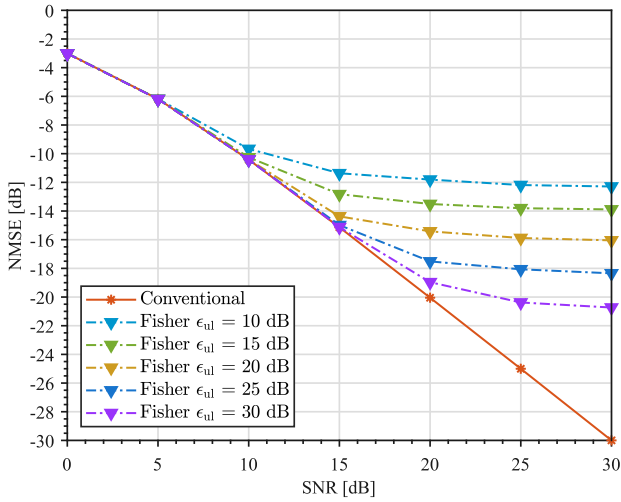


Fig. 7. NMSE of adaptive channel estimation schemes under different ϵ_{ul} and ϵ_{dl} .

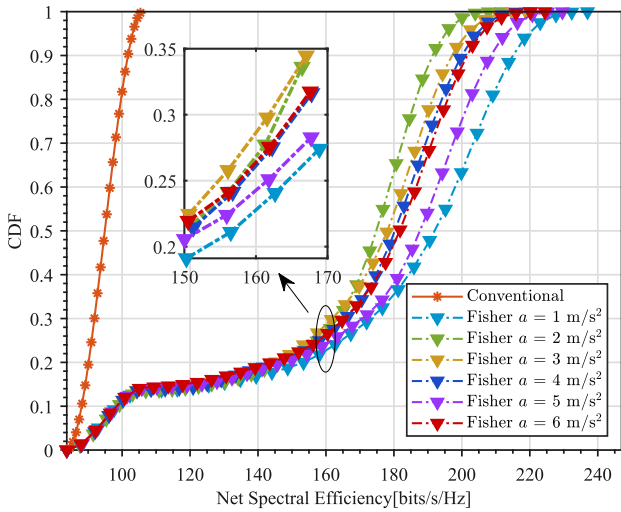


Fig. 8. CDF of net sum spectral efficiency among the conventional estimation scheme and the proposed ones under different accelerations.

in net spectral efficiency. Additionally, the performance gain corresponding to Fisher transformation is higher than that of Pearson except when $\zeta_{ul} = 10$ dB. The underlying reason is that the Pearson correlation coefficient provides an unbiased estimation of the temporal correlation between CSIs. Since the MMSE estimator is employed for channel estimation, it follows from *Lemma 2* that $\rho_{k,u_k}[n] = \frac{\hat{\rho}_{k,u_k}[n]\beta_k}{\gamma_k} \geq \hat{\rho}_{k,u_k}[n]$, so the expected value of Pearson coefficient is only 0.9 times the true channel correlation when the uplink training SNR is 10 dB. Consequently, the Pearson correlation coefficient is prone to underestimating the temporal correlation. (19) and (21) further indicate that the confidence lower bound produced by the Fisher transformation is smaller than the Pearson sample correlation coefficient. Therefore, Fisher transformation yields a severe underestimation of temporal correlation, which forces the proposed channel estimation scheme to adopt a conservative CSI reuse strategy, while Pearson and other algorithms reach the CSI reuse threshold and significantly improve system performance. As ζ_{ul} increases to 15 dB and

20 dB, Pearson may overestimate the temporal correlation, in which case the adaptive channel estimation scheme assigns a channel estimation interval longer than coherence time. The increase in estimation interval introduces the unit channel aging error that outweighs the unit CSI reuse gain, thereby decreasing performance. In contrast, Fisher transformation effectively mitigates the overestimation and achieves a better balance between channel aging errors and CSI reuse gains, as demonstrated in Fig. 4(b) and Fig. 4(c). We can also see from Fig. 4 that the Levenberg-Marquardt algorithm and Bayesian estimation exhibit only marginal performance discrepancies while demonstrating inferior spectral efficiency as compared to Fisher transformation, indicating that Fisher transformation achieves optimal performance among all algorithms in high SNR scenarios where $\zeta_{ul} > 10$ dB.

