Model-Free Learning of Two-Stage Beamformers for Passive IRS-Aided Network Design

Hassaan Hashmi, *Student Member, IEEE*, Spyridon Pougkakiotis, and Dionysis Kalogerias, *Senior Member, IEEE*

***Abstract*—Electronically tunable metasurfaces, or Intelligent Reflecting Surfaces (IRSs), are a popular technology for achieving high spectral efficiency in modern wireless systems by shaping channels using a multitude of tunable passive reflecting elements. Capitalizing on key practical limitations of IRS-aided beamforming pertaining to system modeling and channel sensing/estimation, we propose a novel, fully data-driven Zeroth-order Stochastic Gradient Ascent (ZoSGA) algorithm for general two-stage (i.e., short/long-term), fully-passive IRS-aided stochastic utility maximization. ZoSGA learns long-term optimal IRS beamformers jointly with short-term optimal precoders (e.g., WMMSE-based) via minimal zeroth-order reinforcement and in a strictly model-free fashion, relying solely on the *effective* compound channels observed at the terminals, while being independent of channel models or network/IRS configurations. Another remarkable feature of ZoSGA is being amenable to analysis, enabling us to establish a state-of-the-art (SOTA) convergence rate of the order of $\mathcal{O}(\sqrt{S}\epsilon^{-4})$ under minimal assumptions, where S is the total number of IRS elements, and ϵ is a desired suboptimality target. Our numerical results on a standard MISO downlink IRS-aided sumrate maximization setting establish SOTA empirical behavior of ZoSGA as well, consistently and substantially outperforming standard fully model-based baselines. Lastly, we demonstrate that ZoSGA can in fact operate *in the field*, by directly optimizing the capacitances of a varactor-based electromagnetic IRS model (unknown to ZoSGA) on a multiple user/IRS, link-dense network setting, with essentially no computational overheads or performance degradation.**

***Index Terms*—6G, intelligent reflecting surfaces (IRS/RIS), two-stage stochastic programming, zeroth-order optimization, model-free learning, sumrate maximization, equivalent circuit model.**

I. Introduction

THE radical growth in the number of mobile and numerous other wireless devices, in particular those with communication modalities requiring high bandwidth and low latency connectivity such as virtual and augmented reality, tactile internet, internet of things, industrial automation, etc., have pushed existing wireless communication systems to their performance limits. The forthcoming era requires seamless wireless connectivity, necessitating proactive research beyond 5G communications [2], [3], [4], [5].

Recently, 5G-enabling technologies such as mmWave communication, massive multi-input multi-output (MIMO) and dense networks have been vigorously investigated [6], [7], [8], [9]. Still, deployment of both massive MIMO and dense networks incurs high installation/maintenance costs and energy consumption, while mmWaves exhibit physical limitations such as susceptibility to blockages and high propagation losses. To compensate for such propagation losses, which are a characteristic of higher carrier frequencies, densely packed and highly directional mmWave antennas have been proposed [10], [11]. High directionality, combined with reduced scattering, attenuates mmWave signals in the non-line-of-sight (non-LOS) paths, thus blocking the signals. The reason lies in the physics of signal propagation: mmWaves exhibit a more prismatic propagation, i.e., they diffract less than microwave signals around obstacles [12].

Conventional beamforming, which is an established technique for improving Quality-of-Service (QoS) in wireless communications, cannot fully compensate for such non-LOS losses. One brute force solution would be to deploy ultra dense networks, i.e., networks with a large number of small cells, with service ranges of tens to hundreds of meters, allowing higher frequency reuse rates [13], [14]. With such low ranges, however, ultra-dense network deployment does not demonstrate economy of scale. Indeed, the consumption of energy increases sharply with the number of base stations. Additionally, such dense deployments would also exhibit acute signal interference patterns.

While the aforementioned approaches aim to improve QoS, they may fail to simultaneously satisfy the data-rate, bandwidth, latency, spectral and energy efficiency requirements of 5G-and-beyond technologies. This has necessitated efforts to develop innovative technologies that could meet such requirements, ideally without requiring extra energy and/or deployment or computational costs. An emerging technology for scalably reducing non-LOS losses while circumventing several limitations related to underlying propagation environments is that of *Intelligent Reflecting Surfaces (IRSs, or RISs)*. An IRS is a metasurface comprised of a planar array of passive reflecting elements with tunable parameters, such as phase-shifts and/or amplitude gains

Manuscript received 21 June 2023; revised 16 November 2023; accepted 11 December 2023. Date of publication 25 December 2023; date of current version 24 January 2024. This work was supported by a Microsoft gift and by the NSF under Grant 2242215. An earlier version of this paper was presented at the IEEE ICASSP, 2023 [DOI: 10.1109/ICASSP49357.2023.10095269]. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Foad Sohrabi. *(Corresponding author: Hassaan Hashmi.)*

Hassaan Hashmi and Dionysis Kalogerias are with the Department of Electrical Engineering, Yale University, New Haven, CT 06520 USA (e-mail: hassaan.hashmi@yale.edu; dionysis.kalogerias@yale.edu).

Spyridon Pougkakiotis is with the School of Science and Engineering, University of Dundee, DD1 4HR Scotland, U.K. (e-mail: spougkakiotis001@dundee.ac.uk).

Digital Object Identifier 10.1109/TSP.2023.3346182

Fig. 1. Concept of an IRS-aided wireless network.

of incident signals [15], [16], [17], [18], [19], [20], [21]. A concept network in which users (or terminals) are linked with an Access Point (AP) and enjoy improved QoS from the utilization of IRSs is shown in Fig. 1.

In IRS-aided communications, the goal is to optimally tune the IRS elements along with other resources (such as AP precoders), so as to optimize a certain system utility. A standard setup is that of a weighted sumrate utility in a multi-user multiple-input single-output (MISO) downlink scenario, as depicted in Fig. 1, where the goal is to maximize the total downlink rate of a number of users/terminals actively serviced by an AP, while *passively* aided by one or multiple IRSs [22], [23], [24], [25]. For this standard setting, a core objective is to jointly optimize the IRS parameters and AP precoders under certain power constraints, noting that AP precoders are usually continuous-valued, while IRS phase-shifts can be either quantized [17], or continuously varying [18], [19], [20], [21].

Prior Art on IRS Optimization: Optimizing IRS parameters is a challenging task, particularly due to three major bottlenecks. First, a sufficiently accurate channel model will most certainly be unknown, and even if known, it depends heavily on network structure and the surrounding environment. Second, by realistically assuming that IRSs are passive components of a wireless network, it might not be possible to continuously and reliably estimate the channel-state-information (CSI) of all the intermediate channels, e.g., from APs to IRSs, and from IRSs to users/terminals [26], [27]. While some recent works [28], [29] have investigated methods for cascaded channel estimation with passive IRSs, they typically rely on specific a priori known channel models. It is thus clear that IRS optimization benefits from being free from the need of knowing such detailed CSI. Third, any reasonable IRS phase-shift tuning approach – which should only require *effective* CSI (i.e., the conventional compound channels observed at the end terminals)–, should target as infrequent IRS parameter updates as possible, at least in long-term operation mode. This is in contrast with earlier approaches that consider fully reactive IRS infrastructure demanding resource-wasteful, perpetual IRS control [30], [31]. Addressing these challenges constitutes an active area of research. In fact, one or more of these bottlenecks are persistent in most available strategies for IRS optimization in wireless networks (e.g., [27], [32]).

Recently, model-based methods relying on some flavor of stochastic successive convex optimization (SSCO) [33], [34], [35] for weighted sumrate optimization have gained substantial traction. These methods operate over *two time-scale* customized protocols, where reactive precoding vectors at the AP(s) are optimized on a shorter time-scale and non-reactive (static, "long-term") IRS beamformers are optimized on a longer time-scale [22], Section II-C], [23], [24], [25]. Apart from relying on SSCO, which operates on convex surrogates of the original problem, these model-based approaches require complete knowledge of the network structure and the channel model, along with accurate intermediate CSI (statistic) estimation [36], [37], [38], [39]. Such estimates are difficult to obtain because, as an ultimately passive device, an IRS cannot (or should not) transmit and receive pilot signals [26]. Consequently, these methods must rely on active sensing at the IRSs, demanding expensive and wasteful IRS implementations. Lastly, any change in the network or channel model incurs a high environment remodeling cost, while the modeling complexity increases dramatically with an increase in the number of intermediate channels.

To avoid such limitations, researchers have explored machine learning (ML) methods to optimize IRS-aided networks. As a parallel to channel estimation by hand, offline ML methods have been explored to approximate true CSI models from labelled datasets using function approximators (FAs) [40], [41], [42], [43]. These offline approaches are brittle in that the learned models are unable to adapt to slight changes in the network/channel behavior, induced either by movement of the users/terminals, or due to potential environmental factors. Deep Reinforcement Learning (DRL) methods, on the other hand, are adaptive policy learning methods which have been used for joint beamforming optimization. In particular, deep Q-learning based methods have been explored, assuming quantized IRS phase-shifts [44], [45], [46]. For continuous phase-shifts, off-policy policy gradient methods have also been explored [47], [48], [49], [50], primarily based on the deep deterministic policy gradient (DDPG) algorithm [51]. DRL methods are designed to be *end-to-end*. As such, all intermediate steps such as CSI estimation, channel modeling and beamforming optimization can be *offloaded* to the FAs, the latter approximating state value functions to jointly model the entire optimization task.

Nonetheless, without expert domain knowledge, defining those FAs –e.g., deep neural networks (DNNs)– often increases the problem complexity; most frequently, FAs are considered as black-box data-driven models, resulting in non-interpretability and lack of robustness. On the other hand, explicit utilization of domain knowledge within the context of FAs often results in overfitting, thus limiting the versatility and transferability of DRL models to distinct environments. Further, such learning-based methods incur increased power/resource consumption as they primarily consider reactive IRS operation, where each model output (phase-shift element values) directly depends on the observed CSI, hence resulting in perpetual IRS control. Combinations of the mixed-timescale iterative approach with ML via deep unfolding models [52] also suffer from similar limitations.

Contributions: We develop a *Zeroth-order Stochastic Gradient Ascent (ZoSGA)* algorithm for tackling *fully passive* IRS-assisted utility maximization in a wireless communication setting. We consider two-stage stochastic programming formulations of the problem, in which the first-stage problem consists of an on-average (long-term) optimization of IRS-parameters, while the second-stage problem seeks for optimal instantaneous (short-term) beamformers (e.g., of an AP) associated with a given network instance (occurring every time new –random– CSI is revealed). ZoSGA tacitly exploits WMMSE [53] as a standard method for solving the (deterministic) second-stage problems, while remaining agnostic to the channel dynamics or network topology, thus operating in a completely model-free manner. The algorithm relies on minimal system probing and terminal-end effective CSI, both conventionally available (even approximately) regardless of the number or spatial configuration of the IRSs, in sharp contrast to model-based approaches [22], [23], [24], [25].

ZoSGA does not rely on function approximations (unlike DRL methods [44], [45], [46], [47], [48], [49], [50]), and can be run in real-time since, at each time step, it only requires to probe the network twice (to obtain a sample zeroth-order gradient), greatly improving upon SSCO-based methods, such as TTS-SSCO [23], which utilize internal (model-based) sampling for approximating stochastic gradients, and –prone to error– strongly convex surrogate utilities (despite the inherent nonconvexity of the associated optimization problem). Assuming continuous-valued IRS phase-shifts (which are feasible in practice [18], [19], [20], [21]), ZoSGA treats IRSs as fully passive tunable network elements (without sensing capabilities, extra hardware, or special scheduling requirements), and can be readily applied to a wide range of different wireless network settings. Being model-free, ZoSGA avoids the inherently nonconvex unit modulus constraints associated with the IRS phase-shift elements, by utilizing the (unknown) polar representation of the effective channels. This has far reaching benefits, as we shall see in Section V-C, where we simulate *physical* IRS models.

After discussing the problem of interest, providing our assumptions and some preliminary technical results (Section II), we develop and analyze ZoSGA (Section III) under a set of general assumptions (and regularity conditions), and we establish a state-of-the-art (SOTA) convergence rate of the order of $\mathcal{O}\left(\sqrt{S}\epsilon^{-4}\right)$, where S is the total number of IRS tunable parameters and ϵ is a desirable sub-optimality target. We note that our theoretical analysis is novel and involves a minimal set of assumptions, covering a wide range of realistic settings and shedding light into the practical behavior of ZoSGA, fully characterizing its convergence. Unlike most alternative approaches, we do not consider convex approximations of the associated two-stage stochastic problem, and deal with the inherent nonconvexity of the problem by utilizing key results from optimization theory. At the same time, we bypass (without utilizing any approximation) the nonconvexity associated with unit modulus constraints, obtaining a model-free method operating in a standard Euclidean setting.

Specializing to the case of sumrate maximization on a MISO downlink scenario (Section IV), we show that most of the technical assumptions imposed by the theory are automatically satisfied, except for some mild regularity conditions on the channel which cannot be avoided. We then numerically demonstrate (Section V) that ZoSGA exhibits SOTA performance in a wide-range of scenarios, substantially outperforming the two-time scale method recently proposed in [23], which is a standard *model-based* SSCO-type SOTA baseline for the problem under consideration. Despite the latter method assuming full knowledge of channel models and spatial network configurations, the model-agnostic ZoSGA reliably learns near-optimal solutions yielding significantly better QoS. Importantly, we also demonstrate the applicability of ZoSGA in a realistic *physical* IRS setting, in which the algorithm has only *indirect access* to the amplitudes and phases of the IRS elements, by tuning the capacitances of certain varactor diodes controlling each IRS element; such IRS tuning *in the field* is a particularly unique feature of ZoSGA, not enjoyed by any other IRS tuning approach in the literature. Fully reproducible source code of simulation results can be found here.

Notation: Let $\|\cdot\|$ denote the induced norm of an associated inner-product space, defined as $\|\boldsymbol{x}\| := \sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}$ for any $\boldsymbol{x} \in \mathbb{F}^n$, where $\mathbb{F}$ is a field (assuming that $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$). In case of a complex vector we use the Hermitian inner product. In case of a matrix we assume that the induced norm is utilized. We assume a complete base probability space $(\Xi, \mathscr{F}, P)$, and use "a.e." to denote "almost every(where)". For $p \in [1, \infty)$, we use $\mathcal{Z}_p \equiv \mathcal{L}_p(\Xi, \mathscr{F}, P; \mathbb{R})$ to denote the space of all $\mathscr{F}$-measurable functions $\phi : \Xi \to \mathbb{R}$, such that $\int_\Xi |\phi|^p dP < \infty$. Given $f : \mathbb{R}^n \to \mathbb{R}$ and $\rho > 0$, we say that f is ρ-weakly convex (resp. ρ-weakly concave) if $f(\cdot) + \frac{\rho}{2}\|\cdot\|^2$ (resp. $-f(\cdot) + \frac{\rho}{2}\|\cdot\|^2$) is convex.

II. Problem Formulation and Preliminary Results

A. Problem Formulation

In an IRS-aided network, the effective channels observed at the terminals of the communication task are treated as functions of both the intermediate channels as well the phase-shift vectors of the IRSs. Thus, we may think of the IRSs as *network-defining* parameters. Indeed, each instance of their phase-shift elements produces a different wireless network altogether. Taking this into account, the objective here is to find an optimal instance of such a wireless network which maximizes a given terminal QoS utility, in accordance with all the underlying physical dynamics.

Traditionally, we only optimize the precoding vectors (e.g., at an AP) to maximize a given QoS metric function, and these precoding vectors are optimized for a particular wireless network instance. However, in an IRS-aided wireless network, tuning the IRSs essentially changes the network structure and we have to re-optimize the precoders, responding to this new network.

We assume dynamic (i.e. reactive) precoders $\boldsymbol{W}$, while (realistically) viewing the IRSs as static (i.e. non-reactive during operation) elements with tunable parameters $\boldsymbol{\theta}$ (such as in [22], [23], [24], [25]), which encode any propagation feature of the IRSs that is learnable. For example, $\boldsymbol{\theta}$ can represent amplitudes

and/or phases, or any other tunable element of an IRS (e.g., *physical* varactor tunable capacitances, see [18], [19]).

Under this setting, beamforming optimization can be formulated as a *stochastic two-stage problem*. The first-stage problem seeks for an optimal (say in expectation) wireless network by tuning the (static) IRSs' parameters $\boldsymbol{\theta}$ assuming optimal instantaneous precoders on random effective channels. The second-stage problem seeks those optimal precoders $\boldsymbol{W}$ given a (possibly optimal) network instance set by already fixing the IRSs.

We hereafter assume that the IRSs and the precoders maximize the same QoS utility function. This is an intuitive and standard choice in practical applications, which enables the development of very efficient solution methods. Formally, we are interested in two-stage problems of the form

$$\max_{\boldsymbol{\theta}\in\Theta} \mathbb{E}\left\{\max_{\boldsymbol{W}\in\mathcal{W}} F\left(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta},\omega)\right)\right\}, \tag{2SP}$$

where $\mathcal{W}$ is a (known) compact set of feasible dynamic beamformers $\boldsymbol{W}$, and $\Theta \subset \mathbb{R}^S$, is a (known) convex and compact set of feasible IRS parameter values, where S denotes the number of real-valued parameters of the complex-valued phase-shift elements (e.g., amplitudes and phases). The utility function $F\colon \mathbb{C}^{M_U} \times \mathbb{C}^{M_U} \to \mathbb{R}$ is a (known) function of the precoding vectors $\boldsymbol{W}$ as well as the (unknown) observed effective channels $\boldsymbol{H}\colon \mathbb{R}^S \times \Omega \to \mathbb{C}^{M_U}$, which in turn are functions of both the IRS parameters and any intermediate random channels (a "state of nature"), denoted as $\omega\colon \Xi \to \Omega$ (the statistics of which are unknown). The random vector ω represents anything that is unknown about the underlying communication system, such as propagation or (compound) interference patterns, internal channel states, or in general the underlying intermediate communication channels. The (observed) effective channels $\boldsymbol{H}(\cdot,\omega)$ are assumed to have unknown dynamics, and we are only allowed to evaluate them at specific IRS parameter instances $\boldsymbol{\theta} \in \Theta$.

