Configuring Intelligent Reflecting Surface with Performance Guarantees: Blind Beamforming

Shuyi Ren, *Student Member, IEEE*, Kaiming Shen, *Member, IEEE*,
Yaowen Zhang, *Student Member, IEEE*, Xin Li, Xin Chen, Zhi-Quan Luo, *Fellow, IEEE*

Abstract—This work gives a blind beamforming strategy for intelligent reflecting surface (IRS), aiming to boost the received signal-to-noise ratio (SNR) by coordinating phase shifts across reflective elements in the absence of channel information. While the existing methods of IRS beamforming typically first estimate channels and then optimize phase shifts, we propose a conditional sample mean based statistical approach that explores the wireless environment via random sampling without performing any channel estimation. Remarkably, the new method just requires a polynomial number of random samples to yield an SNR boost that is quadratic in the number of reflective elements, whereas the standard random-max sampling algorithm can only achieve a linear boost under the same condition. Moreover, we gain additional insight into blind beamforming by interpreting it as a least squares problem. Field tests demonstrate the significant advantages of the proposed blind beamforming algorithm over the benchmark algorithms in enhancing wireless transmission.

Index Terms—Intelligent reflecting surface (IRS), blind beamforming without channel estimation, statistical approach, signal-to-noise (SNR) boost, field tests.

I. INTRODUCTION

THE MAIN PROBLEM in configuring intelligent reflecting surface (IRS) is that of coordinating phase shifts across reflective elements, namely beamforming. The conventional beamforming paradigm of first estimating channels and then optimizing phase shifts has been developed and analyzed extensively for IRS in the literature to date, but there is little work addressing the challenge of *blind beamforming* without channel estimation. This paper devises a statistical approach to blind beamforming in an IRS-assisted system, which has provable performance of boosting the signal-to-noise ratio (SNR) quadratically. Aside from theoretical justifications, field tests in a real-world environment validate the practical effectiveness of the proposed blind beamforming algorithm.

With IRS treated as an unsourced multi-antenna device, the existing studies mostly follow the traditional model-driven beamforming approach that requires channel informa-

tion. However, channel acquisition for IRS poses formidable challenges in engineering practice, mainly in the following three respects:

i. Each reflected channel alone can be easily overwhelmed by the much stronger background channel and noise;
ii. It entails modifying the current networking protocol to incorporate the channel estimation for IRS;
iii. Channel estimation would increase the costs of deploying IRS.

Actually, the authors of [1] have realized the above issues as well when building up their prototype. Similar to [1], this work pursues a blind treatment of the IRS beamforming problem in order to sidestep channel estimation. In particular, one must distinguish between the neural networks based approach in [2]–[6] and the statistics based approach with performance guarantees as considered in this paper, although both are data-driven.

There has been an explosion of research interests in the IRS beamforming over the past few years. According to the survey in [7], semidefinite relaxation (SDR) [8] and fractional programming (FP) [9], [10] are among the most popular standard optimization tools in this area, e.g., [11]–[17] find SDR useful because of the quadratic form of the IRS beamforming problem, while [18]–[25] rely on FP to deal with the signal-to-noise-plus-interference ratio (SINR) terms. Other approaches include successive convex approximation [26], [27], minorization maximization [28], and alternating direction method of multipliers (ADMM) [17], [29].

A real technical challenge facing the above methods is that their performance all heavily depends upon the accurate knowledge of channels—which is quite costly to acquire in practice because of the three reasons as noted earlier. A simple idea from [12] is to sequentially activate each individual reflective element at a time so as to measure the corresponding reflected channel. Apart from being inefficient, such channel estimation is prone to large error since each reflected channel alone can be much weaker than the background channel. Some more recent works promote this ON-OFF policy by activating a group of reflective elements simultaneously. Another line of studies [30]–[33] focus on the pilot sequence design based on the discrete Fourier transform (DFT) matrix; the Cramér-Rao lower bound to the least squares channel estimation can be achieved under certain assumptions [31]. Moreover, [34]–[37] view the channel estimation for IRS as a compressed sensing problem. Notice that all the above methods are only tested in simulations with artificially generated channels, but

the wireless environments in the real world can be far more complex as observed in [1] and in our field tests.

The idea of configuring IRS without channel information, i.e., blind beamforming, has worked its way into a new frontier [1], [6]. As opposed to [2]–[5], [35], [38] that apply neural networks to either channel estimation or beamforming for IRS, the recent work [6] suggests merging the two parts into directly learning how to perform beamforming based on the received signal. It is argued in [6] that such combined deep learning approach is capable of extracting more pertinent information from the raw data. Differing from these neural net-based methods, the beamforming algorithm proposed in [1] exploits the statistics of the received signal. Inspired by [1], this work also pursues a statistical approach and shows that the newly proposed algorithm provides better performance guarantees.

The primary idea behind the proposed blind beamforming algorithm is to utilize the conditional sample mean of the received signal power. We examine how the number of random samples affects the performance of blind beamforming. It is shown that the proposed algorithm achieves an SNR boost that is quadratic in the number of reflective elements, by merely using a polynomial number of random samples, whereas the standard random-max sampling method can only achieve a linear boost under the same condition. We further extend this blind beamforming scheme to account for a general utility function.

The rest of the paper is organized as follows. Section II describes the system model and problem formulation. Section III discusses the proposed blind beamforming algorithm, along with the comparison to the standard random-max sampling algorithm and the connection to the least squares problems. Section IV is devoted to an enhanced conditional sample mean method. Section V presents the field test results. Section VI concludes the paper.

Throughout the paper, we use the bold lower-case letter to denote a vector, the bold upper-case letter a matrix, and the calligraphy upper-case letter a set. For a matrix A, A^T refers to the transpose, A^H the conjugate transpose, and A^{-1} the inverse. For a vector a, $\|a\|$ refers to the Euclidean norm, a^T the transpose, and a^H the conjugate transpose. The cardinality of a set $\mathcal{Q}$ is denoted as $|\mathcal{Q}|$. The set of real numbers is denoted as $\mathbb{R}$. The set of complex numbers is denoted as $\mathbb{C}$. For a complex number u, $\mathrm{Re}\{u\}$, $\mathrm{Im}\{u\}$, and $\mathrm{Arg}(u)$ refer to the real part, the imaginary part, and the principal argument of u, respectively. For an event $\mathcal{E}$, let $\mathbb{P}\{\mathcal{E}\}$ be its probability and let $\mathcal{E}^c$ be its complement. Let $\mathbb{E}[X]$ be the expectation of the random variable X, and let $\hat{\mathbb{E}}[X]$ be the sample mean. We use the Bachmann-Landau notation: $f(n) = O(g(n))$ if there exists some $c > 0$ such that $|f(n)| \le cg(n)$ for n sufficiently large; $f(n) = o(g(n))$ if there exists some $c > 0$ such that $|f(n)| < cg(n)$ for n sufficiently large; $f(n) = \Omega(g(n))$ if there exists some $c > 0$ such that $f(n) \ge cg(n)$ for n sufficiently large; $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$ hold simultaneously.

II. System Model

Consider a pair of transmitter and receiver, along with an IRS that facilitates the data transmission between them. The IRS consists of N passive reflective elements. Let $h_n \in \mathbb{C}$, $n = 1, \ldots, N$, be the cascaded channel from the transmitter to the receiver that is induced by the nth reflective element; let $h_0 \in \mathbb{C}$ be the superposition of all those channels from the transmitter to the receiver that are not related to the IRS, namely the *background channel*. Each channel can be rewritten in an exponential form as

$$h_n = \beta_n e^{j\alpha_n}, \ n = 0, \ldots, N, \tag{1}$$

with the magnitude $\beta_n \in (0, 1)$ and the phase $\alpha_n \in [0, 2\pi)$.

Denote the IRS beamformer as $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N)$, where each $\theta_n \in [0, 2\pi)$ refers to the phase shift of the nth reflective element. The choice of each θ_n is restricted to the discrete set

$$\Phi_K = \{\omega, 2\omega, \ldots, K\omega\} \tag{2}$$

with the distance parameter

$$\omega = \frac{2\pi}{K}. \tag{3}$$

Let $X \in \mathbb{C}$ be the transmit signal with the mean power P, i.e., $\mathbb{E}[|X|^2] = P$. The received signal $Y \in \mathbb{C}$ is given by

$$Y = \left(h_0 + \sum_{n=1}^{N} h_n e^{j\theta_n}\right) X + Z, \tag{4}$$

where an i.i.d. random variable $Z \sim \mathcal{CN}(0, \sigma^2)$ models the additive thermal noise. The received SNR can be computed as

$$\mathsf{SNR} = \frac{\mathbb{E}[|Y - Z|^2]}{\mathbb{E}[|Z|^2]} \tag{5a}$$

$$= \frac{P\left|\beta_0 e^{j\alpha_0} + \sum_{n=1}^{N} \beta_n e^{j(\alpha_n + \theta_n)}\right|^2}{\sigma^2}. \tag{5b}$$

The baseline SNR without IRS is

$$\mathsf{SNR}_0 = \frac{P\beta_0^2}{\sigma^2}. \tag{6}$$

The *SNR boost* is defined as

$$f(\boldsymbol{\theta}) = \frac{\mathsf{SNR}}{\mathsf{SNR}_0} \tag{7a}$$

$$= \frac{1}{\beta_0^2}\left|\beta_0 e^{j\alpha_0} + \sum_{n=1}^{N} \beta_n e^{j(\alpha_n + \theta_n)}\right|^2. \tag{7b}$$

We seek the optimal $\boldsymbol{\theta}$ to maximize the SNR boost:

$$\underset{\boldsymbol{\theta}}{\text{maximize}} \quad f(\boldsymbol{\theta}) \tag{8a}$$

$$\text{subject to} \quad \theta_n \in \Phi_K, \ \forall n = 1, \ldots, N. \tag{8b}$$

The difficulties of the above problem are two-fold: first, it is numerically difficult to optimize the discrete variable $\boldsymbol{\theta}$ for large N; second, it is costly to acquire the channel information $\{h_0, \ldots, h_N\}$ in practice.

III. Blind Beamforming

We begin with the notion of the *average per element reflection gain*:

$$\delta^2 = \frac{1}{N}\sum_{n=1}^{N} \beta_n^2. \tag{9}$$

We can readily obtain an order upper bound on the achievable SNR boost as stated in the following proposition.

