DOA-Based Hybrid Reconfigurable Intelligent Surface Phase-Shift Design for Achievable Rate Enhancement

Taehee Choi, *Student Member, IEEE,* Jaehong Kim, *Student Member, IEEE,* Han-Gyeol Lee, *Student Member, IEEE,* Koichi Adachi, *Senior Member, IEEE,* and Jingon Joung, *Senior Member, IEEE*

***Abstract*—This study proposes a direction-of-arrival (DOA)-based phase-shift method for hybrid reconfigurable intelligent surface (RIS)-assisted communication systems. The hybrid RIS can estimate the DOA of the incident signal by utilizing its hybrid elements, which can receive a portion of the incident signal and reflect the remainder. Then, the optimal phase shifts of the hybrid RIS are calculated based on the estimated DOA without any phase-shift information feedback. Here, a tradeoff between the DOA estimation accuracy and the signal-reflecting power can be balanced by optimizing the sensing factor of the hybrid RIS, i.e., the power ratio of the received (or sensing) signal to the total incident signal at the hybrid element. To this end, the optimal sensing factor that maximizes the effective ergodic achievable rate is analytically derived. The simulation results demonstrate that the proposed hybrid RIS system with optimal sensing factor outperforms both conventional passive RIS systems and channel-estimation-based hybrid RIS systems when the number of RIS elements is large and the line-of-sight channel is dominant.**

***Index Terms*—Hybrid reconfigurable intelligent surface, direction-of-arrival (DOA), sensing factor, effective ergodic achievable rate.**

I. Introduction

A reconfigurable intelligent surface (RIS) is a promising technology for manipulating wireless channels to a desired condition by adjusting the electromagnetic characteristics of incident signals [1]. In the last decade, passive RISs that can reflect the incident signal after phase shifts have been vigorously investigated to improve the performance of various wireless communication systems [2], [3], [4], [5], [6]. However, due to the absence of signal reception or amplification capabilities in passive RIS, phase shifts should be calculated at the base station (BS) and fed back to the passive RIS, resulting in substantial channel estimation (CE) and feedback overheads in every channel update [7], [8]. To reduce CE overhead, a compressed sensing-based sparse CE method has been proposed for passive RIS-aided millimeter wave communication systems [9]. However, even with the sparse CE method, phase-shift feedback overhead remains substantial. On the other hand, passive RIS phase-shift methods based on statistical channel state information (CSI) have been investigated in [10], [11] to address this issue. Since statistical CSI remains constant over a long period, phase-shift methods based on statistical CSI can significantly reduce both feedback and CE overhead [10]. Specifically, in [11], it is shown that the optimal passive RIS phase-shift vector that maximizes the ergodic achievable rate can be determined by the direction-of-arrival (DOA) of the line-of-sight (LoS) component in the Rician channel model. Accordingly, the passive RIS phase-shift vector can be optimized by estimating the DOA of the LoS component at the BS. However, DOA estimation accuracy at the BS can be significantly degraded by the severe power attenuation due to the double-path loss of passive RIS-reflected signals.

The concept of a hybrid RIS, which comprises hybrid elements that partially receive and reflect incident signals, has recently been introduced [12]. Here, the power ratio of the received (or sensing) signal to the total incident signal in the hybrid element is called *sensing factor*, which varies depending on the hardware structure of the hybrid element[1]. Upon receiving incident signals, the RIS controller can estimate the DOA of the incident LoS signal and thereby calculate optimal phase shifts without the BS's intervention. Thus, the proposed hybrid RIS system can eliminate feedback overhead and improve DOA estimation accuracy by avoiding the double-path loss inherent in passive RIS systems. Due to the energy conservation law, the total sum of the receiving and reflecting powers at the hybrid element is limited by the power of the incident signal [13]. Therefore, a tradeoff exists between DOA estimation accuracy and signal-reflecting power. For instance, as the sensing factor increases, the DOA estimation accuracy improves while the signal-reflecting power decreases. Therefore, it is crucial to investigate an optimal sensing factor that balances the DOA estimation accuracy and the signal-reflecting power.

In this study, we propose a DOA-based hybrid RIS phase-shift design to enhance the effective ergodic achievable rate, considering both the received signal-to-noise ratio (SNR) of the user equipment (UE) and the signaling overhead associated with the RIS phase-shift design. The proposed DOA-based hybrid RIS system can significantly reduce the amount of signaling overhead, thereby achieving a higher effective achievable ergodic rate compared to conventional passive RIS systems. Moreover, the optimal sensing factor that maximizes the effective ergodic achievable rate is derived analytically. Numerical simulations verify that the proposed hybrid RIS

Manuscript received April 2, 2025; revised September 3, 2025; accepted November 13, 2025. This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT (RS-2024-00405510), in part by the Basic Science Research Program through the NRF funded by the Ministry of Education (RS-2025-25397301), and in part by the Chung-Ang University Graduate Research Scholarship in 2024. (*Corresponding author: Jingon Joung.*)

T. Choi, J. Kim, H.-G. Lee, and J. Joung are with the Department of Electrical and Electronics Engineering, Chung-Ang University, Seoul 06974, South Korea (e-mail: {chlxogml00, kjhct9606, forener, jgjoung}@cau.ac.kr).

K. Adachi is with the Department of Information and Computer Science, Keio University, Yokohama, Japan (e-mail: koichi.adachi@keio.jp).

Digital Object Identifier 10.1109/TVT.2025.0000000

[1]A hybrid element can be implemented by coupling a waveguide to a conventional RIS element, where the waveguide is connected to a radio-frequency chain [12]. The sensing factor is determined by the waveguide size or the annular ring surrounding the coupling via. Therefore, it is typically assumed that the coupling is static after manufacturing.