As we discuss in Section II-C, the resulting stochastic bilevel program assumes a common function for the inner- and outer-level programs, allowing for first-order optimization, without the need of computing any second-order information (e.g. as in [54]).

B. Assumptions

In what follows, we provide certain regularity assumptions on (2SP) and subsequently prove certain core technical results, allowing us to derive the proposed optimization scheme.

Second-stage problem: Given some realization $\omega \in \Omega$, and some $\boldsymbol{\theta} \in \Theta$, the second-stage problem reads

$$\max_{\boldsymbol{W}\in\mathcal{W}} \left\{G(\boldsymbol{W},\boldsymbol{\theta},\omega) \triangleq F\left(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta},\omega)\right)\right\}. \tag{SSP}$$

Notice that (SSP) is deterministic, since we are required to solve this after the state of nature ω has been revealed.

First-stage problem: The first-stage problem, which is equivalently given in (2SP), can be provisionally written as

$$\max_{\boldsymbol{\theta}\in\Theta} \left\{f(\boldsymbol{\theta}) \triangleq \mathbb{E}\left\{F\left(\boldsymbol{W}^*(\boldsymbol{\theta},\omega), \boldsymbol{H}(\boldsymbol{\theta},\omega)\right)\right\}\right\} \tag{FSP}$$

for some $\boldsymbol{W}^*(\boldsymbol{\theta},\omega) \in \arg\max_{\boldsymbol{W}\in\mathcal{W}} F\left(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta},\omega)\right)$. In what follows, we enforce certain regularity conditions on (2SP).

Assumption A: The following conditions are in effect:

- **(A1)** The function $F\colon \mathbb{C}^{M_U} \times \mathbb{C}^{M_U} \to \mathbb{R}$ is twice continuously (real) differentiable;
- **(A2)** The sets Θ and $\mathcal{W}$ are compact, and Θ is also convex;
- **(A3)** The function $\boldsymbol{H}(\cdot,\omega)$ is B_H-uniformly bounded on Θ and twice continuously differentiable on an open set $\mathcal{U} \supset \Theta$, for a.e. $\omega \in \Omega$. Moreover, there exist numbers $L_{H,0}, L_{H,1}$, such that $\boldsymbol{H}(\cdot,\omega)$ is $L_{H,0}$-Lipschitz continuous with $L_{H,1}$-Lipschitz gradients on Θ for a.e. $\omega \in \Omega$;
- **(A4)** There exists a positive function $\widetilde{\rho}(\cdot) \in \mathcal{Z}_1$, such that $\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W}, \boldsymbol{H}(\cdot,\omega))$ is $\widetilde{\rho}(\omega)$-weakly concave on Θ;
- **(A5)** $F(\boldsymbol{W}^*(\boldsymbol{\theta},\cdot), \boldsymbol{H}(\boldsymbol{\theta},\cdot)) \in \mathcal{Z}_2$ is bounded below for all $\boldsymbol{\theta} \in \Theta$ and any $\boldsymbol{W}^*(\boldsymbol{\theta},\omega) \in \arg\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta},\omega))$, and we can draw independent and identically distributed (i.i.d.) samples from the law of ω.

Remark 1: Let us observe that Assumption A is very mild, and is informed by our application, i.e. two-stage beamforming for passive IRS-aided network design. Conditions **(A1)–(A2)** are standard and are most often met in practical settings involving IRS-aided wireless communication systems. In particular, twice-continuous (real) differentiability of $F(\cdot,\cdot)$ is standard and subsumes several utility functions of interest (three popular examples are the weighted sumrate utility, the proportional fairness utility, or the harmonic-rate utility; see [55]). Note that *real differentiability* refers to differentiability of the real and imaginary parts of F (see [56, Section 3.2]). Furthermore, as already noted, by utilizing the *polar form* of the function $\boldsymbol{H}(\cdot,\omega)$ (which we do in this work), Θ (typically) represents a set of phases and amplitudes which can be chosen to be real, compact and convex, without loss of generality. Finally, compactness of $\mathcal{W}$ is also standard, since it typically reflects constraints relating to the available power of the wireless communication system.

Condition **(A3)** ensures that the compositional function of interest is well-defined and retains its properties on an open set containing Θ, while the bound and Lipschitz random functions associated with $\boldsymbol{H}(\cdot,\omega)$ are uniformly bounded in ω. The latter condition could potentially be relaxed (e.g. assuming bounded variance of random functions $B_{H,0}(\omega)$, $L_{H,0}(\omega)$, $L_{H,1}(\omega)$), but is imposed for brevity in exposition.

(A4) is a very general condition since weak concavity subsumes a large class of functions (e.g. all Lipschitz smooth functions or all twice continuously differentiable functions on a compact set are weakly concave; see also [57, Section 2] for some additional examples). In particular, it implies that for a.e. $\omega \in \Omega$, there exists $\widetilde{\rho}(\omega) > 0$ such that $\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W}, \boldsymbol{H}(\cdot,\omega)) - \frac{\widetilde{\rho}(\omega)}{2}\|\cdot\|^2$ is concave on Θ. Specialized, albeit technical, conditions ensuring that this holds for the objective in (FSP) will be discussed later in Section IV.

Finally, condition **(A5)** ensures that f is well-defined and bounded on Θ. Notice that the same would hold under the weaker condition $F(\boldsymbol{W}^*(\boldsymbol{\theta},\cdot), \boldsymbol{H}(\boldsymbol{\theta},\cdot)) \in \mathcal{Z}_1$, however **(A5)** is utilized later in Lemma 5.

C. Technical Results

Since $F(\boldsymbol{W},\cdot)$ takes a complex input, we need to devise an appropriate gradient generalization for it. To that end, we utilize the so-called *Wirtinger calculus* (see [56]). In the following lemma we derive the full *compositional* gradient of $F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ by following the developments in [56, Section 4].

Lemma 1: For every $\boldsymbol{\theta}\in\Theta$, $\boldsymbol{W}\in\mathcal{W}$ and a.e. $\omega\in\Omega$, the gradient of $F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ with respect to $\boldsymbol{\theta}$ reads

$$\begin{aligned}&\nabla_{\boldsymbol{\theta}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega)) \qquad (1)\\ &= 2\nabla_{\boldsymbol{\theta}}\Re(\boldsymbol{H}(\boldsymbol{\theta},\omega))\,\Re\left(\frac{\partial^{\circ}}{\partial\boldsymbol{z}}F(\boldsymbol{W},\boldsymbol{z})\Big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{\theta},\omega)}\right)^{\top}\\ &\quad+2\nabla_{\boldsymbol{\theta}}\Im(\boldsymbol{H}(\boldsymbol{\theta},\omega))\,\Re\left(j\frac{\partial^{\circ}}{\partial\boldsymbol{z}}F(\boldsymbol{W},\boldsymbol{z})\Big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{\theta},\omega)}\right)^{\top},\end{aligned}$$

where $\frac{\partial^{\circ}}{\partial\boldsymbol{z}}(\cdot)$ is the Wirtinger cogradient operator.

Proof: For a complete proof, see Appendix A. □

We proceed by proving that, for a.e. $\omega\in\Omega$, the Wirtinger cogradient of F, evaluated at some $\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{\theta},\omega)$ for any $\boldsymbol{\theta}\in\Theta$, is bounded by a positive constant, independent of ω.

Lemma 2: Given Assumption A, there exists a constant $B_F>0$ such that, for all $(\boldsymbol{\theta},\boldsymbol{W})\in\Theta\times\mathcal{W}$, and a.e. $\omega\in\Omega$,

$$\left\|\frac{\partial^{\circ}}{\partial\boldsymbol{z}}F(\boldsymbol{W},\boldsymbol{z})\big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{\theta},\omega)}\right\|\le B_F.$$

Proof: From **(A3)** we obtain that the sets $\boldsymbol{H}(\Theta,\omega)$, for $\omega\in\Omega$, are uniformly (in ω) compact. Thus, we can find a compact set $\mathcal{V}$, such that $\boldsymbol{H}(\Theta,\omega)\subset\mathcal{V}$ for a.e. $\omega\in\Omega$. From **(A2)** we also have that $\mathcal{W}$ is compact. Since F is continuously (real) differentiable on the compact set $\mathcal{V}\times\mathcal{W}$, it follows from the extreme value theorem applied to $\frac{\partial}{\partial\Re(\boldsymbol{z})}F(\boldsymbol{W},\boldsymbol{z})$ and to $\frac{\partial}{\partial\Im(\boldsymbol{z})}F(\boldsymbol{W},\boldsymbol{z})$, respectively, with $\boldsymbol{z}\in\mathcal{V}$, that there exists some constant B_F such that

$$\left\|\frac{\partial}{\partial\Re(\boldsymbol{z})}F(\boldsymbol{W},\boldsymbol{z})\big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{\theta},\omega)}\right\|\le B_F,$$

$$\left\|\frac{\partial}{\partial\Im(\boldsymbol{z})}F(\boldsymbol{W},\boldsymbol{z})\big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{\theta},\omega)}\right\|\le B_F,$$

for a.e. $\omega\in\Omega$ and any $(\boldsymbol{\theta},\boldsymbol{W})\in\Theta\times\mathcal{W}$. To complete the proof, we simply bound the Wirtinger cogradient directly by its definition (adjusting the constant upper bound). □

Next, we note that under Assumption A, the second-stage problem (as a function of θ, for a.e. $\omega\in\Omega$) is a *lower-C^2* function (see [58] for a precise definition). We formalize the consequences of this fact in the following lemma.

Lemma 3: Let Assumption A hold. For a.e. $\omega\in\Omega$, there exists a constant $\widehat{\rho}(\omega)>0$ such that the mapping $\theta\mapsto\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ is $\widehat{\rho}(\omega)$-weakly convex on Θ.

Proof: From Conditions **(A1)–(A3)**, it follows easily that the mapping $\theta\mapsto\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ is lower-C^2 (see [58]). Then, [58, Theorem 6] yields local weak convexity of this mapping on Θ. Compactness of Θ then yields the result. □

We are now ready to show that f in (FSP) is actually continuously differentiable and weakly concave on Θ.

Theorem 4: Let Assumption A hold. Then, for any $\boldsymbol{\theta}\in\Theta$, the function f is well-defined and differentiable, with

$$\nabla_{\boldsymbol{\theta}}f(\boldsymbol{\theta})=\mathbb{E}\left\{\nabla_{\boldsymbol{\theta}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))\big|_{\boldsymbol{W}=\boldsymbol{W}^*(\boldsymbol{\theta},\omega)}\right\}, \qquad (2)$$

for any $\boldsymbol{W}^*(\boldsymbol{\theta},\omega)\in\arg\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$. Moreover, if $\widehat{\rho}(\cdot)\in\mathcal{Z}_1$, where $\widehat{\rho}$ is given in Lemma 3, then there exists a positive constant $\rho\triangleq\max\{\mathbb{E}\{\widetilde{\rho}(\omega)\},\mathbb{E}\{\widehat{\rho}(\omega)\}\}$, such that f is ρ-weakly concave on Θ.

Proof: Condition **(A5)** ensures that f is well-defined and finite for any $\boldsymbol{\theta}\in\Theta$. On the other hand, by combining Condition **(A4)** with Lemma 3, we obtain that the map $\boldsymbol{\theta}\mapsto\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ is continuously Fréchet differentiable (the proof of this fact can be found in [59, Corollary 4.9]). Since this mapping is *subdifferentially regular*, and its (generalized) subdifferential is a singleton, it follows that

$$\nabla_{\boldsymbol{\theta}}\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))=\nabla_{\boldsymbol{\theta}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))\big|_{\boldsymbol{W}=\boldsymbol{W}^*(\boldsymbol{\theta},\omega)},$$

for any $\boldsymbol{W}^*(\boldsymbol{\theta},\omega)\in\arg\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ (e.g. see [60]). Similarly, by utilizing the integrability of $\widetilde{\rho}(\cdot)$ and $\widehat{\rho}(\cdot)$, we obtain that f is also continuously Fréchet differentiable, (again invoking [59, Corollary 4.9], by noting that f is both ρ-weakly convex and ρ-weakly concave, with $\rho\triangleq\max\{\mathbb{E}\{\widetilde{\rho}(\omega)\},\mathbb{E}\{\widehat{\rho}(\omega)\}\}$). Using [61, Theorem 7.44], the gradient of f reads

$$\nabla_{\boldsymbol{\theta}}f(\boldsymbol{\theta})=\mathbb{E}\left\{\nabla_{\boldsymbol{\theta}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))\big|_{\boldsymbol{W}=\boldsymbol{W}^*(\boldsymbol{\theta},\omega)}\right\},$$

for any $\boldsymbol{W}^*(\boldsymbol{\theta},\cdot)\in\arg\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\cdot))$. □

Remark 2: We observe that the integrability of the weak convexity constants $\widehat{\rho}(\omega)$ (given in Lemma 3) is a very mild condition and is almost always met in practice (where one usually has a finite collection of scenarios). As such, for the rest of this article we make the implicit assumption that this holds, i.e. $\widehat{\rho}(\cdot)\in\mathcal{Z}_1$.

D. Zeroth-Order Gradient Approximation of the Channel

Assuming the lack of availability of any first-order information of $\boldsymbol{H}(\cdot,\omega)$ for any $\omega\in\Omega$, we will employ a zeroth-order scheme in order to obtain a gradient estimate of $\nabla_{\boldsymbol{\theta}}\boldsymbol{H}(\cdot,\omega)$, using which we can solve (FSP) via a stochastic projected gradient ascent scheme. The proposed method will be based on gradient estimates arising from a two-point stochastic evaluation of $\boldsymbol{H}(\cdot,\omega)$ (similar to, among others, [62], [63], [64]). From **(A3)**, we have

$$\nabla_{\boldsymbol{\theta}}\boldsymbol{H}(\boldsymbol{\theta},\omega)=\nabla_{\boldsymbol{\theta}}\Re(\boldsymbol{H}(\boldsymbol{\theta},\omega))+j\nabla_{\boldsymbol{\theta}}\Im(\boldsymbol{H}(\boldsymbol{\theta},\omega)).$$

We would like to approximate each of the above parts of the gradient using only function evaluations of $\boldsymbol{H}(\cdot,\omega)$. We let $\boldsymbol{U}\sim\mathcal{N}(0,\boldsymbol{I})$ be a normal random vector, where $\boldsymbol{I}$ is the identity matrix of size S. Given a smoothing parameter $\mu>0$, we consider the following gradient estimate

$$\begin{aligned}\nabla_{\boldsymbol{\theta}}\boldsymbol{H}_{\mu}(\boldsymbol{\theta},\omega)&\triangleq\frac{1}{2\mu}\mathbb{E}\{(\boldsymbol{H}(\boldsymbol{\theta}+\mu\boldsymbol{U},\omega)-\boldsymbol{H}(\boldsymbol{\theta}-\mu\boldsymbol{U},\omega))\,\boldsymbol{U}^{\top}\}^{\top}\\&\equiv\nabla_{\boldsymbol{\theta}}\boldsymbol{H}^{R}_{\mu}(\boldsymbol{\theta},\omega)+j\nabla_{\boldsymbol{\theta}}\boldsymbol{H}^{I}_{\mu}(\boldsymbol{\theta},\omega), \qquad (3)\end{aligned}$$

where $\nabla_{\boldsymbol{\theta}} \boldsymbol{H}_{\mu}^{R}(\boldsymbol{\theta},\omega) \triangleq \Re\left(\nabla_{\boldsymbol{\theta}} \boldsymbol{H}_{\mu}(\boldsymbol{\theta},\omega)\right)$ and $\nabla_{\boldsymbol{\theta}} \boldsymbol{H}_{\mu}^{I}(\boldsymbol{\theta},\omega) \triangleq \Im\left(\nabla_{\boldsymbol{\theta}} \boldsymbol{H}_{\mu}(\boldsymbol{\theta},\omega)\right)$. Let us notice that given condition **(A3)**, there exists an open set $\mathcal{U} \supset \Theta$ such that (3) is still well-defined. Thus, the gradient approximation is valid, even if $\boldsymbol{\theta}$ is a point in the boundary of Θ, assuming that an appropriately small μ is chosen. Observe that the smaller the value of μ is, the better the aforementioned zeroth-order approximation is. There is a trade-off between approximation accuracy and numerical stability, but in practice we observe that μ can be chosen to be quite small.

The assumption that the dynamics of $\boldsymbol{H}(\cdot,\omega)$ are unknown has multiple benefits. It allows us to bypass any modelling assumptions about the underlying communication channels, which typically incur modelling errors. At the same time, it enables the evaluation of $\boldsymbol{H}(\cdot,\omega)$ using polar coordinates. That is, we can (and we do) assume that the IRS parameters $\boldsymbol{\theta}$ represent any real-valued parameters determining the complex-valued phase-shift elements of the IRSs (e.g. amplitudes and phases). This allows us to bypass the typical nonconvex unit-modulus constraints that arise when optimizing over complex IRS phase-shifts.

The proposed zeroth-order gradient approximation of (1), based on the zeroth-order approximation given in (3), reads

$$\begin{aligned}\nabla_{\boldsymbol{\theta}}^{\mu} F\left(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta},\omega)\right) &\triangleq 2\nabla_{\boldsymbol{\theta}} \boldsymbol{H}_{\mu}^{R}(\boldsymbol{\theta},\omega)\left(\Re\left(\frac{\partial^{\circ}}{\partial \boldsymbol{z}} F(\boldsymbol{W},\boldsymbol{z})\Big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{\theta},\omega)}\right)\right)^{\top} \\ &\quad + 2\nabla_{\boldsymbol{\theta}} \boldsymbol{H}_{\mu}^{I}(\boldsymbol{\theta},\omega)\left(\Re\left(j\frac{\partial^{\circ}}{\partial \boldsymbol{z}} F(\boldsymbol{W},\boldsymbol{z})\Big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{\theta},\omega)}\right)\right)^{\top},\end{aligned} \tag{4}$$

where $\boldsymbol{H}_{\mu}^{R}(\boldsymbol{\theta},\omega)$ and $\boldsymbol{H}_{\mu}^{I}(\boldsymbol{\theta},\omega)$ are defined in (3). It should be noted here that the zeroth-order gradient approximation relies on small perturbations in the IRS parameters $\boldsymbol{\theta}$. This implies that ZoSGA is not applicable when optimizing IRSs with discrete phases/states. We identify this limitation of ZoSGA as a direction for future investigation.

