RIScatter: Unifying Backscatter Communication and Reconfigurable Intelligent Surface

Yang Zhao, *Member, IEEE,* and Bruno Clerckx, *Fellow, IEEE*

Abstract

Backscatter Communication (BackCom) nodes harvest energy from and modulate information over an external electromagnetic wave. Reconfigurable Intelligent Surface (RIS) adapts its phase shift response to enhance or attenuate channel strength in specific directions. In this paper, we show how those two seemingly different technologies (and their derivatives) can be unified to leverage their benefits simultaneously into a single architecture called RIScatter. RIScatter consists of multiple dispersed or co-located scatter nodes, whose reflection states can be adapted to partially engineer the wireless channel of the existing link and partially modulate their own information onto the scattered wave. This contrasts with BackCom (resp. RIS) where the reflection pattern is exclusively a function of the information symbol (resp. Channel State Information (CSI)). The key principle in RIScatter is to render the probability distribution of reflection states (i.e., backscatter channel input) as a joint function of the information source, CSI, and Quality of Service (QoS) of the coexisting active primary and passive backscatter links. This enables RIScatter to softly bridge, generalize, and outperform BackCom and RIS; boil down to either under specific input distribution; or evolve in a mixed form for heterogeneous traffic control and universal hardware design. For a single-user multi-node RIScatter network, we characterize the achievable primary-(total-)backscatter rate region by optimizing the input distribution at the nodes, the active beamforming at the Access Point (AP), and the backscatter detection regions at the user. Simulation results demonstrate RIScatter nodes can exploit the additional propagation paths to smoothly transition between backscatter modulation and passive beamforming, and the proposed practical receiver accommodates the double fading, double modulation, and symbol period difference of active-passive coexisting networks without Successive Interference Cancellation (SIC).

Index Terms

Backscatter communication, reconfigurable intelligent surface, ambient backscatter communication, symbiotic radio.

The authors are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail:{yang.zhao18, b.clerckx}@imperial.ac.uk). B. Clerckx is also with Silicon Austria Labs (SAL), Graz A-8010, Austria.

I. INTRODUCTION

Future wireless network is envisioned to provide high throughput, uniform coverage, pervasive connectivity, heterogeneous control, and cognitive intelligence for trillions of portable devices. As a mature low-power communication technique, Backscatter Communication (BackCom) separates a conventional transmitter into a Radio-Frequency (RF) carrier emitter with power-hungry elements (e.g., synthesizer and amplifier) and an information-bearing node with power-efficient components (e.g., harvester and modulator) [1]. In particular, the node harvests energy from impinging wave and embeds information over scattered signal, and the backscatter reader can be either co-located or separated with the carrier emitter, known as Monostatic BackCom (MBC) in Fig. 1(a) and Bistatic BackCom (BBC) in Fig. 1(b). Relevant applications such as Radio-Frequency Identification (RFID) [2], [3] and passive sensor network [4], [5] have been extensively researched, standardized, and commercialized in the era of Internet of Everything (IoE). However, conventional BackCom nodes only respond when externally inquired by a nearby reader. To tackle this, [6] proposed Ambient Backscatter Communication (AmBC) where battery-free nodes recycle ambient signals (e.g., radio, television and Wi-Fi) to harvest energy and establish connection in between. As shown in Fig. 1(c), it eliminates the need of dedicated power supply, carrier emitter and frequency spectrum for BackCom, but the decoding is subject to a strong interference from the primary (legacy) link. Cooperative AmBC was then proposed where a co-located receiver decodes both active primary and passive backscatter links, and the rate and error performance are evaluated for various detection schemes [7]. The concept was further refined by [8] as Symbiotic Radio (SR) in Fig. 1(d) that cognitively incorporates AmBC with existing systems. In a SR system, the active transmitter generates RF wave carrying primary information, the passive node creates additional propagation path and ride its information over scattered signal, while the cooperative receiver jointly, individually, or sequentially decodes both links. Those BackCom applications[1] employ scatter nodes as pure information sources, and the active primary transmission (if it exists) can be influenced by the randomness of backscatter modulation. On the other hand, Reconfigurable Intelligent Surface (RIS) in Fig. 1(e) is a smart planar signal reflector that consists of numerous compact passive elements with adjustable amplitude and phase responses, whose reflection pattern over time is deterministic and adaptively optimized based on Channel State Information (CSI). RIS smartly reflects the surrounding RF waves to customize the wireless propagation environment for signal enhancement, interference suppression, scattering enrichment and non-line-of-sight bypassing [9]. In summary, MBC/BBC (resp. RIS) involves a *single* passive backscatter (resp. active primary) link where

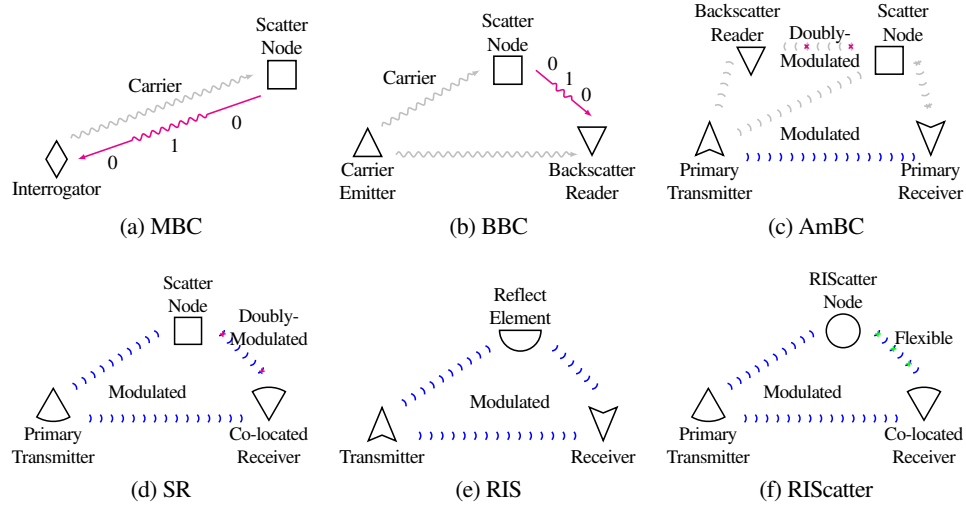[1]In this paper, BackCom refers to MBC, BBC, AmBC and SR.

Fig. 1. Illustration of scattering applications. For each, the blue flow(s) constitute the active primary link while the magenta/green flow denotes the passive backscatter link.

TABLE I
COMPARISON OF SCATTERING APPLICATIONS

	MBC/BBC	AmBC	SR	RIS	RIScatter
Information link(s)	Passive backscatter	Active-passive co-existing (competitive)	Active-passive coexisting (collaborative)	Active primary	Active-passive coexisting (collaborative)
Primary impact on backscatter detection	Carrier	Multiplicative interference	Spreading code or multiplicative interference	—	Energy uncertainty
Backscatter impact on primary detection	—	Multiplicative interference	CSI uncertainty or multiplicative interference	Passive beamforming	Dynamic passive beamforming
Collaborative devices based on CSI	—	—	Co-located receiver (potentially primary transmitter)	—	Primary transmitter, scatter nodes, and co-located receiver
Cooperative decoding strategy	—	No	Usually SIC from primary to backscatter	—	No SIC, backscatter detection as part of primary channel training
Primary detection	—	Semi-coherent	Semi-coherent	Coherent	Coherent
Backscatter detection	Coherent	Semi-coherent	Coherent	—	Semi-coherent
Reflection pattern over time depends on	Information source	Information source	Information source	CSI	Information source, CSI, and QoS
Reflection state (i.e., backscatter input) distribution	Equiprobable	Equiprobable	Equiprobable or Gaussian	Degenerate	Flexible
Load-switching speed	Fast	Fast or slow	Very slow	Quasi-static	Arbitrary

the passive device modulates information (resp. enhances propagation) and the active device generates carrier wave (resp. modulated signal). On the other hand, AmBC (resp. SR) involves *coexisting* active primary and passive backscatter links in a competitive (resp. collaborative) relationship. In AmBC, the backscatter link only needs to know the basic information (e.g., operating frequency) of the primary link, while the primary link may be completely unaware of the backscatter link. In SR, relevant CSI is exploited at the co-located receiver for cooperative decoding, and may also be exploited at the primary transmitter for cooperative active beamforming. A comparison of scattering applications is summarized in Table I.

Similar to Cognitive Radio (CR), active-passive coexisting networks can be classified into commensal,

parasitic, and competitive relationships, whose instantaneous rates, power schemes, and outage probabilities were acquired in [10], [11]. Joint Maximum-Likelihood (ML) decoding provides the best error performance but requires prohibitive computational complexity especially for high-order modulations [7]. Indeed, active-passive coexisting networks have three special and important properties: 1) primary and backscatter symbols are superimposed by *double modulation* (i.e., multiplication coding); 2) backscatter signal strength is much weaker than primary due to the *double fading* effect; 3) backscatter *symbol period* is longer than primary due to the load switching constraint. The second property motivated [7], [8], [10]–[18] to view SR as a multiplicative Non-Orthogonal Multiple Access (NOMA) and perform SIC from primary to backscatter link. During primary decoding, the randomness of backscatter modulation can be modelled as either channel uncertainty or multiplicative interference, when the spreading factor (i.e., backscatter-over-primary symbol period ratio) N is large or small, respectively[2]. Such a SIC receiver is operation-intensive and CSI-sensitive because it involves N re-encoding, precoding, subtraction together with a time-domain Maximal Ratio Combining (MRC) to decode each backscatter symbol. When the spreading factor is sufficiently large (i.e., commensal SR), the average primary achievable rate under semi-coherent detection[3] asymptotically approaches its coherent counterpart, and both links may be decoded in an interference-free manner [12]. On the other hand, efficient backscatter multiple access design remains an open issue for BackCom. [15] proposed a NOMA-based SR where the SIC order depends on the backscatter channel strength, but the performance deteriorates fast as the number of nodes increases. Time-Division Multiple Access (TDMA)-based SR was also considered in [16] where each node transmits information during its dedicated slot and harvests energy otherwise. It enables adaptive transmission time and reflection ratio design but introduces extra coordination cost. [20] modifies the load-switching speed to shift the scattered signal to the desired frequency band. This enables backscatter Frequency-Division Multiple Access (FDMA) at the cost of larger bandwidth and higher power consumption at the scatter nodes. To reduce the coordination cost, [17] proposed a random code-assisted multiple access for SR and evaluated the asymptotic Signal-to-Interference-plus-Noise Ratio (SINR) using random matrix theory. This code-domain solution suffers from the near-far problem and imperfect synchronization, both being severe in BackCom.

Conventional RIS design with fixed reflection coefficients during each channel block has been extensively studied in communication, sensing, and power literatures [21]–[26]. On the other hand,

---

[2]For state-of-the-art low-power RFID tags, the load-switching speed usually varies between 100s of kHz to 10s of MHz [19], corresponding to a typical spreading factor between 10 and 1000.