Furthermore, we depict the net sum spectral efficiency against the K in Fig. 5. We set $\zeta_{ul} = \zeta_{dl} = 15$ dB, with K in Fig. 5a-c as 10, 40, and 70, respectively. From Fig. 5, we observe that the system performance of all algorithms improves as we increase K , with Fisher transformation and Pearson consistently outperforming Bayesian estimation and Levenberg-Marquardt algorithm in terms of net sum spectral efficiency. We can also see from Fig. 5a that the net sum spectral efficiency of all algorithms is lower than that of the conventional scheme, with Fisher transformation exhibiting the smallest performance loss. The reason is that few spectrum resources are allocated for channel estimation when K is small, in which case the unit channel aging error outweighs the unit CSI reuse gain. Although threshold b is designed to prevent CSI reuse in such scenarios, the ideal propagation conditions assumed in *Corollary 1* are not fully satisfied due to the environmental noise, making the threshold derived from (32) suboptimal for real scenarios. Consequently, the adaptive scheme still reuses CSI with degraded performance, and Fisher transformation minimizes channel aging errors due to the tendency to underestimate. As shown in Fig. 5b, the CSI reuse gain exceeds the channel aging error when K increases to 40, in which case the net sum spectral efficiencies of all algorithms are mainly in the range of [170, 200] bit/s/Hz, achieving a 30%-40% gain over the conventional scheme, with Pearson outperforming the other algorithms. As discussed in Section IV, the Fisher transformation is inherently conservative and aims to mitigate performance degradation caused by overestimation. Moreover, fewer users in Fig. 5b correspond to a higher threshold according to (32) and thus a more conservative CSI reuse policy as compared to the scenario in Fig. 4b, which prevents Fisher transformation from fully exploiting the redundant channel estimation resources, resulting in worse performance than Pearson. Moreover, the performance gains in Fig. 5c are akin to the ones in Fig. 4c. So we conclude from Fig. 4 and Fig. 5 that in scenarios with numerous users and high SNR, the Fisher transformation-based channel estimation scheme can achieve a significant improvement in net sum spectral efficiency and consistently outperform other algorithms.

Now we illustrate the impact of channel aging by comparing the NMSE performance of adaptive schemes using Levenberg-Marquardt algorithm, Fisher transformation, Bayesian

estimation, and Pearson with respect to ζ_{ul} . The results are exhibited in Fig. 6, from which we can observe that Fisher transformation achieves the lowest NMSE performance among the algorithms. Additionally, the NMSE of all algorithms has a declining trend as ζ_{ul} increases at the beginning, and then gradually converges to a fixed value. This is consistent with the theoretical analysis in Lemma 1 that NMSE comprises the channel aging error and the estimation error. As ζ_{ul} increases, the channel estimation error approaches zero, while the channel aging error is independent of ζ_{ul} , leading to the NMSE converging to a constant value. Note that the channel aging error is decided by b according to Theorem 1. To further validate that the phenomenon is caused by channel aging, we compare NMSE performance under different threshold b , as depicted in the Fig. 7, with ϵ_{ul} set to 10, 15, 20, 25, and 30 dB. Fig. 7 indicates that the converged NMSE value decreases as ϵ_{ul} increases, with an approximate 5 dB reduction in channel aging error for every 5 dB increase in ϵ_{ul} . The phenomenon suggests that higher ϵ_{ul} leads to a more conservative threshold b , which effectively mitigates the channel aging error. In conclusion, the Fisher transformation-based adaptive scheme introduces fewer channel aging errors as compared to other algorithms. Furthermore, our proposed method for setting threshold b allows for quantitative control of the channel aging error.

In the proposed channel estimation scheme, the length of stationary blocks is determined by the hyperparameter N_s . In practice, however, variations in user mobility induce changes in the length of stationary blocks, which may affect the performance of the proposed scheme. Since mobility variations are manifested through user acceleration, we use acceleration as a metric of mobility dynamics and evaluate the performance of the proposed scheme under different acceleration levels. Fig. 8 compares the CDF of the net sum spectral efficiency among the conventional channel estimation scheme and the proposed schemes under different accelerations. We set $\zeta_{ul} = \zeta_{dl} = 15$ dB and $K = 70$. The acceleration of mobile users denoted by a is set to be $\{1, 2, 3, 4, 5, 6\}$ m/s² in this scenario, respectively. From Fig. 8, we can observe that the net sum spectral efficiencies obtained by the proposed adaptive channel estimation scheme are mainly in the range of [180, 200] bits/s/Hz, while those of the conventional scheme are primarily in the range of [90, 100] bits/s/Hz. Thus, the adaptive scheme approximately achieves a gain of 90 bit/s/Hz, equivalently 25.7 Mbit/s for each user in a communication system of 20 MHz bandwidth with 70 users. Furthermore, both the range and the CDF growth trends of the net sum spectral efficiency remain approximately consistent across different accelerations. These results indicate that the proposed scheme can accommodate variations in the stationary block caused by changes in user mobility, and can effectively utilize outdated CSI to improve the performance gains of massive MIMO systems under diverse communication scenarios.