III. Algorithm and Convergence Analysis

We now derive a zeroth-order projected stochastic gradient method for the solution of (FSP). To that end, we assume the availability of an oracle solving the deterministic problem (SSP). The method treats the unknown function $\boldsymbol{H}(\cdot,\omega)$ as a black-box, utilizing samples of the gradient approximation given in (3).

Assumption B: Given any $\boldsymbol{\theta} \in \Theta$ and for a.e. $\omega \in \Omega$, we have access to an oracle yielding a (measurable) optimal solution to (SSP), and Assumption A holds.

A. A Zeroth-Order Projected Stochastic Gradient Ascent

Let us briefly present the proposed (channel-agnostic) zeroth-order projected stochastic gradient ascent. From condition **(A5)**, we have available i.i.d. samples $\omega \in \Omega$. Thus, from (4), at every $(\boldsymbol{\theta}, \boldsymbol{W}) \in \Theta \times \mathcal{W}$, and for a.e. $\omega \in \Omega$, we can utilize the following sample gradient approximation

Algorithm ZoSGA Zeroth-order Stochastic Gradient Ascent

Input: $\boldsymbol{\theta}_0 \in \Theta$, $\{\eta_t\}_{t\geq 0} \subset \mathbb{R}_+$, $\mu > 0$, and $T > 0$.
for $(t = 0, 1, 2, \ldots, T)$ **do**
  Sample (i.i.d.) $\omega_t \in \Omega$, $\boldsymbol{U}_t \sim \mathcal{N}(0, \boldsymbol{I})$.
  Find $\boldsymbol{W}^* \in \arg\max_{\boldsymbol{W}\in\mathcal{W}} F\left(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta}_t, \omega_t)\right)$.
  Set $\boldsymbol{D}_{\mu}(\boldsymbol{\theta}_t, \omega_t, \boldsymbol{U}_t) \equiv \boldsymbol{D}_{\mu}(\boldsymbol{\theta}_t, \omega_t, \boldsymbol{U}_t; \boldsymbol{W}^*)$ as in (5).
  $\boldsymbol{\theta}_{t+1} = \mathbf{proj}_{\Theta}\left(\boldsymbol{\theta}_t + \eta_t \boldsymbol{D}_{\mu}\left(\boldsymbol{\theta}_t, \omega_t, \boldsymbol{U}_t\right)\right)$.
end for
Sample $t^* \in \{0, \ldots, T\}$ according to $\mathbb{P}(t^* = t) = \frac{\eta_t}{\sum_{i=0}^{T} \eta_i}$.
return $\boldsymbol{\theta}_{t^*}$.

$$\begin{aligned}\boldsymbol{D}_{\mu}(\boldsymbol{\theta},\omega,\boldsymbol{U};\boldsymbol{W}) &\equiv \boldsymbol{D}_{\mu}(\boldsymbol{\theta},\omega,\boldsymbol{U}) \\ &\triangleq \boldsymbol{\Delta}_{\mu}^{R}\left(\Re\left(\frac{\partial^{\circ}}{\partial \boldsymbol{z}} F(\boldsymbol{W},\boldsymbol{z})\Big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{\theta},\omega)}\right)\right)^{\top} \\ &\quad + \boldsymbol{\Delta}_{\mu}^{I}\left(\Re\left(j\frac{\partial^{\circ}}{\partial \boldsymbol{z}} F(\boldsymbol{W},\boldsymbol{z})\Big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{\theta},\omega)}\right)\right)^{\top},\end{aligned} \tag{5}$$

with

$$\left(\boldsymbol{\Delta}_{\mu}^{R}, \boldsymbol{\Delta}_{\mu}^{I}\right) \triangleq \frac{1}{\mu}\Big[\left(\Re\left(\boldsymbol{\Delta}_{\mu}(\boldsymbol{\theta},\omega,\boldsymbol{U})\right)\boldsymbol{U}^{\top}\right)^{\top}\left(\Im\left(\boldsymbol{\Delta}_{\mu}(\boldsymbol{\theta},\omega,\boldsymbol{U})\right)\boldsymbol{U}^{\top}\right)^{\top}\Big],$$

where $\mu > 0$ is a smoothing parameter, $\boldsymbol{U} \sim \mathcal{N}(0, \boldsymbol{I})$ and $\boldsymbol{\Delta}_{\mu}(\boldsymbol{\theta},\omega,\boldsymbol{U}) \triangleq \boldsymbol{H}\left(\boldsymbol{\theta} + \mu\boldsymbol{U}, \omega\right) - \boldsymbol{H}\left(\boldsymbol{\theta} - \mu\boldsymbol{U}, \omega\right)$. Note that this is simply a sample from (4) and is obtained by probing the wireless network twice with the perturbed IRS parameters $\boldsymbol{\theta} + \mu\boldsymbol{U}$ and $\boldsymbol{\theta} - \mu\boldsymbol{U}$, with the induced overhead due to the associated channel estimation effort, implicitly assumed to be within the coherence time of the channel; note that this a standard assumption in the related literature on learning resource policies in wireless systems; see, e.g., the seminal work [65].

At this point, it may be also worth noting that ZoSGA requires exactly three effective channels to be estimated, per operational iteration: one to communicate (on which the optimal short-term precoders $\boldsymbol{W}^*$ are calculated), and two more pertaining to the required function evaluations (system probes) for constructing the sample gradient approximations outlined above. Therefore, any conventional scheme can be employed for channel estimation, *as if no IRS is present in the system*. The proposed method is summarized in Algorithm ZoSGA.

B. Convergence Analysis

We proceed by proving the convergence of Algorithm ZoSGA. Let us start by proving certain technical results.

Lemma 5: Let Assumption B hold, and fix any $\boldsymbol{\theta} \in \Theta$. For a.e. $\omega \in \Omega$ let $\boldsymbol{W}$ be the output of the oracle at $(\boldsymbol{\theta}, \omega)$. Then, for any $\mu \geq 0$, and any $\boldsymbol{U} \sim \mathcal{N}(0, \boldsymbol{I})$, the following holds

$$\mathbb{E}\left\{\left\|\boldsymbol{D}_{\mu}(\boldsymbol{\theta},\omega,\boldsymbol{U};\boldsymbol{W})\right\|^2\right\} \leq 4B_F^2 L_{H,0}^2 (S^2 + 2S), \tag{6}$$

where $\boldsymbol{U}$ and ω are assumed to be statistically independent.

Proof: By Lemma 2, we have that for any $(\boldsymbol{\theta}, \boldsymbol{W}) \in \Theta \times \mathcal{W}$ and a.e. $\omega \in \Omega$, $\left\| \frac{\partial^\circ}{\partial \boldsymbol{z}} F(\boldsymbol{W}, \boldsymbol{z})\big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{\theta},\omega)} \right\| \leq B_F$. Using the Cauchy-Schwartz inequality, we obtain

$$\begin{aligned}\mathbb{E}\left\{\|\boldsymbol{D}_\mu(\boldsymbol{\theta},\omega,\boldsymbol{U})\|^2\right\} &\leq \frac{1}{\mu^2} B_F^2 \mathbb{E}\left\{\left\|\left(\boldsymbol{\Delta}_\mu(\boldsymbol{\theta},\omega,\boldsymbol{U})U^\top\right)^\top\right\|^2\right\} \\ &\leq \frac{1}{\mu^2} B_F^2 \mathbb{E}\left\{\mathbb{E}\left\{\|\boldsymbol{\Delta}_\mu(\boldsymbol{\theta},\omega,\boldsymbol{U})\|^2 \|\boldsymbol{U}\|^2 \big| \boldsymbol{U}\right\}\right\} \\ &\leq 4B_F^2 L_{H,0}^2 \mathbb{E}\left\{\|\boldsymbol{U}\|^4\right\} = 4B_F^2 L_{H,0}^2 (S^2+2S),\end{aligned}$$

where in the last inequality we used Lipschitz continuity of $\boldsymbol{H}(\cdot,\omega)$ (from **(A3)**) while in the last equality we evaluated the 4-th moment of the χ-distribution. Since $F(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta},\cdot)) \in \mathcal{Z}_2$ from **(A5)**, we observe that all of the above expectations are well-defined and finite and hence the proof is complete. □

Lemma 6: Let Assumption B hold, fix any $\boldsymbol{\theta} \in \Theta$, and for a.e. $\omega \in \Omega$, let $\boldsymbol{W}$ be the output of the oracle. Then, the sample gradient approximation given in (5) satisfies

$$\begin{aligned}&\mathbb{E}\left\{\boldsymbol{D}_\mu(\boldsymbol{\theta},\omega,\boldsymbol{U};\boldsymbol{W})\right] = \mathbb{E}\left\{\nabla_{\boldsymbol{\theta}}^\mu F(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta},\omega))\right\} \triangleq \widehat{\nabla} f(\boldsymbol{\theta}), \\ &\left\|\widehat{\nabla} f(\boldsymbol{\theta}) - \nabla f(\boldsymbol{\theta})\right\| \leq 2\mu B_F L_{H,1}\sqrt{MS}.\end{aligned}$$

Proof: Firstly, by utilizing Fubini's theorem we obtain that

$$\mathbb{E}\left\{\boldsymbol{D}_\mu(\boldsymbol{\theta},\omega,\boldsymbol{U};\boldsymbol{W})\right] = \mathbb{E}\left\{\nabla_x^\mu F(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta},\omega))\right\},$$

where the first expectation is taken with respect to the product measure of the two random variables ω and U. Furthermore,

$$\begin{aligned}&\|\nabla_{\boldsymbol{\theta}}^\mu F(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta},\omega)) - \nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta},\omega))\| \\ &= \left\| 2\Re\left(\frac{\partial^\circ}{\partial \boldsymbol{z}} F(\boldsymbol{W},\boldsymbol{z})\Big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{\theta},\omega)} \left(\boldsymbol{\varepsilon}^R + j\boldsymbol{\varepsilon}^I\right)\right)\right\| \\ &\leq 2B_F \left\|\boldsymbol{\varepsilon}^R + j\boldsymbol{\varepsilon}^I\right\|\end{aligned}$$

where $\boldsymbol{\varepsilon}^R \triangleq \nabla_{\boldsymbol{\theta}} \Re\left(\boldsymbol{H}(\boldsymbol{\theta},\omega)\right) - \nabla_{\boldsymbol{\theta}} \boldsymbol{H}_\mu^R(\boldsymbol{\theta},\omega)$ and $\boldsymbol{\varepsilon}^I \triangleq \nabla_{\boldsymbol{\theta}} \Im\left(\boldsymbol{H}(\boldsymbol{\theta},\omega)\right) - \nabla_{\boldsymbol{\theta}} \boldsymbol{H}_\mu^I(\boldsymbol{\theta},\omega)$, and we used Lemma 2. Next, we observe (from equivalence of norms) that

$$\begin{aligned}\|\boldsymbol{\varepsilon}^R\| &\leq \sqrt{M} \max_i \left\| \left(\nabla_{\boldsymbol{\theta}} \Re\left(\boldsymbol{H}(\boldsymbol{\theta},\omega)\right) - \nabla_{\boldsymbol{\theta}} \boldsymbol{H}_\mu^R(\boldsymbol{\theta},\omega)\right)_i \right\| \\ &\leq \frac{\mu}{2} L_{H,1}\sqrt{MS},\end{aligned}$$

where, given a matrix $\boldsymbol{A}$, $(\boldsymbol{A})_i$ denotes the i-th column vector, and the last bound follows as in the proof of [66, Theorem 1]. The same procedure can be repeated for bounding $\|j\boldsymbol{\varepsilon}^I\|$ and hence a simple application of the triangle inequality and subsequently Jensen's inequality yield the desired result. □

The Moreau envelope: Let us write the objective function of (FSP) as $\phi(\boldsymbol{\theta}) \triangleq -f(\boldsymbol{\theta}) + \delta_\Theta(\boldsymbol{\theta})$, where $\delta_\Theta(\boldsymbol{\theta})$ is the indicator function for the convex set Θ. Given some penalty parameter $\lambda > 0$, we define the proximity operator as

$$\mathbf{prox}_{\lambda\phi}(\boldsymbol{u}) \triangleq \arg\min_{\boldsymbol{\theta}\in\mathbb{R}^S} \left\{\phi(\boldsymbol{\theta}) + \frac{1}{2\lambda}\|\boldsymbol{u}-\boldsymbol{\theta}\|^2\right\},$$

and the corresponding Moreau envelope as

$$\phi^\lambda(\boldsymbol{u}) \triangleq \min_{\boldsymbol{\theta}\in\mathbb{R}^S} \left\{\phi(\boldsymbol{\theta}) + \frac{1}{2\lambda}\|\boldsymbol{u}-\boldsymbol{\theta}\|^2\right\}.$$

It is well-known that the Moreau envelope with parameter $\lambda > \rho$ (the weak convexity constant) is smooth even if $\phi(\cdot)$ is not, and the magnitude of its gradient can be used as a near-stationarity measure of the non-smooth problem of interest. Indeed, if a point $\boldsymbol{\theta}$ is ϵ-stationary for the Moreau envelope, then it is close to an near-stationary point of (2SP). This (standard) approach is adopted in this work, following [57], [64], among others.

Theorem 7: Let Assumption B be in effect and assume that $\{\boldsymbol{\theta}_t\}_{t=0}^T$, $T > 0$, is generated by ZoSGA, where $\boldsymbol{\theta}_{t^*}$ is the point that the algorithm returns. For any $\bar{\rho} > \rho$, it holds that

$$\begin{aligned}&\mathbb{E}\left\{\left\|\nabla \phi^{1/\bar{\rho}}(\boldsymbol{\theta}_{t^*})\right\|^2\right\} \qquad (7)\\ &\leq \frac{\bar{\rho}}{\bar{\rho}-\rho}\left(\frac{\phi^{1/\bar{\rho}}(x_0) - \min\phi(x) + C_2\bar{\rho}\sum_{t=0}^T \eta_t^2}{\sum_{t=0}^T \eta_t} + C_1\bar{\rho}\mu\right)\end{aligned}$$

where, letting Δ_Θ be the diameter of Θ,

$$C_1 \triangleq 2\Delta_\Theta B_F L_{H,1}\sqrt{MS}, \quad C_2 \triangleq 2B_F^2 L_{H,0}^2 (S^2+2S).$$

Moreover, if we set $\bar{\rho} = 2\rho$, and

$$\eta_t = \sqrt{\frac{\Delta_f}{2C_2\rho(T+1)}}, \qquad \text{for all } t \geq 0,$$

for some $\Delta_f \geq \phi^{1/(2\rho)}(\boldsymbol{\theta}_0) - \min\phi(\boldsymbol{\theta})$, then it holds that

$$\mathbb{E}\left\{\left\|\nabla \phi^{1/(2\rho)}(\boldsymbol{\theta}_{t^*})\right\|^2\right\} \leq 8\left(\sqrt{\frac{\Delta_f \rho C_2}{2(T+1)}} + C_1\rho\mu\right). \qquad (8)$$