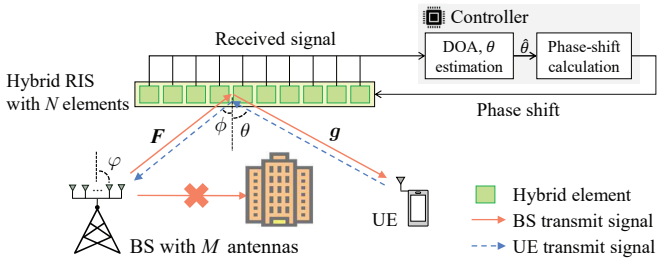

Fig. 1. Hybrid RIS-assisted communication system using DOA-based hybrid RIS phase-shift method.

system, with the optimal sensing factor, outperforms both the conventional passive RIS and the CE-based hybrid RIS systems. Furthermore, the performance improvement of the proposed method over the passive RIS increases as the number of RIS elements increases, especially when the LoS component is dominant and the coherence time is short.

Notations: For any scalar, vector, or matrix, the superscripts T, H, and $*$ denote transposition, Hermitian transposition, and complex conjugate, respectively. $|x|$ and $\|\boldsymbol{x}\|$ represent the absolute value of x and the Euclidean norm of vector $\boldsymbol{x}$, respectively. $\odot$ denotes Hadamard product; $\mathrm{diag}\,(\boldsymbol{x})$ returns a diagonal matrix whose main diagonal elements are $\boldsymbol{x}$; $\mathrm{tri}\,(x)$ denotes the triangular function, i.e., $\mathrm{tri}\,(x) \triangleq \max\,(1-|x|\,,0)$; $\mathrm{erf}\,(x)$ represents a Gaussian error function, i.e., $\mathrm{erf}\,(x) \triangleq \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt$; $x \sim \mathcal{N}(\mu,\sigma^2)$ represents that a random variable x conforms to a normal distribution with mean μ and variance σ^2; $x \sim \mathcal{CN}\left(0,\sigma^2\right)$ means that a complex random variable x conforms to a complex normal distribution with zero mean and variance σ^2. $\mathbb{E}\,[x]$ represents the expectation of a random variable x.

II. System, Channel, and Signal Models

This study considers a hybrid RIS-assisted multiple-input single-output communication system, as illustrated in Fig. 1. BS and UE have M and single antennas, respectively. The direct path between BS and UE is assumed to be blocked by obstacles. To construct a bypass link between the BS and the UE, a hybrid RIS with N hybrid elements is deployed [14]. Each hybrid element partially receives the power of the incident signal with a sensing factor of $\rho \in [0,1]$. Simultaneously, the hybrid elements adjust the phase of the incident signal by $\chi_n \in [0,2\pi)$ and then reflect the remaining $1-\rho$ portion of the total signal power [15]. For simplicity, it is assumed that both the BS antennas and the hybrid elements of the RIS are arranged in uniform linear arrays. Furthermore, since the RIS is initially deployed to guarantee the LoS link with the BS, the DOA at the hybrid RIS from the BS, denoted by ϕ, and the DOA at the BS from the hybrid RIS, denoted by φ, are considered known at the hybrid RIS.

The channel between BS and hybrid RIS is modeled as a Rician fading channel as follows:

$$\boldsymbol{F} = \sqrt{\eta_{\mathrm{B,R}}}\left(\sqrt{\frac{\kappa_F}{\kappa_F+1}}\bar{\boldsymbol{F}} + \sqrt{\frac{1}{\kappa_F+1}}\tilde{\boldsymbol{F}}\right), \tag{1}$$

where $\eta_{\mathrm{B,R}}$ and κ_F represent a distance-dependent path loss and the Rician factor between BS and hybrid RIS, respectively. The non-LoS (NLoS) channel matrix is denoted by $\tilde{\boldsymbol{F}} \in \mathbb{C}^{N\times M}$ whose elements are independent and identically distributed complex Gaussian random variables with zero mean and unit variance, i.e., $\mathcal{CN}(0,1)$; and $\bar{\boldsymbol{F}} = \boldsymbol{b}\,(\phi)\,\boldsymbol{a}^T\,(\varphi) \in \mathbb{C}^{N\times M}$ denotes the LoS components where $\boldsymbol{b}\,(\phi) \in \mathbb{C}^{N\times 1}$ and $\boldsymbol{a}\,(\varphi) \in \mathbb{C}^{M\times 1}$ represent the steering vectors of RIS and BS, respectively. Specifically, the nth element of $\boldsymbol{b}\,(\phi)$ is given by $e^{-\frac{j2\pi\Delta}{\lambda}(n-1)\sin\phi}$, and the mth element of $\boldsymbol{a}\,(\varphi)$ is $e^{-\frac{j2\pi\Delta}{\lambda}(m-1)\sin\varphi}$. Here, Δ and λ denote the antenna spacing and the carrier wavelength, respectively. Similarly, a channel between the hybrid RIS and the UE can be written as follows:

$$\boldsymbol{g} = \sqrt{\eta_{\mathrm{R,U}}}\left(\sqrt{\frac{\kappa_g}{\kappa_g+1}}\bar{\boldsymbol{g}} + \sqrt{\frac{1}{\kappa_g+1}}\tilde{\boldsymbol{g}}\right), \tag{2}$$

where $\eta_{\mathrm{R,U}}$ and κ_g represent a path loss and the Rician factor of the channel $\boldsymbol{g}$. The LoS component of the channel $\boldsymbol{g}$ is denoted as $\bar{\boldsymbol{g}} = \boldsymbol{b}\,(\theta) \in \mathbb{C}^{N\times 1}$ where θ denotes the DOA of the channel $\bar{\boldsymbol{g}}$ observed at the hybrid RIS, and $\tilde{\boldsymbol{g}} \in \mathbb{C}^{N\times 1}$ represents the NLoS component whose element follows $\mathcal{CN}(0,1)$.