VI. CONCLUSION

In this paper, we investigated how to set appropriate channel estimation intervals to utilize the redundant spectral resources in the uplink training phase. We developed an adaptive channel

estimation scheme according to the temporal channel correlation and derived the net sum spectral efficiency of the proposed scheme. Moreover, we employed Fisher transformation on CSI samples to obtain high-confidence temporal channel correlation interval estimates and determined the channel estimation interval by comparing the lower bound of the interval estimates to a predefined threshold. To control the channel aging error from CSI reuse, we derived the asymptotic achievable rate of the system and proposed a threshold design method accordingly. Numerical results demonstrate that our proposed scheme effectively enhances net sum spectral efficiency over conventional one and achieves a good balance between the channel aging error and CSI overhead reduction as compared to common parameter estimation methods. Furthermore, our proposed scheme is adaptable to various communication scenarios. For future work, extensions to more general communication scenarios involving pilot contamination and spatially correlated channels will be investigated to further enhance the robustness and efficiency of the proposed channel estimation scheme.

APPENDIX A PROOF OF LEMMA 1

Note that the relationship between $\mathbf{g}_k[n - u_k]$ and $\mathbf{g}_k[n]$ is $\mathbf{g}_k[n] = \rho_{k,u_k}[n]\mathbf{g}_k[n - u_k] + \boldsymbol{\varepsilon}_{k,u_k}[n]$, where $\rho_{k,u_k}[n]$ is the temporal correlation coefficient between $\mathbf{g}_k[n]$ and $\mathbf{g}_k[n - u_k]$, $\boldsymbol{\varepsilon}_{k,u_k}[n] \sim \mathcal{CN}(\mathbf{0}, \beta_k(1 - \rho_{k,u_k}[n]^2)\mathbf{I}_M)$ represents channel aging noise independent of $\mathbf{g}_k[n - u_k]$. The m th component of $\mathbf{g}_k[n]$ can be expressed as

$$g_k^m[n] = \rho_{k,u_k}[n]g_k^m[n - u_k] + \varepsilon_{k,u_k}^m[n], \quad (35)$$

where $\varepsilon_{k,u_k}^m[n] \sim \mathcal{CN}(0, \beta_k(1 - \rho_{k,u_k}[n]^2))$. Then the variance of the channel estimate can be derived as

$$\begin{aligned} & \mathbb{E}\{|e_{k,u_k}^m[n]|^2\} \\ &= \mathbb{E}\{|\rho_{k,u_k}[n]\hat{g}_k^m[n - u_k] - g_k^m[n]|^2\} \\ &= \mathbb{E}\{|\rho_{k,u_k}[n](\hat{g}_k^m[n - u_k] - g_k^m[n - u_k]) - \varepsilon_{k,u_k}^m[n]|^2\}. \end{aligned} \quad (36)$$

From (4), we have

$$\hat{g}_k^m[n] = \frac{\tau_{ul}p_{ul}\beta_k}{\tau_{ul}p_{ul}\beta_k + 1}g_k^m[n] + \frac{\sqrt{\tau_{ul}p_{ul}\beta_k}}{\tau_{ul}p_{ul}\beta_k + 1}w_{p,k}^m[n], \quad (37)$$

which allows (36) to be expressed as

$$\begin{aligned} & \mathbb{E}\{|e_{k,u_k}^m[n]|^2\} \\ &= \mathbb{E}\{|\rho_{k,u_k}[n](\hat{g}_k^m[n - u_k] - g_k^m[n - u_k]) - \varepsilon_{k,u_k}^m[n]|^2\} \\ &= \mathbb{E}\left\{\left|\frac{-\rho_{k,u_k}[n]}{\tau_{ul}p_{ul}\beta_k + 1}g_k^m[n - u_k] + \frac{\rho_{k,u_k}[n]\sqrt{\tau_{ul}p_{ul}\beta_k}}{\tau_{ul}p_{ul}\beta_k + 1}w_{p,k}^m[n - u_k] - \varepsilon_{k,u_k}^m[n]\right|^2\right\}. \end{aligned} \quad (38)$$