[3]In this paper, semi-coherent detection refers to the primary/backscatter decoding based on known CSI and unknown backscatter/primary symbols.

dynamic RIS with independent reflection patterns in different time slots has gained recent attentions in multi-user and multi-purpose wireless networks. The concept was first proposed in [27] to fine-tune the Orthogonal Frequency-Division Multiplexing (OFDM) resource blocks, then extended to the downlink power and uplink information phases of Wireless Powered Communication Network (WPCN) [28]–[30]. It introduces artificial channel diversity and enables flexible resource allocation, but misses the opportunity to encode its own message. RIS can also be used as a transmitter when placed in the near field of a carrier emitter, and prototypes have been implemented for Phase Shift Keying (PSK) [31] and Quadrature Amplitude Modulation (QAM) [32]. From an information-theoretic perspective, [33] reported that using RIS as a naive passive beamformer to maximize the receive Signal-to-Noise Ratio (SNR) is generally rate-suboptimal for finite-input constellations. Instead, the capacity of RIS-aided channel is achieved by joint transmitter-RIS encoding, and even layered encoding with pilots can outperform passive beamforming at high SNR. It inspired RIS-empowered BackCom [34]–[43] to combine passive beamforming and backscatter modulation in the overall reflection pattern. In particular, *symbol level precoding* maps the information symbols to the optimized RIS coefficient sets [34], [35], *overlay modulation* superposes the information symbols over a common auxiliary matrix [36]–[39], *spatial modulation* switches between the reflection coefficient sets that maximize SNR at different receive antennas [40]–[42], and *index modulation* employs dedicated reflection elements (resp. information elements) for passive beamforming (resp. backscatter modulation) [43]. However, those joint designs incur advanced hardware architecture and high optimization complexity. Relevant literatures also consider either *Gaussian* codebook [10]–[16], [38] that is impractical for low-power nodes, or finite *equiprobable* inputs [7], [8], [17], [18], [34]–[37], [39]–[43] that does not fully exploit the CSI and signal properties of active-passive coexisting networks.

First, we propose RIScatter as a novel scatter protocol that unifies BackCom and RIS by adaptive reflection state (i.e., backscatter input) distribution design. As shown in Fig. 1(f), one or more passive RIScatter nodes ride over an active transmission and smartly scatters the incident wave to simultaneously engineer the primary channel and modulate its own information, while a co-located receiver cooperatively decodes both coexisting links. The reflection pattern over time is neither fully random like BackCom nor fully deterministic like RIS, but semi-randomly chosen with the guidance of reflection state distribution, whose design accommodates the impact of information source, CSI, and QoS of both links. Such an adaptive backscatter channel coding boils down to the degenerate distribution of RIS when the primary link is prioritized, and outperforms the uniform distribution of line-coded BackCom (by accounting the CSI) when the backscatter link is prioritized. Besides, when multiple RIScatter nodes are co-located, joint encoding can further increase the total backscatter rate. RIScatter exploits the transmitter-scatter-receiver

path for simultaneous backscatter modulation and dynamic passive beamforming, and can smoothly transition in between under different QoS. However, such a novel scatter protocol encounters two major challenges, namely the cooperative receiver design and the adaptive backscatter input distribution design. This is the first paper to unify BackCom and RIS from the perspective of reflection state distribution.

Second, we address the first challenge by proposing a practical receiver that accommodates the double fading, double modulation, and symbol period difference of active-passive coexisting networks. Since joint decoding involves high complexity while individual decoding can be inefficient (neither cooperative nor exploiting signal properties), most relevant literatures consider a sequential primary-backscatter decoding based on SIC. Specifically, the decoding of each backscatter symbol requires N re-encoding, precoding, subtraction together with a time-domain MRC, which can be operation-intensive and CSI-sensitive. In contrast, the proposed RIScatter receiver semi-coherently decodes the backscatter symbol from the accumulated receive energy, re-encodes to recover the exact reflection pattern, and combine it with relevant CSI to determine the transmitter-scatter-receiver path response. Thanks to the double modulation and symbol period difference, the backscatter symbol can be modelled within primary channel (similar to RIS) rather than being re-precoded and cancelled, and the semi-coherent energy-based backscatter detection indeed becomes part of primary channel training. The strategy is suitable for arbitrary primary sampling rate and backscatter load-switching speed (i.e., arbitrary spreading factor). This is the first paper to consider such a sequential backscatter-primary decoding and treat semi-coherent backscatter detection as part of primary channel training.

Third, we address the second challenge by considering a specific scenario where multiple co-located/dispersed RIScatter nodes ride over an active point-to-point Multiple-Input Single-Output (MISO) transmission to perform backscatter modulation and dynamic passive beamforming towards a nearby user. For a given QoS, we first express the primary and total backscatter rates as functions of *input distribution* at RIScatter nodes, *active beamforming* at the Access Point (AP), and *backscatter decision regions* at the user. Then, we characterize the achievable rate region by optimizing those variables under different QoS. Since the original problem is highly non-convex, we propose a suboptimal Block Coordinate Descent (BCD) algorithm where the Karush-Kuhn-Tucker (KKT) input distribution is numerically evaluated by limit of sequences, the active beamforming is iteratively updated by Projected Gradient Descent (PGD) accelerated by Backtracking Line Search (BLS), and the decision regions are refined by existing sequential quantization methods for Discrete Memoryless Thresholding Channel (DMTC). Uniquely, we account for CSI, QoS, and backscatter constellation in those optimizations, and the proposed backscatter input distribution design is applicable to more general scenarios. This is also the first paper to reveal the importance

of backscatter input distribution and decision region designs in active-passive coexisting networks.

Fourth, we provide numerical results to demonstrate the benefits of RIScatter and proposed algorithms. The observations include: 1) adaptive reflection state distribution design can flexibly transition between backscatter modulation and passive beamforming; 2) when the primary link is prioritized, input distribution becomes degenerate and RIScatter nodes coincide with discrete RIS; 3) when the backscatter link is prioritized, adaptive RIScatter encoding achieves higher backscatter rate than conventional line-coded BackCom with equiprobable inputs; 4) co-located RIScatter nodes can further leverage total backscatter rate by joint encoding; 5) the proposed receiver maintains the passive beamforming benefit and provides comparable backscatter rate to SIC-based SR, with re-encoding costs reduced to $1/N$ and no re-precoding/cancellation; 6) it also supports scatter nodes with faster load-switching speed for potentially higher throughput; 7) PGD active beamformer effectively increases the primary (resp. backscatter) rate by boosting the receive SNR (resp. widening the energy gap under different reflection states), and can smoothly transition in between under different QoS; 8) distribution-aware backscatter detectors provide higher backscatter rate than the conventional ML detector.

Notations: Italic, bold lower-case, and bold upper-case letters denote scalars, vectors and matrices, respectively. $\mathbf{0}$ and $\mathbf{1}$ denote zero and one array of appropriate size, respectively. $\mathbb{I}^{x \times y}$, $\mathbb{R}_+^{x \times y}$, and $\mathbb{C}^{x \times y}$ denote the real unit, real nonnegative, and complex spaces of dimension $x \times y$, respectively. j denotes the imaginary unit. $\mathrm{diag}(\cdot)$ returns a square matrix with the input vector on its main diagonal and zeros elsewhere. $\mathrm{card}(\cdot)$ returns the cardinality of a set. $(\cdot)^*$, $(\cdot)^{\mathsf{T}}$, $(\cdot)^{\mathsf{H}}$, $|\cdot|$, and $\|\cdot\|$ denote the conjugate, transpose, conjugate transpose, absolute value, and Euclidean norm operators, respectively. $(\cdot)^{(r)}$ and $(\cdot)^\star$ denote the r-th iterated and optimal results, respectively. The distribution of a Circularly Symmetric Complex Gaussian (CSCG) random variable with zero mean and variance σ^2 is denoted by $\mathcal{CN}(0, \sigma^2)$, and $\sim$ means "distributed as".

II. SCATTERING PRINCIPLES

RF wave scattering and reflecting can be described by a unified model and realized by variable-load antennas or programmable metamaterial [44]. An antenna-based scatterer usually consists of an integrated antenna, a load-switching modulator, an energy harvester, and on-chip components (e.g., microcontroller and sensors) [2]. It first receives the impinging signals, then reradiates some back to the space and dissipates the remaining. In comparison, a typical metamaterial-based scatterer comprises an outer metamaterial layer of numerous sub-wavelength metallic/dielectric patches (with tunable permittivity/permeability), a middle copper plate layer that avoids leakage, an inner circuit board layer

that adjusts patch responses, and an integrated microcontroller/FPGA for coordination and control [45]. Ideally, it reflects the wave at the space-metamaterial boundary without receiving them and mainly applies a phase shift. In practice, both types of scatterers have finite reflection states with non-zero reflection loss. The scattered signal can be decomposed into a structural mode component that consistently contributes to environment multipath and can be modelled within CSI, and an *antenna* mode component that depends on the impedance mismatch and can be used for backscatter modulation and/or passive beamforming [46]. For an antenna-based (resp. metamaterial-based) scatter node with M reflection states, the reflection coefficient at state $m \in \mathcal{M} \triangleq \{1,...,M\}$ is

$$\Gamma_m = \frac{Z_m - Z^*}{Z_m + Z}, \tag{1}$$

where Z_m is the antenna load (resp. metamaterial unit) impedance at state m and Z is the antenna input (resp. medium characteristic) impedance. BackCom employ scatter nodes as information sources that *randomly switching* between different states. For M-ary QAM, constellation point c_m maps to reflection coefficient Γ_m by [47]

$$\Gamma_m = \alpha \frac{c_m}{\max_{m'} |c_{m'}|}, \tag{2}$$

where $\alpha \in \mathbb{I}$ is the common amplitude scattering ratio at the direction of interest. In contrast, RIS employ reflect elements as channel adaptors that *deterministically choosing* the reflection state based on CSI. For a RIS element with M available states, phase shift θ_m maps to reflection coefficient Γ_m by [21]

$$\Gamma_m = \beta_m \exp(j\theta_m), \tag{3}$$

where $\beta_m \in \mathbb{I}$ is the overall amplitude scattering ratio at state m.[4]

III. RIScatter

We propose RIScatter in Fig. 1(f) as a generalization of BackCom and RIS, which involves coexisting active primary and passive backscatter links in a flexible and mutualistic manner. RIScatter nodes leverage CSI- and QoS-based reflection state distribution design to smoothly transition between backscatter modulation and passive beamforming. Fig. 2 illustrates the block diagram, equivalent circuit, and scatter model of a RIScatter node. It can be implemented by adding an integrated receiver [48] and adaptive encoder [49] to off-the-shelf passive RFID tags. Fig. 3 compares the reflection state distribution and reflection pattern of RIScatter and other applications. Instead of using fully random or fully deterministic reflection pattern over time, each RIScatter node *semi-randomly* chooses the reflection state for each backscatter block with *guidance of input probability* $P(\Gamma_m)$ at state m. Such an adaptive backscatter channel coding includes

[4]Most existing RIS literatures assume lossless reflection $\beta_m = 1, \forall m$.