Proposition 1: The SNR boost is at most quadratic in the number of reflective elements, i.e.,

$$f(\boldsymbol{\theta}) = \frac{\delta^2}{\beta_0^2} \cdot O(N^2). \tag{10}$$

Proof: For the best-case scenario, we are able to align every $h_n e^{j\theta_n}$ with h_0 exactly; the resulting ideal SNR boost is $f(\boldsymbol{\theta}) = 1/\beta_0^2 \cdot (\beta_0 + \sum_{n=1}^{N} \beta_n)^2 \le (N+1)/\beta_0^2 \cdot (\beta_0^2 + N\delta^2) = \delta^2/\beta_0^2 \cdot O(N^2)$. ∎

As the main result of this section, we show that the above upper bound can be achieved by a conditional sample mean based statistical method with a polynomial number of random samples, whereas the standard method of random-max sampling can only yield a linear boost.

A. Random-Max Sampling Algorithm

We start with a natural baseline method called the *Random-Max Sampling (RMS)* algorithm. Its central idea is to choose the best out of T random samples of $\boldsymbol{\theta}$. Use $t = 1, \ldots, T$ to index the random sample; let θ_{nt} be the phase shift of the nth reflective element for the tth random sample and let $\boldsymbol{\theta}_t = (\theta_{1t}, \ldots, \theta_{Nt})$. Each θ_{nt} is uniformly drawn from Φ_K in an i.i.d. fashion across $n = 1, \ldots, N$ and $t = 1, \ldots, T$. The received signal of the tth random sample is given by

$$Y_t = \left(h_0 + \sum_{n=1}^{N} h_n e^{j\theta_{nt}} \right) X_t + Z_t, \tag{11}$$

where the transmit signal X_t has a fixed power P. The random-max sampling algorithm decides $\boldsymbol{\theta}$ according to the received signal strength, i.e.,

$$\boldsymbol{\theta}^{\text{RMS}} = \boldsymbol{\theta}_{t_0} \text{ where } t_0 = \arg\max_{1 \le t \le T} |Y_t|^2. \tag{12}$$

It can be seen that the performance of random-max sampling is closely related to the number of random samples T. We establish scaling laws of the SNR boost in the following theorem to characterize this relationship.

Theorem 1: Consider T independent and uniformly distributed samples of $\boldsymbol{\theta}$. The expected SNR boost achieved by the random-max sampling algorithm has the following order bounds:

$$\mathbb{E}[f(\boldsymbol{\theta}^{\text{RMS}})] = \frac{\delta^2}{\beta_0^2} \cdot \Omega(N \log T) \text{ if } T = o(\sqrt{N}), \tag{13}$$

$$\mathbb{E}[f(\boldsymbol{\theta}^{\text{RMS}})] = \frac{\delta^2}{\beta_0^2} \cdot O(N \log T) \text{ in general}, \tag{14}$$

where the expectation is taken over random samples of $\boldsymbol{\theta}$.

Proof: See Appendix A. ∎

We immediately obtain a tight order bound by combining (13) and (14), as stated below.

Corollary 1: If $T = o(\sqrt{N})$, the SNR boost achieved by the random-max sampling algorithm satisfies $\mathbb{E}[f(\boldsymbol{\theta}^{\text{RMS}})] = (\delta^2/\beta_0^2) \cdot \Theta(N \log T)$.

Because T is polynomial in N in most practical cases, the scaling rate $\Theta(N \log T) \approx \Theta(N)$, and hence the random-max sampling algorithm gives an approximately linear SNR

Algorithm 1 Conditional Sample Mean (CSM)

1: **input:** Φ_K, N, T
2: **for** $t = 1, 2, \ldots, T$ **do**
3:     generate $\boldsymbol{\theta}_t = (\theta_{1t}, \ldots, \theta_{Nt})$ i.i.d. based on Φ_k
4:     measure received signal power $|Y_t|^2$ with $\boldsymbol{\theta}_t$
5: **end for**
6: **for** $n = 1, 2, \ldots, N$ **do**
7:     **for** $k = 1, 2, \ldots, K$ **do**
8:         compute $\hat{\mathbb{E}}[|Y_t|^2 | \theta_n = \phi_k]$ according to (16)
9:     **end for**
10:     decide θ_n^{CSM} according to (17)
11: **end for**
12: **output:** $\boldsymbol{\theta}^{\text{CSM}} = (\theta_1^{\text{CSM}}, \ldots, \theta_N^{\text{CSM}})$

---

boost. The weakness of random-max sampling lies in that it solely uses the "best" random sample while neglecting the rest. In contrast, the conditional sample mean algorithm as introduced in the next subsection makes full use of all the random samples, thereby leading to a quadratic SNR boost.

B. Conditional Sample Mean Algorithm

We still generate a total of T random samples of $\boldsymbol{\theta}$ in an i.i.d. fashion. Use $\mathcal{Q}_{nk} \subseteq \{1, \ldots, T\}$ to denote the set of all those random samples with $\theta_{nt} = \phi_k$, i.e.,

$$\mathcal{Q}_{nk} = \{t : \theta_{nt} = \phi_k\}, \ \forall(n, k). \tag{15}$$

Compute the following conditional sample mean of the received signal power $|Y_t|^2$ with respect to each (n, k) pair:

$$\hat{\mathbb{E}}[|Y|^2 | \theta_n = \phi_k] = \frac{1}{|\mathcal{Q}_{nk}|} \sum_{t \in \mathcal{Q}_{nk}} |Y_t|^2, \ \forall(n, k). \tag{16}$$

Intuitively, the conditional sample mean $\hat{\mathbb{E}}[|Y_t|^2 | \theta_n = \phi_k]$ reflects the average performance of letting $\theta_n = \phi_k$. Thus, it is natural to choose each θ_n according to the conditional sample mean, i.e.,

$$\theta_n^{\text{CSM}} = \arg\max_{\varphi \in \Phi_K} \hat{\mathbb{E}}[|Y|^2 | \theta_n = \varphi]. \tag{17}$$

The above method is referred to as the *Conditional Sample Mean (CSM)* algorithm and is summarized in Algorithm 1. We examine its performance in the following theorem.

Theorem 2: Consider the same setting as in Theorem 1. When $K \ge 3$, the expected SNR boost achieved by the conditional sample mean algorithm has a tight order bound:

$$\mathbb{E}[f(\boldsymbol{\theta}^{\text{CSM}})] = \frac{\delta^2}{\beta_0^2} \cdot \Theta(N^2) \text{ if } T = \Omega(N^2(\log N)^3), \tag{18}$$

where the expectation is taken over random samples of $\boldsymbol{\theta}$.

Proof: See Appendix B. ∎

Remark 1: In proving Theorem 2, we also show that $\boldsymbol{\theta}^{\text{CSM}} = \boldsymbol{\theta}^{\text{CPP}}$ with high probability if $T = \Omega(N^2(\log N)^3)$.

C. Connection to Least Squares Problems

We show that the conditional sample mean algorithm can be interpreted as solving a least squares problem. With each

θ_n independently and uniformly distributed in Φ_K, the expectation of the received signal power is given by

$$\mathbb{E}[|Y|^2] = \beta_0^2 P + \sum_{m=1}^{N} \beta_m^2 P + \sigma^2, \quad (19)$$

and the mean received signal power conditioned on $\theta_n = \phi_i$ is given by

$$\mathbb{E}[|Y|^2|\theta_n = \phi_k] = P|h_0 + h_n e^{j\phi_k}|^2 + \sum_{m \neq n} \beta_m^2 P + \sigma^2. \quad (20)$$

Defining a new variable[1] θ_n^∞ as

$$\theta_n^\infty = \alpha_0 - \alpha_n, \ n = 1, \ldots, N, \quad (21)$$

we can write the difference between the above two expectations as a function of θ_n^∞:

$$J_{nk}(\theta_n^\infty) = \mathbb{E}[|Y|^2|\theta_n = \phi_k] - \mathbb{E}[|Y|^2]$$
$$= 2\beta_0 \beta_n P \cos(\phi_k - \theta_n^\infty). \quad (22)$$

We also evaluate $J_{nk}(\theta_n^\infty)$ based on random samples as

$$\hat{J}_{nk} = \frac{1}{|\mathcal{Q}_{nk}|} \sum_{t \in \mathcal{Q}_{nk}} |Y_t|^2 - \frac{1}{T} \sum_{t=1}^{T} |Y_t|^2. \quad (23)$$

We then estimate θ_n^∞ by minimizing the gap between J_{nk} and $\hat{J}_{nk}$. In particular, for the *least squared-max* metric

$$\sum_{n=1}^{N} \left| \max_{1 \leq k \leq K} \{J_{nk}(\theta_n^\infty)\} - \max_{1 \leq k' \leq K} \{\hat{J}_{nk'}\} \right|^2,$$

the optimal estimate is given by

$$\hat{\theta}_n^\infty = \phi_{k_0} \text{ where } k_0 = \arg \max_{1 \leq k \leq K} \hat{J}_{nk}. \quad (24)$$

Thus, the best estimation of θ_n^∞ is to maximize the conditional sample mean $\hat{J}_{nk}$ for each n, which can be recognized as the conditional sample mean algorithm.

Moreover, if the distortion metric is

$$\sum_{n=1}^{N} \sum_{k=1}^{K} |J_{nk}(\theta_n^\infty) - \hat{J}_{nk}|^2,$$

then the optimal estimate of θ_n^∞ becomes

$$\hat{\theta}_n^\infty = \begin{cases} -\arctan \dfrac{F_n}{E_n} + \dfrac{\pi}{2}, & \text{if } E_n \geq 0, \\[2mm] -\arctan \dfrac{F_n}{E_n} - \dfrac{\pi}{2}, & \text{if } E_n < 0, \end{cases} \quad (25)$$

where

$$E = \sum_{k=0}^{K-1} \hat{J}_{nk} \sin(\phi_k) \quad \text{and} \quad F = \sum_{k=1}^{K} \hat{J}_{nk} \cos(\phi_k). \quad (26)$$

We could let $\theta_n^{\text{CSM}} = \hat{\theta}_n^\infty$ and thereby devise a variation of the conditional sample mean algorithm.

Actually, we can obtain many variations of the conditional sample mean algorithm by defining the distortion between J_{nk} and $\hat{J}_{nk}$ differently. Further, these variations can be readily

---

[1]In Section IV we show that θ_n^∞ is the optimal phase shift as $K \to \infty$.

extended to the enhanced conditional sample mean algorithm.