Then, we utilize the fact that $g_k^m[n - u_k]$, $w_{p,k}^m[n - u_k]$, and $\varepsilon_{k,u_k}^m[n]$ are mutually independent and observe that

$$\begin{aligned} & \mathbb{E}\{|e_{k,u_k}^m[n]|^2\} \\ &= \frac{1}{(\tau_{ul}p_{ul}\beta_k + 1)^2} \{ \rho_{k,u_k}[n]^2\beta_k \\ &+ \rho_{k,u_k}[n]^2\tau_{ul}p_{ul}\beta_k^2 + \beta_k(1 - \rho_{k,u_k}[n]^2)(\tau_{ul}p_{ul}\beta_k + 1)^2 \} \end{aligned}$$

$$\begin{aligned}
&= \beta_k(1 - \rho_{k,u_k}[n]^2) + \frac{\beta_k \rho_{k,u_k}[n]^2}{\tau_{dl} p_{dl} \beta_k + 1} \\
&= \beta_k(1 - \rho_{k,u_k}[n]^2) + \rho_{k,u_k}[n]^2 \mathbb{E}\{|e_k^m[n]|^2\}. \quad (39)
\end{aligned}$$

APPENDIX B PROOF OF COROLLARY 2

According to the condition of *Corollary 2*, we have $\epsilon_{ul} = K p_{ul} \beta_k$ and $\epsilon_{dl} = \frac{p_{dl} \beta_k}{K}$. Recalling that

$$\mathbf{x} = \mathbf{A} \mathbf{D}_\eta^{1/2} \mathbf{s}, \quad (40)$$

so we can also rewrite $\mathbb{E}\{\|\mathbf{x}\|^2\}$ as

$$\begin{aligned}
\mathbb{E}\{\|\mathbf{x}\|^2\} &= \mathbb{E}\left\{\text{tr}\left\{\mathbf{A} \mathbf{D}_\eta \mathbf{A}^H\right\}\right\} \\
&= (M - K) \mathbb{E}\left\{\text{tr}\left\{\mathbf{D}_\eta \left(\mathbf{H}^T \mathbf{H}^*\right)^{-1}\right\}\right\} \\
&= (M - K) \sum_{k=1}^K \eta_k \mathbb{E}\left\{\left[\left(\mathbf{H}^T \mathbf{H}^*\right)^{-1}\right]_{kk}\right\} \\
&= \sum_{k=1}^K \eta_k = 1, \quad (41)
\end{aligned}$$

where $\mathbf{H} \triangleq [\mathbf{h}_1, \dots, \mathbf{h}_k, \dots, \mathbf{h}_K]$ denotes the normalized channel estimation matrix. Therefore, the downlink effective SINR in conventional scheme with a target propagation environment can be derived as

$$\begin{aligned}
\text{SINR}_k &= \frac{(M - K) \gamma_k \eta_k}{\frac{1}{p_{dl}} + (\beta_k - \gamma_k) \sum_{k=1}^K \eta_k} = \frac{(M - K) \gamma_k \eta_k}{\frac{1}{p_{dl}} + \beta_k - \gamma_k} \\
&= \frac{\left(\frac{M}{K} - 1\right) \frac{\epsilon_{ul}}{1 + \epsilon_{ul}}}{\frac{1}{K \epsilon_{dl}} + \frac{1}{1 + \epsilon_{ul}}} = \frac{\left(\frac{M}{K} - 1\right) \epsilon_{ul}}{\frac{1}{K \epsilon_{dl}} + \frac{\epsilon_{ul}}{K \epsilon_{dl}} + 1} \\
&= \frac{(M - K) \epsilon_{ul} \epsilon_{dl}}{1 + \epsilon_{ul} + K \epsilon_{dl}}, \quad (42)
\end{aligned}$$

and the downlink achievable rate is

$$C_k = \log_2 \left(\frac{(M - K) \epsilon_{ul} \epsilon_{dl}}{1 + \epsilon_{ul} + K \epsilon_{dl}} \right). \quad (43)$$

Then we set the asymptotic net sum spectral efficiency equal to that of the conventional channel estimation scheme with $\epsilon_{ul} = p_{ul} \beta_k$ and $\epsilon_{dl} = \frac{p_{dl} \beta_k}{K}$, which yields

$$\log_2 \left(1 + \frac{\frac{M}{K} - 1}{\frac{1}{b^2} - 1} \right) = \left(1 - \frac{K}{N_s} \right) \log_2 \left(1 + \frac{(M - K) \epsilon_{ul} \epsilon_{dl}}{1 + \epsilon_{ul} + K \epsilon_{dl}} \right). \quad (44)$$