Proof: For any $t \geq 0$, we have $\widehat{\nabla} f(\boldsymbol{\theta}_t) \equiv \mathbb{E}_{[t]}\left\{\boldsymbol{D}_\mu(\boldsymbol{\theta},\omega,\boldsymbol{U})\right\}$, where $\mathbb{E}_{[t]}\{\cdot\} \triangleq \mathbb{E}\{\cdot | \boldsymbol{U}_{t-1}, \omega_{t-1}, \ldots, \boldsymbol{U}_0, \omega_0\}$ (see Lemma 6). We define the point $\hat{\boldsymbol{\theta}}_t \triangleq \mathbf{prox}_{\phi/\bar{\rho}}(\boldsymbol{\theta}_t)$. Then, we obtain

$$\begin{aligned}&\mathbb{E}_{[t]}\left\{\phi^{1/\bar{\rho}}(\boldsymbol{\theta}_{t+1})\right\} \leq \mathbb{E}_{[t]}\left\{f(\bar{\boldsymbol{\theta}}_t) + \frac{\bar{\rho}}{2}\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_{t+1}\|^2\right\} \\ &= \phi(\hat{\boldsymbol{\theta}}_t) + \frac{\bar{\rho}}{2}\mathbb{E}_{[t]}\Big\{\|\mathbf{proj}_\Theta(\hat{\boldsymbol{\theta}}_t) \\ &\qquad - \mathbf{proj}_\Theta\left(\boldsymbol{\theta}_t + \eta_t \boldsymbol{D}_\mu(\boldsymbol{\theta}_t,\omega_t,\boldsymbol{U}_t)\right)\|^2\Big\} \\ &\leq \phi(\hat{\boldsymbol{\theta}}_t) + \frac{\bar{\rho}}{2}\mathbb{E}_{[t]}\left\{\left\|\boldsymbol{\theta}_t + \eta_t \boldsymbol{D}_\mu(\boldsymbol{\theta}_t,\omega_t,\boldsymbol{U}_t) - \hat{\boldsymbol{\theta}}_t\right\|^2\right\} \\ &\leq \phi(\hat{\boldsymbol{\theta}}_t) + \frac{\bar{\rho}}{2}\|\boldsymbol{\theta}_t - \hat{\boldsymbol{\theta}}_t\|^2 \\ &\quad + \bar{\rho}\eta_t \mathbb{E}_{[t]}\left\{\left\langle \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_t, -\boldsymbol{D}_\mu(\boldsymbol{\theta}_t,\omega_t,\boldsymbol{U}_t)\right\rangle\right\} + C_2\bar{\rho}\eta_t^2 \\ &= \phi(\hat{\boldsymbol{\theta}}_t) + \frac{\bar{\rho}}{2}\|\boldsymbol{\theta}_t - \hat{\boldsymbol{\theta}}_t\|^2 + \bar{\rho}\eta_t\left\langle \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_t, -\nabla f(\boldsymbol{\theta})\right\rangle \\ &\quad + \bar{\rho}\eta_t\left\langle \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_t, \nabla f(\boldsymbol{\theta}_t) - \widehat{\nabla} f(\boldsymbol{\theta}_t)\right\rangle + C_2\bar{\rho}\eta_t^2 \\ &\leq \phi^{1/\bar{\rho}}(\boldsymbol{\theta}_t) + \bar{\rho}\eta_t\left\langle \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_t, -\nabla f(\boldsymbol{\theta})\right\rangle \\ &\quad + \bar{\rho}\eta_t\left\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_t\right\|\left\|\widehat{\nabla} f(\boldsymbol{\theta}) - \nabla f(\boldsymbol{\theta})\right\| + C_2\bar{\rho}\eta_t^2 \\ &\leq \phi^{1/\bar{\rho}}(\boldsymbol{\theta}_t) + \bar{\rho}\eta_t\left(f(\boldsymbol{\theta}_t) - f(\hat{\boldsymbol{\theta}}_t) + \frac{\rho}{2}\left\|\boldsymbol{\theta}_t - \hat{\boldsymbol{\theta}}_t\right\|^2\right) \\ &\quad + C_1\bar{\rho}\mu\eta_t + C_2\bar{\rho}\eta_t^2,\end{aligned}$$

where in the second inequality we used the non-expansiveness of the projection, in the fourth inequality we used Cauchy-Schwartz as well as the definition of the Moreau envelope, and

in the fifth inequality we used the weak convexity of $-f(\cdot)$, Lemma 6 and the fact that Θ is assumed to be compact (and hence there exists a constant $\Delta_\Theta > 0$ such that $\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_t\| \leq \Delta_\Theta$). Next, by following exactly the developments in [57, Section 3.1], we notice that the mapping $\boldsymbol{\theta} \mapsto -f(\boldsymbol{\theta}) + \frac{\bar{\rho}}{2}\|\boldsymbol{\theta} - \boldsymbol{\theta}_t\|^2$ is strongly convex with parameter $\bar{\rho} - \rho$, and is minimized at $\hat{\boldsymbol{\theta}}_t$, thus we obtain

$$\begin{aligned} f(\hat{\boldsymbol{\theta}}_t) - f(\boldsymbol{\theta}_t) - \frac{\rho}{2}\|\boldsymbol{\theta}_t - \hat{\boldsymbol{\theta}}_t\|^2 &= \left(-f(\boldsymbol{\theta}_t) + \frac{\bar{\rho}}{2}\left\|\hat{\boldsymbol{\theta}}_t - \hat{\boldsymbol{\theta}}_t\right\|^2\right) \\ &\quad - \left(-f(\hat{\boldsymbol{\theta}}_t) + \frac{\bar{\rho}}{2}\left\|\boldsymbol{\theta}_t - \hat{\boldsymbol{\theta}}_t\right\|^2\right) + \frac{\bar{\rho} - \rho}{2}\left\|\boldsymbol{\theta}_t - \hat{\boldsymbol{\theta}}_t\right\|^2 \\ &\geq (\bar{\rho} - \rho)\left\|\boldsymbol{\theta}_t - \hat{\boldsymbol{\theta}}_t\right\|^2 \equiv \frac{\bar{\rho} - \rho}{\bar{\rho}^2}\left\|\nabla\phi^{1/\bar{\rho}}(\boldsymbol{\theta}_t)\right\|^2, \end{aligned}$$

where the last equivalence follows from [57, Lemma 2.2]. Hence,

$$\begin{aligned} \mathbb{E}_{[t]}\left\{\phi^{1/\bar{\rho}}(\boldsymbol{\theta}_{t+1})\right\} \leq \phi^{1/\bar{\rho}}(\boldsymbol{\theta}_t) - \frac{\eta_t(\bar{\rho} - \rho)}{\bar{\rho}}\left\|\nabla\phi^{1/\bar{\rho}}(\boldsymbol{\theta}_t)\right\|^2 \\ + C_1\bar{\rho}\mu\eta_t + C_2\bar{\rho}\eta_t^2. \end{aligned}$$

Taking expectations with respect to the history ω_0, $\boldsymbol{U}_0$, $\ldots$, ω_{t-1}, $\boldsymbol{U}_{t-1}$ and using the total expectation, yields

$$\begin{aligned} \mathbb{E}\left\{\phi^{1/\bar{\rho}}(\boldsymbol{\theta}_{t+1})\right\} \leq \mathbb{E}\left\{\phi^{1/\bar{\rho}}(\boldsymbol{\theta}_t)\right\} + \bar{\rho}\eta_t(\mu C_1 + C_2\eta_t) \\ - \frac{\eta_t(\bar{\rho} - \rho)}{\bar{\rho}}\mathbb{E}\left\{\left\|\nabla\phi^{1/\bar{\rho}}(\boldsymbol{\theta}_t)\right\|^2\right\}, \end{aligned}$$

Subsequently, we can unfold the latter inequality to obtain

$$\begin{aligned} \mathbb{E}\left\{\phi^{1/\bar{\rho}}(\boldsymbol{\theta}_{T+1})\right\} \leq \phi^{1/\bar{\rho}}(\boldsymbol{\theta}_0) + C_1\bar{\rho}\mu\sum_{t=0}^{T}\eta_t + C_2\bar{\rho}\sum_{t=0}^{T}\eta_t^2 \\ - \frac{\bar{\rho} - \rho}{\bar{\rho}}\sum_{t=0}^{T}\eta_t\mathbb{E}\left\{\left\|\nabla\phi^{1/\bar{\rho}}(\boldsymbol{\theta}_t)\right\|^2\right\}. \end{aligned}$$

Then, we can lower bound the left-hand side by $\phi(\boldsymbol{\theta}^*) \triangleq \min_{\boldsymbol{\theta}\in\Theta} f(\boldsymbol{\theta})$, and rearrange, to obtain

$$\begin{aligned} &\frac{1}{\sum_{t=0}^{T}\eta_t}\sum_{t=0}^{T}\eta_t\mathbb{E}\left\{\left\|\nabla\phi^{1/\bar{\rho}}(\boldsymbol{\theta}_t)\right\|^2\right\} \\ &\quad\leq \frac{\bar{\rho}}{\bar{\rho} - \rho}\left(\frac{\phi^{1/\bar{\rho}}(\boldsymbol{\theta}_0) - \phi(\boldsymbol{\theta}^*) + C_2\bar{\rho}\sum_{t=0}^{T}\eta_t^2}{\sum_{t=0}^{T}\eta_t} + C_1\bar{\rho}\mu\right). \end{aligned}$$

Since the left-hand side is exactly $\mathbb{E}\{\|\nabla\phi^{1/\bar{\rho}}(\boldsymbol{\theta}_{t^*})\|^2\}$, we deduce that (7) holds. Finally, setting $\bar{\rho} = 2\rho$, letting $\Delta_f \geq \phi^{1/\bar{\rho}}(\boldsymbol{\theta}_0) - \min\phi(\boldsymbol{\theta})$, and choosing a constant step size as

$$\eta_t = \sqrt{\frac{\Delta_f}{2C_2(T+1)}}, \qquad \text{for all } t \geq 0,$$

we obtain (8) which completes the proof. □

Remark 3: Note that choosing $\mu = \mathcal{O}\big(1/\sqrt{(MT)}\big)$ yields that $\mathbb{E}\{\|\nabla\varphi^{1/(2\rho)}(\boldsymbol{\theta}_{t^*})\|\} \leq \epsilon$, after $\mathcal{O}(\sqrt{S}\epsilon^{-4})$ iterations.

IV. Case Study: SumRate Maximization

Capitalizing on a standard IRS-aided MISO downlink scenario (see Fig. 2 in Section V), our goal here is to maximize the total downlink rate of K users actively serviced by an AP with M antennas, while passively aided by one or multiple IRSs, arbitrarily spatially placed. As usual, we assume dynamic (reactive) AP precoders, while the IRS beamformers are static (non-reactive) tunable elements. We make no sensing assumptions on the IRSs, i.e., the IRSs are completely passive network elements.

Each of the users $k = 1, \ldots, K$ experiences a random effective channel denoted by $\boldsymbol{h}_k(\boldsymbol{\theta}, \omega)$, indexed by the IRS parameter vector $\boldsymbol{\theta}$ as well as the usual state of nature $\omega \in \Omega$ describing *unobservable* random propagation patterns for each value of $\boldsymbol{\theta}$. In other words, $\boldsymbol{h}_k(\boldsymbol{\theta}, \omega)$ is a random channel field with spatial variable $\boldsymbol{\theta}$. We make the standard assumption that the effective channels $\boldsymbol{h}_k(\boldsymbol{\theta}, \omega), k = 1, \ldots, K$, are known to the AP at the time of transmission [22], [23]. Note that the implementation complexity of estimating effective channels in our setting is exactly the same as that in conventional multi-user downlink beamforming (i.e., involving no IRSs), *regardless* of the number and/or spatial configuration of the IRSs; no extra hardware or customized scheduling schemes are required on either the AP or the IRSs assisting the network.

The QoS of user k is measured by the corresponding SINR,

$$\mathrm{SINR}_k(\boldsymbol{W}, \boldsymbol{h}_k(\boldsymbol{\theta}, \omega)) \triangleq \frac{\left|\boldsymbol{h}_k^{\mathsf{H}}(\boldsymbol{\theta}, \omega)\boldsymbol{w}_k\right|^2}{\sum_{j\in\mathbb{N}_K^+\setminus k}\left|\boldsymbol{h}_k^{\mathsf{H}}(\boldsymbol{\theta}, \omega)\boldsymbol{w}_j\right|^2 + \sigma_k^2},$$

where $\boldsymbol{W} = \mathrm{vec}(\begin{bmatrix}\boldsymbol{w}_1 & \boldsymbol{w}_2 & \cdots & \boldsymbol{w}_K\end{bmatrix}) \in \mathbb{C}^{M_U \triangleq (M\times K)}$, $\boldsymbol{w}_k$ is a transmit precoding vector and σ_k^2 is the noise variance for user k, respectively. Then, the *weighted sumrate utility* of the network is defined as

$$F(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta}, \omega)) \triangleq \sum_{k=1}^{K}\alpha_k \log_2\left(1 + \mathrm{SINR}_k\left(\boldsymbol{W}, \boldsymbol{h}_k(\boldsymbol{\theta}, \omega)\right)\right),$$

with $\boldsymbol{H} = \mathrm{vec}(\begin{bmatrix}\boldsymbol{h}_1 \ldots \boldsymbol{h}_K\end{bmatrix}) \in \mathbb{C}^{M_U}$, and $\alpha_k \geq 0$ the weight associated with user k. We are interested in maximizing the sumrate of the network jointly by selecting instantaneous-optimal dynamic AP precoders $\boldsymbol{W}$, and on-average-optimal static IRS beamformers $\boldsymbol{\theta}$ [22], [23], i.e., we are interested in the problem

$$\max_{\boldsymbol{\theta}\in\Theta}\mathbb{E}\left\{\max_{\boldsymbol{W}:\|\boldsymbol{W}\|^2\leq P} F\left(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta}, \omega)\right)\right\}, \qquad \text{(2SSRM)}$$

where $P > 0$ is a total power budget at the AP, and Θ is a real convex and compact feasible set of amplitudes and phases. Problem (2SSRM) is an instance of (2SP).

Assumption Compatibility: Let us now briefly discuss the compatibility of Assumption A with problem (2SSRM). We firstly note that conditions **(A1)**, **(A2)** are both satisfied. Condition **(A5)** is also satisfied in light of uniform boundedness. Condition **(A3)** of Assumption A imposes regularity that is required for the grounded development of our optimization scheme and for its convergence analysis. Also, observe that the boundedness assumption in condition **(A3)** is natural, since

(IRS-aided) wireless channels are always bounded in practice. While we usually have no information on the analytical properties of the effective channel, this condition is easily satisfied in widely used channel models of IRS-aided systems, see, e.g., [23] or Section V.

We next showcase that the regularity condition **(A4)** also holds under several reasonable circumstances. In particular, we identify three typical situations under which condition **(A4)** (i.e. weak concavity of the sample objective function of the first-stage problem (FSP)) is readily satisfied. These will be stated here for completeness, and the reader is referred to Appendix B for a technical discussion showcasing how weak concavity can be shown in each of the following cases.

Firstly, condition **(A4)** is readily satisfied in cases where the second-stage problem (SSP) admits a unique solution, assuming, of course, the *strong second-order sufficient optimality conditions* for the second-stage problem (we refer the reader to Appendix B for a precise description of these conditions). Indeed, in this case one can invoke the Implicit Function Theorem [67, Theorem 1B.1] to showcase twice-continuous differentiability of $f(\cdot)$ (as well as of $\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$, for a.e. $\omega\in\Omega$), and since f is considered on a compact set, weak concavity then follows immediately. On the other hand, if the solution set of the second-stage problem is connected (instead of a singleton), and instead of the strong second-order sufficient conditions, the problem satisfies the regularity condition given in [68, Assumption 4], we also obtain twice-continuous differentiability and thus weak concavity. Finally, if $\boldsymbol{H}(\cdot,\omega)$ is real-analytic for a.e. $\omega\in\Omega$ (which often holds for channel models appearing in the literature; e.g. see [23] and Section V), then one can show that the function $\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ is *sub-analytic*, and satisfies the *Łojasiewicz inequality* (see [69]) at every $\boldsymbol{\theta}\in\Theta$, with uniform exponent. If the Łojasiewicz constant is uniformly bounded and the second-stage problem satisfies the strong second-order sufficient optimality conditions, weak concavity also follows (see Appendix B for the technical details).

A. Practical Per-Iteration Complexity for Sumrate Maximization

Before proceeding with the associated simulations of our case study, it would be useful to calculate the practical *per-iteration complexity* of ZoSGA. To that end, we assume that the (deterministic here-and-now) second-stage problem is (approximately) solved using T_2 iterations of the well-known WMMSE algorithm [53], which has a documented computational complexity given by $\mathcal{O}(T_2K^2M^3)$. For the first-stage optimization step, the complexity of the zeroth-order stochastic gradient approximations can be shown to be of the order of $\mathcal{O}(K^2M+KM^2+S)$, where S is the number of IRS phase-shift parameters. Thus, the effective per-iteration complexity of ZoSGA is of order $\mathcal{O}(T_2K^2M^3+S)$.

As mentioned in Section I, model-based SSCO methods have been recently considered for tackling problem (2SSRM), mainly due to their fast convergence. Then, it would be appropriate to compare the per-iteration complexity of ZoSGA with that of a state-of-the-art SSCO method, in particular TTS-SSCO [23].

In order to derive the practical *per-iteration* complexity of TTS-SSCO, we need to assume that the effective channel $\boldsymbol{H}(\cdot,\omega)$ follows a specific and common cascaded model [23]. Considering this, the complexity of TTS-SSCO can be shown to be of order $\mathcal{O}(T_H(T_2K^2M^3+K^2S^2M))$, since each TTS-SSCO iteration requires T_H WMMSE runs for gradient statistical approximation. We should also emphasize that this complexity is *in addition to* the (highly nontrivial) computational effort required for estimating *cascaded* CSI statistics (called S-CSI in [23]), which are required by TTS-SSCO. It is evident that the proposed algorithm (i.e. ZoSGA) has a significantly smaller per-iteration complexity.

V. Simulations

Building upon the IRS-aided MISO downlink sumrate maximization case discussed in Section IV, we present a set of detailed simulations to empirically evaluate and confirm the efficacy of the proposed ZoSGA algorithm. Unless stated otherwise, all our empirical results are averaged over 2000 independent simulations, with an additional 4-th order Savitzky–Golay filter [70] of length 500, while 95% confidence intervals are also provided. In what follows, we describe and evaluate three distinct wireless network setups and examine the obtained empirical results. The last set of experiments will involve a physical model for the IRSs, which is briefly described in the beginning of Subsection V-C.

A. Baseline Channel Model

With the practically feasible assumption of insufficient angular spread of signals in a scattering environment, we consider both LOS and non-LOS channels. Following [23], we consider three types of intermediate channels in our simulations, namely, the reflected channel $\boldsymbol{h}_{r,k}$ from an IRS to user k, the channel $\boldsymbol{G}$ from an AP to an IRS, as well as the direct path channel $\boldsymbol{h}_{d,k}$ from an AP to user k. We model these as general spatially correlated Rician fading channels [71]. Additionally, we assume that the second-order statistics of the LOS links are identical for all users, due to the large distances between an AP and its served users.