D. Generalized Conditional Sample Mean Algorithm

Our target is the SNR boost thus far. Nevertheless, sometimes the SNR boost cannot be quantified as a scalar value, e.g., for the multiple-input multiple-ouput (MIMO) channels, and sometimes the SNR boost may not even be the main consideration, e.g., for the ultra-reliable communications.

We propose extending the conditional sample mean algorithm to a general utility function of $\boldsymbol{\theta}$. More specifically, we still generate a total of T random samples, and use $U_t \in \mathbb{R}$ to denote the performance utility of each random sample $\boldsymbol{\theta}_t$, $t = 1, \ldots, T$. For instance, we can let U_t be the spectral efficiency in order to maximize the throughput of MIMO channel.

The conditional sample mean is now computed with respect to the performance utility for every (n, k) pair:

$$\hat{\mathbb{E}}[U|\theta_n = \phi_k] = \frac{1}{|\mathcal{Q}_{nk}|} \sum_{t \in \mathcal{Q}_{nk}} U_t. \quad (27)$$

Following Algorithm 1, we choose each θ_n with the highest conditional sample mean utility related to the nth reflective element, i.e.,

$$\theta_n = \arg \max_{\varphi \in \Phi_K} \hat{\mathbb{E}}[U|\theta_n = \varphi]. \quad (28)$$

The above changes can be immediately carried over to the enhanced conditional sample mean algorithm.

IV. ENHANCED CONDITIONAL SAMPLE MEAN METHOD

The discrete beamforming problem in (8) becomes the continuous as $K \to \infty$, in which case it is optimal to align every h with h_0 exactly by letting $\theta_n = \theta_n^\infty$. In the existing literature [39]–[45], a common way of discrete beamforming is to round the relaxed continuous solution θ_n^∞ to the closest point in the discrete set Φ_K, assuming that channels are known *a priori*, i.e.,

$$\theta_n^{\text{CPP}} = \arg \min_{\theta_n \in \Phi_K} |\theta_n - \theta_n^\infty|, \quad (29)$$

referred to as the *Closest Point Projection (CPP)* algorithm. Thus, the proposed conditional sample mean algorithm can be interpreted as performing the closest point projection with the estimated $\hat{\theta}_n^\infty$ in (24).

As shown in our recent work [46], the closest point projection algorithm has an approximation ratio of $\cos^2(\pi/K)$, while an enhanced closest point projection algorithm proposed in [46] guarantees a higher approximation ratio of $1/2 + \cos(\pi/K)/2$. Having used the conditional sample mean algorithm to mimic the closest point projection, we now wish to mimic the enhanced version.

Toward this end, we need to compute three solution candidates as discussed in [46]: $\boldsymbol{\theta}'$ with each $h_n e^{j\theta_n'} \in \mathcal{S}_1 \cup \mathcal{S}_2$, $\boldsymbol{\theta}''$ with each $h_n e^{j\theta_n''} \in \mathcal{S}_2 \cup \mathcal{S}_3$, and $\boldsymbol{\theta}'''$ with each $h_n e^{j\theta_n'''} \in \mathcal{S}_3 \cup \mathcal{S}_4$, where

$$\mathcal{S}_i = \left\{ u \in \mathbb{C} : \alpha_0 + \frac{(2-i)\omega}{2} \leq \text{Arg}(u) \leq \alpha_0 + \frac{(3-i)\omega}{2} \right\},$$
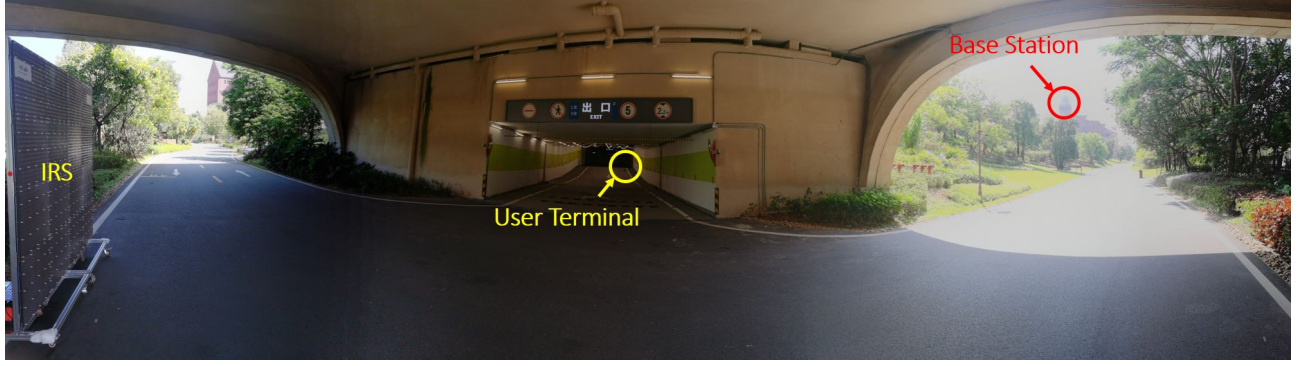$$\forall i = 1, 2, 3, 4. \quad (30)$$

Fig. 1. A panoramic view of the field test site. The base station is located on a 20-meter-high terrace while the user terminal is located inside an underground parking lot. The IRS is placed at the entrance of the parking lot. The IRS is approximately 250 meters away from the base station, and the user terminals are approximately 40 meters away from the IRS.

Notice that the second candidate $\boldsymbol{\theta}''$ is exactly the closest point projection solution, which can be recovered by the conditional sample mean algorithm.

It remains to recover the other two candidates $\boldsymbol{\theta}'$ and $\boldsymbol{\theta}'''$. After shifting all the channels to $\mathcal{S}_2 \cup \mathcal{S}_3$ with $\boldsymbol{\theta}''$, we can obtain $\boldsymbol{\theta}'$ by rotating all those channels in $\mathcal{S}_3$ counterclockwise by an angle of ω. Now we just need to tell which channels are located in $\mathcal{S}_3$ under the beamformer $\boldsymbol{\theta}''$. We assume that $K \geq 3$; the case of $K = 2$ is later dealt with in Remark 3. Consider a particular channel $h_n e^{j\theta_n''}$ that is located in $\mathcal{S}_3$ with θ_n''. If we rotate $h_n e^{j\theta_n''}$ counterclockwise by an angle of ω, the new channel $h_n e^{j(\theta_n''+\omega)}$ would be located in $\mathcal{S}_1$, so $\left| \mathrm{Arg}(h_0) - \mathrm{Arg}(h_n e^{j(\theta_n''+\omega)}) \right| \leq \omega$. In contrast, the clockwise rotation of the channel $h_n e^{j\theta_n''}$ by an angle of ω gives $\left| \mathrm{Arg}(h_0) - \mathrm{Arg}(h_n e^{j(\theta_n''-\omega)}) \right| \geq \omega$. Combining the above results with (20), we conclude that $h_n e^{j\theta_n''} \in \mathcal{S}_3$ if $\mathbb{E}[|Y|^2|\theta_n = \phi_k + \omega] \geq \mathbb{E}[|Y|^2|\theta_n = \phi_k - \omega]$ and $h_n e^{j\theta_n''} \in \mathcal{S}_2$ otherwise. With each $\mathbb{E}[|Y|^2|\theta_n = \phi_k]$ empirically evaluated as $\hat{\mathbb{E}}[|Y|^2|\theta_n = \phi_k]$ in (16), we propose computing $\boldsymbol{\theta}'$ based on $\boldsymbol{\theta}''$ as

$$\theta_n' = \theta_n'' + \omega\Lambda_n, \ \forall n = 1, \ldots, N, \tag{31}$$

with

$$\Lambda_n = \mathbb{1}_+\left(\hat{\mathbb{E}}[|Y|^2|\theta_n = \theta_n''+\omega] - \hat{\mathbb{E}}[|Y|^2|\theta_n = \theta_n''-\omega]\right), \tag{32}$$

where the indicator function $\mathbb{1}_+(x) = 1$ gives 1 if $x \geq 0$ and gives 0 otherwise.

Furthermore, $\boldsymbol{\theta}'''$ can be readily obtained from $\boldsymbol{\theta}'$ by rotating every $h_n e^{j\theta_n'}$ clockwise by an angle of ω, i.e.,

$$\theta_n''' = \theta_n' - \omega, \ \forall n = 1, \ldots, N. \tag{33}$$

Following the approximation algorithm in [46], we choose the best beamformer of the three candidates:

$$\boldsymbol{\theta}^{\mathrm{ECSM}} = \arg\max_{\boldsymbol{\vartheta} \in \{\boldsymbol{\theta}', \boldsymbol{\theta}'', \boldsymbol{\theta}'''\}} \hat{\mathbb{E}}[|Y|^2|\boldsymbol{\theta} = \boldsymbol{\vartheta}]. \tag{34}$$

The above method is referred to as the *Enhanced Conditional Sample Mean (ECSM)* algorithm. The entire procedure is shown in Algorithm 2.

Remark 2: The way we compute $\boldsymbol{\theta}'$ and $\boldsymbol{\theta}'''$ based on $\boldsymbol{\theta}''$ as in (31)–(33) is of practical significance. It only requires measuring the received signal power $|Y_t|^2$ rather than the in-phase and quadrature branches of the received signal.

Algorithm 2 Enhanced Conditional Sample Mean (ECSM)
1: **input:** Φ_K, N, T
2: run Algorithm 1 and let $\boldsymbol{\theta}'' = \boldsymbol{\theta}^{\mathrm{CSM}}$
3: compute $\boldsymbol{\theta}'$ by (31); compute $\boldsymbol{\theta}'''$ by (33)
4: **for** $n = 1, 2, \ldots, N$ **do**
5:     compute $\hat{\mathbb{E}}[|Y_t|^2|\theta_n = \vartheta]$, $\forall \vartheta \in \{\boldsymbol{\theta}', \boldsymbol{\theta}'', \boldsymbol{\theta}'''\}$
6:     decide θ_n^{ECSM} according to (34)
7: **end for**
8: **output:** $\boldsymbol{\theta}^{\mathrm{ECSM}} = (\theta_1^{\mathrm{ECSM}}, \ldots, \theta_N^{\mathrm{ECSM}})$

Remark 3: When $K = 2$, we can no longer use (31)–(33) to compute $\boldsymbol{\theta}'$ and $\boldsymbol{\theta}'''$ because the clockwise rotation by an angle of ω is now identical to the counterclockwise rotation. In this case, we send the same symbol X_0 across the random samples and redefine the indicator variable Λ_n as

$$\Lambda_n = \mathbb{1}_+\left(\hat{\mathbb{E}}[Y|\theta_n = \phi_k]\right). \tag{35}$$

We still compute $\boldsymbol{\theta}'$ as in (31) and compute $\boldsymbol{\theta}'''$ as in (33).