Take the exponential of both sides of (44) and rearranging, we have

$$\begin{aligned}
1 + \frac{\frac{M}{K} - 1}{\frac{1}{b^2} - 1} &= \left(1 + \frac{(M - K) \epsilon_{ul} \epsilon_{dl}}{1 + \epsilon_{ul} + K \epsilon_{dl}} \right)^{1 - \frac{K}{N_s}} \\
\frac{1}{b^2} &= \frac{\frac{M}{K} - 1}{\left(1 + \frac{(M - K) \epsilon_{ul} \epsilon_{dl}}{1 + \epsilon_{ul} + K \epsilon_{dl}} \right)^{1 - \frac{K}{N_s}} - 1} + 1, \quad (45)
\end{aligned}$$

and thereby we set the threshold as

$$b = \left(\frac{\frac{M}{K} - 1}{\left(1 + \frac{(M - K) \epsilon_{ul} \epsilon_{dl}}{1 + \epsilon_{ul} + K \epsilon_{dl}} \right)^{1 - \frac{K}{N_s}} - 1} + 1 \right)^{-\frac{1}{2}}. \quad (46)$$

REFERENCES

- [1] F. Rusek et al., "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [2] Z. Tao, T. Wang, and S. Wang, "Improve downlink rates of FDD massive MIMO systems by exploiting CSI feedback waiting phase," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.
- [3] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.
- [4] Z. Tao and S. Wang, "Improved downlink rates for FDD massive MIMO systems through Bayesian neural networks-based channel prediction," *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 2122–2134, Mar. 2022.
- [5] T. L. Marzetta, "Massive MIMO: An introduction," *Bell Labs Tech. J.*, vol. 20, pp. 11–22, 2015.
- [6] J. Zheng et al., "Mobile cell-free massive MIMO: Challenges, solutions, and future directions," *IEEE Wireless Commun.*, vol. 31, no. 3, pp. 140–147, Jun. 2024.
- [7] L. Ribeiro and M. Juntti, "Spectral efficiency maximization for massive MIMO uplink with intra-cell pilot reuse," *IEEE Wireless Commun. Lett.*, vol. 13, no. 3, pp. 637–641, Mar. 2024.
- [8] J. Li, L. Xiao, and S. Zhou, "Adaptive pilot allocation for multi-cell massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 7, pp. 972–975, Jul. 2020.
- [9] J.-Y. Sohn, S. W. Yoon, and J. Moon, "Pilot reuse strategy maximizing the weighted-sum-rate in massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 8, pp. 1728–1740, Aug. 2017.
- [10] L. You, X. Gao, X.-G. Xia, N. Ma, and Y. Peng, "Pilot reuse for massive MIMO transmission over spatially correlated Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3352–3366, Jun. 2015.
- [11] J. Guo, C.-K. Wen, and S. Jin, "CANet: Uplink-aided downlink channel acquisition in FDD massive MIMO using deep learning," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 199–214, Jan. 2022.
- [12] M. Ke, Z. Gao, Y. Wu, X. Gao, and R. Schober, "Compressive sensing-based adaptive active user detection and channel estimation: Massive access meets massive MIMO," *IEEE Trans. Signal Process.*, vol. 68, pp. 764–779, 2020.
- [13] A. K. Papazafeiropoulos, "Impact of general channel aging conditions on the downlink performance of massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1428–1442, Feb. 2017.
- [14] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1650–1662, Jul. 2005.
- [15] G. Fodor, S. Fodor, and M. Telek, "Performance analysis of a linear MMSE receiver in time-variant Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4098–4112, Jun. 2021.
- [16] H. Abeida, "Data-aided SNR estimation in time-variant Rayleigh fading channels," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5496–5507, Nov. 2010.
- [17] S. Bazzi and W. Xu, "Downlink training sequence design for FDD multiuser massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4732–4744, Sep. 2017.
- [18] H. Kim, S. Kim, H. Lee, C. Jang, Y. Choi, and J. Choi, "Massive MIMO channel prediction: Kalman filtering Vs. machine learning," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 518–528, Jan. 2021.
- [19] Y. Jin, Y. Wu, Y. Gao, S. Zhang, S. Xu, and C.-X. Wang, "LinFormer: A linear-based lightweight transformer architecture for time-aware MIMO channel prediction," *IEEE Trans. Wireless Commun.*, vol. 24, no. 9, pp. 7177–7190, Sep. 2025.
- [20] S. Zhang, S. Zhang, W. Yuan, and T. Q. S. Quek, "Transformer-empowered predictive beamforming for rate-splitting multiple access in non-terrestrial networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 12, pp. 19776–19788, Dec. 2024.
- [21] K. T. Truong and R. W. Heath, "Effects of channel aging in massive MIMO systems," *J. Commun. Netw.*, vol. 15, no. 4, pp. 338–351, Aug. 2013.
- [22] A. K. Papazafeiropoulos and T. Ratnarajah, "Deterministic equivalent performance analysis of time-varying massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5795–5809, Oct. 2015.
- [23] C. Kong, C. Zhong, A. K. Papazafeiropoulos, M. Matthaiou, and Z. Zhang, "Sum-rate and power scaling of massive MIMO systems with channel aging," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4879–4893, Dec. 2015.