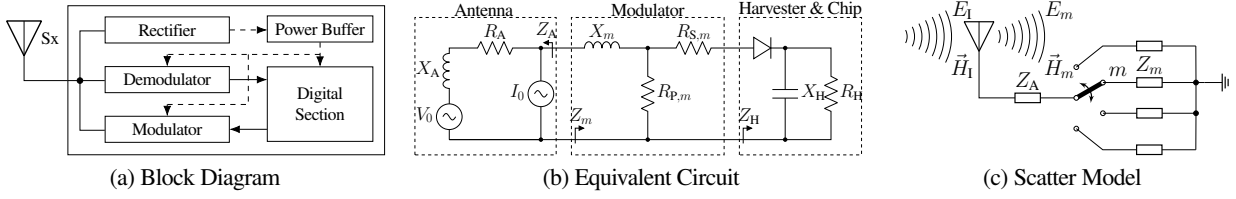(a) Block Diagram   (b) Equivalent Circuit   (c) Scatter Model

Fig. 2. Block diagram, equivalent circuit, and scatter model of a RIScatter node. The solid and dashed vectors represent signal and energy flows. The scatter antenna behaves as a constant power source, where the voltage V_0 and current I_0 are introduced by incident electric field $\vec{E}_\mathrm{I}$ and magnetic field $\vec{H}_\mathrm{I}$ [50].

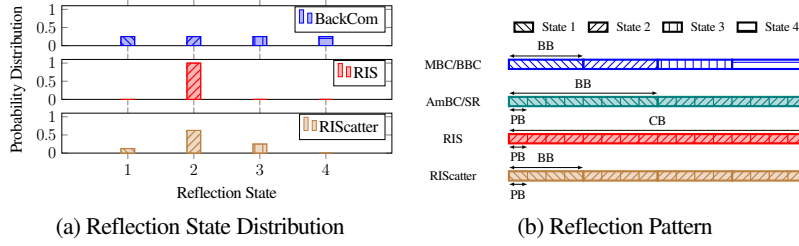(a) Reflection State Distribution   (b) Reflection Pattern

Fig. 3. Reflection state distribution and reflection pattern over time of scattering applications. "PB", "BB", and "CB" refer to primary block, backscatter block, and channel block, respectively. BackCom employs all possible states as information codewords, RIS treats them as beamformer codewords, and RIScatter simultaneously exploits both benefits.
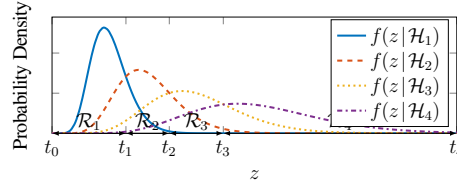

Fig. 4. PDF of accumulated receive energy per backscatter block conditioned on different reflection state hypotheses.

the degenerate distribution of RIS and the uniform distribution of line-coded BackCom as special cases.

Importantly, RIScatter system design only requires the *direct* transmitter-receiver and *cascaded* transmitter-scatter-receiver CSI. For dispersed RIScatter nodes, relevant CSI can be estimated by sequential [51]–[53] or parallel approaches [54] originally proposed for BackCom. For co-located RIScatter nodes, the estimation can be simplified by group-based [55] and hierarchical [56] trainings originally proposed for RIS.

We also propose a novel cooperative receiver that exploits the double modulation, double fading and symbol period difference to reduce the decoding complexity and improve the primary-backscatter tradeoff. As illustrated in Fig. 4, conditioned on different reflection state hypotheses, the accumulated receive energy per backscatter block follows Gamma distribution with different scale parameters [57]. Hence, the receiver can semi-coherently (under primary uncertainty) decode RIScatter nodes from the accumulated receive energy, re-encode to recover the exact reflection patterns, and combine those with cascaded CSI to determine the multipath response and model within the primary channel. Such a backscatter-primary decoding enables simultaneous backscatter modulation and dynamic passive beamforming by only one energy com-
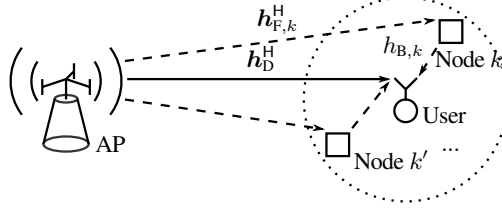
Fig. 5. A single-user multi-node RIScatter system.

parison and re-encoding during each backscatter (instead of primary) block. It enjoys a three-fold benefit compared to conventional primary-backscatter decoding with SIC, namely 1) reduces the re-encoding frequencies to $1/N$; 2) requires no re-precoding, cancellation, or MRC; 3) preserves backscatter modulation and passive beamforming at smaller spreading factors for potentially higher backscatter throughput.

As shown in Fig. 5, we consider a RIScatter system where a Q-antenna AP serves a single-antenna user as well as K nearby dispersed/co-located single-antenna M-state RIScatter nodes. In the primary point-to-point system, the AP transmits information to the user over the multipath channel enhanced by RIScatter nodes. In the backscatter multiple access system, the AP and user become carrier emitter and backscatter reader, and the RIScatter nodes modulate over the scattered RF signals. For simplicity, we consider a quasi-static block fading model and focus on a specific block where all CSI remain constant. Without loss of generality, let the spreading factor N be a positive integer. We also omit the signal reflected by two or more times and ignore the multipath propagation time difference. Denote the AP-user direct channel as $\boldsymbol{h}_{\mathrm{D}}^{\mathsf{H}} \in \mathbb{C}^{1 \times Q}$, the AP-node $k \in \mathcal{K} \triangleq \{1,...,K\}$ forward channel as $\boldsymbol{h}_{\mathrm{F},k}^{\mathsf{H}} \in \mathbb{C}^{1 \times Q}$, the node k-user backward channel as $h_{\mathrm{B},k}$, and the cascaded AP-node k-user channel as $\boldsymbol{h}_{\mathrm{C},k}^{\mathsf{H}} \triangleq h_{\mathrm{B},k} \boldsymbol{h}_{\mathrm{F},k}^{\mathsf{H}} \in \mathbb{C}^{1 \times Q}$. Let $x_k \in \mathcal{X} \triangleq \{c_1,...,c_M\}$ be the coded backscatter symbol of node k and $x_{\mathcal{K}} \triangleq (x_1,...,x_K)$ be the backscatter symbol tuple of all nodes. Due to double modulation, the primary equivalent channel is a function of backscatter symbol tuple[5]

$$\boldsymbol{h}_{\mathrm{E}}^{\mathsf{H}}(x_{\mathcal{K}}) \triangleq \boldsymbol{h}_{\mathrm{D}}^{\mathsf{H}} + \sum_{k} \alpha_k \boldsymbol{h}_{\mathrm{C},k}^{\mathsf{H}} x_k \tag{4a}$$

$$= \boldsymbol{h}_{\mathrm{D}}^{\mathsf{H}} + \boldsymbol{x}^{\mathsf{H}} \mathrm{diag}(\boldsymbol{\alpha}) \boldsymbol{H}_{\mathrm{C}}, \tag{4b}$$

where $\alpha_k \in \mathbb{I}$ is the common amplitude scattering ratio of node k, $\boldsymbol{\alpha} \triangleq [\alpha_1,...,\alpha_K]^{\mathsf{T}} \in \mathbb{I}^K$, $\boldsymbol{x} \triangleq [x_1,...,x_K]^{\mathsf{H}} \in \mathcal{X}^K$, and $\boldsymbol{H}_{\mathrm{C}} \triangleq [\boldsymbol{h}_{\mathrm{C},1},...,\boldsymbol{h}_{\mathrm{C},K}]^{\mathsf{H}} \in \mathbb{C}^{K \times Q}$. In the following context, we focus on one backscatter block (i.e., N primary blocks). The signal received by the user at primary block $n \in \mathcal{N} \triangleq \{1,...,N\}$ is

$$y[n] = \boldsymbol{h}_{\mathrm{E}}^{\mathsf{H}}(x_{\mathcal{K}}) \boldsymbol{w} s[n] + v[n], \tag{5}$$

where $\boldsymbol{w} \in \mathbb{C}^Q$ is the active beamformer satisfying average transmit power constraint $\|\boldsymbol{w}\|^2 \leq P$, $s \sim \mathcal{CN}(0,1)$ is the primary symbol, and $v \sim \mathcal{CN}(0,\sigma_v^2)$ is the Additive White Gaussian Noise

---

[5](4a) and (4b) are often used in BackCom and RIS literatures, respectively.

(AWGN) with average power σ_v^2. Let $m_k \in \mathcal{M} \triangleq \{1,...,M\}$ be the reflection state index of node k and $m_{\mathcal{K}} \triangleq (m_1,...,m_K)$ be the reflection state index tuple of all nodes. The backscatter symbol x_k (resp. symbol tuple $x_{\mathcal{K}}$) is a random variable (resp. variable tuple) that takes value x_{m_k} (resp. value tuple $x_{m_{\mathcal{K}}}$) when state m_k (resp. state tuple $m_{\mathcal{K}}$) is employed. Conditioned on $m_{\mathcal{K}}$, the receive signal at each primary block follows CSCG distribution $\mathcal{CN}(0,\sigma_{m_{\mathcal{K}}}^2)$, where

$$\sigma_{m_{\mathcal{K}}}^2 = |\boldsymbol{h}_{\mathsf{E}}^{\mathsf{H}}(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2 + \sigma_v^2 \tag{6}$$

is the received variance. Let $z = \sum_n |y[n]|^2$ be the accumulated receive energy per backscatter block. Since z is the sum of N independent and identically distributed (i.i.d.) exponential variables, its conditional PDF follows Gamma distribution

$$f(z|\mathcal{H}_{m_{\mathcal{K}}}) = \frac{z^{N-1}\exp(-z/\sigma_{m_{\mathcal{K}}}^2)}{\sigma_{m_{\mathcal{K}}}^{2N}(N-1)!}, \tag{7}$$

where $\mathcal{H}_{m_{\mathcal{K}}}$ denotes hypothesis $m_{\mathcal{K}}$. At the receiver, the energy space is divided into disjoint decision regions associated with different hypotheses, as illustrated in Fig. 4.

Remark 1. The capacity-achieving decision region design for DMTC with non-binary inputs in arbitrary distribution remains an open issue. It was proved deterministic detector can be rate-optimal, but non-convex decision regions (i.e., comprise non-adjacent partitions) are generally involved and the optimal number of thresholds remains unknown [58], [59]. Hence, we limit the backscatter energy detector to convex deterministic decision regions and consider sequential threshold design in the following context.

For the ease of notations, we map the state index tuple $m_{\mathcal{K}}$ to $l \in \mathcal{L} \triangleq \{1,...,L \triangleq M^K\}$, where $\sigma_1^2,...,\sigma_L^2$ is a non-decreasing sequence. Both notations are used interchangeably in the following context. As such, the decision region of backscatter symbol tuple l can be written as

$$\mathcal{R}_l \triangleq [t_{l-1},t_l), \quad 0 \le t_{l-1} \le t_l, \tag{8}$$

where t_l is the energy decision threshold between hypotheses $\mathcal{H}_l$ and $\mathcal{H}_{l+1}$. For a given decision threshold vector $\boldsymbol{t} \triangleq [t_0,...,t_L]^{\mathsf{T}} \in \mathbb{R}_+^{(L+1)}$, we can formulate a Discrete Memoryless Thresholding Multiple Access Channel (DMTMAC) with transition probability from input tuple $x_{m_{\mathcal{K}}}$ to output tuple $\hat{x}_{m'_{\mathcal{K}}}$ as

$$P(\hat{x}_{m'_{\mathcal{K}}}|x_{m_{\mathcal{K}}}) = \int_{\mathcal{R}_{m'_{\mathcal{K}}}} f(z|\mathcal{H}_{m_{\mathcal{K}}})dz, \tag{9}$$

over which adaptive backscatter channel coding can be performed.