In contrast to the conditional sample mean algorithm that guarantees a quadratic SNR boost only when $K \geq 3$, the enhanced conditional sample mean algorithm gives a quadratic SNR boost regardless of K, as stated in the following theorem.

Theorem 3: Consider the same setting as in Theorem 1. The expected SNR boost achieved by the enhanced conditional sample mean algorithm has a tight order bound:

$$\mathbb{E}[f(\boldsymbol{\theta}^{\mathrm{ECSM}})] = \frac{\delta^2}{\beta_0^2} \cdot \Theta(N^2) \ \text{if} \ T = \Omega(N^2(\log N)^3), \tag{36}$$

where the expectation is taken over random samples of $\boldsymbol{\theta}$. If $K = 2$, the condition can be relaxed as $T = \Omega(N^2 \log N)$.

Proof: See Appendix C. ∎

V. FIELD TESTS

We validate the performance of the proposed algorithms in a real-world environment. The field tests are carried out for downlink transmission from a public base station to a user terminal in bandwidth 200 MHz at 2.6 GHz. Throughout our tests, *not any* knowledge of the base station setting is required, and *not any* coordination or customization on the base station side is required. In other words, the IRS works in a plug-

Fig. 2. A satellite image of the field test site. The base station and the IRS are outdoor while the user terminals are indoor.

Fig. 4. The IRS is formed by a 4×4 array of reflecting tiles. Each reflecting tile is 50cm×50cm large and consists of 16 reflective elements.

Fig. 3. The view from the user terminal toward the IRS.

and-play fashion, and its deployment is completely invisible to service provider.

As shown in Fig. 1 and Fig. 2, the base station is located on a 20-meter-high terrace while the user terminal is located in an underground parking lot. There is no line-of-sight propagation from the base station to the user terminal. The IRS is placed outdoors near the entrance of the parking lot. The distance from the base station to the IRS is approximately 250 meters; the distance from the IRS to the user terminal is approximately 40 meters. It is worth remarking that the wireless environment is highly volatile in our case because of the busy traffic in the parking lot, as can be observed from Fig. 3.

As shown in Fig. 4, the IRS is formed by 16 "reflecting tiles"—a tiny IRS prototype that is $50\,\mathrm{cm} \times 50\,\mathrm{cm}$ large—arranged in a 4×4 array. Each reflecting tile consists of 16 reflective elements, so the assembled large IRS consists of 256 reflective elements in total. There are 4 phase shift choices $\{0, \pi/2, \pi, 3\pi/2\}$ for each individual reflective element.

The random-max sampling (RMS) algorithm uses a total of 2560 random samples, i.e., $T = 10N$. We compare RMS with the conditional sample mean (CSM) algorithm and the enhanced conditional sample mean (ECSM) algorithm. In

addition, we include in the tests a baseline method called OFF, which simply deactivates all the reflective elements so that the IRS reduces to a metal surface without beamforming.

We start with the single-input single-output (SISO) transmission, aiming to improve the SNR boost. There are two measurements: Reference Signal Received Power (RSRP) and Signal-to-Interference-plus-Noise Ratio (SINR). Notice that the SNR cannot be measured directly because of co-channel interference. Following the definition of the SNR boost, we let the RSRP (or SINR) boost be the ratio between the achieved RSRP (or SINR) and the baseline RSRP (or SINR) without IRS. Fig. 5 shows the RSRP boosts achieved by the various methods. It can be seen that CSM and ECSM outperform the other methods significantly. For instance, there is an approximately 5 dB gap between ECSM and OFF. As shown in the figure, although ECSM encounters two sharp drops, which are due to the shadowing effect caused by vehicles, its overall performance is still more consistent over time than RMS and OFF. Observe from Fig. 5 that the RSRP boost by OFF is mostly below 0 dB; the reason is that the reflected signals without proper beamforming can result in a destructive superposition. Observe also that RMS yields the worst RSRP performance, even 4 dB lower than not using IRS. This surprising result indicates that, in a complicated wireless environment with interference and noise, the beamformer decision based on the best single sample is not reliable.

We further compare the SINR boosts of the various methods in Fig. 6. It can be seen that the SINR boosts and the RSRP boosts have similar profiles. The average RSRP boosts and the average SINR boosts are summarized in Table I. According to the table, the SINR gain is smaller than the RSRP gain. One reason for this gain reduction is that IRS incurs additional reflected interference. Nevertheless, the constructive effect on the desired signals outweighs that on the interfering signals. As a result, CSM and ECSM can still bring considerable performance gains as compared to the benchmark methods and not using CSI.

Moreover, we consider the MIMO transmission. In our case, the base station has 64 transmit antennas while the user

TABLE I
AVERAGE PERFORMANCE OF THE VARIOUS ALGORITHMS

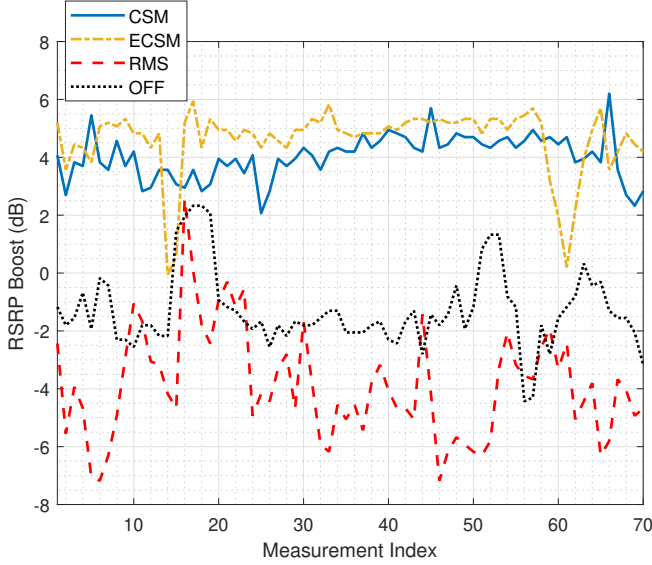|	SISO		MIMO
Algorithm	RSRP Boost (dB)	SINR Boost (dB)	SE Increment (bps/Hz)
CSM	4.02	3.57	2.02
ECSM	4.62	3.81	2.08
RMS	−3.93	−3.84	1.97
OFF	−1.69	−1.69	0.77

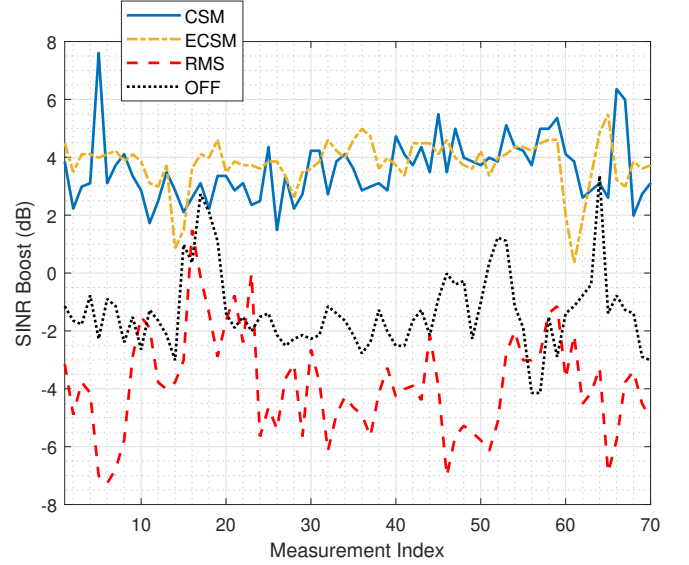Fig. 5. RSRP boost for SISO transmission.

Fig. 6. SINR boost for SISO transmission.

terminal has 4 receive antennas, so at most 4 data streams are supported. Because the base station is a black box to us, how the transmit precoding is performed is unknown. As stated in Section III-D, it is difficult to define a scalar-valued SNR boost in this scenario. We adopt the generalized CSM and ECSM in Section III-D with the *Spectral Efficiency (SE)* utility. Thus, we measure the SE in bps/Hz at the user terminal for each random sample. Fig. 7 shows the SE increments by the various algorithms against the baseline SE without IRS. Observe that all the algorithms can bring improvements, although OFF occasionally gives negative effects. The figure shows that RMS becomes more robust in the MIMO case. Actually, RMS is sometimes even better than CSM and ECSM, but it still has inferior performance on average. The average SE increment results summarized in Table I agree with what we observe from Fig. 7.

VI. CONCLUSION

This paper suggests a statistical approach to the IRS beamforming that is fundamentally different from most of the existing methods in that it does not entail any channel estimation. The proposed algorithm only requires a polynomial number of random samples to yield an SNR boost that is quadratic in the number of reflective elements—which is the highest possible SNR boost. In contrast, the standard method of random-
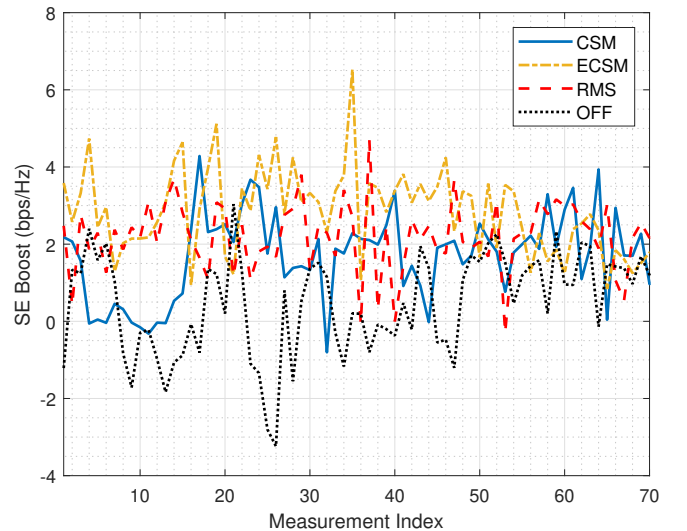

Fig. 7. SE increment for MIMO transmission.

max sampling can only give a linear boost under the same condition. Field experiments show that the proposed algorithm can achieve significant performance gains consistently over time even in a highly volatile wireless environment.