The downlink received signal at the UE is written as follows:

$$r = \sqrt{(1-\rho)}\boldsymbol{g}^H\,\mathrm{diag}\,(\boldsymbol{\psi})\boldsymbol{F}\boldsymbol{w}s + z, \tag{3}$$

where s is the transmit symbol at the BS with a transmit power of P_{B}, i.e., $\mathbb{E}[|s|^2] = P_{\mathrm{B}}$; $\boldsymbol{w} \in \mathbb{C}^{M\times 1}$ is a transmit beamforming vector; and z denotes the additive white Gaussian noise (AWGN) following $\mathcal{CN}\left(0,\sigma_z^2\right)$ where σ_z^2 is the noise power at the UE. Here, $\boldsymbol{\psi} = \left[e^{j\chi_1}\cdots e^{j\chi_N}\right]^T \in \mathbb{C}^{N\times 1}$ represents the hybrid RIS phase-shift vector. By using maximum ratio transmission (MRT) at the BS, i.e., $\boldsymbol{w} = \boldsymbol{F}^H\,\mathrm{diag}(\boldsymbol{\psi})^H\boldsymbol{g}/\|\boldsymbol{F}^H\,\mathrm{diag}(\boldsymbol{\psi})^H\boldsymbol{g}\|$, the instantaneous received SNR for given ρ and $\boldsymbol{\psi}$ is written as follows:

$$\mathrm{SNR}\,(\rho,\boldsymbol{\psi}) = (1-\rho)P_{\mathrm{B}}\sigma_z^{-2}\left\|\boldsymbol{F}^H\,\mathrm{diag}\,(\boldsymbol{\psi})^H\boldsymbol{g}\right\|^2. \tag{4}$$

The instantaneous received SNR can be maximized by optimizing the sensing factor ρ and the phase-shift vector $\boldsymbol{\psi}$. In conventional passive RIS systems that purely reflect the incident signal, i.e., $\rho = 0$, a significant amount of CE and feedback overhead is required to optimize the RIS phase shift based on instantaneous CSI [7].

III. DOA-Based Hybrid RIS Phase-Shift Design with Optimal Sensing Factor

The instantaneous CSI-based RIS phase-shift designs that maximize (4) can significantly degrade the effective ergodic achievable rate due to excessive CE and feedback overheads. To address this issue, we propose the DOA-based hybrid RIS phase-shift method that maximizes the effective ergodic achievable rate, considering practical overhead models in RIS-assisted communication systems. Furthermore, the optimal sensing factor of the hybrid RIS is analytically derived in a closed-form expression.

A. DOA-Based Hybrid RIS Phase-Shift Design

By incorporating signaling overheads into the ergodic achievable rate expression, the *effective* ergodic achievable rate is modeled as follows:

$$R(\rho,\boldsymbol{\psi}) = B\left(1-\tfrac{\tau}{T}\right)\mathbb{E}\left[\log_2\left(1+\mathrm{SNR}\,(\rho,\boldsymbol{\psi})\right)\right], \tag{5}$$

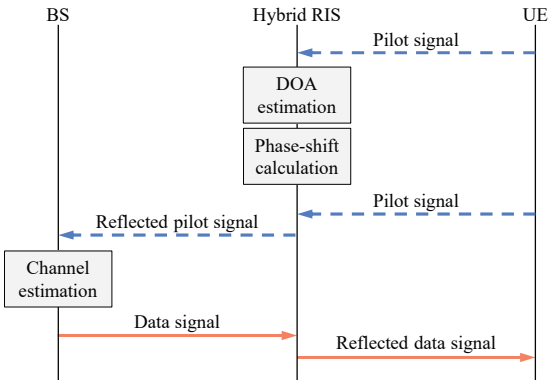

Fig. 2. Framework of the proposed DOA-based hybrid RIS phase-shift design.

where B and T represent the transmission bandwidth and the channel coherence time, respectively. The overhead model τ is given as the total time duration required for every signaling overhead [7], [8]. Here, (5) is intractable to derive further an explicit form that provides an intuitive interpretation of the effective ergodic achievable rate. To convert (5) into a tractable form, the upper bound of the effective ergodic achievable rate is derived from Jensen's inequality as follows:

$$R(\rho,\boldsymbol{\psi}) \le \check{R}(\rho,\boldsymbol{\psi}) = B\left(1-\tfrac{\tau}{T}\right)\log_2\left(1+\mathbb{E}\left[\mathrm{SNR}\left(\rho,\boldsymbol{\psi}\right)\right]\right). \quad (6)$$

Here, (6) offers a tight upper bound of (5) in the high SNR region [11]. The average received SNR is derived as follows:

$$\begin{aligned}&\mathbb{E}\left[\mathrm{SNR}\left(\rho,\boldsymbol{\psi}\right)\right]\\ &= \zeta(1-\rho)\left(\kappa_F\kappa_g\left|\boldsymbol{b}^H(\phi)\,\boldsymbol{\Psi}^H\boldsymbol{b}\left(\theta\right)\right|^2+N(\kappa_F+\kappa_g+1)\right) \quad (7a)\\ &\le \zeta(1-\rho)\left(\kappa_F\kappa_g N^2+N(\kappa_F+\kappa_g+1)\right), \quad (7b)\end{aligned}$$

where $\zeta = \frac{P_{\mathrm{B}}\eta_{\mathrm{B,R}}\eta_{\mathrm{R,U}}M}{\sigma_z^2(\kappa_F+1)(\kappa_g+1)}$ and $\boldsymbol{\Psi} \triangleq \mathrm{diag}\left(\boldsymbol{\psi}\right)$. The upper bound (7b) is derived according to the Cauchy-Schwarz inequality. From the equality condition of the Cauchy-Schwarz inequality, the optimal phase-shift vector that maximizes the average received SNR can be derived as follows [11]:

$$\boldsymbol{\psi}^{\mathrm{o}}(\theta) = \boldsymbol{b}\left(\theta\right)\odot\boldsymbol{b}^*(\phi). \quad (8)$$

From (8), it is shown that the optimal phase-shift vector can be calculated using the DOAs ϕ and θ. To this end, the DOA θ is estimated in the hybrid RIS using Q pilot transmissions from the UE [7], [8]. Assuming that channel reciprocity between the hybrid RIS and the UE is maintained during the coherence time T in time-division duplex mode, the pilot signals received at the hybrid RIS can be written as

$$\boldsymbol{Y} = \sqrt{\rho}\boldsymbol{g}\boldsymbol{x}^T + \boldsymbol{W} \in \mathbb{C}^{N\times Q}, \quad (9)$$

where $\boldsymbol{x} = [x_1\cdots x_q\cdots x_Q]^T \in \mathbb{C}^{Q\times 1}$; x_q is the qth pilot symbol at the UE satisfying $\mathbb{E}\left[|x_q|^2\right] = P_{\mathrm{U}}$; P_{U} denotes the transmit power at the UE; and $\boldsymbol{W} \in \mathbb{C}^{N\times Q}$ is an AWGN matrix at the hybrid RIS whose elements conforms to $\mathcal{CN}(0,\sigma_w^2)$.