Denote the probability of node k choosing reflection state m_k as $P_k(x_{m_k})$, and let $\boldsymbol{p}_k \triangleq [P_k(c_1),...,P_k(c_M)]^{\mathsf{T}} \in \mathbb{I}^M$. For dispersed nodes with independent encoding, the probability of backscatter symbol value tuple

$x_{m_\mathcal{K}}$ is

$$P_\mathcal{K}(x_{m_\mathcal{K}}) = \prod_{k \in \mathcal{K}} P_k(x_{m_k}). \tag{10}$$

Following [60], we define the backscatter information function between input value tuple $x_{m_\mathcal{K}}$ and output variable tuple $\hat{x}_\mathcal{K}$ as

$$I_\mathrm{B}(x_{m_\mathcal{K}}; \hat{x}_\mathcal{K}) \triangleq \sum_{m'_\mathcal{K}} P(\hat{x}_{m'_\mathcal{K}}|x_{m_\mathcal{K}}) \log \frac{P(\hat{x}_{m'_\mathcal{K}}|x_{m_\mathcal{K}})}{P_\mathcal{K}(\hat{x}_{m'_\mathcal{K}})}, \tag{11}$$

where $P_\mathcal{K}(\hat{x}_{m'_\mathcal{K}}) = \sum_{m_\mathcal{K}} P_\mathcal{K}(x_{m_\mathcal{K}}) P(\hat{x}_{m'_\mathcal{K}}|x_{m_\mathcal{K}})$ is the probability of output value tuple $\hat{x}_{m'_\mathcal{K}}$. We also define the backscatter marginal information of x_{m_k} as

$$I_{\mathrm{B},k}(x_{m_k}; \hat{x}_\mathcal{K}) \triangleq \sum_{m_{\mathcal{K}\setminus\{k\}}} P_{\mathcal{K}\setminus\{k\}}(x_{m_{\mathcal{K}\setminus\{k\}}}) I_\mathrm{B}(x_{m_\mathcal{K}}; \hat{x}_\mathcal{K}), \tag{12}$$

where $P_{\mathcal{K}\setminus\{k\}}(x_{m_{\mathcal{K}\setminus\{k\}}}) = \prod_{q \in \mathcal{K}\setminus\{k\}} P_q(x_{m_q})$. Hence, the backscatter mutual information can be written as

$$I_\mathrm{B}(x_\mathcal{K}; \hat{x}_\mathcal{K}) = \sum_{m_\mathcal{K}} P_\mathcal{K}(x_{m_\mathcal{K}}) I_\mathrm{B}(x_{m_\mathcal{K}}; \hat{x}_\mathcal{K}). \tag{13}$$

Once nodes are successfully decoded, we can re-encode and determine the primary equivalent channel by (4). Therefore, the primary information function conditioned on $x_{m_\mathcal{K}}$ is

$$I_\mathrm{P}(s; y | x_{m_\mathcal{K}}) \triangleq \log\left(1 + \frac{|\boldsymbol{h}_\mathrm{E}^\mathsf{H}(x_{m_\mathcal{K}})\boldsymbol{w}|^2}{\sigma_v^2}\right), \tag{14}$$

the primary marginal information of x_{m_k} is

$$I_{\mathrm{P},k}(s; y | x_{m_k}) \triangleq \sum_{m_{\mathcal{K}\setminus\{k\}}} P_{\mathcal{K}\setminus\{k\}}(x_{m_{\mathcal{K}\setminus\{k\}}}) I_\mathrm{P}(s; y | x_{m_\mathcal{K}}), \tag{15}$$

and the average primary mutual information is

$$I_\mathrm{P}(s; y | x_\mathcal{K}) = \sum_{m_\mathcal{K}} P_\mathcal{K}(x_{m_\mathcal{K}}) I_\mathrm{P}(s; y | x_{m_\mathcal{K}}). \tag{16}$$

With a slight abuse of notation, we define the corresponding weighted sum information function, marginal information, and mutual information as

$$I(x_{m_\mathcal{K}}) \triangleq \rho I_\mathrm{P}(s; y | x_{m_\mathcal{K}}) + (1-\rho) I_\mathrm{B}(x_{m_\mathcal{K}}; \hat{x}_\mathcal{K}), \tag{17}$$

$$I_k(x_{m_k}) \triangleq \rho I_{\mathrm{P},k}(s; y | x_{m_k}) + (1-\rho) I_{\mathrm{B},k}(x_{m_k}; \hat{x}_\mathcal{K}), \tag{18}$$

$$I(x_\mathcal{K}) \triangleq \rho I_\mathrm{P}(s; y | x_\mathcal{K}) + (1-\rho) I_\mathrm{B}(x_\mathcal{K}; \hat{x}_\mathcal{K}), \tag{19}$$

where $\rho \in \mathbb{I}$ is the relative QoS of the primary link. We notice the average primary rate (16) depends on the backscatter input distribution and active beamforming, while the total backscatter rate depends on the input distribution and DMTMAC (9) that relates to the active beamforming and decision thresholds.

IV. RATE-REGION CHARACTERIZATION

To characterize the achievable primary-(total-)backscatter rate region for the RIScatter system in Fig. 5, we aim to maximize the weighted sum rate with respect to input distribution $\{\boldsymbol{p}_k\}_{k \in \mathcal{K}}$, active beamforming $\boldsymbol{w}$, and decision thresholds $\boldsymbol{t}$ by

$$\max_{\{\boldsymbol{p}_k\}_{k \in \mathcal{K}}, \boldsymbol{w}, \boldsymbol{t}} \quad I(x_{\mathcal{K}}) \tag{20a}$$

$$\text{s.t.} \quad \mathbf{1}^{\mathsf{T}} \boldsymbol{p}_k = 1, \quad \forall k, \tag{20b}$$

$$\boldsymbol{p}_k \geq \mathbf{0}, \quad \forall k, \tag{20c}$$

$$\|\boldsymbol{w}\|^2 \leq P \quad, \tag{20d}$$

$$0 \leq t_{l-1} \leq t_l, \quad \forall l. \tag{20e}$$

Problem (20) generalizes conventional BackCom by allowing CSI- and QoS-adaptive input distribution and detection region design. It also generalizes the discrete RIS phase shift selection by allowing stochastic reflection (i.e., relaxing the feasible domain from the *vertices* of M-dimensional probability simplex to the simplex itself). Since problem (20) is highly non-convex, we propose a BCD algorithm that iteratively updates $\{\boldsymbol{p}_k\}_{k \in \mathcal{K}}$, $\boldsymbol{w}$ and $\boldsymbol{t}$ until convergence.

A. Input Distribution

For any given $\boldsymbol{w}$ and $\boldsymbol{t}$, we can construct the equivalent DMTMAC by (9) and simplify (20) to

$$\max_{\{\boldsymbol{p}_k\}_{k \in \mathcal{K}}} \quad I(x_{\mathcal{K}}) \tag{21a}$$

$$\text{s.t.} \quad (20b), (20c), \tag{21b}$$

which is convex when $K = 1$ or joint encoding[6] over $K > 1$ co-located nodes is available. When the nodes are dispersed, problem (21) involves coupled term $\prod_{k \in \mathcal{K}} P_k(x_{m_k})$ and is non-convex. Following [60], we first recast the KKT conditions to their equivalent forms, then propose a numerical method that guarantees those conditions on convergence of sequences.

Remark 2. As demonstrated in [61], KKT conditions are generally necessary but insufficient for total rate maximization of discrete Multiple Access Channel (MAC). We will show by numerical results that, for a moderate K, the average achievable rate regions of KKT and global-optimal input distributions completely coincide with each other.

---

[6]Joint encoding formulates an equivalent source of M^K codewords, such that one can directly design $P_{\mathcal{K}}(x_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}$ instead of $P_k(x_{m_k})$, $\forall k, m_k$.

Algorithm 1: Numerical KKT Input Distribution Evaluation by Limits of Sequence

---

Input: $K, N, h_{\text{D}}^{\text{H}}, H_{\text{C}}, \alpha, \mathcal{X}, \sigma_v^2, \rho, w, t, \epsilon$
Output: $\{p_k^\star\}_{k \in \mathcal{K}}$
  1: Set $h_{\text{E}}^{\text{H}}(x_{m_{\mathcal{K}}}), \forall m_{\mathcal{K}}$ by (4)
  2:      $\sigma_{m_{\mathcal{K}}}^2, \forall m_{\mathcal{K}}$ by (6)
  3:      $f(z | \mathcal{H}_{m_{\mathcal{K}}}), \forall m_{\mathcal{K}}$ by (7)
  4:      $P(\hat{x}_{m_{\mathcal{K}}'} | x_{m_{\mathcal{K}}}), \forall m_{\mathcal{K}}, m_{\mathcal{K}}'$ by (9)
  5: Initialize $r \leftarrow 0$
  6:        $p_k^{(0)} > 0, \forall k$
  7: Get $P_{\mathcal{K}}^{(r)}(x_{m_{\mathcal{K}}}), \forall m_{\mathcal{K}}$ by (10)
  8:      $I^{(r)}(x_{m_{\mathcal{K}}}), \forall m_{\mathcal{K}}$ by (11), (14), (17)
  9:      $I_k^{(r)}(x_{m_k}), \forall k, m_k$ by (12), (15), (18)
10:      $I^{(r)}(x_{\mathcal{K}})$ by (13), (16), (19)
11: **Repeat**
12:      Update $r \leftarrow r+1$
13:         $p_k^{(r)}, \forall k$ by (23)
14:      Redo step 7–10
15: **Until** $I^{(r)}(x_{\mathcal{K}}) - I^{(r-1)}(x_{\mathcal{K}}) \leq \epsilon$

Proposition 1. The KKT optimality conditions for problem (21) are equivalent to, $\forall k, m_k$,

$$I_k^\star(x_{m_k}) = I^\star(x_{\mathcal{K}}), \quad P_k^\star(x_{m_k}) > 0, \tag{22a}$$

$$I_k^\star(x_{m_k}) \leq I^\star(x_{\mathcal{K}}), \quad P_k^\star(x_{m_k}) = 0. \tag{22b}$$

Proof. Please refer to Appendix A. $\square$

For each node, (22a) suggests each probable state should produce the same marginal information (averaged over all states of other nodes), while (22b) suggests any state with potentially less marginal information should not be used.