APPENDIX A
PROOF OF THEOREM 1

A. *Lower Order Bound in (13)*

We first introduce a multidimensional version of the Berry-Esséen theorem for non-i.i.d. random vectors.

Theorem 4 (Theorem 1.1 in [47]): Consider N independent (but not necessarily identically distributed) random vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N \in \mathbb{R}^d$ with zero mean $\mathbb{E}[\boldsymbol{x}_n] = \boldsymbol{0}$, $n = 1, \ldots, N$. Write $\boldsymbol{s} = \sum_{n=1}^{N} \boldsymbol{x}_n$, $\boldsymbol{W}^2 = \mathbb{E}[\boldsymbol{s}\boldsymbol{s}^\mathsf{T}]$, $\boldsymbol{W}$ the positive root of $\boldsymbol{W}^2$ (assuming that $\boldsymbol{W}^2 \succ \boldsymbol{0}$), $\xi_n = \mathbb{E}[\|\boldsymbol{W}^{-1}\boldsymbol{x}_n\|^3]$, and $\xi = \sum_{n=1}^{N} \xi_n$. Letting $\mathcal{C}$ be the class of all convex subsets of $\mathbb{R}^d$, there exists an absolute positive constant $\rho > 0$ such that

$$\sup_{\mathcal{A} \in \mathcal{C}} \left| \mathbb{P}\{\boldsymbol{s} \in \mathcal{A}\} - \mathbb{P}\{\boldsymbol{u} \in \mathcal{A}\} \right| \leq \sqrt[4]{d} \rho \xi, \qquad (37)$$

where $\boldsymbol{u}$ is a d-dimensional Gaussian vector with $\mathbb{E}[\boldsymbol{u}] = \boldsymbol{0}$ and $\mathbb{E}[\boldsymbol{u}\boldsymbol{u}^\mathsf{T}] = \boldsymbol{W}^2$.

The following corollary specializes the above theorem to our problem case.

Corollary 2: If every phase shift θ_n is drawn from the set Φ_K uniformly and independently, there exists an absolute positive constant $\rho > 0$ such that the complementary cumulative distribution function (CCDF) of the received signal power $|Y|^2$ is bounded as

$$\left| \mathbb{P}\{|Y|^2 \geq \gamma\} - e^{-\lambda\gamma} \right| \leq \sqrt[4]{2} \rho \xi \qquad (38)$$

given any threshold $\gamma > 0$, where

$$\lambda = \frac{1}{\sum_{n=0}^{N} \beta_n^2 P + \sigma^2} = \frac{1}{h_0 P + N\delta^2 P + \sigma^2} \qquad (39)$$

and

$$\xi = (2\lambda)^{3/2} \left(\sum_{n=0}^{N} \beta_n^3 P^{3/2} + \sigma^3 \right) \cdot \mathbb{E}[|u|^3] \qquad (40)$$

with a unit complex Gaussian random variable $u \sim \mathcal{CN}(0, 1)$.

Proof: The main idea is to rewrite each complex number as a real 2-dimensional vector in $\mathbb{R}^2$. Recall that $Y = h_0 X + \sum_{n=1}^{N} h_n e^{j\theta_n} X + Z$. Let $\boldsymbol{x}_0 = (\mathrm{Re}\{h_0 X\}, \mathrm{Im}\{h_0 X\})^\mathsf{T}$, $\boldsymbol{x}_n = (\mathrm{Re}\{h_n e^{j\theta_n} X\}, \mathrm{Im}\{h_n e^{j\theta_n} X\})^\mathsf{T}$ for $n = 1, \ldots, N$, and $\boldsymbol{x}_{N+1} = (\mathrm{Re}\{Z\}, \mathrm{Im}\{Z\})^\mathsf{T}$. Notice that $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N+1}$ are zero-mean independent random vectors, so Theorem 4 applies. With $\boldsymbol{s} = \sum_{n=0}^{N+1} \boldsymbol{x}_n$, we compute $\boldsymbol{W}^2$ as $\mathbb{E}[\boldsymbol{s}\boldsymbol{s}^\mathsf{T}] = \frac{1}{2}(\sum_{n=0}^{N} \beta_n^2 P + \sigma^2)\boldsymbol{I}$, and thus $\boldsymbol{W} = \frac{1}{\sqrt{2}}(\sum_{n=0}^{N} \beta_n^2 P + \sigma^2)^{\frac{1}{2}} \boldsymbol{I}$. We then obtain ξ as in (40).

Consider a convex subset

$$\mathcal{A} = \left\{ \boldsymbol{v} \in \mathbb{R}^2 : \|\boldsymbol{v}\|^2 < \gamma \right\}. \qquad (41)$$

For a Gaussian random variable $\boldsymbol{u} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{W}^2)$, we have $\mathbb{P}\{\boldsymbol{u} \in \mathcal{A}\} = 1 - e^{-\lambda\gamma}$ with $\lambda = (\sum_{n=0}^{N} \beta_n^2 P + \sigma^2)^{-1}$. Applying Theorem 4, we arrive at

$$\left| \mathbb{P}\{\boldsymbol{s} \in \mathcal{A}\} - \mathbb{P}\{\boldsymbol{u} \in \mathcal{A}\} \right| \leq \sqrt[4]{2} \rho \xi, \qquad (42)$$

which can be rewritten as

$$\left| \mathbb{P}\{\boldsymbol{s} \notin \mathcal{A}\} - \mathbb{P}\{\boldsymbol{u} \notin \mathcal{A}\} \right| \leq \sqrt[4]{2} \rho \xi \qquad (43)$$

since $\mathbb{P}\{\boldsymbol{s} \in \mathcal{A}\} = 1 - \mathbb{P}\{\boldsymbol{s} \notin \mathcal{A}\}$ and $\mathbb{P}\{\boldsymbol{u} \in \mathcal{A}\} = 1 -$

$\mathbb{P}\{\boldsymbol{u} \notin \mathcal{A}\}$. Substituting $\mathbb{P}\{\boldsymbol{u} \notin \mathcal{A}\} = e^{-\lambda\gamma}$ and $\mathbb{P}\{\boldsymbol{s} \notin \mathcal{A}\} = \mathbb{P}\{|Y|^2 \geq \gamma\}$ in (43) gives the bound in (38). ∎

The following lemma characterizes the behavior of the noise power $|Z|^2$ under random-max sampling.

Lemma 1: Given a sequence of i.i.d. noise $Z_t \sim \mathcal{CN}(0, \sigma^2)$, $t = 1, \ldots, T$, we have

$$\mathbb{E}\left[\max_{1 \leq t \leq T} |Z_t|^2 \right] = \sigma^2 \cdot \Theta(\log T). \qquad (44)$$

Proof: Each $|Z_t|^2$ has an exponential distribution with the scale parameter σ^2, so

$$\mathbb{P}\{|Z_t|^2 \leq \tau\} = 1 - e^{-\tau/\sigma^2} \qquad (45)$$

and hence

$$\mathbb{P}\left\{ \max_{1 \leq t \leq T} |Z_t|^2 > \tau \right\} = 1 - (1 - e^{-\tau/\sigma^2})^T. \qquad (46)$$

With $\tau = \sigma^2 \log T$, we use Markov's inequality to show that

$$\begin{aligned}
\mathbb{E}\left[\max_{1 \leq t \leq T} |Z_t|^2 \right] &\geq \tau \cdot \mathbb{P}\left\{ \max_{1 \leq t \leq T} |Z_t|^2 > \tau \right\} \\
&= \sigma^2 \log T \cdot \left(1 - (1 - 1/T)^T \right) \\
&\geq \sigma^2 \log T \cdot (1 - e^{-1}) \\
&= \sigma^2 \cdot \Omega(\log T).
\end{aligned} \qquad (47)$$

Moreover, by Jensen's inequality, we have

$$\begin{aligned}
\exp\left(\frac{1}{2\sigma^2} \mathbb{E}\left[\max_{1 \leq t \leq T} |Z_t|^2 \right] \right) &\leq \mathbb{E}\left[\exp\left(\frac{1}{2\sigma^2} \cdot \max_{1 \leq t \leq T} |Z_t|^2 \right) \right] \\
&\leq \sum_{t=1}^{T} \mathbb{E}\left[\exp\left(\frac{1}{2\sigma^2} |Z_t|^2 \right) \right] \\
&= 2T.
\end{aligned} \qquad (48)$$

Taking the logarithm of both sides yields

$$\mathbb{E}\left[\max_{1 \leq t \leq T} |Z_t|^2 \right] \geq 2\sigma^2 \log(2T) = \sigma^2 \cdot O(\log T). \qquad (49)$$

Combining (47) and (49) gives the order bound in (44). ∎

In light of Corollary 2 and Lemma 1, we now turn to showing that $\mathbb{E}[f(\boldsymbol{\theta}^{\mathrm{RMS}})] = (\delta^2/\beta_0^2) \cdot \Omega(N \log T)$ conditioned on $T = o(\sqrt{N})$. With the threshold γ specified as

$$\gamma_0 = \frac{\log T}{\lambda}, \qquad (50)$$

it follows from Corollary 2 that

$$\begin{aligned}
\mathbb{P}\{|Y|^2 \geq \gamma_0\} &\geq e^{-\lambda\gamma_0} - \sqrt[4]{2} \rho \xi \\
&= \frac{1}{T} - O\left(\frac{1}{\sqrt{N}} \right) \\
&\geq \frac{c}{T}
\end{aligned} \qquad (51)$$

x with some positive constant $c > 0$. The last step in (51) is due to the condition that $T = o(\sqrt{N})$. We further extend the

above bound to the T random samples:

$$\mathbb{P}\left\{\max_{1\le t\le T}|Y_t|^2 \ge \gamma_0\right\} = 1 - \mathbb{P}\left\{\max_{1\le t\le T}|Y_t|^2 < \gamma_0\right\}$$
$$= 1 - \mathbb{P}\left\{|Y_t|^2 < \gamma_0, \forall t\right\}$$
$$= 1 - (1 - c/T)^T$$
$$\ge 1 - e^{-c}. \tag{52}$$