The hybrid RIS can estimate θ from (9) using existing DOA estimation methods. In this study, the multiple signal classification (MUSIC) algorithm is adopted, considering its high angular resolution and robustness against interferences [16][2]. The mean squared error (MSE) of the MUSIC algorithm under a high SNR assumption is derived as follows [18]:

$$\epsilon = \frac{3\lambda^2\left(\rho Q\eta_{\mathrm{R,U}}P_{\mathrm{U}}+(\kappa_g+1)\sigma_w^2\right)}{2\pi^2\Delta^2\kappa_g Q\eta_{\mathrm{R,U}}P_{\mathrm{U}}\cos^2\theta}\frac{1}{\rho(N^3-N)}. \quad (10)$$

From (10), it is verified that the DOA estimation accuracy improves as ρ increases, thereby enabling the accurate computation of the optimal phase-shift vector. In contrast, the upper bound of the average received SNR decreases as ρ increases, as shown in (7b). Consequently, ρ is the key parameter that balances the tradeoff between the accuracy of the DOA estimation and the signal-reflecting power.

Using the DOA estimates, $\hat{\theta}$, the hybrid RIS calculates the optimal phase-shift vector $\boldsymbol{\psi}^{\mathrm{o}}(\hat{\theta})$ in (8). Next, the BS estimates the cascaded channel, i.e., $\boldsymbol{g}^H\,\mathrm{diag}(\boldsymbol{\psi}^{\mathrm{o}}(\hat{\theta}))\boldsymbol{F}$, by using a single pilot signal transmitted from the UE. In downlink data transmission, the BS employs MRT. The overall framework of the proposed DOA-based hybrid RIS phase-shift design is described in Fig. 2. Accordingly, the overhead model τ of the proposed method is given by the sum of the time durations required for the DOA estimation and the cascaded CE, i.e., $(Q+1)T_s$, where T_s denotes the symbol duration.

B. Derivation of Optimal Sensing Factor

The sensing factor optimization problem that maximizes the upper bound of the effective ergodic achievable rate in (6) can be formulated as follows:

$$\rho^{\mathrm{o}} = \underset{\rho\in[0,1]}{\arg\max}\quad \check{R}\left(\rho,\boldsymbol{\psi}^{\mathrm{o}}\left(\hat{\theta}\right)\right). \quad (11)$$

Since the overhead model τ in (6) is independent of ρ, and the logarithm function is monotonically increasing, (11) is equivalent to the problem that maximizes the average received SNR, which is written as follows:

$$\rho^{\mathrm{o}} = \underset{\rho\in[0,1]}{\arg\max}\quad \mathbb{E}\left[\mathrm{SNR}\left(\rho,\boldsymbol{\psi}^{\mathrm{o}}\left(\hat{\theta}\right)\right)\right]. \quad (12)$$

By treating $\hat{\theta}$ as a random variable following $\mathcal{N}\left(\theta,\epsilon\right)$ [19], the average received SNR using $\boldsymbol{\psi}^{\mathrm{o}}\left(\hat{\theta}\right)$ can be approximated as

$$\begin{aligned}&\mathbb{E}\left[\mathrm{SNR}\left(\rho,\boldsymbol{\psi}^{\mathrm{o}}\left(\hat{\theta}\right)\right)\right]\\ &\approx \zeta(1-\rho)N\left(\kappa_F\kappa_g N+\tfrac{\kappa_F}{2}+\kappa_g+1-\tfrac{\kappa_F(\kappa_g+1)\sigma_w^2}{2\rho Q\eta_{\mathrm{R,U}}P_{\mathrm{U}}}\right)\\ &\triangleq \overline{\mathrm{SNR}}\left(\rho\right). \quad (13)\end{aligned}$$

The detailed derivation of (13) is given in Appendix.

The closed-form solution for the optimal sensing factor is now analytically derived. It can be shown directly in (13) that ρ satisfying $\partial\overline{\mathrm{SNR}}\left(\rho\right)/\partial\rho = 0$ is unique and provides the maximum average received SNR for $\rho\in[0,1]$. Therefore, the optimal solution of (12) can be derived as follows:

$$\rho^{\mathrm{o}} = \sqrt{\frac{\kappa_F(\kappa_g+1)\sigma_w^2}{2Q\eta_{\mathrm{R,U}}P_{\mathrm{U}}(\kappa_F\kappa_g N+\kappa_F/2+\kappa_g+1)}}. \quad (14)$$

[2]The various DOA estimation algorithms, such as MUSIC, conventional beamforming, and estimation of signal parameters via rotational invariance technique (ESPRIT) [17], can also be employed for the proposed method. From the omitted MSE evaluation results, it is verified that the MUSIC and conventional beamforming methods achieve 2.67×10^{-6} and 2.65×10^{-6} of MSE, respectively. The MSE of ESPRIT is degraded to 4.98×10^{-3} under the simulation parameters used in Section III-B.