Proposition 2. For any strictly positive initializer $\{p_k^{(0)}\}_{k \in \mathcal{K}}$, the KKT input probability of node k at state m_k is given by the converging point of the sequence

$$P_k^{(r+1)}(x_{m_k}) = \frac{P_k^{(r)}(x_{m_k}) \exp\left(\frac{\rho}{1-\rho} I_k^{(r)}(x_{m_k})\right)}{\sum_{m_k'} P_k^{(r)}(x_{m_k'}) \exp\left(\frac{\rho}{1-\rho} I_k^{(r)}(x_{m_k'})\right)}, \tag{23}$$

where r is the iteration index.

Proof. Please refer to Appendix B. $\square$

For (23) at iteration $r+1$, the input distribution of node k is updated over $\left\{\{p_q^{(r+1)}\}_{q=1}^{k-1}, \{p_q^{(r)}\}_{q=k}^{K}\right\}$. The KKT input distribution design is summarized in Algorithm 1.

B. Active Beamforming

For any given $\{\boldsymbol{p}_k\}_{k\in\mathcal{K}}$ and $\boldsymbol{t}$, problem (20) reduces to

$$\max_{\boldsymbol{w}} \quad I(x_{\mathcal{K}}) \tag{24a}$$

$$\text{s.t.} \quad (20\text{d}), \tag{24b}$$

which is still non-convex due to the integration and entropy terms. To tackle this, we rewrite the DMTMAC transition probability (9) from input index tuple $m_{\mathcal{K}}$ to output index l as a regularized incomplete Gamma function in the series representation [62, Theorem 3]

$$Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)=\frac{\int_{t_{l-1}/\sigma^2_{m_{\mathcal{K}}}}^{t_l/\sigma^2_{m_{\mathcal{K}}}} z^{N-1}\exp(-z)dz}{(N-1)!}=\exp\Big(-\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}}\Big)\sum_{n=0}^{N-1}\frac{\big(\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}}\big)^n}{n!}-\exp\Big(-\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)\sum_{n=0}^{N-1}\frac{\big(\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\big)^n}{n!}. \tag{25}$$

Its gradient with respect to $\boldsymbol{w}^*$ can be derived as

$$\nabla_{\boldsymbol{w}^*}Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)=\frac{\boldsymbol{h}_{\mathrm{E}}(x_{m_{\mathcal{K}}})\boldsymbol{h}_{\mathrm{E}}^{\mathsf{H}}(x_{m_{\mathcal{K}}})\boldsymbol{w}}{(\sigma^2_{m_{\mathcal{K}}})^2}g_{m_{\mathcal{K}}}(t_{l-1},t_l), \tag{26}$$

where $g_{m_{\mathcal{K}}}(t_{l-1},t_l)\triangleq g_{m_{\mathcal{K}}}(t_l)-g_{m_{\mathcal{K}}}(t_{l-1})$ and $g_{m_{\mathcal{K}}}(t_l)=t_l\exp\Big(-\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)\Big(-1+\sum_{n=1}^{N-1}\frac{\big(n-\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\big)\big(\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\big)^{n-1}}{n!}\Big)$.

On top of (25) and (26), we explicitly express the objective function (24a) and its gradient as

$$I(x_{\mathcal{K}})=\sum_{m_{\mathcal{K}}}P_{\mathcal{K}}(x_{m_{\mathcal{K}}})\Bigg(\rho\log\Big(1+\frac{|\boldsymbol{h}_{\mathrm{E}}^{\mathsf{H}}(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2}{\sigma^2_v}\Big)$$
$$+(1-\rho)\sum_l Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)\log\frac{Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)}{\sum_{m'_{\mathcal{K}}}P_{\mathcal{K}}(x_{m'_{\mathcal{K}}})Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m'_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m'_{\mathcal{K}}}}\Big)}\Bigg), \tag{27}$$

$$\nabla_{\boldsymbol{w}^*}I(x_{\mathcal{K}})=\sum_{m_{\mathcal{K}}}P_{\mathcal{K}}(x_{m_{\mathcal{K}}})\Bigg(\rho\frac{\boldsymbol{h}_{\mathrm{E}}(x_{m_{\mathcal{K}}})\boldsymbol{h}_{\mathrm{E}}^{\mathsf{H}}(x_{m_{\mathcal{K}}})\boldsymbol{w}}{\sigma^2_{m_{\mathcal{K}}}}+(1-\rho)\sum_l\Bigg(\log\frac{Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)}{\sum_{m'_{\mathcal{K}}}P_{\mathcal{K}}(x_{m'_{\mathcal{K}}})Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m'_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m'_{\mathcal{K}}}}\Big)}+1\Bigg)$$
$$\times\nabla_{\boldsymbol{w}^*}Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)-\frac{Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)\sum_{m'_{\mathcal{K}}}P_{\mathcal{K}}(x_{m'_{\mathcal{K}}})\nabla_{\boldsymbol{w}^*}Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m'_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m'_{\mathcal{K}}}}\Big)}{\sum_{m'_{\mathcal{K}}}P_{\mathcal{K}}(x_{m'_{\mathcal{K}}})Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m'_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m'_{\mathcal{K}}}}\Big)}\Bigg). \tag{28}$$

They allows problem (24) to be solved by the PGD method, where any unregulated beamformer $\bar{\boldsymbol{w}}$ can be projected onto the feasible domain of average transmit power constraint (20d) by

$$\boldsymbol{w}=\sqrt{P}\frac{\bar{\boldsymbol{w}}}{\max(\sqrt{P},\|\bar{\boldsymbol{w}}\|)}. \tag{29}$$

The PGD active beamforming optimization with step size by BLS [63, Section 9.2] is summarized in Algorithm 2.

Algorithm 2: Iterative Active Beamforming Optimization by PGD with BLS

Input: Q, N, $\boldsymbol{h}_{\mathrm{D}}^{\mathsf{H}}$, $\boldsymbol{H}_{\mathrm{C}}$, $\boldsymbol{\alpha}$, $\mathcal{X}$, P, σ_v^2, ρ, $\{\boldsymbol{p}_k\}_{k\in\mathcal{K}}$, $\boldsymbol{t}$, α, β, γ, ϵ

Output: $\boldsymbol{w}^\star$

1: Set $\boldsymbol{h}_{\mathrm{E}}^{\mathsf{H}}(x_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}$ by (4)

2:      $P_{\mathcal{K}}(x_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}$ by (10)

3: Initialize $r \leftarrow 0$

4:        $\boldsymbol{w}^{(0)}$, $\|\boldsymbol{w}^{(0)}\|^2 \leq P$

5: Get $(\sigma_{m_{\mathcal{K}}}^{(r)})^2$, $\forall m_{\mathcal{K}}$ by (6)

6:      $Q^{(r)}\big(N, \frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}}^2}, \frac{t_l}{\sigma_{m_{\mathcal{K}}}^2}\big)$, $\forall m_{\mathcal{K}},l$ by (25)

7:      $I^{(r)}(x_{\mathcal{K}})$ by (27)

8:      $\nabla_{\boldsymbol{w}^*} Q^{(r)}\big(N, \frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}}^2}, \frac{t_l}{\sigma_{m_{\mathcal{K}}}^2}\big)$, $\forall m_{\mathcal{K}},l$ by (26)

9:      $\nabla_{\boldsymbol{w}^*} I^{(r)}(x_{\mathcal{K}})$ by (28)

10: **Repeat**

11:      Update $r \leftarrow r+1$

12:        $\gamma^{(r)} \leftarrow \gamma$

13:        $\bar{\boldsymbol{w}}^{(r)} \leftarrow \boldsymbol{w}^{(r-1)} + \gamma \nabla_{\boldsymbol{w}^*} I^{(r-1)}(x_{\mathcal{K}})$

14:        $\boldsymbol{w}^{(r)}$ by (29)

15:      Redo step 5–7

16:      **While** $I^{(r)}(x_{\mathcal{K}}) < I^{(r-1)}(x_{\mathcal{K}}) + \alpha\gamma\|\nabla_{\boldsymbol{w}^*} I^{(r-1)}(x_{\mathcal{K}})\|^2$

17:        Set $\gamma^{(r)} \leftarrow \beta\gamma^{(r)}$

18:        Redo step 13–15

19:      **End While**

20:      Redo step 8, 9

21: **Until** $\|\boldsymbol{w}^{(r)} - \boldsymbol{w}^{(r-1)}\| \leq \epsilon$

C. Decision Threshold

For any given $\{\boldsymbol{p}_k\}_{k\in\mathcal{K}}$ and $\boldsymbol{w}$, problem (20) reduces to

$$\max_{\boldsymbol{t}} \quad I(x_{\mathcal{K}}) \tag{30a}$$

$$\text{s.t.} \quad (20e), \tag{30b}$$

which is still non-convex because variable $\boldsymbol{t}$ appears on the limits of integration (9). Fortunately, we can further simplify problem (30) as a point-to-point rate-optimal quantizer design for a discrete-input continuous-output memoryless channel, thanks to Remark 3 and 4.

Remark 3. Backscatter decision design has no impact on the primary achievable rate, because the primary equivalent channel (4) at each backscatter block can always be determined upon successful backscatter decoding. It suggests any thresholding scheme that maximize the total backscatter rate (13) is also optimal for problem (30).

Remark 4. In terms of total backscatter rate, the potentially dispersed nodes with known input distribution can be viewed as an equivalent source with symbol tuples as codewords. As such, the DMTMAC (9) becomes a DMTC and problem (30) reduces to the rate-optimal quantization design for a discrete-input continuous-output memoryless channel.
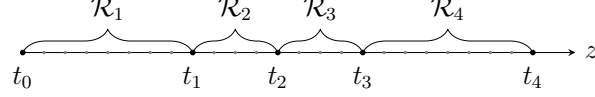
Fig. 6. The decision thresholds are selected from fine-grained discrete energy levels instead of continuous space, and each decision region consists of at least one neighbor energy bins.

Next, we constrain the feasible domain of problem (30) from continuous space $\mathbb{R}_+^{L+1}$ to finite candidate set (i.e., fine-grained discrete energy levels) $\mathcal{T}^{L+1}$. As shown in Fig. 6, by introducing an extra analog-to-digital conversion, we can group adjacent high-resolution energy bins to construct backscatter decision regions. Thus, problem (30) can be recast as

$$\max_{\boldsymbol{t} \in \mathcal{T}^{L+1}} \quad I_{\mathrm{B}}(x_{\mathcal{K}}; \hat{x}_{\mathcal{K}}) \tag{31a}$$

$$\text{s.t.} \quad (20e), \tag{31b}$$

which is solvable using existing rate-optimal sequential quantizer designs for DMTC. To obtain global optimal solution, [64] started from the quadrangle inequality and proposed a Dynamic Programming (DP) method accelerated by the Shor-Moran-Aggarwal-Wilber-Klawe (SMAWK) algorithm with computational complexity $\mathcal{O}\big(L^2(\mathrm{card}(\mathcal{T}) - L)\big)$, while [65] started from the optimality condition for three neighbor thresholds and presented a traverse-then-bisect algorithm with complexity $\mathcal{O}\big(\mathrm{card}(\mathcal{T})L\log(\mathrm{card}(\mathcal{T})L)\big)$. In Section V, both schemes will be compared with the ML scheme [66]

$$t_l^{\mathrm{ML}} = N \frac{\sigma_{l-1}^2 \sigma_l^2}{\sigma_{l-1}^2 - \sigma_l^2} \log \frac{\sigma_{l-1}^2}{\sigma_l^2}, \quad l \in \mathcal{L} \backslash \{L\}, \tag{32}$$

which is suboptimal for problem (30) except when all nodes are with equiprobable inputs.