Concretely, for each user k, we define the path loss adjusted versions of these channels, respectively, as

$$\begin{aligned}\breve{\boldsymbol{h}}_{r,k}&\triangleq\sqrt{\frac{\beta_{Iu}}{1+\beta_{Iu}}}\breve{\boldsymbol{v}}_{r,k}+\sqrt{\frac{1}{1+\beta_{Iu}}}\boldsymbol{\Phi}_{r,k}^{1/2}\boldsymbol{v}_{r,k},\\ \breve{\boldsymbol{G}}&\triangleq\sqrt{\frac{\beta_{AI}}{1+\beta_{AI}}}\breve{\boldsymbol{F}}+\sqrt{\frac{1}{1+\beta_{AI}}}\boldsymbol{\Phi}_{r}^{1/2}\boldsymbol{F}\boldsymbol{\Phi}_{d}^{1/2},\\ \breve{\boldsymbol{h}}_{d,k}&\triangleq\sqrt{\frac{\beta_{Au}}{1+\beta_{Au}}}\breve{\boldsymbol{v}}_{d,k}+\sqrt{\frac{1}{1+\beta_{Au}}}\boldsymbol{\Phi}_{d}^{1/2}\boldsymbol{v}_{d,k},\end{aligned}\tag{9}$$

where $\boldsymbol{v}_{r,k}\in\mathbb{C}^{N\times1}$, $\boldsymbol{F}\in\mathbb{C}^{N\times M}$, and $\boldsymbol{v}_{d,k}\in\mathbb{C}^{M\times1}$ are the instantaneous components (I-CSI), and $\breve{\boldsymbol{v}}_{r,k}\in\mathbb{C}^{N\times1}$, $\breve{\boldsymbol{F}}\in\mathbb{C}^{N\times M}$, and $\breve{\boldsymbol{v}}_{d,k}\in\mathbb{C}^{M\times1}$ are the statistical components (S-CSI) of the above channels, all having i.i.d. circularly

symmetric complex Gaussian (CSCG) entries with zero mean and unit variance, and with S-CSI being sampled once per simulation. The dimensions M and N denote the number of AP antennas and the number of passive reflecting elements in an IRS, respectively. The matrices $\boldsymbol{\Phi}_{r,k} \in \mathbb{C}^{N\times N}$, $\boldsymbol{\Phi}_d \in \mathbb{C}^{M\times M}$, and $\boldsymbol{\Phi}_r \in \mathbb{C}^{N\times N}$ are, in order, the reflected channel correlation matrix, the AP transmit correlation matrix, and the IRS receive correlation matrix. We assume an exponential correlation model for $\boldsymbol{\Phi}_d$ [72], expressed as

$$\boldsymbol{\Phi}_d(i,j) = \begin{cases} r_d^{j-i}, & \text{if} \quad i \leq j, \\ \boldsymbol{\Phi}_d(j,i), & \text{if} \quad i > j, \end{cases} \tag{10}$$

where $r_d \in (0,1)$ is the correlation coefficient. The matrices $\boldsymbol{\Phi}_r$ and $\boldsymbol{\Phi}_{r,k}$ are modeled as Kronecker products [73] as

$$\boldsymbol{\Phi}_r = \boldsymbol{\Phi}_r^h \otimes \boldsymbol{\Phi}_r^v, \qquad \boldsymbol{\Phi}_{r,k} = \boldsymbol{\Phi}_{r,k}^h \otimes \boldsymbol{\Phi}_{r,k}^v,$$

where $\boldsymbol{\Phi}_r^h$, $\boldsymbol{\Phi}_{r,k}^h$, and $\boldsymbol{\Phi}_r^v$, $\boldsymbol{\Phi}_{r,k}^v$, for $k \in \mathbb{N}_K^+$, represent the spatial correlation matrices of the horizontal and vertical domains, respectively, and are all modeled similar to $\boldsymbol{\Phi}_d$ in (10), with spatial correlation coefficients $r_r \in (0,1)$ and $r_{r,k} \in (0,1)$ for $\boldsymbol{\Phi}_r$, and $\boldsymbol{\Phi}_{r,k}$, $k \in \mathbb{N}_K^+$, respectively.

The deterministic components, i.e, $\breve{\boldsymbol{v}}_{r,k}, \breve{\boldsymbol{F}}$, and $\breve{\boldsymbol{v}}_{d,k}$ determine the moments of their respective CSIs. Lastly, the real scalar β_{Au} denotes the Rician fading factor for the LOS channel, while β_{Iu} and β_{AI} are the same for the reflected channels. These factors define the relative dominance of I-CSIs and S-CSIs in their respective combined CSIs. Moreover, all intermediate channels in (9) suffer from an exponential path loss proportional to the path distance. We model this loss as $L_\alpha(d) = \sqrt{C_0 d^{-\alpha}}$, where d is path distance in meters, α is the path loss exponent depending upon the channel being considered, and C_0 is the common path-loss for when the path distance is one meter.

Considering both path loss and fading, we may adopt the standard simplified baseline model for the effective channel $\boldsymbol{h}_k(\boldsymbol{\theta},\omega)$ of user k in the presence of an AP and *one IRS*, i.e.,

$$\boldsymbol{h}_k(\boldsymbol{\theta},\omega) \triangleq \underbrace{\boldsymbol{G}^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{A}\circ e^{j\boldsymbol{\phi}})\boldsymbol{h}_{r,k}}_{\text{non-LOS link}} + \underbrace{\boldsymbol{h}_{d,k}}_{\text{LOS link}},$$

where $\boldsymbol{h}_{r,k} = L_{\alpha_{Iu}}(d_{Iu,k})\breve{\boldsymbol{h}}_{r,k}$, $\boldsymbol{G} = L_{\alpha_{AI}}(d_{AI})\breve{\boldsymbol{G}}$, and $\boldsymbol{h}_{d,k} = L_{\alpha_{Au}}(d_{Au,k})\breve{\boldsymbol{h}}_{d,k}$. We may take $\omega = \{\boldsymbol{G}, \boldsymbol{h}_{r,k}, \boldsymbol{h}_{d,k}, k \in \mathbb{N}_K^+\}$, while the IRS parameters $\boldsymbol{\theta}$ are represented by amplitude and phase vectors $\boldsymbol{A} \in [0,1]^N$ and $\boldsymbol{\phi} \in [-2\pi, 2\pi]^N$, respectively [74]. Adding more IRSs to the system increases the non-LOS terms comprising the effective channel of each user accordingly.

For slow moving users, as is generally the case, we assume that the values of S-CSI and the spatial correlation matrices remain fixed throughout the duration of the AP service. Further, we assume that IRS-to-IRS links do not contribute to the signal or the interference in the presented multi-IRS cases. Of course, the latter assumption is only made for ease of presentation.

In Subsections V-B and V-C, having defined the channel model, we present simulated results with ideal and physical IRSs, where physical IRSs are constrained both in terms of amplitude and phase, as well as non-linear sensitivity relative to the action space, i.e., the ranges of varactor diode capacitances [75].

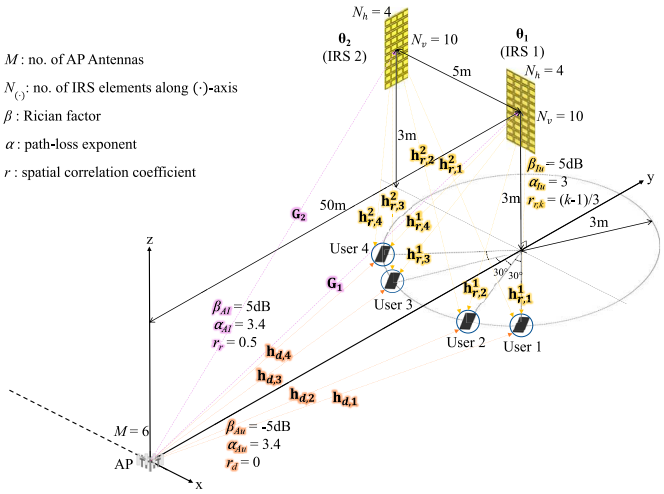


Fig. 2. First IRS-aided network configuration (ideal IRSs).

B. Networks With Ideal IRSs

In the first set of simulations, we assume that we have full control over the amplitudes and the phases of IRS phase-shift elements, i.e. there is no constraint on achievable phase-amplitude pairs, and we can control them directly. We start by discussing the wireless network setting as shown in Fig. 2.

To highlight the efficacy of the presence of IRSs in a wireless network, we assume a more acute signal attenuation in the LOS links from the AP to the users. Thus, we set $\alpha_{Au} = 3.4$, $\alpha_{Iu} = 3$ and $\alpha_{AI} = 2.2$, where the largest α_{Au} is the LOS path loss exponent while the remaining two are path loss exponents of IRS-User and AP-IRS links, respectively. Moreover, since the distances between the IRS and its served users are relatively small, IRS elements reflect signals with a finite angular spread and a user-location dependent mean angle in practice [76].

We consider two IRSs, as shown in Fig. 2, equipped with $N{=}N_h{\times}N_v$ rectangular phase-shift elements where $N_h(=4)$ and $N_v(=10)$ denote the number of columns and rows, respectively ($N = 40$). For the ideal IRS case we define the controllable parameters of each IRS as a vector $\boldsymbol{\theta}_i{=}\left[\boldsymbol{\phi}_i^{\mathsf{T}}\ \boldsymbol{A}_i^{\mathsf{T}}\right]^{\mathsf{T}}$, for $i \in \{1,2\}$, where $\boldsymbol{\phi}_i \in [-2\pi, 2\pi]^N$ and $\boldsymbol{A}_i \in [0,1]^N$ are phases and amplitudes of the IRS elements, respectively. We do not consider the relative orientations of the IRSs, AP and users, as those can be incorporated via orientation offsets, if needed. The effective channel for a user k can, thus, be expressed as

$$\boldsymbol{h}_k(\boldsymbol{\theta},\omega) = \sum_{i=1}^{2} \underbrace{\boldsymbol{G}_i^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{A}_i\circ e^{j\boldsymbol{\phi}_i})\boldsymbol{h}_{r,k}^i}_{\boldsymbol{\theta}_i\text{-non-LoS link}} + \underbrace{\boldsymbol{h}_{d,k}}_{\text{LoS link}},$$

where $\boldsymbol{h}_{r,k}^i = L_{\alpha_{Iu}^i}(d_{Iu,k,i})\breve{\boldsymbol{h}}_{r,k}^i$, $\boldsymbol{G}^i = L_{\alpha_{AI}^i}(d_{AI,i})\breve{\boldsymbol{G}}^i$, and $\boldsymbol{h}_{d,k}^i = L_{\alpha_{Au}^i}(d_{Au,k,i})\breve{\boldsymbol{h}}_{d,k}^i$ for $i \in \{1,2\}$. Also, for brevity, we take $\boldsymbol{\theta}{=}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ and $\omega = \{\boldsymbol{G}_1, \boldsymbol{G}_2, \boldsymbol{h}_{r,k}^1, \boldsymbol{h}_{r,k}^2, \boldsymbol{h}_{d,k}, k \in \mathbb{N}_K^+\}$.

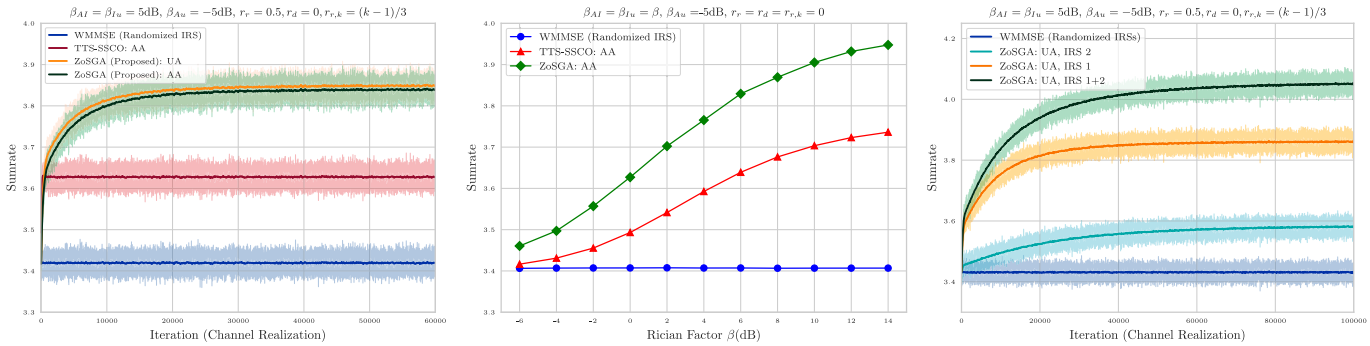


Fig. 3. (a) Average sumrates achieved by WMMSE [53] (random IRS phase-shifts), TTS-SSCO [23], and ZoSGA, with only IRS 1 present (AA: Adjustable Amplitude | UA: Unit Amplitude); (b) Corresponding average sumrates versus the Rician factor; (c) Average sumrates achieved by WMMSE (random IRS) and ZoSGA, with both IRSs present.

After defining the wireless network setting, we move forward with our simulations. In the first simulated comparison we compare the proposed algorithm (ZoSGA) with the well-documented Stochastic Successive Convex Optimization (SSCO) method, specifically a version of it proposed in [23], which we shall refer to as Two-Time Scale SSCO (TTS-SSCO). Both ZoSGA and TTS-SSCO employ the WMMSE algorithm to optimize the precoding vectors. To keep the comparison justified, we let WMMSE optimize for 20 iterations per channel instance for both of these methods, and have also included WMMSE with randomized IRS parameters as a reference. All parameters pertaining to the TTS-SSCO are taken from [23, Section V], i.e., $T_H = 10$, $\tau = 0.01$, $\rho_t = t^{-0.8}$ and $\gamma_t = t^{-1}$. There are four users, so $k \in \mathbb{N}_4^+$, weighted uniformly i.e., $\alpha_k = 1$; the AP has six antennas ($M = 6$), and the noise variance is $\sigma_k = -80$dBm for all k. The reference path loss is $C_0 = -30$dB, and the total allocated power budget is $P = 5$dBm. The LOS Rician factor is $\beta_{Au} = -5$dB, while $\beta_{Iu} = \beta_{AI} = 5$dB, unless specified otherwise. We let the smoothing parameter $\mu = 10^{-12}$ in Algorithm ZoSGA, and choose separate initial step-sizes $\eta_\phi^0 = 0.4$ and $\eta_A^0 = 0.01$ for updating phases and amplitudes, respectively, scaled by 0.9972^t for $t \in \mathbb{N}_{10^3}^+$, keeping them constant for $t > 10^3$.

The comparison of ZoSGA and TTS-SSCO in terms of achieved sumrate is shown in Fig. 3(a). The comparison is done with only IRS 1 present, matching the TTS-SSCO setting in [23], which requires exact channel and network models. We observe that ZoSGA, although requiring more iterations to converge, substantially outperforms TTS-SSCO solely on the basis of effective CSI, while having no access to the statistical model of the channel or the spatial configuration of the network. There are two main reasons for this gain in performance. Firstly, TTS-SSCO evaluates gradients by utilizing internally sampled I-CSI and the corresponding optimal precoding vectors (via WMMSE), with a frequency of ten samples per iteration (T_H). This can greatly limit the convergence of TTS-SSCO if the internal channel model is not accurate. Secondly, the surrogate objective optimization employed by TTS-SSCO, which simplifies the problem by decoupling the IRS phase-shift elements, allowing the computation of their optimal values in a closed form, is not equivalent to the original nonconvex sumrate optimization problem. As shown by the convergence curves provided in Fig. 3(a), this introduces additional errors, preventing TTS-SSCO from realizing a competitive QoS gain, compared with ZoSGA.

In Fig. 3(b), we discern the effect of the Rician factor for spatially uncorrelated channels, i.e., we set $r_d = r_r = r_{r,k} = 0$, $\forall\ k \in \mathbb{N}_4^+$. We observe that the relative gain of ZoSGA in the achievable sumrate increases with respect to the Rician factors pertaining to $\boldsymbol{\theta}_1$-reflected links, i.e., as we move from I-CSI to S-CSI dominated channels. The performance gap between ZoSGA and TTS-SSCO increases in β, since in a close-to-deterministic effective channel zeroth-order gradient approximations approach the true gradients, while TTS-SSCO is optimizing an approximate surrogate objective. We also observe that WMMSE with a randomized IRS remains insensitive to changes in the Rician factors; this is expected as the reflected channel is not optimized to gain any performance improvements.

We can now verify that ZoSGA can also optimize networks with multiple IRSs without any model knowledge. We do so by tuning both IRSs as shown in Fig. 2, while keeping the same hyper-parameters for ZoSGA as well as for WMMSE. Fig. 3(c) shows that ZoSGA succeeds in optimizing both IRSs simultaneously, without any information about their spatial configuration. Fig. 3(c) also visualizes the improved performance gains when optimizing $\boldsymbol{\theta}_1$ versus optimizing the more distant $\boldsymbol{\theta}_2$, showing that ZoSGA not only scales well to unknown system/channel models, but is also robust with respect to η_ϕ and η_A.

In practical scenarios, IRSs may have multiple phase-shift elements, frequently on the order of hundreds. To demonstrate that ZoSGA can scale well with the number of IRS parameters θ, we increase N_v to 100 rows of elements, and compare the sumrates achieved by ZoSGA and TTS-SSCO in Fig. 4. It is evident that ZoSGA is able to scale well and also retains the gain in performance over TTS-SSCO despite the substantial increase in the number of IRS phase-shift elements.

Lastly, while in general ZoSGA does take more iterations to converge as compared with model-based methods (here, TTS-SSCO; this is in line with the related literature on model-free stochastic resource allocation: see, e.g., [65]), we may also

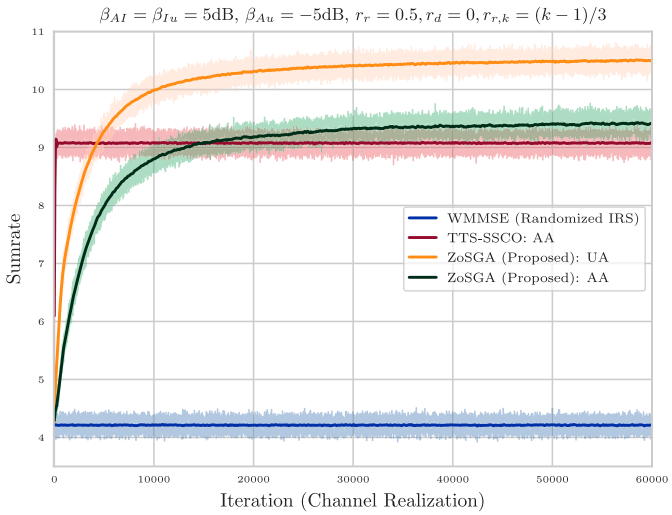


Fig. 4. Average sumrates achieved by WMMSE (random IRS), TTS-SSCO, and ZoSGA, with only IRS 1 and $N_v = 400$.