According to Markov's inequality, we have

$$\mathbb{E}\left[\max_{1\le t\le T}|Y_t|^2\right] \ge \gamma_0 \cdot \mathbb{P}\left\{\max_{1\le t\le T}|Y_t|^2 \ge \gamma_0\right\}$$
$$\ge \frac{\log T}{\lambda} \cdot (1 - e^{-c})$$
$$= \left(\sum_{n=0}^{N}\beta_n^2 P + \sigma^2\right)(1 - e^{-c})\log T$$
$$= \delta^2 P \cdot \Omega(N\log T), \tag{53}$$

where the expectation is taken over random samples of $\boldsymbol{\theta}$. The superposition of all the channels from the transmitter to the receiver is denoted as

$$g = h_0 + \sum_{n=1}^{N}h_n e^{j\theta_n}. \tag{54}$$

Let g_t be the value of g for the tth random sample and let

$$t_0 = \arg\max_{1\le t\le T}|Y_t|^2. \tag{55}$$

We now explore the relationship between $\mathbb{E}[\max_{1\le t\le T}|Y_t|^2]$ and $\mathbb{E}[\max_{1\le t\le T}|g_t|^2]$ as follows:

$$\mathbb{E}\left[|Y_{t_0}|^2\right] = \mathbb{E}\left[|g_{t_0}X_{t_0} + Z_{t_0}|^2\right]$$
$$\le \mathbb{E}\left[2|g_{t_0}X_{t_0}|^2 + 2|Z_{t_0}|^2\right]$$
$$\le \mathbb{E}\left[2|g_{t_0}X_{t_0}|^2 + \max_{1\le t'\le T}2|Z_{t'}|^2\right]$$
$$= 2P\mathbb{E}\left[|g_{t_0}|^2\right] + 2\mathbb{E}\left[\max_{1\le t'\le T}|Z_{t'}|^2\right]$$
$$= 2P\mathbb{E}\left[|g_{t_0}|^2\right] + 2\sigma^2\Theta(\log T), \tag{56}$$

where the last equality follows by Lemma 1. We incorporate the previous bound $\mathbb{E}[\max_{1\le t\le T}|Y_t|^2] \ge \delta^2 P \cdot \Omega(N\log T)$ into (56), obtaining

$$\mathbb{E}\left[|g_{t_0}|^2\right] \ge \frac{1}{2}\delta^2 \cdot \Omega(N\log T) - \sigma^2 \cdot \Theta(\log T)$$
$$= \delta^2 \cdot \Omega(N\log T). \tag{57}$$

It immediately follows that $\mathbb{E}[f(\boldsymbol{\theta}^{\mathrm{RMS}})] = 1/\beta_0^2 \cdot \mathbb{E}[|g_{t_0}|^2] = (\delta^2/\beta_0^2) \cdot \Omega(N\log T)$.

B. Upper Order Bound in (14)

We first bound the tail probability of the power of the channel superposition g in (54).

Lemma 2: With each θ_n drawn from Φ_K uniformly and independently, the CCDF to the overall channel strength is bounded as

$$\mathbb{P}\{|g|^2 > \tau\} \le 4e^{-\nu\tau/4}, \tag{58}$$

where

$$\nu = \frac{1}{\sum_{n=0}^{N}\beta_n^2} = \frac{1}{\beta_0 + N\delta^2}. \tag{59}$$

Proof: Let g_{Re} be the real part of g, so $g_{\mathrm{Re}} = \mathrm{Re}\{h_0\} + \sum_{n=1}^{N}\mathrm{Re}\{h_n e^{j\theta_n}\}$. Note that $|\mathrm{Re}\{h_0\}| \le \beta_0$ and $|\mathrm{Re}\{h_n e^{j\theta_n}\}| \le \beta_n$ for each n; note also that $\mathrm{Re}\{h_0\}, \mathrm{Re}\{h_1 e^{j\theta_1}\}, \ldots, \mathrm{Re}\{h_N e^{j\theta_N}\}$ are statistically independent. As a result, g_{Re} constitutes a *sub-Gaussian* random variable with the variance proxy $1/\nu$. Then

$$\mathbb{P}\{g_{\mathrm{Re}} \ge \tau\} \le 2e^{-\nu\tau/2}. \tag{60}$$

By symmetry, the tail probability of the imaginary part of g, g_{Im}, can be bounded as

$$\mathbb{P}\{g_{\mathrm{Im}} \ge \tau\} \le 2e^{-\nu\tau/2}. \tag{61}$$

Accordingly, the tail probability of $|g|^2$ can be limited to

$$\mathbb{P}\{|g|^2 \ge \tau\} = \mathbb{P}\{g_{\mathrm{Re}}^2 + g_{\mathrm{Im}}^2 \ge \tau\}$$
$$\le \mathbb{P}\{g_{\mathrm{Re}}^2 \ge \tau/2 \text{ or } g_{\mathrm{Im}}^2 \ge \tau/2\}$$
$$\le \mathbb{P}\{g_{\mathrm{Re}}^2 \ge \tau/2\} + \mathbb{P}\{g_{\mathrm{Im}}^2 \ge \tau/2\}$$
$$= 4e^{-\nu\tau/4}. \tag{62}$$

The proof is then completed. ∎

We now verify that $\mathbb{E}[f(\boldsymbol{\theta}^{\mathrm{RMS}})] = (\delta^2/\beta_0^2) \cdot O(N\log T)$ in general. With

$$s = \frac{\nu}{8} \tag{63}$$

and t_0 defined in (55), we get

$$\exp\left(s \cdot \mathbb{E}\left[|g_{t_0}|^2\right]\right) \le \exp\left(s \cdot \mathbb{E}\left[\max_{1\le t\le T}|g_t|^2\right]\right)$$
$$\overset{(a)}{\le} \mathbb{E}\left[\exp\left(s \cdot \max_{1\le t\le T}|g_t|^2\right)\right]$$
$$\le \sum_{t=1}^{T}\mathbb{E}\left[e^{s|g_t|^2}\right]$$
$$= T \cdot \mathbb{E}\left[e^{s|g|^2}\right]$$
$$\overset{(b)}{=} T + T\int_1^\infty \mathbb{P}\left\{e^{s|g|^2} \ge \tau\right\}d\tau$$
$$= T + T\int_1^\infty \mathbb{P}\left\{|g|^2 \ge \frac{\log\tau}{s}\right\}d\tau$$
$$\overset{(c)}{\le} T + T\int_1^\infty \frac{4}{\tau^2}d\tau$$
$$= 5T, \tag{64}$$

where (a) follows by Jensen's inequality, (b) follows by the identity $\mathbb{E}[X] = a + \int_a^\infty \mathbb{P}\{X \ge x\}dx$ for $X \ge a$, and (c) follows by Lemma 2. Furthermore, we take the logarithm of both sides in (64), and thus have

$$\mathbb{E}\left[|g_{t_0}|^2\right] \le \frac{8}{\lambda}\log(5T).$$
$$= \delta^2 \cdot O(N\log T). \tag{65}$$

With the above result, we can readily bound the SNR boost as $\mathbb{E}[f(\boldsymbol{\theta}^{\mathrm{RMS}})] = (\delta^2/\beta_0^2) \cdot O(N\log T)$. Unlike the lower order bound case, the above proof does not impose any constraint on the number of random samples T, so this upper order bound

holds in general.

APPENDIX B
PROOF OF THEOREM 2

Let χ_{n1} and χ_{n2} be the largest and the second largest values of $\cos(k\omega - \theta_n^\infty)$, respectively, across $k = 1, \ldots, K$. The difference between χ_{n1} and χ_{n2} is defined as

$$\epsilon_n = \chi_{n1} - \chi_{n2}. \tag{66}$$

If the gap between the conditional sample mean and the conditional expectation satisfies

$$\left|\hat{\mathbb{E}}[|Y|^2|\theta_n = \phi_k] - \mathbb{E}[|Y|^2|\theta_n = \phi_k]\right| \leq 2\beta_0\beta_n P\epsilon_n, \ \forall k, \tag{67}$$

then we have

$$\arg\max_{\varphi \in \Phi_K} \hat{\mathbb{E}}[|Y|^2|\theta_n = \varphi] = \arg\max_{\varphi \in \Phi_K} \mathbb{E}[|Y|^2|\theta_n = \varphi], \tag{68}$$

namely $\theta_n^{\text{CSM}} = \theta_n^{\text{CCP}}$. Hence, to decide the condition for $\boldsymbol{\theta}^{\text{CSM}} = \boldsymbol{\theta}^{\text{CCP}}$, it suffices to examine in what regimes of (N, T) the inequality in (67) holds for every $n = 1, \ldots, N$.

Without loss of generality, we focus on a particular conditional subset $\mathcal{Q}_{nk}$ of random samples with $\theta_n = \phi_k$. Let T_{nk} be the number of random samples in $\mathcal{Q}_{nk}$. With each θ_n drawn from Φ_K uniformly and independently, we have

$$T_{nk} = \frac{T}{K} \quad \text{with high probability.} \tag{69}$$

Denote the noise-free received signal power under $\theta_n = \phi_k$ as

$$\eta_{nk} = P \cdot \left| h_0 + h_n e^{j\phi_k} + \sum_{m=1,m\neq n}^{N} h_m e^{j\theta_m} \right|^2. \tag{70}$$

We use η_{nkt} to denote the value of η_{nk} for the tth random sample. The sample mean of η_{nk} is computed as

$$\bar{\eta}_{nk} = \frac{1}{T_{nk}} \sum_{t \in \mathcal{Q}_{nk}} \eta_{nkt}. \tag{71}$$

The expectation of η_{nk} is given by

$$\mathbb{E}[\eta_{nk}] = P\left|h_0 + h_n e^{j\theta_n}\right|^2 + \sum_{m\neq n} \beta_m^2 P. \tag{72}$$

Moreover, let ζ_{nk} be the sample mean of noise power over the subset $\mathcal{Q}_{nk}$:

$$\zeta = \frac{1}{T_{nk}} \sum_{t \in \mathcal{Q}_{nk}} |Z_t|^2. \tag{73}$$

Consider the following three events:

$$\mathcal{E}_{nk,1} = \left\{\eta_{nk} > qP/\nu\right\} \text{ given some } q > 0, \tag{74}$$

$$\mathcal{E}_{nk,2} = \left\{\left|\bar{\eta}_{nk} - \mathbb{E}[\eta_{nk}]\right| \geq \beta_0\beta_n P\epsilon_n\right\}, \tag{75}$$

$$\mathcal{E}_{nk,3} = \left\{\left|\zeta_{nk} - \sigma^2\right| \geq \beta_0\beta_n P\epsilon_n\right\}, \tag{76}$$

where ν is defined in (59). We now bound the probability of each event. According to Lemma 2, we have

$$\mathbb{P}\{\mathcal{E}_{nk,1}\} \leq 4e^{-q/4}. \tag{77}$$

Assuming that every $\eta_{nk,t} \in \mathcal{E}_{nk,1}^c$ so that

$$0 \leq \eta_{nk,t} \leq qP/\nu, \ \forall t \in \mathcal{Q}_{nk}, \tag{78}$$

we derive the following bound from Hoeffding's inequality:

$$\mathbb{P}\left\{\mathcal{E}_{nk,2}|\mathcal{E}_{nk,1}^c\right\}$$
$$\leq \exp\left(-T_{nk}(\beta_0\beta_n P\epsilon_n)^2 \frac{\nu^2}{q^2P^2}\right)$$
$$= \exp\left(-\frac{(\beta_0\beta_n\epsilon_n\nu)^2 T}{q^2 K}\right) \quad \text{with high probability.} \tag{79}$$