V. Simulation Results

In this section, we provide numerical results to evaluate the proposed input distribution, active beamforming, and backscatter decision designs for the considered RIScatter system. We assume the AP-user distance is $10\,\mathrm{m}$ and at least one RIScatter nodes are randomly dropped in a disk centered at the user with radius r. The AP is with an average transmit power budget $P = 36\mathrm{dBm}$ and all nodes employs M-QAM with common amplitude scattering ratio $\alpha = 0.5$. For all channels involved, we consider a distance-dependent path loss model

$$L(d) = L_0 \left(\frac{d_0}{d} \right)^{\gamma}, \tag{33}$$

together with a Rician fading model

$$\boldsymbol{H} = \sqrt{\frac{\kappa}{1+\kappa}} \bar{\boldsymbol{H}} + \sqrt{\frac{1}{1+\kappa}} \tilde{\boldsymbol{H}}, \tag{34}$$
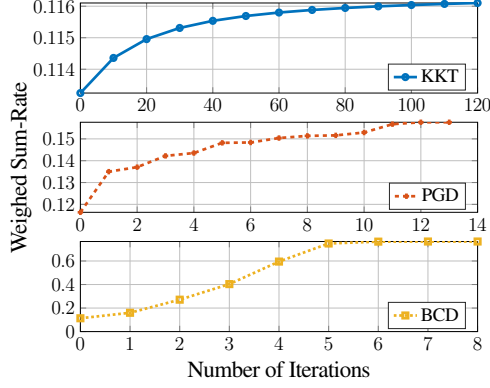
Fig. 7. Typical convergence curves at $\rho=0$ for $Q=4$, $K=8$, $M=2$, $N=20$, $\sigma_v^2=-40$dBm and $r=2$m.

where d is the transmission distance, $L_0=-30$dB is the reference path loss at $d_0=1$m, κ is the Rician K-factor, $\bar{H}$ is the deterministic line-of-sight component with unit-magnitude entries, and $\tilde{H}$ is the Rayleigh fading component with standard i.i.d. CSCG entries. We choose $\gamma_D=2.6$, $\gamma_F=2.4$, $\gamma_B=2$, and $\kappa_D=\kappa_F=\kappa_B=5$ for direct, forward and backward links. The finite decision threshold domain $\mathcal{T}$ is obtained by b-bit uniform discretization over the critical interval defined by the confidence bounds of edge hypotheses (i.e., lower bound of $\mathcal{H}_1$ and upper bound of $\mathcal{H}_L$) with confidence $1-\epsilon$, and we choose $b=9$ and $\epsilon=10^{-3}$. All achievable rate points/regions are averaged over 1000 channel realizations.

A. Evaluation of Proposed Algorithms

1) Initialization: To characterize each achievable rate region, we progressively obtain all boundary points by successively increasing the primary QoS ρ and solving problem (20). For $\rho=0$ where the backscatter link is prioritized, we initialize Algorithm 1 and 2 by uniform input distribution and Maximum Ratio Transmission (MRT) towards the sum cascaded channel $\sum_k h_{C,k}^H$, respectively. At the following points, both algorithms are initialized by the final solutions at the previous point.

2) Convergence: In Fig. 7, we plot the weighted sum of primary and total backscatter rates at $\rho=0$ for KKT, PGD and BCD algorithms on the first call. For $K=8$ and $M=2$, Algorithm 1 typically takes around 100 fast iterations by (23) to converge to the KKT input distribution. For $Q=4$, around 10 iterations are required for Algorithm 2 to converge, where the gradient is computed by (28) and the step size is refined by BLS. Overall, the BCD algorithm initially requires at most 5 iterations to converge. At the following points (not presented here), the convergence of all three algorithms are much faster thanks to the progressive initialization. Hence, we conclude the proposed algorithms have good convergence performances.

B. *Comparison of Scattering Applications*

On top of the setup in Fig. 5, we consider RIScatter and the following benchmark applications:

- *Legacy:* Conventional active transmission without antenna mode scattering, $\alpha = 0$.

- *BBC:* The primary symbol becomes deterministic $s[n] = 1$ and the receive signal at each primary block is

$$y^{\text{BBC}}[n] = \left(\boldsymbol{h}_{\text{D}}^{\text{H}} + \sum_k \alpha_k \boldsymbol{h}_{\text{C},k}^{\text{H}} x_k\right)\boldsymbol{w} + v[n], \tag{35}$$

which follows non-zero mean complex Gaussian distribution $\mathcal{CN}\left((\boldsymbol{h}_{\text{D}}^{\text{H}} + \sum_k \alpha_k \boldsymbol{h}_{\text{C},k}^{\text{H}} x_{m_k})\boldsymbol{w}, \sigma_v^2\right)$ under hypothesis $\mathcal{H}_{m_\mathcal{K}}$. The corresponding PDF of accumulated receive energy over N primary blocks is

$$f^{\text{BBC}}(z|\mathcal{H}_{m_\mathcal{K}}) = \frac{(z - \mu_{m_\mathcal{K}}^{\text{BBC}})^{N-1}\exp\left(-(z - \mu_{m_\mathcal{K}}^{\text{BBC}})/\sigma_v^2\right)}{\sigma_v^{2N}(N-1)!}, \tag{36}$$

where $\mu_{m_\mathcal{K}}^{\text{BBC}} \triangleq N\left|\left(\boldsymbol{h}_{\text{D}}^{\text{H}} + \sum_k \alpha_k \boldsymbol{h}_{\text{C},k}^{\text{H}} x_{m_k}\right)\boldsymbol{w}\right|^2$. The ML decision threshold is derived as, $\forall l \in \mathcal{L} \setminus \{L\}$,

$$t_l^{\text{BBC}} = \frac{\mu_{l-1}^{\text{BBC}}\exp\left((\mu_{l-1}^{\text{BBC}} - \mu_l^{\text{BBC}})/\sigma_v^2(N-1)\right) - \mu_l^{\text{BBC}}}{\exp\left((\mu_{l-1}^{\text{BBC}} - \mu_l^{\text{BBC}})/\sigma_v^2(N-1)\right) - 1}. \tag{37}$$

- *AmBC:* The user decodes both links independently and semi-coherently by treating the other as interference. Hence, the primary achievable rate is approximately[7]

$$I_{\text{P}}^{\text{AmBC}}(s;y) \approx \log\left(1 + \frac{|\boldsymbol{h}_{\text{D}}^{\text{H}}\boldsymbol{w}|^2}{\sum_k |\alpha_k \boldsymbol{h}_{\text{C},k}^{\text{H}}\boldsymbol{w}|^2 + \sigma_v^2}\right), \tag{38}$$

while the total backscatter rate follows (13) with uniform input distribution.

- *SR:* For a sufficiently large N, the average primary rate under semi-coherent detection asymptotically approaches (16) with uniform input distribution [12]. When $s[n]$ is successfully decoded and the direct interference $\boldsymbol{h}_{\text{D}}^{\text{H}}\boldsymbol{w}s[n]$ is perfectly cancelled, the intermediate signal is

$$\hat{y}^{\text{SR}}[n] = \sum_k \alpha_k \boldsymbol{h}_{\text{C},k}^{\text{H}} x_k \boldsymbol{w} s[n] + v[n], \tag{39}$$

which only involves noise uncertainty under hypothesis $\mathcal{H}_{m_\mathcal{K}}$. During backscatter detection, the primary symbols $s[1],...,s[N]$ function as a spreading code and the receiver performs MRC on $\hat{y}^{\text{SR}}[n]$ over N primary blocks. This essentially formulates a discrete-input continuous-output memoryless channel and the total backscatter achievable rate is maximized with equiprobable inputs as [67]

$$I_{\text{B}}(x_\mathcal{K}; \hat{y}^{\text{SR}}) = K \log M - \frac{\iota}{M^K}, \tag{40}$$

where $\iota \triangleq \sum_{m_\mathcal{K}} \mathbb{E}_{\hat{v}} \log \sum_{m'_\mathcal{K}} \exp(-|x_{m_\mathcal{K}} - x_{m'_\mathcal{K}} + \hat{v}|^2 / 2\sigma^2)$ and $\hat{v} \sim \mathcal{CN}(0, \sigma_v^2/N)$.

- *RIS:* Since the backscatter symbol tuple $x_\mathcal{K}$ is deterministic, the total backscatter rate is zero and

---

[7]To provide a preliminary benchmark, we consider the (correlated) scattered signal from finite-input backscatter sources as independent interference from Gaussian sources during primary decoding [12].
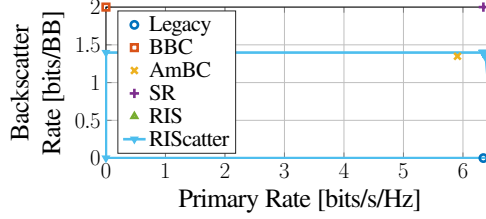
Fig. 8. Typical achievable rate region/points of scattering applications for $Q=1$, $K=1$, $M=4$, $N=1000$, $\sigma_v^2=-40$dBm and $r=2$m.

the primary achievable rate becomes a special case of (16)

$$I_{\mathrm{P}}^{\mathrm{RIS}}(s;y|x_{\mathcal{K}})=I_{\mathrm{P}}(s;y|x_{m_{\mathcal{K}}^{\star}})=\log\Big(1+\frac{|\boldsymbol{h}_{\mathrm{E}}^{\mathsf{H}}(x_{m_{\mathcal{K}}^{\star}})\boldsymbol{w}|^2}{\sigma_v^2}\Big),\tag{41}$$

where $m_{\mathcal{K}}^{\star}=\mathrm{argmax}_{m_{\mathcal{K}}}I_{\mathrm{P}}(s;y|x_{m_{\mathcal{K}}})$.