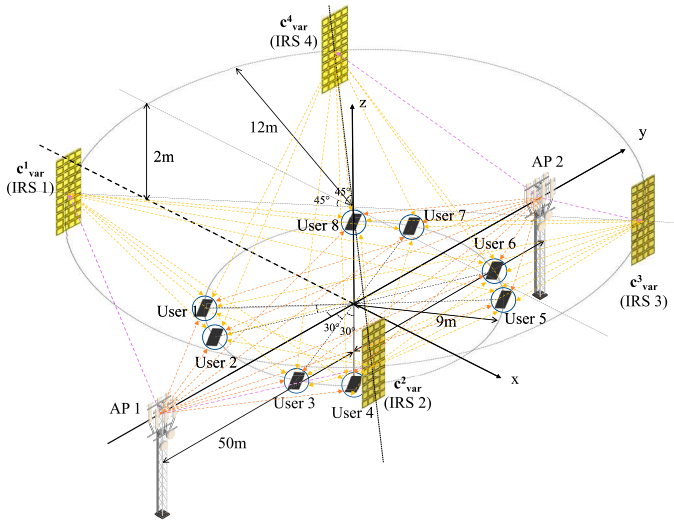


Fig. 5. Second IRS-aided network configuration (physical IRSs).

readily observe that it exhibits a high performance ceiling. Therefore, it can provide very competitive IRS beamformers way before its actual convergence.

C. Networks With Physical IRSs

In purely simulated environments, ZoSGA outperforms TTS-SSCO, allowing us to claim it to be a new SOTA for IRS-aided sumrate optimization. Nonetheless, we would also like to evaluate its robustness in a practical setting with *physically modeled* IRSs. To that end, we first define the IRSs using a practically feasible Transmission Line (TL) equivalent of an electromagnetic (EM) model, as presented in [75]. This TL model accounts for the geometrical and electrical properties of the IRS elements, also referred to as patches. Specifically, it considers reconfigurability, changes in response due to different angles of wave incidence, mutual coupling among closely spaced cells, and reflection losses.

An IRS can be classified as a spatially dispersive device, the reconfigurability of which is achieved by incorporating varactor diodes in its periodic structure. That is, the beamforming angles (of reflection) are controlled by changing DC voltages; thus tuning the capacitances of these diodes. Though there can be two scenarios when an EM wave impinges on an IRS surface, namely, the traverse magnetic (TM) incidence and the traverse electric (TE) incidence, here we only consider the former for simplicity, mostly focusing on the empirical performance analysis. Moreover, on a side note, we employ the Floquet theorem (as assumed in [75]) so as to consider the periodic patches placed in an infinite array, with each element/patch behaving identically.

Given a non-reconfigurable impedance surface, the TL model usually consists of a parallelly connected surface impedance of a reflecting surface, Z_{surf}, and inductive impedance of the grounded dielectric slab Z_d. Then, the input impedance is $Z_v = Z_{surf} \parallel Z_d$. For an IRS, however, the Z_{surf} is further comprised of a parallel connection between the unloaded surface impedance of the patch array Z_{patch} and the lumped impedance of a varactor diode Z_{var}, which is represented as a series of a resistor, inductance and capacitance as

$$Z_{var} = R_{var} + j\omega L_{var} + j\frac{1}{\omega C_{var}},$$

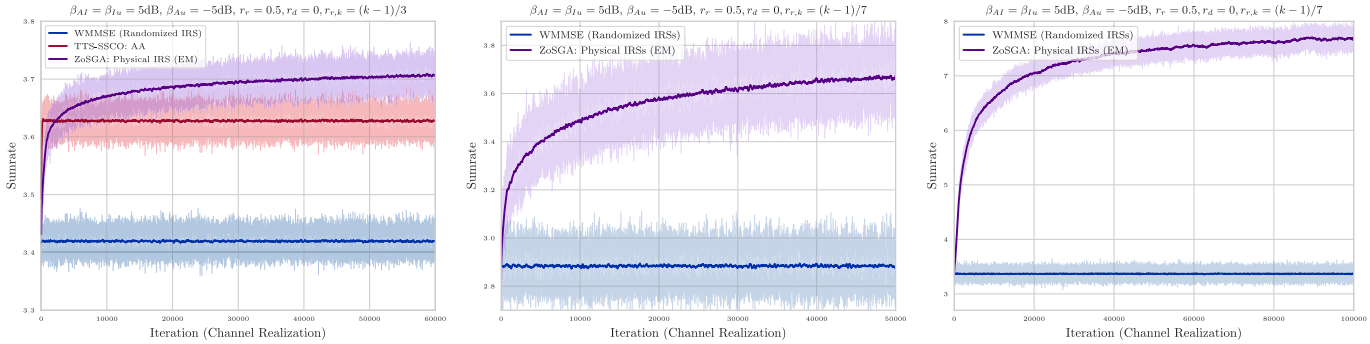
where the inductance L_{var} depends on the size of the lumped component and must be included in the varactor model to take into account the self-resonance of the component. The resistance R_{var} is included to account for the losses of the varactor. The variable capacitance C_{var} of the diode is used to vary Z_{var}.

The other two impedances, i.e. Z_d and Z_{patch}, depend on the properties of the substrate, the dielectric, and the angle of incidence. A detailed description of these is provided in [75, Section 4]. Once all the considered impedances of a patch have been evaluated, its phase-shift coefficient $\theta(C_{var})$ is given by

$$\theta(C_{var}) = a(C_{var})e^{j\phi(C_{var})} = \frac{Z_v(C_{var}) - \zeta_0}{Z_v(C_{var}) + \zeta_0},$$

where ζ_0 is the free space impedance. Thus, by changing the varactor capacitance C_{var}, we can change the varactor impedance Z_{var}, which changes the surface impedance Z_{surf}, which, in turn, changes the input impedance of an IRS element Z_v, finally causing a change in the value of the phase-shift coefficient $\theta(C_{var})$. Due to our assumption that the Floquet theorem holds, this dependency flow is identical for all elements of an IRS. Thus, we may replicate the above relation for all IRS elements, say q in number, and define a vector function $\boldsymbol{\theta}(\cdot)$ of varactor capacitances $\boldsymbol{c}_{var} = [C^1_{var}\, C^2_{var}\, C^3_{var}\, \cdots\, C^q_{var}]^\top$ as $\boldsymbol{\theta}(\boldsymbol{c}_{var}) = [\theta(C^1_{var})\, \theta(C^2_{var})\, \theta(C^3_{var})\, \cdots\, \theta(C^q_{var})]^\top$.

Now that we have a model for a practical IRS, we would like to evaluate our first wireless network setting, with one IRS, to compare the relative performance drop, if any. As shown in Fig. 6(a), the performance of the ZoSGA does drop when constrained in terms of the physical IRS model; this is very much expected, due to a decreased number of degrees of freedom in



(a) (b) (c)

Fig. 6. (a) Average sumrates achieved by WMMSE (random IRS), TTS-SSCO with ideal IRS, and ZoSGA with a physical EM IRS model (network in Fig. 2); Average sumrates achieved by WMMSE (random IRS), and ZoSGA, with four physical (EM model) IRSs with (b) 40 and (c) 400 phase-shift elements, respectively (network in Fig. 5).

tuning the IRS parameters. However, the performance gain relative to the random phase-shifts is still substantial. More interestingly, ZoSGA outperforms TTS-SSCO –the latter optimizing both amplitudes and phases in an unconstrained manner– even in the presence of appreciable practical IRS constraints.

To complete our empirical study, which supports our claim of enabling totally model-free optimization of the IRS amplitudes and phase-shifts, we lastly consider an elaborate wireless network setting, as shown in Fig. 5, consisting of two APs and four IRSs serving a total of eight users. We consider a MISO downlink scenario where both APs transmit a common symbol to each user, i.e., the two APs are different only in their position in space. We consider the same channel model and network environment parameters as given in Subsections V-A and V-B, respectively. Then, the effective received channel by user k is expressed as

$$\boldsymbol{h}_k(\boldsymbol{C}_{var},\omega)=\begin{bmatrix}\sum_{i=1}^{2}\underbrace{\boldsymbol{G}_i^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\theta}(\boldsymbol{c}_{var}^i))\boldsymbol{h}_{r,k}^i}_{\boldsymbol{\theta}_i\text{-non-LoS link}}+\underbrace{\boldsymbol{h}_{d,k,1}}_{\text{AP1 LoS link}}\\ \sum_{j=3}^{4}\underbrace{\boldsymbol{G}_i^{\mathsf{H}}\mathrm{Diag}(\boldsymbol{\theta}(\boldsymbol{c}_{var}^i))\boldsymbol{h}_{r,k}^i}_{\boldsymbol{\theta}_i\text{-non-LoS link}}+\underbrace{\boldsymbol{h}_{d,k,2}}_{\text{AP2 LoS link}}\end{bmatrix},$$

where, again, $\boldsymbol{h}_{r,k}^i=L_{\alpha_{Iu}^i}(d_{Iu,k,i})\breve{\boldsymbol{h}}_{r,k}^i$, $\boldsymbol{G}^i=L_{\alpha_{AI}^i}(d_{AI,i})\breve{\boldsymbol{G}}^i$, and $\boldsymbol{h}_{d,k,j}=L_{\alpha_{Au}^j}(d_{Au,k,j})\breve{\boldsymbol{h}}_{d,k,j}$ for $i\in\mathbb{N}_4^+$ and $j\in\{1,2\}$. Here, the varactor capacitances of the i-th IRS are denoted by the vector $\boldsymbol{c}_{var}^i$, and the matrix $\boldsymbol{C}_{var}=[\boldsymbol{c}_{var}^1\,\boldsymbol{c}_{var}^2\,\boldsymbol{c}_{var}^3\,\boldsymbol{c}_{var}^4]$ combines all the varactor capacitances of the four IRSs. Using the same learning rate scheme as above and a smoothing parameter $\mu=10^{-12}$, we optimize the system sumrate using ZoSGA.

We averaged the results of 40 different simulations in Fig. 6(b) to show not only the performance gain due to IRS capacitance tuning, but also the robustness of the approach under different realizations of the channels. The fact that the proposed approach is able to optimize a complicated network such as the one shown in Fig. 5, without any model information, verifies our claim of true model-free optimization capability of ZoSGA. We conjecture that the proposed optimization scheme can tackle a wide-range of problems arising in practical applications, with little to no additional input from the user.

VI. Conclusion

In this paper we introduced a zeroth-order stochastic gradient ascent (ZoSGA) method for the solution of two-stage stochastic programs with applications to model-free optimal beamforming for passive IRS-assisted stochastic network utility maximization. ZoSGA is amenable to rigorous convergence analysis and achieves state-of-the-art convergence rate under very general assumptions, capturing a wide range of realistic scenarios. By specializing to the case of sumrate maximization, we numerically demonstrated that ZoSGA outperforms current state-of-the-art model-based methodologies on three distinct network settings, yielding solutions of substantially higher quality and in a computationally efficient manner, while evading practical limitations that are inherent in current methods. Our numerical results confirmed that ZoSGA learns (near-)optimal passive IRS beamformers based solely on conventional effective CSI and in the absence of channel models and spatial network configuration information, also verifying our theoretical findings.

Appendix

A. Wirtinger Gradient Derivation

In light of conditions **(A1)**–**(A3)** of Assumption A, we can easily show that $F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ (i.e., the objective function in (SSP)) admits an explicit usual (Fréchet) gradient, for all $\boldsymbol{\theta}\in\mathcal{U}$ and a.e. $\omega\in\Omega$. We do this by utilizing elements of *Wirtinger calculus* (see [56, Section 4] for a detailed exposition).

Indeed, in order to evaluate the gradient of $F(\boldsymbol{W},\boldsymbol{H}(\cdot,\omega))$, i.e., $\nabla_{\boldsymbol{\theta}}F(\boldsymbol{W},\boldsymbol{H}(\cdot,\omega)):\mathcal{U}\to\mathbb{R}^S$, we consider its *Wirtinger cogradient* (a row vector; see [56, Section 4.2]), defined as

$$\frac{\partial^\circ}{\partial\boldsymbol{z}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{z},\omega))\triangleq\frac{1}{2}\left(\frac{\partial}{\partial\Re(\boldsymbol{z})}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{z},\omega))-j\frac{\partial}{\partial\Im(\boldsymbol{z})}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{z},\omega))\right),$$

noting that $\boldsymbol{H}(\boldsymbol{z},\omega)$ is constant relative to $\Im(\boldsymbol{z})$, i.e.

$$\boldsymbol{H}(\boldsymbol{z},\omega)=\boldsymbol{H}(\boldsymbol{x}+j\boldsymbol{y},\omega)=\boldsymbol{H}(\boldsymbol{x},\omega),\quad\forall(\boldsymbol{x},\boldsymbol{y})\in\mathcal{U}\times\mathcal{U},$$

and hence so is $F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{z},\omega)\right)$. It then follows that

$$\frac{\partial^\circ}{\partial \boldsymbol{z}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{z},\omega)\right)=\frac{1}{2}\left(\nabla_{\boldsymbol{x}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{x},\omega)\right)\right)^\top.$$

Using the Wirtinger chain rule (see [56, Eq. (32)]), we obtain

$$\begin{aligned}\frac{\partial^\circ}{\partial \boldsymbol{z}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{z},\omega)\right)&=\frac{\partial^\circ}{\partial \boldsymbol{z}}F\left(\boldsymbol{W},\boldsymbol{z}\right)\Big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{x},\omega)}\frac{\partial^\circ}{\partial \boldsymbol{z}}\boldsymbol{H}(\boldsymbol{z},\omega)\\&\quad+\frac{\partial^\circ}{\partial \overline{\boldsymbol{z}}}F\left(\boldsymbol{W},\boldsymbol{z}\right)\Big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{x},\omega)}\frac{\partial^\circ}{\partial \boldsymbol{z}}\overline{\boldsymbol{H}(\boldsymbol{z},\omega)},\end{aligned}$$

where $\overline{\boldsymbol{z}}$ denotes the complex conjugate of $\boldsymbol{z}$. It follows that

$$\begin{aligned}\frac{\partial^\circ}{\partial \boldsymbol{z}}\boldsymbol{H}(\boldsymbol{z},\omega)&=\frac{1}{2}\left(\frac{\partial}{\partial \Re(\boldsymbol{z})}\boldsymbol{H}(\boldsymbol{z},\omega)-j\frac{\partial}{\partial \Im(\boldsymbol{z})}\boldsymbol{H}(\boldsymbol{z},\omega)\right)\\&=\frac{1}{2}\left(\left(\nabla_{\boldsymbol{x}}\Re\left(\boldsymbol{H}(\boldsymbol{x},\omega)\right)\right)^\top+j\left(\nabla_{\boldsymbol{x}}\Im\left(\boldsymbol{H}(\boldsymbol{x},\omega)\right)\right)^\top\right),\end{aligned}$$

and $\frac{\partial^\circ}{\partial \boldsymbol{z}}\overline{\boldsymbol{H}(\boldsymbol{z},\omega)}=\overline{\left(\frac{\partial^\circ}{\partial \boldsymbol{z}}\boldsymbol{H}(\boldsymbol{z},\omega)\right)}$. From the real-valuedness of $F(\boldsymbol{W},\boldsymbol{H}(\cdot,\omega))$, we have

$$\begin{aligned}\frac{\partial^\circ}{\partial \boldsymbol{z}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{x},\omega)\right)&=\frac{\partial^\circ}{\partial \boldsymbol{z}}F\left(\boldsymbol{W},\boldsymbol{z}\right)\Big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{x},\omega)}\frac{\partial^\circ}{\partial \boldsymbol{z}}\boldsymbol{H}(\boldsymbol{z},\omega)\\&\quad+\overline{\left(\frac{\partial^\circ}{\partial \boldsymbol{z}}F\left(\boldsymbol{W},\boldsymbol{z}\right)\Big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{x},\omega)}\frac{\partial^\circ}{\partial \boldsymbol{z}}\boldsymbol{H}(\boldsymbol{z},\omega)\right)}\\&=2\Re\left(\frac{\partial^\circ}{\partial \boldsymbol{z}}F(\boldsymbol{W},\boldsymbol{z})\Big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{x},\omega)}\frac{\partial^\circ}{\partial \boldsymbol{z}}\boldsymbol{H}(\boldsymbol{z},\omega)\right).\end{aligned}$$

Thus, for any $\boldsymbol{\theta}\in\mathcal{U}$ and a.e. $\omega\in\Omega$, we obtain

$$\begin{aligned}&\nabla_{\boldsymbol{\theta}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega)\right)\\&\quad=2\nabla_{\boldsymbol{\theta}}\Re\left(\boldsymbol{H}(\boldsymbol{\theta},\omega)\right)\Re\left(\frac{\partial^\circ}{\partial \boldsymbol{z}}F(\boldsymbol{W},\boldsymbol{z})\Big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{\theta},\omega)}\right)^\top\\&\qquad+2\nabla_{\boldsymbol{\theta}}\Im\left(\boldsymbol{H}(\boldsymbol{\theta},\omega)\right)\Re\left(j\frac{\partial^\circ}{\partial \boldsymbol{z}}F(\boldsymbol{W},\boldsymbol{z})\Big|_{\boldsymbol{z}=\boldsymbol{H}(\boldsymbol{\theta},\omega)}\right)^\top.\end{aligned}$$

B. Weak Concavity of the First-Stage Objective Function

In what follows, we present three typical situations under which condition **(A4)** of Assumption A is satisfied for problem (2SP). To simplify the discussion, we assume that $\mathcal{W}=\{\boldsymbol{W}\colon\|\boldsymbol{W}\|^2\leq P\}$, which is the constraint utilized in (2SSRM), noting that this is done without loss of generality.