The following bound follows by Chebyshev's inequality:

$$\mathbb{P}\{\mathcal{E}_{nk,3}\} \leq \frac{\sigma^4}{(\beta_0\beta_n P\epsilon_n)^2 T_{nk}}$$
$$= \frac{\sigma^4}{(\beta_0\beta_n P\epsilon_n)^2 T} \quad \text{with high probability.} \tag{80}$$

For the error event

$$\mathcal{E}_{nk} = \left\{\left|\hat{\mathbb{E}}[|Y|^2|\theta_n = \phi_k] - \mathbb{E}[|Y|^2|\theta_n = \phi_k]\right| \geq \epsilon\right\}. \tag{81}$$

The above results give rise to the following bound:

$$\mathbb{P}\{\mathcal{E}_{nk}\} = \mathbb{P}\left\{\left|\bar{\eta}_{nk} + \zeta - \mathbb{E}[\eta_{nk}] - \sigma^2\right| > \epsilon\right\}$$
$$\leq \mathbb{P}\left\{\left|\bar{\eta}_{nk} - \mathbb{E}[\eta_{nk}]\right| > \epsilon/2 \text{ or } \left|\zeta_{nk} - \sigma^2\right| > \epsilon/2\right\}$$
$$= \mathbb{P}\{\mathcal{E}_{nk,2} \cup \mathcal{E}_{nk,3}\}$$
$$\leq \mathbb{P}\{\mathcal{E}_{nk,2}\} + \mathbb{P}\{\mathcal{E}_{nk,3}\}$$
$$= \mathbb{P}\{\mathcal{E}_{nk,1} \cap \mathcal{E}_{nk,2}\} + \mathbb{P}\{\mathcal{E}_{nk,1}^c \cap \mathcal{E}_{nk,2}\} + \mathbb{P}\{\mathcal{E}_{nk,3}\}$$
$$\leq \mathbb{P}\{\mathcal{E}_{nk,1}\} + \mathbb{P}\{\mathcal{E}_{nk,2}|\mathcal{E}_{nk,1}^c\} + \mathbb{P}\{\mathcal{E}_{nk,3}\}$$
$$\leq 4e^{-q/4} + \exp\left(-\frac{(\beta_0\beta_n\epsilon_n\nu)^2 T}{q^2 K}\right)$$
$$+ \frac{\sigma^4}{(\beta_0\beta_n P\epsilon_n)^2 T} \quad \text{with high probability.} \tag{82}$$

The overall error event is defined as

$$\mathcal{E} = \bigcup_{\substack{n=1,\ldots,N \\ k=1,\ldots,K}} \mathcal{E}_{nk}. \tag{83}$$

We upper bound the overall error probability via union bound:

$$\mathbb{P}\{\mathcal{E}\} \leq \sum_{k=1}^{K} \sum_{n=1}^{N} \mathbb{P}\{\mathcal{E}_{nk}\}$$
$$\leq 4NKe^{-q/4} + NK\exp\left(-\frac{(\beta_0\rho\nu)^2 T}{q^2 K}\right)$$
$$+ \frac{\sigma^4 NK}{(\beta_0\rho P)^2 T} \quad \text{with high probability,} \tag{84}$$

where

$$\rho = \min_{1 \leq n \leq N} \beta_n\epsilon_n. \tag{85}$$

For any $0 < p_0 \leq \frac{1}{4}$, we have

$$4NKe^{-q/4} \leq p_0 \text{ if } q = \Omega(\log N),$$
$$NK\exp\left(-\frac{(\beta_0\rho\nu)^2 T}{q^2 K}\right) \leq p_0 \text{ if } T = \Omega(N^2 q^2 \log N),$$
$$\frac{\sigma^4 NK}{(\beta_0\rho P)^2 T} \leq p_0 \text{ if } T = \Omega(N).$$

Combining the above results, we obtain that $\mathbb{P}\{\mathcal{E}\} \leq 3p_0$ and $\mathbb{P}\{\mathcal{E}^c\} \geq 1 - 3p_0 \geq \frac{1}{4}$ given $T = \Omega(N^2(\log N)^3)$, and consequently

$$\mathbb{E}[f(\boldsymbol{\theta}^{\mathrm{CSM}})] \geq f(\boldsymbol{\theta}^{\mathrm{CPP}}) \cdot \mathbb{P}\{\mathcal{E}^c\} = (\delta^2/\beta_0^2) \cdot \Omega(N^2). \quad (86)$$

Combining the above result with the order upper bound in Proposition 1, we have $\mathbb{E}[f(\boldsymbol{\theta}^{\mathrm{CSM}})] = (\delta^2/\beta_0^2) \cdot \Theta(N^2)$.

APPENDIX C
PROOF OF THEOREM 3

When $K \geq 3$, the enhanced conditional sample mean algorithm is based on the statistics of the received signal power $|Y|^2$. It can be shown that $\boldsymbol{\theta}^{\mathrm{ECSM}} = \boldsymbol{\theta}^{\mathrm{APX}}$ if the power gap $\left|\hat{\mathbb{E}}[|Y|^2]|\theta_n = \phi_k] - \mathbb{E}[|Y|^2|\theta_n = \phi_k]\right|$ is smaller than the threshold ϵ_n in (66). The proof in this case follows that of Theorem 2 closely.

When $K = 2$, as shown in Remark 3, the enhanced conditional sample mean algorithm is based on the statistics of the received signal Y. Inspection of δ_n in (35) shows that $\boldsymbol{\theta}^{\mathrm{ECSM}} = \boldsymbol{\theta}^{\mathrm{APX}}$ if $\mathrm{Im}\{\hat{\mathbb{E}}[Y|\theta_n = \phi_k]\}$ and $\mathrm{Im}\{\mathbb{E}[Y|\theta_n = \phi_k]\}$ have the same sign. Consider the error event

$$\mathcal{E}_{nk} = \left\{ \mathrm{sgn}(\hat{\mathbb{E}}[Y|\theta_n = \phi_k]) \neq \mathrm{sgn}(\mathbb{E}[Y|\theta_n = \phi_k]) \right\}. \quad (87)$$

Recall that in Remark 3 we let $X_t = X_0$ for all t. Without loss of generality, let $X_0 = \sqrt{P}$. For ease of notation, we let

$$\tilde{Y}_{n,t} = h_0 X_0 + \sum_{m=1,m\neq n}^{N} h_m e^{j\theta_{mt}} X_0 \quad (88)$$

and let

$$\begin{aligned}
\psi_{nk} &= \mathrm{Im}\{\mathbb{E}[Y|\theta_n = \phi_k]\} \\
&= \sqrt{P}\beta_0 \sin(\alpha_0) + \sqrt{P}\beta_n \sin(\alpha_n + \phi_k).
\end{aligned} \quad (89)$$

The error probability can be bounded as

$$\begin{aligned}
\mathbb{P}\{\mathcal{E}_{nk}\} &= \mathbb{P}\left\{ \left| \mathrm{Im}\{\hat{\mathbb{E}}[Y|\theta_n = \phi_k]\} - \psi_{nk} \right| > \psi_{nk} \right\} \\
&= \mathbb{P}\left\{ \left| \sum_{t\in\mathcal{Q}_{nk}} \left(\mathrm{Im}\{\tilde{Y}_{n,t}\} + \mathrm{Im}\{Z_t\} \right) \right| > \psi_{nk} \right\} \\
&\leq \mathbb{P}\left\{ \left| \sum_{t\in\mathcal{Q}_{nk}} \mathrm{Im}\{\tilde{Y}_{n,t}\} \right| > \frac{\psi_{nk}}{2} \right\} \\
&\quad + \mathbb{P}\left\{ \left| \sum_{t\in\mathcal{Q}_{nk}} \mathrm{Im}\{Z_t\} \right| > \frac{\psi_{nk}}{2} \right\}.
\end{aligned} \quad (90)$$

Since $\mathrm{Im}\{Z_t\} \sim \mathcal{N}(0, \sigma^2/2)$ i.i.d., we can obtain the following bound from Chebyshev's inequality:

$$\mathbb{P}\left\{ \left| \sum_{t\in\mathcal{Q}_{nk}} \mathrm{Im}\{Z_t\} \right| > \frac{\psi_{nk}}{2} \right\} \leq \frac{2\sigma^2}{T_{nk}\psi_{nk}^2}. \quad (91)$$

Notice that $|\mathrm{Im}\{\tilde{Y}_{n,t}\}| \leq \sqrt{P}\sum_{m=0}^{N}\beta_m$. By means of Ho-effding's inequality, we establish that

$$\begin{aligned}
&\mathbb{P}\left\{ \left| \sum_{t\in\mathcal{Q}_{nk}} \mathrm{Im}\{\tilde{Y}_{n,t}\} \right| > \frac{\psi_{nk}}{2} \right\} \\
&\leq \exp\left(-\frac{(\psi_{nk}/2)^2 T_{nk}}{4P(\sum_{m=0}^{N}\beta_m)^2} \right) \\
&\leq \exp\left(-\frac{(\psi_{nk}/2)^2 T_{nk}}{8PN(\beta_0^2 + N\delta^2)} \right) \\
&= \exp\left(-\frac{(\psi_{nk}/2)^2 T}{8PN(\beta_0^2 + N\delta^2)} \right) \text{ with high probability.} \quad (92)
\end{aligned}$$

Again, the overall error event is given by (83). The overall error probability can be bounded as

$$\begin{aligned}
\mathbb{P}\{\mathcal{E}\} &\leq \sum_{k=1}^{K}\sum_{n=1}^{N} \mathbb{P}\{\mathcal{E}_{nk}\} \\
&\leq NK \exp\left(-\frac{(\psi_{\min}/2)^2 T}{8PN(\beta_0^2 + N\delta^2)} \right) + \frac{2NK\sigma^2}{T\psi_{\min}}, \quad (93)
\end{aligned}$$

where $\psi_{\min} = \min_{n,k}\psi_{nk}$. For any $0 < p_0 \leq \frac{1}{4}$, we have

$$NK \exp\left(-\frac{(\psi_{\min}/2)^2 T}{8PN(\beta_0^2 + N\delta^2)} \right) \leq p_0 \text{ if } T = \Omega(N^2\log N),$$

$$\frac{2NK\sigma^2}{T\psi_{\min}} \leq p_0 \text{ if } T = \Omega(N).$$

Thus, we can ensure that $\mathbb{P}\{\mathcal{E}\} \leq 2p_0$ and thus $\mathbb{P}\{\mathcal{E}^c\} \geq 1 - 2p_0 \geq \frac{1}{2}$. The remainder of the proof follows the part of Appendix B starting from (86).