Fig. 8 compares the typical achievable rate region/points of RIScatter and above. *First,* we observe both BBC and SR almost ensure noise-free backscatter transmission when the spreading factor N is sufficiently large. For BBC with coherent energy detection, the conditional PDF of accumulated receive energy (36) is more skewed at a large N, such that the equivalent DMTMAC (9) has lower error probabilities. For SIC-based SR, the effective backscatter SNR is increased by N times and the penalty term ι becomes insignificant. However, the primary rate is only maintained at a very large spreading factor, which significantly constrains the backscatter *throughput*. *Second,* the average primary rate slightly increases/decreases in the presence of a AmBC/RIS node, and the multipath benefit of SR is unobvious. This is because the cascaded channel can be orders of magnitude weaker than the direct channel due to the double fading effect. RIS always ensures constructive superposition of direct and scattered components, while SR only creates a quasi-static rich-scattering environment that *stochastically* enhances the average primary rate. When N is moderate, the randomly scattered signals should be modelled as primary interference rather than multipath components, and the SR point will move towards the AmBC point. *Third,* RIScatter enables a flexible primary-backscatter tradeoff with adaptive input distribution design. In terms of maximum primary achievable rate, RIScatter coincides with RIS and outperforms the others with passive beamforming provided by deterministic reflection pattern. On the other hand, for a large N, the maximum backscatter achievable rate of RIScatter is higher than AmBC but lower than BBC and SR. This is because both RIScatter and AmBC (resp. BBC and SR) employ semi-coherent (resp. coherent) energy detection, while RIScatter with adaptive channel coding can achieve higher backscatter rate than AmBC with equiprobable inputs. Importantly, such a practical RIScatter detection is feasible for arbitrary spreading factor, which unleashes the potential of fast-switching nodes.
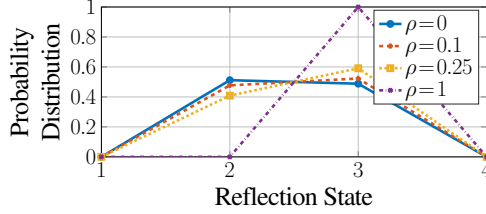
Fig. 9. Typical RIScatter reflection state distribution at different ρ for $Q=1$, $K=1$, $M=4$, $N=20$, $\sigma_v^2=-40$dBm and $r=2$m.

(a) Input Distribution, $Q=1$ (b) Active Beamforming, $Q=4$ (c) Decision Threshold, $Q=4$
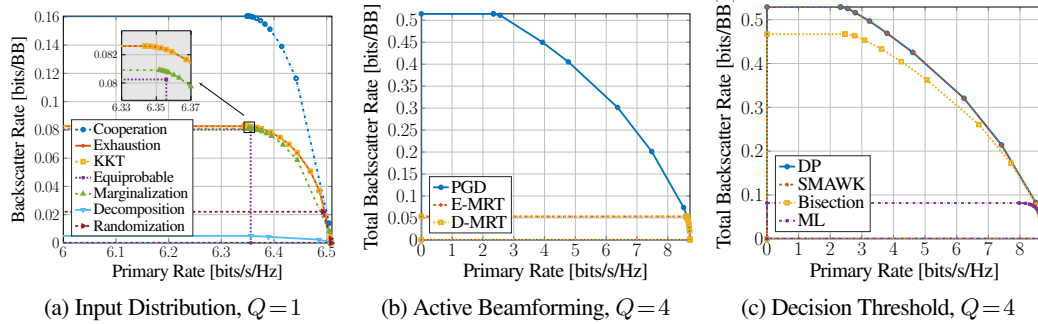
Fig. 10. Average primary-total-backscatter rate regions by different input distribution, active beamforming, and decision threshold schemes for $K=2$, $M=4$, $N=20$, $\sigma_v^2=-40$dBm and $r=2$m.

C. Input Distribution under Different QoS

The objective of this study is to demonstrate RIScatter nodes can leverage CSI- and QoS-adaptive input distribution design to balance backscatter modulation and passive beamforming. For one RIScatter node with $M=4$, we evaluate the KKT input distribution[8] at different primary QoS and present the result in Fig. 9. At $\rho=0$ where the backscatter performance is prioritized, the optimal input distribution is zero on two states and nearly uniform on the other two. This is because, due to the weak scattered signal, the conditional energy PDF under different hypotheses can be closely spaced as illustrated in Fig. 4. In such cases, the extreme states producing the lowest/highest energy are always assigned with non-zero probability, while the middle ones may not provide enough energy difference and thus end up unused. At $\rho=1$ where the primary performance is prioritized, the optimal input distribution is 1 at the state that maximizes the primary SNR and 0 at the others. That is, the reflection pattern becomes deterministic and the RIScatter node boils down to a discrete RIS element. Increasing ρ from 0 to 1 provides a smooth transition from backscatter modulation to passive beamforming, which suggests RIScatter unifies BackCom and RIS from a probabilistic perspective.

D. Rate Region by Different Schemes

1) Input Distribution: We compare these input distribution designs for problem (21):

[8]Since problem (21) is convex when $K=1$, the KKT solution is always global optimal in this case.

- *Cooperation:* Joint node encoding with a K-dimensional probability array $P_{\mathcal{K}}(x_{m_{\mathcal{K}}})$ by Algorithm 1;
- *Exhaustion:* Exhaustive search over M-dimensional probability simplex with step $\Delta p = 10^{-2}$;
- *KKT:* Numerical KKT result evaluation by Algorithm 1;
- *Equiprobable:* Uniform input distribution for all nodes.

For dispersed nodes, joint encoding is inapplicable and we consider these independent input distribution recovery methods (from the joint probability array):

- *Marginalization:* Marginal probability distributions;
- *Decomposition:* Normalized rank-1 Canonical Polyadic (CP) decomposition tensors by `Tensor Toolbox` [68];
- *Randomization:* Gaussian randomization with the guidance of correlation matrix [69].

Fig. 10(a) shows the average achievable rate regions for those designs. We observe *Cooperation* provides the outer bound since joint encoding is always beneficial. The *average* rate performance of *Exhaustion* and *KKT* completely coincide with each other, demonstrating KKT input distribution is favorable for a moderate K as stated in Remark 2. *Equiprobable* experiences minor backscatter and major primary rate losses without CSI- and QoS-adaptive backscatter encoding. Both gaps should become larger when M and K increase. For the recovery methods, the simple *Marginalization* provides a close result to *KKT*, but *Randomization* and *Decomposition* fail our expectations for most channel realizations. Those observations emphasize the importance of (joint) adaptive RIScatter encoding and demonstrate the advantages of the proposed KKT input distribution design.

2) Active Beamforming: We consider three typical active beamforming schemes for problem (24):

- *PGD:* Iterative PGD optimization by Algorithm 2;
- *E-MRT:* MRT towards the ergodic primary equivalent channel $\sum_{m_{\mathcal{K}}} P_{\mathcal{K}}(x_{m_{\mathcal{K}}}) \boldsymbol{h}_{\mathrm{E}}^{\mathsf{H}}(x_{m_{\mathcal{K}}})$;
- *D-MRT:* MRT towards the direct channel $\boldsymbol{h}_{\mathrm{D}}^{\mathsf{H}}$.

Fig. 10(b) presents the average achievable rate regions for those schemes. In the low-ρ regime, the proposed *PGD* beamformer significantly outperforms both MRT schemes in terms of total backscatter rate. This is because backscatter detection relies on the *relative difference* of accumulated receive energy under different symbol tuples. Such an energy diversity is enhanced by *PGD* that effectively exploits backscatter constellation and input distribution knowledge, rather than simply maximizing the direct/ergodic equivalent SNR. As ρ increases, the primary equivalent SNR outweighs the backscatter energy difference in (27), and *PGD* beamformer becomes closer to both MRT schemes. At $\rho = 1$, *PGD* and *E-MRT* boil down to MRT towards the deterministic primary equivalent channel as in RIS literatures

[70]. Besides, the difference between *E-MRT* and *D-MRT* can be insignificant for dispersed RIScatter nodes. Those observations prove the proposed *PGD* active beamforming design can exploit the CSI, QoS, and backscatter constellation to balance the primary equivalent SNR and backscatter energy difference and enlarge the achievable rate region for RIScatter.

3) Decision Threshold: We evaluate the following decision threshold strategies for problem (31):

- *DP:* Benchmark DP method for sequential quantizer [64];
- *SMAWK: DP* accelerated by the SMAWK algorithm [64];
- *Bisection:* The traverse-then-bisect algorithm [65];
- *ML:* Maximum likelihood detector (32) [66].

Fig. 10(c) reveals the average achievable rate region for those strategies. The distribution-aware schemes *DP*, *SMAWK* and *Bisection* ensure higher total backscatter rate than the non-adaptive *ML*. This is because the total backscatter rate (13) is a function of both input distribution and decision regions, and the rate-optimal threshold design heavily depends on input distribution. For example, the backscatter symbol tuples with zero input probability should be assigned with empty decision regions, in order to increase the success detection chances of other hypotheses. It highlights the importance of joint input distribution and decision threshold design in rate maximization problems.

E. Rate Region under Different Configurations

In this study, we choose $Q=4$, $K=8$, $M=2$, $N=20$, $\sigma_v^2=-40$dBm and $r=2$m as a reference.

1) Number of Nodes: Fig. 11(a) reveals how the number RIScatter nodes K influence the primary-backscatter tradeoff. Interestingly, we observe that increasing K has a larger benefit on the total backscatter rate than primary. This is because each RIScatter node not only affects the primary equivalent SNR but also influences the relative energy difference that other nodes can create. To maximize the total backscatter rate, some nodes closer to the user may need to sacrifice their own rate and use the state that *minimizes* the primary equivalent channel strength, in order to increase the backscatter rate of other nodes. This accounts for the significant primary rate decrease in the low-ρ regime. On the other hand, when the primary link is prioritized, the RIScatter nodes boil down to RIS elements and enjoy a classic passive array gain of N^2.

2) Number of States: Fig. 11(b) shows the relationship between available reflection states (i.e., QAM order) M and achievable rate regions. We notice increasing the reflection states has a marginal effect on both primary and total backscatter rates. This is because once the scope of reflection coefficient is determined, using denser constellation points may not create enough phase resolution and energy diversity for primary and backscatter links. Due to the maximum amplitude normalization in (2), the average
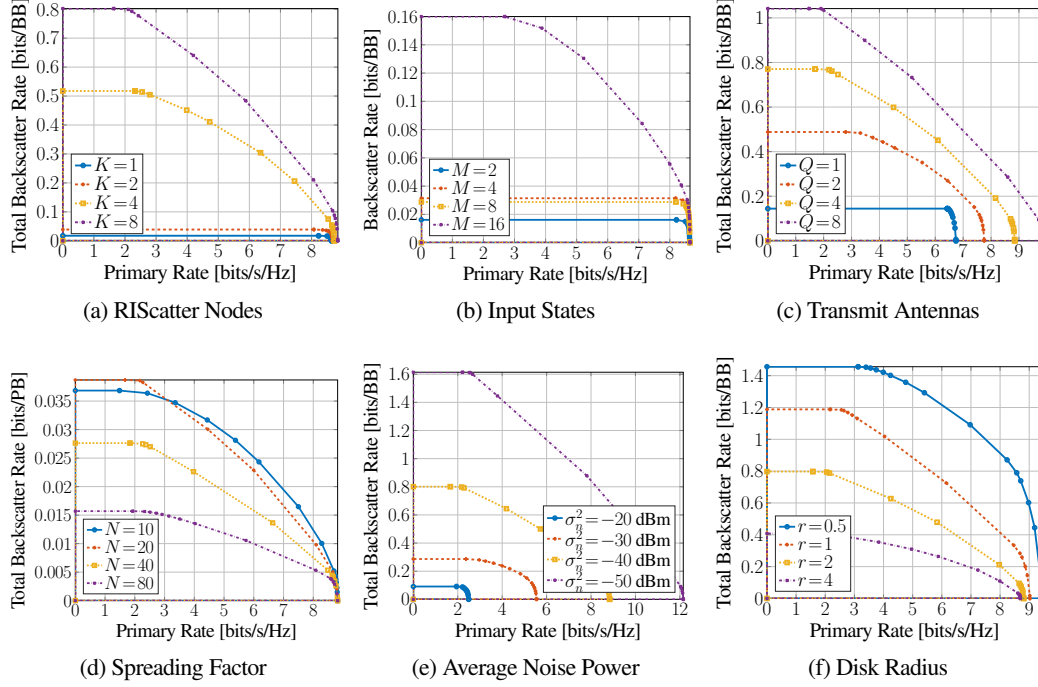
Fig. 11. Average primary-total-backscatter rate regions for different system configurations.

rate region of 8-QAM (points on rectangle border) is smaller than that of 4-QAM (points on square vertices), and the inner constellation points with smaller amplitudes are less frequently used. On the other hand, 16-QAM provides significant higher backscatter rate benefit than primary. It motivates one to use high-resolution metamaterial units as RIScatter nodes.