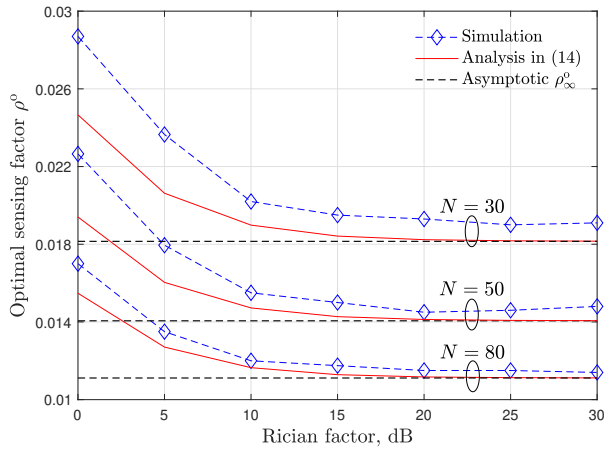

Fig. 3. Optimal sensing factor over the Rician factors, κ_F and κ_g, when $N \in \{30, 50, 80\}$.

Remark 1: (Asymptotic Analysis of ρ) As $\kappa_F \to \infty$ and $\kappa_g \to \infty$, i.e., in the LoS dominant channel, ρ^{o} approaches $\rho^{\mathrm{o}}_{\infty} \triangleq \sqrt{\sigma_w^2/(2NQ\eta_{\mathrm{R,U}}P_{\mathrm{U}})}$. Here, subscript ∞ denotes that the Rician factors approach infinity on the decibel scale. On the other hand, as $\kappa_F \to 0$ and $\kappa_g \to 0$, i.e., the Rayleigh channel, ρ^{o} approaches zero. It can be interpreted that if the power of the LoS channels is relatively weak in the rich scattering channel, the hybrid RIS should abandon its sensing capability and reflect the incident signal power with random phase shifts.

Remark 2: (*Asymptotic SNR Analysis*) As $\kappa_F \to \infty$ and $\kappa_g \to \infty$, $\overline{\mathrm{SNR}}(\rho)$ in (13) asymptotically converges to $\overline{\mathrm{SNR}}_{\infty} \triangleq \frac{P_{\mathrm{B}}}{\sigma_z^2}\eta_{\mathrm{B,R}}\eta_{\mathrm{R,U}}MN^2\left(1-\rho^{\mathrm{o}}_{\infty}\right)^2$. On the other hand, as $\kappa_F \to 0$ and $\kappa_g \to 0$, $\overline{\mathrm{SNR}}(\rho)$ asymptotically converges to $\overline{\mathrm{SNR}}_{-\infty} \triangleq \frac{P_{\mathrm{B}}}{\sigma_z^2}\eta_{\mathrm{B,R}}\eta_{\mathrm{R,U}}MN$. Here, subscript $-\infty$ denotes that the Rician factors approach minus infinity on the decibel scale.

Fig. 3 confirms the validity of the optimal sensing factor ρ^{o} in (14) by comparing it with the true ρ^{o} obtained from numerical simulation. Specifically, the optimal ρ of (12) is numerically found by using the exhaustive grid search method with a grid size of 5×10^{-5}. The simulation is conducted under the following parameters: $M = 32$, $N = 80$, $P_{\mathrm{B}} = 43$ dBm, $P_{\mathrm{U}} = 10$ dBm, data transmission bandwidth for hybrid RIS $B = 5$ MHz, $f_c = 5$ GHz, $\Delta = \lambda/2 = 0.03$ m, $Q = 2$, $\theta = 30°$, $\phi = 40°$, $\varphi = 50°$, $\sigma_z^2 = \sigma_w^2 = -107$ dBm, $T_s = 1/B$, $d_{\mathrm{B,R}} = 200$ m, $d_{\mathrm{R,U}} = 1000$ m, and 3GPP TR 36.814 urban macro LoS path loss model, $\eta_{\mathrm{a,b}} = -18 - 20\log_{10}(f_c) - 22\log_{10}(d_{\mathrm{a,b}})$, is used [20]. The analytically derived ρ^{o} is observed to align with the simulation results and the asymptotic analysis, thereby validating the derivation, including the approximation of the average received SNR in (13). In addition, it is verified that ρ^{o} decreases as the Rician factors κ_F and κ_g increase. This is because hybrid RIS can accurately estimate the DOA of the LoS channel with a lower ρ, since the impact of the NLoS component of the channel $\boldsymbol{g}$, a disturbance to the DOA estimation, is diminished.

IV. Simulation Results and Discussion

This section evaluates the effective ergodic achievable rate performance of the proposed DOA-based hybrid RIS phase-shift design. The three benchmark systems are presented for the performance comparisons. Specifically, two conventional passive RIS systems proposed in [7], [8], and one CE-based hybrid RIS system [13] are presented as follows:

- *Basic passive RIS*: In the passive RIS systems, the BS estimates the instantaneous CSI and calculates the phase-shift vector using the modified block coordinate descent method [4]. The phase-shift information is then fed back to the passive RIS. Here, the CE overhead is proportional to N because the channel coefficient corresponding to each passive RIS element should be estimated at the BS [7]. Therefore, the CE overhead is NT_s. In addition, the phase-shift feedback requires $b_{\mathrm{f}}N$ bits, where b_{f} is the quantization resolution [7]. The feedback information is transmitted through a designated feedback channel between the BS and the passive RIS controller with a single antenna. Here, the feedback channel $\boldsymbol{h}_{\mathrm{f}} \in \mathbb{C}^{M\times 1}$ is modeled as a Rician channel with a Rician factor of 10 dB and the path loss of $\eta_{\mathrm{B,R}}$, providing an achievable rate $R_{\mathrm{f}} = b_{\mathrm{f}}\log_2\left(1 + P_{\mathrm{B}}\sigma_{\mathrm{f}}^{-2}\|\boldsymbol{h}_{\mathrm{f}}\|^2\right)$ where b_{f} and σ_{f}^2 denote the feedback bandwidth and the AWGN at the RIS controller, respectively. Accordingly, the phase-shift feedback overhead is $b_{\mathrm{f}}N/R_{\mathrm{f}}$ [7].
- *Tensor decomposition (TD)-based passive RIS*: To reduce the amount of feedback information, [8] proposed a factorization approach based on TD. The phase-shift vector is factorized by a factor of L, thereby demanding a total $b_{\mathrm{f}}\sum_{l=1}^{L}N_l$ bits of feedback, where N_l denotes the length of the lth factorized vector. Therefore, the phase-shift feedback overhead becomes $\left(b_{\mathrm{f}}\sum_{l=1}^{L}N_l\right)/R_{\mathrm{f}}$ [8].
- *CE-based hybrid RIS*: The CE-based hybrid RIS calculates the phase-shift vector as $\hat{\boldsymbol{g}} \odot \boldsymbol{b}^*(\phi)$ at the hybrid RIS itself, where $\hat{\boldsymbol{g}}$ is the least squares estimate of the RIS-UE channel and the DOA ϕ is assumed to be known at the hybrid RIS. Therefore, phase-shift feedback is not required, unlike in passive RIS systems. In the CE-based hybrid RIS system, K pilot symbols are required to estimate the channel $\boldsymbol{g}$. Subsequently, another single pilot symbol is used to estimate the cascaded channel and perform MRT at the BS. Consequently, the CE overhead is $(K+1)T_s$.