Cases 1 and 2 - Strong second-order sufficient optimality: The first two cases rely on the strong second-order sufficient optimality conditions for (SSP). For any $\boldsymbol{\theta}\in\Theta$ and a.e. $\omega\in\Omega$, the Lagrangian associated with (SSP) reads

$$L(\boldsymbol{W},\lambda;\boldsymbol{\theta},\omega)=F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))+\lambda\left(\|\boldsymbol{W}\|^2-P\right),$$

where the admissible Lagrange multipliers are nonnegative, i.e. $\lambda\geq 0$. Let any $\boldsymbol{W}^*(\boldsymbol{\theta},\omega)\in\arg\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$. Then, the strong second-order sufficient optimality conditions require *strict complementarity slackness* (i.e. $\lambda^*>0$ if $\|\boldsymbol{W}^*(\boldsymbol{\theta},\boldsymbol{H}(\boldsymbol{\theta},\omega))\|^2=P$, and $\lambda^*=0$ otherwise, where λ^* is an optimal Lagrange multiplier associated to $\boldsymbol{W}^*(\boldsymbol{\theta},\boldsymbol{H}(\boldsymbol{\theta},\omega))$), as well as that $\nabla^2_{\boldsymbol{W}}L(\boldsymbol{W},\lambda;\boldsymbol{\theta},\omega)|_{(\boldsymbol{W},\lambda)=(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\lambda^*)}$ is nonsingular. We note that for problem (2SSRM), strict complementarity slackness holds without any additional assumptions.

In the first case (**Case 1**), we assume that the second-stage problem (SSP) admits a unique solution for each $\boldsymbol{\theta}\in\Theta$ and a.e. $\omega\in\Omega$. It then follows that the mapping $\boldsymbol{\theta}\mapsto\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ is twice continuously differentiable on Θ by utilizing [68, Lemma 2.2] (which, in turn, utilizes the Implicit Function Theorem, e.g. see [67, Theorem 1B.1]). Twice continuously differentiable functions on a compact set (in this case Θ) are, in fact, Lipschitz smooth on that set. However, Lipschitz smooth functions are both weakly convex and weakly concave (see [59, Proposition 4.12]), and thus we are done.

In the second case (**Case 2**), instead of assuming that the solution set of the second-stage problem is a singleton, we assume that $\boldsymbol{H}(\cdot,\omega)$ is real analytic (which, for example, is true in Section V). Then, assuming that $F(\cdot,\cdot)$ is the sumrate (as in (2SSRM)), it follows that the mapping $\boldsymbol{\theta}\mapsto\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ is *sub-analytic* on an open bounded subset $\Theta'\supset\Theta$ (e.g. see [77, Example 4]). In turn, this implies (e.g. see [78, Theorem 2.3]) that the function $\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))-F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ satisfies the Łojasiewicz inequality with uniform exponent, i.e. for a.e. $\omega\in\Omega$ and for each $\boldsymbol{\theta}\in\Theta'$, there exists $\eta>0$ and a subanalytic function $C(\boldsymbol{\theta})>0$, such that for every $(\boldsymbol{\theta},\boldsymbol{W})\in\Theta'\times\mathcal{W}$, we have

$$\begin{aligned}&\operatorname{dist}\left(\boldsymbol{W},\arg\max_{\boldsymbol{W}\in\mathcal{W}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega)\right)\right)\\&\quad\leq C(\boldsymbol{\theta})\left(\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))-F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))\right)^\eta.\end{aligned}\tag{11}$$

In what follows we make the reasonable assumption that $C(\boldsymbol{\theta})$ is uniformly bounded on Θ'.

Lemma 8: Let $\Theta''\supset\Theta'\supset\Theta$, be a compact set, with Θ' some open set. Given conditions **(A1)–(A3)** of Assumption A, the multifunction $\boldsymbol{\theta}\mapsto\arg\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ is nonempty and compact-valued on Θ'', for a.e. $\omega\in\Omega$.

Proof: We have that $F(\cdot,\cdot)$ is jointly (real) continuously differentiable, and that for any $\boldsymbol{\theta}\in\Theta''$ and a.e. $\omega\in\Omega$, $\arg\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ is non-empty. Additionally, both Θ'' and $\mathcal{W}$ are assumed to be compact, and $\boldsymbol{H}(\cdot,\omega)$ is continuously differentiable and thus has compact range. In turn, we obtain that $F(\cdot,H(\cdot,\omega))$ also has compact range on $\Theta''\times\mathcal{W}$ and is jointly continuous. We complete the proof by applying Berge's *maximum theorem* (see [79]). □

Lemma 9: Let conditions **(A1)–(A3)** of Assumption A hold, along with the aforementioned conditions of **Case 2**. Then, for any $\boldsymbol{\theta}\in\Theta$ and a.e. $\omega\in\Omega$, and each selection

$$\boldsymbol{W}^*(\boldsymbol{\theta},\omega)\in\arg\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega)),$$

there exists a sequence of selections $\left\{\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z}_k,\omega)\right\}_{k=0}^\infty$, where $\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z}_k,\omega)\in\arg\max_{\boldsymbol{W}\in\mathcal{W}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z}_k,\omega))$, for some sequence $\boldsymbol{\theta}+\{\boldsymbol{z}_k\}_{k=0}^\infty\subset\Theta'$, such that

$$\lim_{\|\boldsymbol{z}_k\|\to 0}\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z}_k,\omega)=\boldsymbol{W}^*(\boldsymbol{\theta},\omega).$$

Proof: We fix some $(\boldsymbol{\theta},\omega)\in\Theta\times\Omega$ and a bounded open set $\Theta'\supset\Theta$. From our assumptions, there must exist two positive constants C, η, such that for every $(\boldsymbol{W},\boldsymbol{\theta}+\boldsymbol{z})\in\mathcal{W}\times\Theta'$, (11) holds with $C(\boldsymbol{\theta})=C$. From Lemma 8 (in particular, from the closed-valuedness of $\arg\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ on the compact set $\Theta''\supset\Theta'$), for every $\boldsymbol{W}\in\mathcal{W}$ and any $\boldsymbol{z}$ such that $\boldsymbol{\theta}+\boldsymbol{z}\in\Theta'$, there exists a selection $\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z},\omega)$ such that

$$\left\|\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z},\omega)-\boldsymbol{W}\right\| = \operatorname{dist}\left(\boldsymbol{W},\arg\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z},\omega))\right).$$

Continuity of $\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\cdot,\omega))-F(\boldsymbol{W},\boldsymbol{H}(\cdot,\omega))$ then yields the desired result, since we can consider a sequence $\boldsymbol{\theta}+\{\boldsymbol{z}_k\}_{k=0}^{\infty}\subset\Theta'$ such that $\|\boldsymbol{z}_k\|\to 0$. □

Lemma 10: Let conditions **(A1)–(A3)** of Assumption A hold, along with the aforementioned conditions of **Case 2**. For a.e. $\omega\in\Omega$, the function $\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\cdot,\omega))$ is Fréchet differentiable on Θ with

$$\nabla_{\boldsymbol{\theta}}\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))=F(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\boldsymbol{H}(\boldsymbol{\theta},\omega)),\quad \boldsymbol{\theta}\in\Theta,$$

where $\boldsymbol{W}^*(\cdot,\omega)\in\arg\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\cdot,\omega))$ is an arbitrary selection on Θ.

Proof: Fix some $(\boldsymbol{\theta},\omega)\in\Theta\times\Omega$ and some $\boldsymbol{W}^*(\boldsymbol{\theta},\omega)\in\arg\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$. Firstly, we note that the function $\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ is well-defined and finite on an open set $\Theta'\supset\Theta$. Additionally, we have that $F(\cdot,\boldsymbol{H}(\boldsymbol{\theta},\omega))$ is real-analytic on $\mathcal{W}$ for any $\boldsymbol{\theta}\in\Theta'$, and $F(\boldsymbol{W},\boldsymbol{H}(\cdot,\omega))$ is continuously differentiable on Θ'. By definition, we obtain that

$$\begin{aligned}&\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z},\omega))-\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))\\ &\quad\le F\left(\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z},\omega),\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z},\omega)\right)\\ &\qquad-F\left(\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z},\omega),\boldsymbol{H}(\boldsymbol{\theta},\omega)\right),\end{aligned}$$

for all $\boldsymbol{z}$ such that $\boldsymbol{\theta}+\boldsymbol{z}\in\Theta'$, where $\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z},\omega)$ can be chosen as in Lemma 9. Furthermore, we observe that

$$\begin{aligned}&\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z},\omega))-\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))\\ &\quad\ge F(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z},\omega))-F(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\boldsymbol{H}(\boldsymbol{\theta},\omega)).\end{aligned}$$

In other words, we have

$$\begin{aligned}&F(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z},\omega))-F(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\boldsymbol{H}(\boldsymbol{\theta},\omega))\\ &\quad\le\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z},\omega))-\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))\\ &\quad\le F\left(\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z},\omega),\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z},\omega)\right)\\ &\qquad-F\left(\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z},\omega),\boldsymbol{H}(\boldsymbol{\theta},\omega)\right)\end{aligned}$$

Since $F(\boldsymbol{W},\boldsymbol{H}(\cdot,\omega))$ is differentiable on Θ', for any $\boldsymbol{W}\in\mathcal{W}$, we may use the mean value theorem to show that for every $\boldsymbol{z}$ with $\boldsymbol{\theta}+\boldsymbol{z}\in\Theta'$, there exists $c\equiv c(\boldsymbol{\theta}+\boldsymbol{z},\boldsymbol{W},\omega)\in(0,1)$ such that

$$\begin{aligned}&F\left(\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z},\omega),\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z},\omega)\right)\\ &\quad-F\left(\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z},\omega),\boldsymbol{H}(\boldsymbol{\theta},\omega)\right)\\ &\quad=\left\langle\nabla_{\boldsymbol{\theta}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta}+c\boldsymbol{z},\omega)\right),\boldsymbol{z}\right\rangle\Big|_{\boldsymbol{W}=\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z},\omega)}.\end{aligned}$$

Given the previous inequalities, the latter yields

$$\begin{aligned}&F(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z},\omega))-F(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\boldsymbol{H}(\boldsymbol{\theta},\omega))\\ &\quad-\left\langle\nabla_{\boldsymbol{\theta}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega)\right),\boldsymbol{z}\right\rangle\Big|_{\boldsymbol{W}=\boldsymbol{W}^*(\boldsymbol{\theta},\omega)}\\ &\le F\left(\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z},\omega),\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z},\omega)\right)\\ &\quad-F\left(\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z},\omega),\boldsymbol{H}(\boldsymbol{\theta},\omega)\right)\\ &\quad-\left\langle\nabla_{\boldsymbol{\theta}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega)\right),\boldsymbol{z}\right\rangle\Big|_{\boldsymbol{W}=\boldsymbol{W}^*(\boldsymbol{\theta},\omega)}\\ &\le\Big\|\nabla_{\boldsymbol{\theta}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega)\right)\Big|_{\boldsymbol{W}=\boldsymbol{W}^*(\boldsymbol{\theta},\omega)}\\ &\quad-\nabla_{\boldsymbol{\theta}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta}+c\boldsymbol{z},\omega)\right)\Big|_{\boldsymbol{W}=\widetilde{\boldsymbol{W}}^*(\boldsymbol{\theta}+\boldsymbol{z},\omega)}\Big\|\|\boldsymbol{z}\|.\end{aligned}$$

Utilizing again the mean value theorem, we also have that there exists $c'\equiv c'(\boldsymbol{\theta}+\boldsymbol{z},\boldsymbol{W},\omega)\in(0,1)$ such that

$$\begin{aligned}&F\left(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z},\omega)\right)-F\left(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\boldsymbol{H}(\boldsymbol{\theta},\omega)\right)\\ &\quad=\left\langle\nabla_{\boldsymbol{\theta}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta}+c'\boldsymbol{z},\omega)\right),\boldsymbol{z}\right\rangle\Big|_{\boldsymbol{W}=\boldsymbol{W}^*(\boldsymbol{\theta},\omega)}.\end{aligned}$$

As before, this implies that

$$\begin{aligned}&F(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z},\omega))-F(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\boldsymbol{H}(\boldsymbol{\theta},\omega))\\ &\quad-\left\langle\nabla_{\boldsymbol{\theta}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega)\right),\boldsymbol{z}\right\rangle\Big|_{\boldsymbol{W}=\boldsymbol{W}^*(\boldsymbol{\theta},\omega)}\\ &\ge-\Big\|\nabla_{\boldsymbol{\theta}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta}+c'\boldsymbol{z},\omega)\right)\Big|_{\boldsymbol{W}=\boldsymbol{W}^*(\boldsymbol{\theta},\omega)}\\ &\quad-\nabla_{\boldsymbol{\theta}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega)\right)\Big|_{\boldsymbol{W}=\boldsymbol{W}^*(\boldsymbol{\theta},\omega)}\Big\|\|\boldsymbol{z}\|.\end{aligned}$$

Combining the previous inequalities with Lemma 9, yields

$$\begin{aligned}&\lim_{\|\boldsymbol{z}\|\to 0}\frac{1}{\|\boldsymbol{z}\|}\Big(F(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\boldsymbol{H}(\boldsymbol{\theta}+\boldsymbol{z},\omega))\\ &\qquad-F(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\boldsymbol{H}(\boldsymbol{\theta},\omega))\\ &\qquad-\left\langle\nabla_{\boldsymbol{\theta}}F\left(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega)\right),\boldsymbol{z}\right\rangle\Big|_{\boldsymbol{W}=\boldsymbol{W}^*(\boldsymbol{\theta},\omega)}\Big)=0,\end{aligned}$$

which implies that $\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\cdot,\omega))$ is (Fréchet) differentiable at $\boldsymbol{\theta}\in\Theta$, for a.e. $\omega\in\Omega$, and its gradient reads as

$$\nabla_{\boldsymbol{\theta}}\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))=\nabla_{\boldsymbol{\theta}}F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))\big|_{\boldsymbol{W}=\boldsymbol{W}^*(\boldsymbol{\theta},\omega)}$$

for any $\boldsymbol{W}^*(\boldsymbol{\theta},\omega)\in\arg\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$. □

Theorem 11: Let conditions **(A1)–(A3)** of Assumption A hold, along with the aforementioned conditions of **Case 2**. For a.e. $\omega\in\Omega$, the function $\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\cdot,\omega))$ is Lipschitz smooth and thus weakly concave.

Proof: Under our assumptions, we know from Lemma 10 that the function $\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$ is Frechét differentiable on Θ and its gradient reads as

$$\nabla_{\boldsymbol{\theta}}\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))=F(\boldsymbol{W}^*(\boldsymbol{\theta},\omega),\boldsymbol{H}(\boldsymbol{\theta},\omega)),$$

for any $\boldsymbol{W}^*(\boldsymbol{\theta},\omega)\in\arg\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W},\boldsymbol{H}(\boldsymbol{\theta},\omega))$. At the same time, from the strong second-order sufficient conditions, we can apply the Implicit Function Theorem (as in [68, Lemma 2.1]; see also [67, Theorem 1B.1]) to the Lagrangian of the second-stage problem, which implies that for every $\boldsymbol{\theta}_\circ\in\Theta$, and any arbitrary selection $\boldsymbol{W}^*(\boldsymbol{\theta}_\circ,\omega)\in\arg\max_{\boldsymbol{W}\in\mathcal{W}}$

$F(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta}_\circ, \omega))$, there exists a neighbourhood $\Theta_\circ \ni \boldsymbol{\theta}_\circ$, on which there exists a unique continuously differentiable mapping $\boldsymbol{W}_\circ^*(\cdot, \omega) : \boldsymbol{\theta} \mapsto \arg\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta}, \omega))$, such that $\boldsymbol{W}_\circ^*(\boldsymbol{\theta}_\circ, \omega) = \boldsymbol{W}^*(\boldsymbol{\theta}_\circ, \omega)$. Let us fix this neighbourhood $\Theta_\circ$, and consider two arbitrary points $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \widetilde{\Theta}_\circ \subset \Theta_\circ$ (assuming that $\widetilde{\Theta}_\circ$ is a compact set), along with some arbitrary selections $\boldsymbol{W}_1^* \triangleq \boldsymbol{W}^*(\boldsymbol{\theta}_1, \omega)$, $\boldsymbol{W}_2^* \triangleq \boldsymbol{W}^*(\boldsymbol{\theta}_2, \omega)$. Then, letting $\boldsymbol{W}_{\circ,i}^* \triangleq \boldsymbol{W}_\circ^*(\boldsymbol{\theta}_i, \omega)$, for $i = 1, 2$, we have

$$\begin{aligned}
&\|\nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_1^*, \boldsymbol{H}(\boldsymbol{\theta}_1, \omega)) - \nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_2^*, \boldsymbol{H}(\boldsymbol{\theta}_2, \omega))\| \\
&\quad = \left\|\nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_{\circ,1}^*, \boldsymbol{H}(\boldsymbol{\theta}_1, \omega)) - \nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_{\circ,2}^*, \boldsymbol{H}(\boldsymbol{\theta}_2, \omega))\right\| \\
&\quad = \Big\|\nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_{\circ,1}^*, \boldsymbol{H}(\boldsymbol{\theta}_1, \omega)) - \nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_{\circ,1}^*, \boldsymbol{H}(\boldsymbol{\theta}_2, \omega)) \\
&\qquad + \nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_{\circ,1}^*, \boldsymbol{H}(\boldsymbol{\theta}_2, \omega)) - \nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_{\circ,2}^*, \boldsymbol{H}(\boldsymbol{\theta}_2, \omega))\Big\| \\
&\quad \le \left\|\nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_{\circ,1}^*, \boldsymbol{H}(\boldsymbol{\theta}_1, \omega)) - \nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_{\circ,1}^*, \boldsymbol{H}(\boldsymbol{\theta}_2, \omega))\right\| \\
&\qquad + \left\|\nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_{\circ,1}^*, \boldsymbol{H}(\boldsymbol{\theta}_2, \omega)) - \nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_{\circ,2}^*, \boldsymbol{H}(\boldsymbol{\theta}_2, \omega))\right\|,
\end{aligned}$$

where we used the fact that $\nabla_{\boldsymbol{\theta}} \max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta}, \omega))$ has the same value independently of the optimum selection, while $\boldsymbol{W}_\circ^*(\boldsymbol{\theta}_i, \omega) \in \arg\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{\theta}_i, \omega))$, for $i = 1, 2$. However, Lipschitz smoothness of the function $F(\boldsymbol{W}, \boldsymbol{H}(\cdot, \omega))$ implies that there exists some constant $L_1 > 0$ such that

$$\begin{aligned}
&\left\|\nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_{\circ,1}^*, \boldsymbol{H}(\boldsymbol{\theta}_1, \omega)) - \nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_{\circ,1}^*, \boldsymbol{H}(\boldsymbol{\theta}_2, \omega))\right\| \\
&\qquad \le L_1 \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|.
\end{aligned}$$

Then, we observe that

$$\begin{aligned}
&\left\|\nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_{\circ,1}^*, \boldsymbol{H}(\boldsymbol{\theta}_2, \omega)) - \nabla_{\boldsymbol{\theta}} F(\boldsymbol{W}_{\circ,2}^*, \boldsymbol{H}(\boldsymbol{\theta}_2, \omega))\right\| \\
&\quad \le L_2 \|\boldsymbol{W}_\circ^*(\boldsymbol{\theta}_1, \omega) - \boldsymbol{W}_\circ^*(\boldsymbol{\theta}_2, \omega)\| \le L_2 L_3(\widetilde{\Theta}_\circ) \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|,
\end{aligned}$$

where we used the L_2-Lipschitz smoothness of $F(\cdot, \boldsymbol{H}(\boldsymbol{\theta}, \omega))$, and the fact that $\boldsymbol{W}_\circ^*(\cdot, \omega)$ is continuously differentiable on the compact set $\widetilde{\Theta}_\circ$, and thus $L_3(\widetilde{\Theta}_\circ)$-Lipschitz continuous. In other words, the function $\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W}, \boldsymbol{H}(\boldsymbol{W}, \omega))$ is locally Lipschitz smooth. However, since Θ is compact, this is equivalent to saying that it is globally Lipschitz smooth, which implies that it is both weakly convex and weakly concave (see [59, Proposition 4.12]). This completes the proof. □

Case 3 - Lack of Hessian Invertibility: Finally, we should mention that if the solution set of the second-stage problem (SSP) is not a singleton, while it also does not satisfy the strong second-order sufficient optimality conditions utilized in **Case 2**, one can still show that condition **(A4)** of Assumption A holds, by utilizing the analysis of [68, Section 4]. Indeed, under some regularity conditions, coupled with an assumption of connectedness of the solution set of the second-stage problem, it follows that the function $\max_{\boldsymbol{W}\in\mathcal{W}} F(\boldsymbol{W}, \boldsymbol{\Theta}(\cdot, \omega))$ is twice continuously differentiable on Θ, and thus weakly concave. The details are omitted, but the reader is referred to [68, Section 4] for a detailed analysis of a simplified case.