REFERENCES

[1] V. Arun and H. Balakrishnan, "RFocus: Beamforming using thousands of passive antennas," in *USENIX Symp. Netw. Sys. Design Implementation (NSDI)*, Feb. 2020, pp. 1047–1061.

[2] J. Gao, C. Zhong, X. Chen, H. Lin, and Z. Zhang, "Unsupervised learning for passive beamforming," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1052–1056, May 2020.

[3] C. Liu, X. Liu, D. W. K. Ng, and J. Yuan, "Deep residual network empowered channel estimation for IRS-assisted multi-user communication systems," in *IEEE Int. Conf. Commun. (ICC)*, Jun. 2021.

[4] C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1839–1850, Aug. 2020.

[5] K. Feng, Q. Wang, X. Li, and C.-K. Wen, "Deep reinforcement learning based intelligent reflecting surface optimization for MISO communication systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 745–749, May 2020.

[6] T. Jiang, H. V. Cheng, and W. Yu, "Learning to reflect and to beamform for intelligent reflecting surface with implicit channel estimation," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1913–1945, Jul. 2021.

[7] S. Gong, X. Lu, D. T. Hoang, D. Niyato, L. Shu, D. I. Kim, and Y.-C. Liang, "Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2283–2314, 2020.

[8] Z.-Q. Luo, W.-K. Ma, A. M. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[9] K. Shen and W. Yu, "Fractional programming for communication systems—Part I: Power control and beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616–2630, May 2018.

[10] ——, "Fractional programming for communication systems—Part II: Uplink scheduling via matching," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2631–2644, May 2018.

[11] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

[12] D. Mishra and H. Johansson, "Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer," in *IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2019.

[13] B. Zheng, C. You, and R. Zhang, "Double-IRS assisted multi-user MIMO: Cooperative passive beamforming design," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4513–4526, Jul. 2021.

[14] G. Zhou, C. Pan, H. Ren, K. Wang, M. D. Renzo, and A. Nallanathan, "Robust beamforming design for intelligent reflecting surface aided MISO communication systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1658–1662, Oct. 2020.

[15] M. Zeng, X. Li, G. Li, W. Hao, and O. A. Dobre, "Sum rate maximization for IRS-assisted uplink NOMA," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 234–238, Jan. 2021.

[16] H. Xie, J. Xu, and Y.-F. Liu, "Max-min fairness in IRS-aided multi-cell MISO systems with joint transmit and reflective beamforming," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 1379–1393, Feb. 2021.

[17] S. Huang, Y. Ye, M. Xiao, H. V. Poor, and M. Skoglund, "Decentralized beamforming design for intelligent reflecting surface-enhanced cell-free networks," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 673–677, Mar. 2021.

[18] Y. Cao, T. Lv, and W. Ni, "Intelligent reflecting surface aided multi-user mmWave communications for coverage enhancement," in *IEEE Ann. Int. Symp. Personal Indoor Mobile Radio Commun.*, Aug. 2021.

[19] K. Feng, X. Li, Y. Han, S. Jin, and Y. Chen, "Physical layer security enhancement exploiting intelligent reflecting surface," *IEEE Commun. Lett.*, vol. 25, no. 3, pp. 734–738, Mar. 2021.

[20] J. Zhu, Y. Huang, J. Wang, K. Navaie, and Z. Ding, "Power efficient IRS-assisted NOMA," *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 900–913, Feb. 2021.

[21] T. Shafique, H. Tabassum, and E. Hossain, "Optimization of wireless relaying with flexible UAV-borne reflecting surfaces," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 309–325, Jan. 2021.

[22] Y. Cao, T. Lv, Z. Lin, and W. Ni, "Delay-constrained joint power control, user detection and passive beamforming in intelligent reflecting surface-assisted uplink mmWave system," *IEEE Trans. Cognitive Commun. Netw.*, vol. 7, no. 2, pp. 482–495, Jun. 2021.

[23] Z. Zhang, L. Dai, X. Chen, C. Liu, F. Yang, R. Schober, and H. V. Poor, "Active RIS vs. passive RIS: Which will prevail in 6G?" 2021, [Online]. Available: https://arxiv.org/abs/2103.15154.

[24] Z. Zhang and L. Dai, "A joint precoding framework for wideband reconfigurable intelligent surface-aided cell-free network," *IEEE Trans. Signal Process.*, vol. 69, pp. 4085–4101, Jun. 2021.

[25] J. Hu, Y.-C. Liang, and Y. Pei, "Reconfigurable intelligent surface enhanced multi-user MISO symbiotic radio system," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2359–2371, Apr. 2021.

[26] X. Mu, Y. Liu, L. Guo, J. Lin, and N. Al-Dhahir, "Exploiting intelligent reflecting surfaces in NOMA networks: Joint beamforming optimization," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6884–2020, Oct. 2020.

[27] M.-M. Zhao, A. Liu, Y. Wan, and R. Zhang, "Two-timescale beamforming optimization for intelligent reflecting surface aided multiuser communication with QoS constraints," *IEEE Trans. Signal Process.*, 2021.

[28] H. Shen, W. Xu, S. Gong, C. Zhao, and D. W. K. Ng, "Beamforming optimization for IRS-aided communications with transceiver hardware impairments," *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 1214–1227, Feb. 2021.

[29] B. Ning, Z. Chen, W. Chen, and J. Fang, "Beamforming optimization for intelligent reflecting surface assisted MIMO: A sum-path-gain maximization approach," *IEEE Wireless Commun. Lett.*, vol. 9, no. 7, pp. 1105–1109, Jul. 2020.

[30] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 518–522, Apr. 2020.

[31] T. L. Jensen and E. de Carvalho, "An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator," in *IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2020.

[32] G. T. de Araújo, A. L. F. de Almeida, and R. Boyer, "Channel estimation for intelligent reflecting surface assisted MIMO systems: A tensor modeling approach," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 3, pp. 789–801, Apr. 2021.

[33] Q.-U.-A. Nadeem, H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Intelligent reflecting surface-assisted multi-user MISO communication: Channel estimation and beamforming design," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 661–680, May 2020.

[34] P. Wang, J. Fang, H. Duan, and H. Li, "Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems," *IEEE Signal Process. Lett.*, vol. 27, pp. 905–909, May 2020.

[35] S. Liu, Z. Gao, J. Zhang, M. D. Renzo, and M.-S. Alouini, "Deep denoising neural network assisted compressive channel estimation for mmWave intelligent reflecting surfaces," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 9223–9228, Aug. 2020.

[36] L. Wei, C. Huang, G. C. Alexandropoulos, C. Yuen, Z. Zhang, and M. Debbah, "Channel estimation for RIS-empowered multi-user MISO wireless communications," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4144–4157, Jun. 2021.

[37] J. Chen, Y.-C. Liang, H. V. Cheng, and W. Yu, "Channel estimation for reconfigurable intelligent surface aided multi-user MIMO systems," *IEEE Trans. Wireless Commun.*, 2021.

[38] A. M. Elbir, A. Papazafeiropoulos, P. Kourtessis, and S. Chatzinotas, "Deep channel learning for large intelligent surfaces aided mm-Wave massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 9, pp. 1447–1451, Sep. 2020.

[39] L. You, J. Xiong, D. W. K. Ng, C. Yuen, W. Wang, and X. Gao, "Energy efficiency and spectral efficiency tradeoff in RIS-aided multiuser MIMO uplink transmission," *IEEE Trans. Signal Process.*, vol. 69, pp. 1407–1421, Mar. 2021.

[40] W. Zhang, J. Xu, W. Xu, D. W. K. Ng, and H. Sun, "Cascaded channel estimation for IRS-assisted mmWave multi-antenna with quantized beamforming," *IEEE Commun. Lett.*, vol. 25, no. 2, pp. 593–597, Feb. 2021.

[41] C. You, B. Zheng, and R. Zhang, "Channel estimation and passive beamforming for intelligent reflecting surface: Discrete phase shift and progressive refinement," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2604–2620, Nov. 2020.

[42] X. Yu, D. Xu, and R. Schober, "Optimal beamforming for MISO communications via intelligent reflecting surfaces," in *IEEE Int. Workshop Sig. Process. Advances Wireless Commun. (SPAWC)*, May 2020.

[43] H. Wang, N. Shlezinger, Y. C. Eldar, S. Jin, M. F. Imani, I. Yoo, and D. R. Smith, "Dynamic metasurface antennas for MIMO-OFDM receivers with bit-limited ADCs," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2643–2659, Apr. 2021.

[44] S. Abeywickrama, R. Zhang, Q. Wu, and C. Yuen, "Intelligent reflecting surface: Practical phase shift model and beamforming optimization," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5849–5863, Sep. 2020.

[45] Q. Wu and R. Zhang, "Beamforming optimization for wireless networks aided by intelligent reflecting surface with discrete phase shifts," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838–1851, Mar. 2020.

[46] Y. Zhang, K. Shen, S. Ren, X. Li, X. Chen, and Z.-Q. Luo, "Configuring intelligent reflecting surface with performance guarantees: Optimal beamforming," 2021, [Online]. Available: https://kaimingshen.github.io/doc/OptimalBeamforming.pdf.

[47] V. Bentkus, "A Lyapunov-type bound in $\boldsymbol{R}^d$," *Theory Probab. Appl.*, vol. 49, no. 2, pp. 311–323, 2005.