TABLE I
Overhead Models of Passive and Hybrid RIS Systems.

Overhead models	Passive RIS		Hybrid RIS	
	Basic [7]	TD-based [8]	CE-based [13]	DOA-based Proposed
Channel estimation (τ_{c})	NT_s	NT_s	$(K+1)Ts$	T_s
DOA estimation (τ_{d})	-	-	-	QT_s
Phase-shift feedback (τ_{f})	$\frac{b_{\mathrm{f}}N}{R_{\mathrm{f}}}$	$\frac{b_{\mathrm{f}}\sum_{l=1}^{L}N_l}{R_{\mathrm{f}}}$	-	-

In summary, the overhead τ in (5) is given by the sum of the following three types of overhead models: 1) CE overhead τ_{c}, 2) DOA estimation overhead τ_{d}, 3) phase-shift feedback overhead τ_{f}, as listed in Table I.

The simulation parameters are the same as those used for

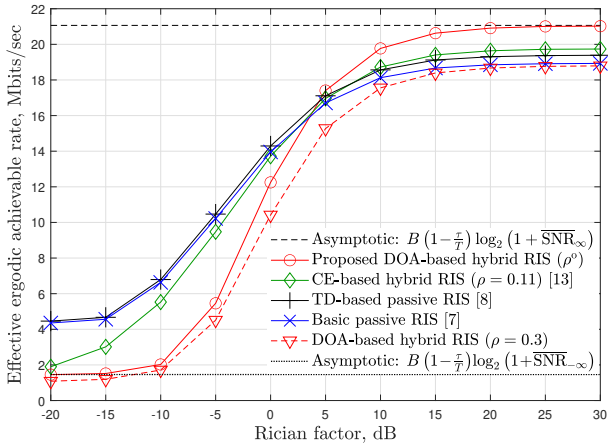

Fig. 4. Effective ergodic achievable rate over Rician factors, κ_F and κ_g, where $N = 80$ and $T = 2$ ms.

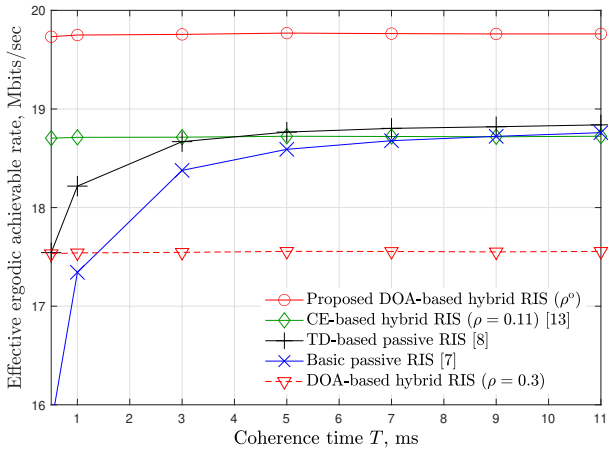

Fig. 5. Effective ergodic achievable rate over coherence time, T, when $\kappa_F = 10$ dB, $\kappa_g = 10$ dB, and $N = 80$.

the simulation in Section III-B. Additionally, the simulation parameters for the passive RIS systems are as follows: $b_{\mathrm{f}} = 12$, $L=3$, $N_1=\frac{N}{4}$, $N_2=N_3=2$, $\sigma_{\mathrm{f}}^2=-117$ dBm, $b_{\mathrm{f}}=0.5$ MHz; and the data transmission bandwidth B for passive RIS is set to 4.5 MHz. The pilot length K in the CE-based hybrid RIS system is set to two in the simulations. For simplicity, the performance of the passive RIS systems is evaluated using their upper bound, assuming that the CE error at the BS, phase-shift vector quantization noise in feedback, and feedback information reconstruction errors in TD feedback are neglected.

Fig. 4 shows the effective ergodic achievable rate over the Rician factors κ_F and κ_g, when $N = 80$ and $T = 2$ ms. The effective ergodic achievable rate increases with the Rician factors. Furthermore, it is observed that the performance of the proposed hybrid RIS system follows asymptotic analysis in Remark 2: it converges to $B\left(1-\frac{\tau}{T}\right)\log_2\left(1+\overline{\mathrm{SNR}}_{\infty}\right)$ as Rician factors increase, and to $B\left(1-\frac{\tau}{T}\right)\log_2\left(1+\overline{\mathrm{SNR}}_{-\infty}\right)$ as Rician factors decrease. When the Rician factors are smaller than 0 dB, i.e., in the NLoS dominant channel, both passive RIS and CE-based hybrid RIS systems outperform the proposed DOA-based hybrid RIS system because they exploit both LoS and NLoS components in their phase-shift design. However, the