ACKNOWLEDGMENT

The authors would like to thank the AE and the three anonymous reviewers for their diligent comments, which helped us improve the technical exposition and overall quality of the paper.

REFERENCES

[1] H. Hashmi, S. Pougkakiotis, and D. S. Kalogerias, "Model-free learning of optimal beamformers for passive IRS-assisted sumrate maximization," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2023, pp. 1–5.

[2] F. Guo, F. R. Yu, H. Zhang, X. Li, H. Ji, and V. C. Leung, "Enabling massive IoT toward 6G: A comprehensive survey," *IEEE Internet Things J.*, vol. 8, no. 15, pp. 11891–11915, Aug. 2021.

[3] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6G: A comprehensive survey," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 334–366, 2021.

[4] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, Jul. 2021.

[5] X. You et al., "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Sci. China Inf. Sci.*, vol. 64, pp. 1–74, Jan. 2021.

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

[7] A. Gotsis, S. Stefanatos, and A. Alexiou, "UltraDense networks: The new wireless frontier for enabling 5G access," *IEEE Veh. Technol. Mag.*, vol. 11, no. 2, pp. 71–78, Jun. 2016.

[8] J. G. Andrews et al., "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[9] M. Shafi et al., "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, Jun. 2017.

[10] M. Elkashlan, T. Q. Duong, and H.-H. Chen, "Millimeter-wave communications for 5G: Fundamentals: Part I [Guest Editorial]," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 52–54, Sep. 2014.

[11] X. Wang et al., "Millimeter wave communication: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 1616–1653, Third Quart. 2018.

[12] F. Al-Ogaili and R. M. Shubair, "Millimeter-wave mobile communications for 5G: Challenges and opportunities," in *Proc. IEEE Int. Symp. Antennas Propag. (APSURSI)*, 2016, pp. 1003–1004.

[13] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-dense networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2522–2545, Fourth Quart. 2016.

[14] J. F. Valenzuela-Valdés, A. Palomares, J. C. González-Macías, A. Valenzuela-Valdés, P. Padilla, and F. Luna-Valero, "On the ultra-dense small cell deployment for 5G networks," in *Proc. IEEE 5G World Forum (5GWF)*, Piscataway, NJ, USA: IEEE Press, 2018, pp. 369–372.

[15] D. F. Sievenpiper, J. H. Schaffner, H. J. Song, R. Y. Loo, and G. Tangonan, "Two-dimensional beam steering using an electrically tunable impedance surface," *IEEE Trans. Antennas Propag.*, vol. 51, no. 10, pp. 2713–2722, Oct. 2003.

[16] S. V. Hum, M. Okoniewski, and R. J. Davies, "Realizing an electronically tunable reflectarray using varactor diode-tuned elements," *IEEE Microw. Wireless Compon. Lett.*, vol. 15, no. 6, pp. 422–424, Jun. 2005.

[17] H. Kamoda, T. Iwasaki, J. Tsumochi, T. Kuki, and O. Hashimoto, "60-GHz electronically reconfigurable large reflectarray using single-bit phase shifters," *IEEE Trans. Antennas Propag.*, vol. 59, no. 7, pp. 2524–2531, Jul. 2011.

[18] J. Zhao, Q. Cheng, J. Chen, M. Q. Qi, W. X. Jiang, and T. J. Cui, "A tunable metamaterial absorber using varactor diodes," *New J. Phys.*, vol. 15, no. 4, p. 043049, 2013.

[19] A. Araghi et al., "Reconfigurable intelligent surface (RIS) in the sub-6 GHz band: Design, implementation, and real-world demonstration," *IEEE Access*, vol. 10, pp. 2646–2655, 2022.

[20] B. Rana, S.-S. Cho, and I.-P. Hong, "Review paper on hardware of reconfigurable intelligent surfaces," *IEEE Access*, vol. 11, pp. 29614–29634, 2023.

[21] S. P. Chepuri, N. Shlezinger, F. Liu, G. C. Alexandropoulos, S. Buzzi, and Y. C. Eldar, "Integrated sensing and communications with reconfigurable intelligent surfaces: From signal modeling to processing," *IEEE Signal Process. Mag.*, vol. 40, no. 6, pp. 41–62, Sep. 2023.

[22] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3064–3076, May 2020.

[23] M.-M. Zhao, Q. Wu, M.-J. Zhao, and R. Zhang, "Intelligent reflecting surface enhanced wireless networks: Two-timescale beamforming optimization," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 2–17, Jan. 2021.

[24] 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. Wireless Commun.*, vol. 20, no. 9, pp. 6179–6194, Sep. 2021.

[25] Z. Yang et al., "Energy-efficient wireless communications with distributed reconfigurable intelligent surfaces," *IEEE Trans. Wireless Commun.*, vol. 21, no. 1, pp. 665–679, Jan. 2022.

[26] B. Zheng, C. You, W. Mei, and R. Zhang, "A survey on channel estimation and practical passive beamforming design for intelligent reflecting surface aided wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 1035–1071, Second Quart. 2022.

[27] M. Munochiveyi, A. C. Pogaku, D.-T. Do, A.-T. Le, M. Voznak, and N. D. Nguyen, "Reconfigurable intelligent surface aided multi-user communications: State-of-the-art techniques and open issues," *IEEE Access*, vol. 9, pp. 118584–118605, 2021.

[28] H. Liu, X. Yuan, and Y.-J. A. Zhang, "Matrix-calibration-based cascaded channel estimation for reconfigurable intelligent surface assisted multiuser MIMO," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2621–2636, Nov. 2020.

[29] H. Guo and V. K. N. Lau, "Uplink cascaded channel estimation for intelligent reflecting surface assisted multiuser MISO systems," *IEEE Trans. Signal Process.*, vol. 70, pp. 3964–3977, 2022.

[30] 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.

[31] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020.

[32] K. Faisal and W. Choi, "Machine learning approaches for reconfigurable intelligent surfaces: A survey," *IEEE Access*, vol. 10, pp. 27343–27367, 2022.

[33] M. Hong, Q. Li, and Y.-F. Liu, "Decomposition by successive convex approximation: A unifying approach for linear transceiver design in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1377–1392, Feb. 2016.

[34] G. Scutari, F. Facchinei, P. Song, D. P. Palomar, and J.-S. Pang, "Decomposition by partial linearization: Parallel optimization of multi-agent systems," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 641–656, Feb. 2014.

[35] Y. Yang, G. Scutari, D. P. Palomar, and M. Pesavento, "A parallel decomposition method for nonconvex stochastic multi-agent optimization problems," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2949–2964, Jun. 2016.

[36] M. Jian and Y. Zhao, "A modified off-grid SBL channel estimation and transmission strategy for RIS-assisted wireless communication systems," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, 2020, pp. 1848–1853.

[37] X. Chen, J. Shi, Z. Yang, and L. Wu, "Low-complexity channel estimation for intelligent reflecting surface-enhanced massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 10, no. 5, pp. 996–1000, May 2021.

[38] Y. Lin, S. Jin, M. Matthaiou, and X. You, "Tensor-based algebraic channel estimation for hybrid IRS-assisted MIMO-OFDM," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3770–3784, Jun. 2021.

[39] A. L. Swindlehurst, G. Zhou, R. Liu, C. Pan, and M. Li, "Channel estimation with reconfigurable intelligent surfaces—A general framework," *Proc. IEEE*, vol. 110, no. 9, pp. 1312–1338, Sep. 2022.

[40] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Deep learning for large intelligent surfaces in millimeter wave and massive MIMO systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2019, pp. 1–6.

[41] E. Balevi, A. Doshi, A. Jalal, A. Dimakis, and J. G. Andrews, "High dimensional channel estimation using deep generative networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 18–30, Jan. 2021.

[42] B. Yang, X. Cao, C. Huang, C. Yuen, L. Qian, and M. Di Renzo, "Intelligent spectrum learning for wireless networks with reconfigurable intelligent surfaces," *IEEE Trans. Veh. Technol.*, vol. 70, no. 4, pp. 3920–3925, Apr. 2021.

[43] S. Zhang, S. Zhang, F. Gao, J. Ma, and O. A. Dobre, "Deep learning optimized sparse antenna activation for reconfigurable intelligent surface assisted communication," *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 6691–6705, Oct. 2021.

[44] F. B. Mismar, B. L. Evans, and A. Alkhateeb, "Deep reinforcement learning for 5G networks: Joint beamforming, power control, and interference coordination," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1581–1592, Mar. 2020.

[45] A. Taha, Y. Zhang, F. B. Mismar, and A. Alkhateeb, "Deep reinforcement learning for intelligent reflecting surfaces: Towards standalone operation," in *Proc. 21st Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, 2020, pp. 1–5.

[46] G. Lee, M. Jung, A. T. Z. Kasgari, W. Saad, and M. Bennis, "Deep reinforcement learning for energy-efficient networking with reconfigurable intelligent surfaces," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2020, pp. 1–6.

[47] 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.

[48] Z. Yang, Y. Liu, Y. Chen, and J. T. Zhou, "Deep reinforcement learning for RIS-aided non-orthogonal multiple access downlink networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Piscataway, NJ, USA: IEEE Press, 2020, pp. 1–6.

[49] X. Xie, S. Jiao, and Z. Ding, "A reinforcement learning approach for an IRS-assisted NOMA network," 2021, *arXiv:2106.09611*.

[50] S. Evmorfos, A. P. Petropulu, and H. V. Poor, "Deep reinforcement learning for IRS phase shift design in spatiotemporally correlated environments," 2022, *arXiv:2211.09726*.

[51] T. P. Lillicrap et al., "Continuous control with deep reinforcement learning," 2015, *arXiv:1509.02971*.

[52] Y. Liu, Q. Hu, Y. Cai, G. Yu, and G. Y. Li, "Deep-unfolding beamforming for intelligent reflecting surface assisted full-duplex systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 4784–4800, Jul. 2022.

[53] Q. Shi, M. Razaviyayn, Z. Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, pp. 4331–4340, Sep. 2011.

[54] S. Dempe, *Bilevel Programming: Implicit Function Approach*. Boston, MA, USA: Springer-Verlag, 2009, pp. 260–266.

[55] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[56] K. Kreutz-Delgado, "The complex gradient operator and the $\mathcal{CR}$-calculus," 2009, *arXiv:0906.4835*.

[57] D. Davis and D. Drusvyatskiy, "Stocahstic model-based minimization of weakly convex functions," *SIAM J. Optim.*, vol. 29, no. 1, pp. 207–239, 2019.

[58] R. T. Rockafellar, "Favorable classes of Lipschitz-continuous functions in subgradient optimization," in *Progress in Nondifferentiable Optimization*, E. Nurminski, Ed., Laxenburg, Austria: IIASA, pp. 125–143.

[59] J.-P. Vial, "Strong and weak convexity of sets and functions," *Math. Oper. Res.*, vol. 8, no. 2, pp. 231–259, 1983.

[60] J. M. Danskin, "The theory of max-min, with applications," *SIAM J. Appl. Math.*, vol. 14, no. 4, pp. 641–664, 1966.

[61] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory* (MOS-SIAM Series on Optimization). Philadelphia, PA, USA: SIAM, 2014.

[62] D. S. Kalogerias and W. B. Powell, "Zeroth-order stochastic compositional algorithms for risk-aware learning," *SIAM J. Optim.*, vol. 32, no. 2, pp. 386–416, Apr. 2022.

[63] Y. Nesterov and V. Spokoiny, "Random gradient-free minimization of convex functions," *Found. Comput. Math.*, vol. 17, pp. 527–566, 2017.

[64] S. Pougkakiotis and D. S. Kalogerias, "A zeroth-order proximal stochastic gradient method for weakly convex stochastic optimization," *SIAM J. Sci. Comput.*, vol. 45, no. 5, pp. A2679–A2702, 2023.

[65] M. Eisen, C. Zhang, L. F. O. Chamon, D. D. Lee, and A. Ribeiro, "Learning optimal resource allocations in wireless systems," *IEEE Trans. Signal Process.*, vol. 67, no. 10, pp. 2775–2790, May 2019.

[66] H. Kumar, D. S. Kalogerias, G. J. Pappas, and A. Ribeiro, "Zeroth-order deterministic policy gradient," 2020, *arXiv:2006.07314*.

[67] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings* (Springer Monographs in Mathematics). New York, NY, USA: Springer-Verlag, 2009.

[68] A. Shapiro, "Second-order derivatives of extremal-value functions and optimality conditions for semi-infinite programs," *Math. Oper. Res.*, vol. 10, no. 2, pp. 207–219, 1985.

[69] S. Łojasiewicz, "Sur la géométrie semi-et sous-analytique," *Ann. Inst. Fourier*, vol. 43, no. 5, pp. 1575–1595, 1993.

[70] A. Savitzky and M. J. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, 1964.
[71] M. R. McKay and I. B. Collings, "General capacity bounds for spatially correlated Rician MIMO channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3121–3145, Sep. 2005.
[72] S. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, Sep. 2001.
[73] J. Choi and D. J. Love, "Bounds on eigenvalues of a spatial correlation matrix," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1391–1394, Aug. 2014.
[74] 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.
[75] F. Costa and M. Borgese, "Electromagnetic model of reflective intelligent surfaces," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1577–1589, 2021.
[76] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.
[77] J. Bolte, A. Daniilidis, and A. Lewis, "The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems," *SIAM J. Optim.*, vol. 17, no. 4, pp. 1205–1223, 2007.
[78] A. Denkowska and M. P. Denkowski, "UPC condition with parameter for subanalytic sets," 2018, *arXiv:1804.10567*.
[79] E. A. Ok, *Continuity II*. Princeton, NJ, USA: Princeton Univ. Press, 2007, pp. 283–354.

Hassaan Hashmi (Student Member, IEEE) received the B.E. degree in mechatronics engineering from NUST, Pakistan, in 2016, and the M.S. degree in electrical engineering from KAIST, South Korea, in 2018. He has been working toward the Ph.D. degree with the Department of Electrical Engineering, Yale University, since September 2021. Before joining Yale, he worked as a Research Associate with the National Center of AI, Pakistan, from 2019 to 2020. His research interests are in the areas of reinforcement learning, wireless autonomy, and risk-aware learning.

Spyridon Pougkakiotis received the Ph.D. degree in optimization and operational research from the University of Edinburgh, in 2022. He is a Lecturer in mathematics with the School of Science and Engineering, University of Dundee. Before joining the University of Dundee, he was a Postdoctoral Research Associate with the Electrical Engineering Department, Yale University. His research revolves around optimization, risk-aware decision-making, and computational mathematics, with applications in operational research, data science, and engineering. He has received several awards, including the 2022 OR Society Doctoral Award for the most distinguished body of research leading to the award of a doctorate in the field of operational research.

Dionysis Kalogerias (Senior Member, IEEE) received the M.Eng. degree in computer engineering and informatics from the University of Patras, Greece, in 2010, the M.Sc. degree in signal processing and communications from the University of Patras, Greece, in 2012, and the Ph.D. degree in electrical and computer engineering from Rutgers, The State University of New Jersey, in 2017. He is an Assistant Professor in electrical engineering at Yale University. Before joining Yale, he was an Assistant Professor with the Department of Electrical and Computer Engineering, Michigan State University, from 2020 to 2021. His research interests are in the areas of mathematical optimization, risk, statistical machine learning, decision under uncertainty and signal processing, and relevant applications in autonomous systems, wireless networking and communications, security and privacy, and system trustworthiness. He has received several awards, including the Best Paper Award at ICASSP 2020 and the Best Student Paper of the Special Sessions Award at ICASSP 2016.