- [24] J. Zheng, J. Zhang, E. Björnson, and B. Ai, "Impact of channel aging on cell-free massive MIMO over spatially correlated channels," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6451–6466, Oct. 2021.
- [25] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge, U.K.: Cambridge Univ. Press, 2016.
- [26] T. X. Vu, T. A. Vu, S. Chatzinotas, and B. Ottersten, "Spectral-efficient model for multiuser massive MIMO: Exploiting user velocity," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017, pp. 1–6.
- [27] S. E. Hajri, M. Assaad, and M. Larrañaga, "Enhancing massive MIMO: A new approach for uplink training based on heterogeneous coherence times," in *Proc. 25th Int. Conf. Telecommun. (ICT)*, France, Jun. 2018, pp. 361–366.
- [28] S. Fodor, G. Fodor, D. Gürgünoğlu, and M. Telek, "Optimizing pilot spacing in MU-MIMO systems operating over aging channels," *IEEE Trans. Commun.*, vol. 71, no. 6, pp. 3708–3720, Jun. 2023.
- [29] R. Deng, Z. Jiang, S. Zhou, and Z. Niu, "Intermittent CSI update for massive MIMO systems with heterogeneous user mobility," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 4811–4824, Jul. 2019.
- [30] J. Yuan, H. Q. Ngo, and M. Matthaiou, "Machine learning-based channel prediction in massive MIMO with channel aging," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 2960–2973, May 2020.
- [31] W. C. Jakes and D. C. Cox, *Microwave Mobile Communications*. Hoboken, NJ, USA: Wiley, 1994.
- [32] J. Zheng, J. Zhang, H. Du, D. Niyato, D. I. Kim, and B. Ai, "Rate-splitting for CF massive MIMO systems with channel aging," *IEEE Trans. Veh. Technol.*, vol. 73, no. 1, pp. 1485–1490, Jan. 2024.
- [33] Z. Tao and S. Wang, "How often do we need to estimate wireless channels in massive MIMO with channel aging?" in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Madrid, Spain, Dec. 2021, pp. 1–6.
- [34] R. A. Fisher, "Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population," *Biometrika*, vol. 10, no. 4, pp. 507–521, May 1915.
- [35] S. Sesia, I. Toufik, and M. Baker, *LTE-The UMTS Long Term Evolution: From Theory to Practice*. Hoboken, NJ, USA: Wiley, 2011.
- [36] L. Liu et al., "The COST 2100 MIMO channel model," *IEEE Wireless Commun.*, vol. 19, no. 6, pp. 92–99, Dec. 2012.
- [37] M. T. Hagan, H. B. Demuth, and M. Beale, *Neural Network Design*. Boston, MA, USA: PWS Publishing Co., 1997.
- [38] D. J. C. MacKay, "Bayesian interpolation," *Neural Comput.*, vol. 4, no. 3, pp. 415–447, May 1992.



Zhouyi Qian received the B.S. degree from Hohai University, Nanjing, China, in 2023. He is currently pursuing the M.S. degree with the School of Electronic Science and Engineering, Nanjing University. His research interests include massive MIMO and machine learning.



Shaowei Wang (Senior Member, IEEE) received the Ph.D. degree from Wuhan University, Wuhan, China, in 2006. He joined the School of Electronic Science and Engineering, Nanjing University, Nanjing, China, as a Faculty Member in 2006, where he is currently a Full Professor. From 2012 to 2013, he was a Visiting Scholar/Professor with Stanford University, Stanford, CA, USA, and The University of British Columbia, Vancouver, BC, Canada. His research interests include communications and networking, operations research, and machine learning.