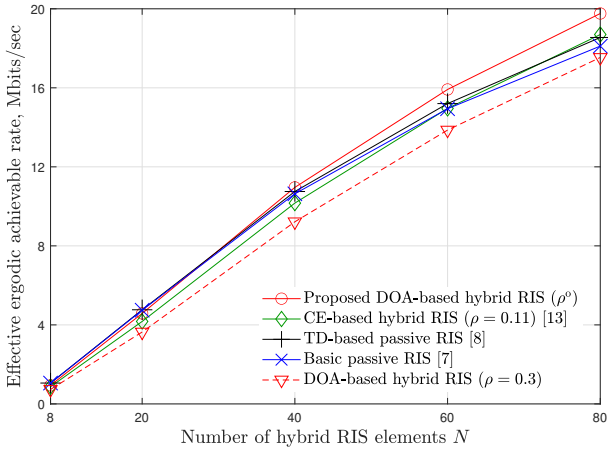

Fig. 6. Effective ergodic achievable rate over number of RIS elements, N, when $\kappa_F = 10$ dB, $\kappa_g = 10$ dB, and $T = 2$ ms.

performance of the proposed hybrid RIS system is significantly improved, outperforming the benchmark systems when the Rician factors exceed 5 dB. This is because the portion of the reflection power, i.e., $1-\rho^{\mathrm{o}}$, increases, as verified in Fig. 3.

In Fig. 5, the effective ergodic achievable rate is evaluated over the coherence time T, when $\kappa_F = 10$ dB, $\kappa_g = 10$ dB, and $N = 80$. It is verified that the proposed DOA-based hybrid RIS phase-shift method outperforms all benchmark systems and maintains its performance almost constant regardless of T. However, the performance of passive RIS systems is severely degraded when T is small. This is because the ratio of data transmission time to coherence time, $1-\frac{\tau}{T}$, remains constant in the hybrid RIS system, even as T decreases, since τ is small. On the other hand, in the passive RIS system, $1-\frac{\tau}{T}$ decreases drastically as T decreases because τ is relatively high. As T increases, the impact of overhead becomes negligible, i.e., $\lim_{T\to\infty}\frac{\tau}{T}=0$. Although the passive RIS systems utilizing both LoS and NLoS components achieve a higher received SNR than the DOA-based hybrid RIS systems, the proposed DOA-based hybrid RIS system still outperforms the passive RIS systems because hybrid RIS systems can use a wider bandwidth for data transmission. Specifically, hybrid RIS systems utilize the entire bandwidth for data transmission, whereas passive RIS systems should dedicate a portion of the bandwidth, B_{f}, for the feedback link. Moreover, the proposed DOA-based hybrid RIS system outperforms the CE-based hybrid RIS system as the required sensing factor for accurate DOA estimation is lower than that of CE, thereby utilizing a higher reflecting power $1-\rho$ in the downlink data transmission.

Fig. 6 evaluates the effective ergodic achievable rate over the number of RIS elements, when $\kappa_F = 10$ dB, $\kappa_g = 10$ dB, and $T = 2$ ms. The proposed hybrid RIS system outperforms the CE-based hybrid RIS system across the entire range of the number of RIS elements, and surpasses passive RIS systems when $N \geq 40$. Moreover, performance improvement compared to the passive RIS systems increases as N increases, showing an 1.2 Mbits/sec performance gap compared to TD-based passive RIS system when $N = 80$. This performance improvement can be achieved through a significant reduction in signaling

overhead in the proposed hybrid RIS system compared to the passive RIS systems. Specifically, the amount of signaling overhead in the passive RIS system increases as the number of RIS elements increases. In contrast, the signaling overhead of the proposed hybrid RIS system remains constant.

V. Conclusion

This study proposed a DOA-based hybrid RIS phase-shift design and derived a closed-form solution for the optimal sensing factor that balances the DOA estimation accuracy and signal-reflecting power. The proposed method significantly reduces the CE and feedback overhead by directly calculating the RIS phase-shift vector at the hybrid RIS, thereby maximizing the effective ergodic achievable rate. The simulation results verified that the proposed hybrid RIS system with optimal sensing factor outperforms passive RIS and CE-based hybrid RIS systems, especially when the LoS channel is dominant and the RIS has a large number of elements.

Appendix
Derivation of Approximated Average Received SNR in (13)

By substituting (8) into (7a), the average received SNR of the hybrid RIS system is rewritten as follows:

$$\mathbb{E}\left[\mathrm{SNR}\left(\rho,\boldsymbol{\psi}^{\circ}\left(\hat{\theta}\right)\right)\right]=\zeta(1-\rho)\left(\kappa_F\kappa_g\beta+N(\kappa_F+\kappa_g+1)\right),$$

where $\beta \triangleq \mathbb{E}\left[\left|\boldsymbol{b}^H\left(\hat{\theta}\right)\boldsymbol{b}\left(\theta\right)\right|^2\right]$. Since $\hat{\theta}\sim\mathcal{N}\left(\theta,\epsilon\right)$ [19], it can be approximated as follows:

$$\begin{aligned}
\beta &= \mathbb{E}_{\hat{\theta}}\left[\left|\sum_{n=1}^{N} e^{-j(n-1)2\pi\Delta\left(\sin\theta-\sin\hat{\theta}\right)/\lambda}\right|^2\right]\\
&\overset{(b)}{\approx} \mathbb{E}_{\delta}\left[\left|\frac{1-e^{-jN2\pi\Delta\delta/\lambda}}{1-e^{-j2\pi\Delta\delta/\lambda}}\right|^2\right]=\mathbb{E}_{\delta}\left[\frac{\sin^2\left(N\pi\Delta\delta/\lambda\right)}{\sin^2\left(\pi\Delta\delta/\lambda\right)}\right]\\
&\overset{(c)}{\approx} \mathbb{E}_{\delta}\left[\left(\frac{\sin\left(N\pi\Delta\delta/\lambda\right)}{\pi\Delta\delta/\lambda}\right)^2\right]\\
&= \frac{N^2}{\sqrt{2\pi\epsilon\cos^2\theta}}\int_{-\infty}^{\infty}\left(\frac{\sin\left(N\pi\Delta\delta/\lambda\right)}{N\pi\Delta\delta/\lambda}\right)^2 e^{-\frac{\delta^2}{2\epsilon\cos^2\theta}}\,d\delta\\
&\overset{(d)}{=} \frac{N^2}{2\pi}\int_{-\infty}^{\infty}\frac{\mathrm{tri}\left(\frac{\omega\lambda}{2\pi\Delta N}\right)}{\Delta N/\lambda}e^{-\frac{\omega^2\epsilon\cos^2\theta}{2}}\,d\omega\\
&= N\sqrt{\tfrac{\pi(N^3-N)}{3\Omega}}\mathrm{erf}\left(N\sqrt{\tfrac{3\Omega}{N^3-N}}\right)+\tfrac{N^3-N}{3\Omega}\left(e^{-\frac{3N\Omega}{N^2-1}}-1\right)\\
&\overset{(e)}{\approx} N^2-N\Omega/2,
\end{aligned}$$

where $\Omega=\frac{1}{\kappa_g}+\frac{\sigma_w^2(\kappa_g+1)}{\rho\eta_{\mathrm{R,U}}QP_{\mathrm{U}}\kappa_g}$. Here, (b) comes from the first-order Taylor approximation of $\sin\theta-\sin\hat{\theta}\approx\left(\theta-\hat{\theta}\right)\cos\theta=\delta$, thereby $\delta\sim\mathcal{N}\left(0,\epsilon\cos^2\theta\right)$. The approximation (c) comes from the first-order Taylor approximation of $\sin x\approx x$, and (d) is derived by using Parseval's theorem. Furthermore, (e) is from the approximation $\frac{N^2}{N^2-1}\approx 1$, the third-order Taylor approximation of error function, i.e., $\mathrm{erf}\left(x\right)\approx 2/\sqrt{\pi}\left(x-x^3/3\right)$, and second-order Taylor approximation of exponential function, i.e., $e^x-1\approx x+x^2/2$. Therefore, the approximated average received SNR is derived as (13).

References

[1] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1546–1577, 3rd Quart., 2021.

[2] J. Yuan, Y.-C. Liang, J. Joung, G. Feng, and E. G. Larsson, "Intelligent reflecting surface-assisted cognitive radio system," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 675–687, Jan. 2021.

[3] J. Wang, Y.-C. Liang, J. Joung, X. Yuan, and X. Wang, "Joint beamforming and reconfigurable intelligent surface design for two-way relay networks," *IEEE Trans. Commun.*, vol. 69, no. 8, pp. 5620–5633, Aug. 2021.

[4] J. Kim, J. Choi, J. Joung, and Y.-C. Liang, "Modified block coordinate descent method for intelligent reflecting surface-aided space-time line coded systems," *IEEE Wireless Commun. Lett.*, vol. 11, no. 9, pp. 1820–1824, Sep. 2022.

[5] Y. Zhu, J. Zhang, E. Shi, Z. Liu, C. Yuen, D. Niyato, and B. Ai, "Joint precoding and phase shift design for RIS-aided cell-free massive MIMO with heterogeneous-agent trust region policy," *IEEE Trans. Veh. Technol.*, vol. 74, no. 1, pp. 1794–1799, Jan. 2025.

[6] J. Kim, H.-G. Lee, J. Joung, and Y. Zeng, "Coverage partitioning for aerial intelligent reflecting surface phase mapping methods," *IEEE Trans. Veh. Technol.*, 2025, (early access articles).

[7] A. Zappone, M. Di Renzo, F. Shams, X. Qian, and M. Debbah, "Overhead-aware design of reconfigurable intelligent surfaces in smart radio environments," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 126–141, Jan. 2021.

[8] B. Sokal, P. R. B. Gomes, A. L. F. de Almeida, B. Makki, and G. Fodor, "Reducing the control overhead of intelligent reconfigurable surfaces via a tensor-based low-rank factorization approach," *IEEE Trans. Wireless Commun.*, vol. 22, no. 10, pp. 6578–6593, Oct. 2023.

[9] Z. Chen, J. Tang, X. Y. Zhang, D. K. C. So, S. Jin, and K.-K. Wong, "Hybrid evolutionary-based sparse channel estimation for IRS-assisted mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 1586–1601, Mar. 2022.

[10] D. Gunasinghe and G. A. A. Baduge, "Achievable rate analysis for multi-cell RIS-aided massive MIMO with statistical CSI-based optimizations," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 8117–8135, Aug. 2024.

[11] Y. Han, W. Tang, S. Jin, C.-K. Wen, and X. Ma, "Large intelligent surface-assisted wireless communication exploiting statistical CSI," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8238–8242, Aug. 2019.

[12] G. C. Alexandropoulos, N. Shlezinger, I. Alamzadeh, M. F. Imani, H. Zhang, and Y. C. Eldar, "Hybrid reconfigurable intelligent metasurfaces: Enabling simultaneous tunable reflections and sensing for 6G wireless communications," *IEEE Veh. Technol. Mag.*, vol. 19, no. 1, pp. 75–84, Mar. 2024.

[13] J. Choi and J. H. Cho, "A joint optimization of pilot and phase shifts in uplink channel estimation for hybrid RIS-aided multi-user communication systems," *IEEE Trans. Veh. Technol.*, vol. 73, no. 4, pp. 5197–5212, Apr. 2024.

[14] G. Zhou, C. Pan, H. Ren, P. Popovski, and A. L. Swindlehurst, "Channel estimation for RIS-aided multiuser millimeter-wave systems," *IEEE Trans. Signal Process.*, vol. 70, pp. 1478–1492, 2022.

[15] Z. Liu, H. Zhang, T. Huang, F. Xu, and Y. C. Eldar, "Hybrid RIS-assisted MIMO dual-function radar-communication system," *IEEE Trans. Signal Process.*, vol. 72, pp. 1650–1665, 2024.

[16] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[17] Z. Chen, G. Gokeda, and Y. Yu, *Introduction to Direction-of-Arrival Estimation*. Norwood, MA, USA: Artech House, 2010.

[18] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: Unification, simplification, and observations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 4, pp. 1170–1184, Oct. 1993.

[19] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 5, pp. 720–741, May 1989.

[20] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 9)," 3GPP, Tech. Rep. TR 36.814 V9.2.0, Mar. 2017.