3) Number of Transmit Antennas: Fig. 11(c) illustrates the impact of transmit antennas Q on the average performance. As Q increases, more transmitter-scatter-receiver paths become available to *PGD* and their diversity can be further exploited to enhance the primary equivalent SNR and backscatter energy difference, which leads to a larger achievable rate region. It emphasizes the importance of multi-antenna RIScatter systems and demonstrate the effectiveness of the proposed PGD design.

4) Spreading factor: Fig. 11(d) presents how the spreading factor N affects the achievable rate region, where the unit of total backscatter rate is bits per primary block (to indicate throughput). Apparently, using a very large N can severely constrain the backscatter throughput, since the gain in energy certainty (by law of large numbers) cannot withstand the loss in gross rate. As $N \to \infty$, RIScatter nodes boil down to RIS elements with fixed reflection pattern during whole channel block and the total backscatter rate approaches 0. On the other hand, when N is too small, the equivalent DMTMAC (9) has very high error probabilities and the energy detection can thus be unreliable. It explains the observation that $N = 10$ provides lower backscatter achievable rate than $N = 20$, although nodes are encoding at higher gross

rate. Therefore, we conclude the design of spreading factor N in RIScatter systems should account for multiple factors (e.g., data rate requirements, load switching constraints at the nodes, and signal processing capability at the user).

5) Average Noise Power: Fig. 11(e) depicts the impact of average noise power σ_v^2 on average rate regions. It shows the proposed practical semi-coherent backscatter energy detection is suitable for a wide range of noise levels. When σ_v^2 relatively high, one can choose a larger N to maintain the backscatter SNR for better detection performance.

6) Coverage Disk Radius: Fig. 11(f) shows the relationship between disk radius r and achievable rate region. We observe both primary and backscatter performance are enhanced when nodes are located closer to the user. This is because the double fading effect is less severe for near-far setups. In a multi-user RIScatter system with dispersed nodes, each node may be allocated to the nearest user to guarantee uniformly good performance for both links.

VI. CONCLUSION

This paper introduced RIScatter as a low-power scatter protocol that unifies backscatter modulation and dynamic passive beamforming by smart input distribution and practical receiver design. Starting from scattering principles, we showed how RIScatter nodes include information node of BackCom and reflect element of RIS as special cases, how they can be built over existing passive scatter devices, and how they simultaneously encode self information and assist legacy transmission. We also propose a practical RIScatter receiver that preserves the benefits of backscatter modulation and dynamic passive beamforming without SIC. The achievable primary-total-backscatter rate region is then studied for a single-user multi-node RIScatter system, where the input distribution, active beamforming, and decision thresholds are iteratively updated. Numerical results not only validated the proposed algorithms, but also emphasized the importance of adaptive input distribution and cooperative decoding on both primary and backscatter links.

One possible direction is to consider backscatter detection over the received signal domain rather than energy domain, where multi-antenna and learning-based approaches can be promising. Another interesting question is how to design RIScatter node and receiver in a multi-user system to fully exploit the dynamic passive beamforming that naturally origins from backscatter modulation.

APPENDIX

A. *Proof of Proposition 1*

Denote the Lagrange multipliers associated with (20b) and (20c) as $\{\nu_k\}_{k\in\mathcal{K}}$ and $\{\lambda_{k,m_k}\}_{k\in\mathcal{K},m_k\in\mathcal{M}}$, respectively. The KKT conditions of problem (21) are, $\forall k, m_k$,

$$-\nabla_{P_k^\star(x_{m_k})}I^\star(x_\mathcal{K})+\nu_k^\star-\lambda_{k,m_k}^\star=0, \tag{42a}$$

$$\lambda_{k,m_k}^\star=0, \quad P_k^\star(x_{m_k})>0, \tag{42b}$$

$$\lambda_{k,m_k}^\star\geq0, \quad P_k^\star(x_{m_k})=0, \tag{42c}$$

where $\nabla_{P_k^\star(x_{m_k})}I^\star(x_\mathcal{K})=I_k^\star(x_{m_k})-(1-\rho)$. Therefore, we have

$$I_k^\star(x_{m_k})=\nu_k^\star+(1-\rho), \quad P_k^\star(x_{m_k})>0, \tag{43a}$$

$$I_k^\star(x_{m_k})\leq\nu_k^\star+(1-\rho), \quad P_k^\star(x_{m_k})=0, \tag{43b}$$

such that $\sum_{m_k}P_k^\star(x_{m_k})I_k^\star(x_{m_k}) = \nu_k^\star + (1-\rho)$. On the other hand, by definition (18) we have $\sum_{m_k}P_k^\star(x_{m_k})I_k^\star(x_{m_k})=I^\star(x_\mathcal{K})$, where the right-hand side is irrelevant to k. This complete the proof.

B. *Proof of Proposition 2*

We first prove sequence (23) is non-decreasing in weighted sum mutual information. Let $P_\mathcal{K}(x_{m_\mathcal{K}})=\prod_{q\in\mathcal{K}}P_q(x_{m_q})$ and $P_\mathcal{K}'(x_{m_\mathcal{K}})=P_k'(x_{m_k})\prod_{q\in\mathcal{K}\setminus\{k\}}P_q(x_{m_q})$ be two probability distributions with potentially different marginal for tag k at state m_k, and define an intermediate function

$$J\big(P_\mathcal{K}(x_{m_\mathcal{K}}),P_\mathcal{K}'(x_{m_\mathcal{K}})\big)\triangleq\sum_{m_\mathcal{K}}P_\mathcal{K}(x_{m_\mathcal{K}})\bigg(\rho\log\Big(1+\frac{|\boldsymbol{h}_\mathsf{E}^\mathsf{H}(x_{m_\mathcal{K}})\boldsymbol{w}|^2}{\sigma_v^2}\Big)$$

$$+(1-\rho)\sum_{m_\mathcal{K}'}P(\hat{x}_{m_\mathcal{K}'}|x_{m_\mathcal{K}})\log\frac{P(\hat{x}_{m_\mathcal{K}'}|x_{m_\mathcal{K}})P_\mathcal{K}'(x_{m_\mathcal{K}})}{P_\mathcal{K}'(\hat{x}_{m_\mathcal{K}'})P_\mathcal{K}(x_{m_\mathcal{K}})}\bigg). \tag{44}$$

It is straightforward to verify $J\big(P_\mathcal{K}(x_{m_\mathcal{K}}),P_\mathcal{K}(x_{m_\mathcal{K}})\big)=I(x_\mathcal{K})$ and $J\big(P_\mathcal{K}(x_{m_\mathcal{K}}),P_\mathcal{K}'(x_{m_\mathcal{K}})\big)$ is a concave function for a fixed $P_\mathcal{K}'(x_{m_\mathcal{K}})$. Setting $\nabla_{P_k^\star(x_{m_k})}J\big(P_\mathcal{K}(x_{m_\mathcal{K}}),P_\mathcal{K}'(x_{m_\mathcal{K}})\big)=0$, we have

$$S_k'(x_{m_k})-S_k'(x_{i_k})+(1-\rho)\log\frac{P_k(x_{i_k})}{P_k^\star(x_{m_k})}=0, \tag{45}$$

where $S_k'(x_{m_k}) \triangleq I_k'(x_{m_k}) + (1-\rho)\sum_{m_{\mathcal{K}\setminus\{k\}}}P_{\mathcal{K}\setminus\{k\}}(x_{m_{\mathcal{K}\setminus\{k\}}})\sum_{m_\mathcal{K}'}P(\hat{x}_{m_\mathcal{K}'}|x_{m_\mathcal{K}})\log P_\mathcal{K}'(x_{m_\mathcal{K}})$ and $i_k\neq m_k$ is the reference state. Evidently, $\forall m_k\neq i_k$, (45) boils down to

$$P_k^\star(x_{m_k})=\frac{P_k'(x_{m_k})\exp\Big(\frac{\rho}{1-\rho}I_k'(x_{m_k})\Big)}{\sum_{m_k'}P_k'(x_{m_k'})\exp\Big(\frac{\rho}{1-\rho}I_k'(x_{m_k'})\Big)}. \tag{46}$$

Since $P_k(x_{i_k}) = 1 - \sum_{m_k \neq i_k} P_k^\star(x_{m_k})$ has exactly the same form as (46), the choice of reference state i_k does not matter and (46) is indeed optimal $\forall m_k \in \mathcal{M}$. That is, for a fixed $P_\mathcal{K}'(x_{m_\mathcal{K}})$, choosing $P_k(x_{m_k})$ by (46) ensures $J\big(P_\mathcal{K}(x_{m_\mathcal{K}}), P_\mathcal{K}'(x_{m_\mathcal{K}})\big) \geq I'(x_\mathcal{K})$. On the other hand, we also have

$$\Delta \triangleq I(x_\mathcal{K}) - J\big(P_\mathcal{K}(x_{m_\mathcal{K}}), P_\mathcal{K}'(x_{m_\mathcal{K}})\big) \tag{47a}$$

$$= (1-\rho) \sum_{m_k} \frac{P_k'(x_{m_k}) f_k'(x_{m_k})}{\sum_{m_k'} P_k'(x_{m_k'}) f_k'(x_{m_k'})} \sum_{m_\mathcal{K}''} P(\hat{x}_{m_\mathcal{K}''} | x_{m_k}) \log \frac{\sum_{m_k'} P_k'(x_{m_k'}) P(\hat{x}_{m_\mathcal{K}''} | x_{m_k'}) f_k'(x_{m_k})}{\sum_{m_k'} P_k'(x_{m_k'}) P(\hat{x}_{m_\mathcal{K}''} | x_{m_k'}) f_k'(x_{m_k'})} \tag{47b}$$

$$\geq (1-\rho) \sum_{m_k} \frac{P_k'(x_{m_k}) f_k'(x_{m_k})}{\sum_{m_k'} P_k'(x_{m_k'}) f_k'(x_{m_k'})} \sum_{m_\mathcal{K}''} P(\hat{x}_{m_\mathcal{K}''} | x_{m_k}) \left(1 - \frac{\sum_{m_k'} P_k'(x_{m_k'}) P(\hat{x}_{m_\mathcal{K}''} | x_{m_k'}) f_k'(x_{m_k'})}{\sum_{m_k'} P_k'(x_{m_k'}) P(\hat{x}_{m_\mathcal{K}''} | x_{m_k'}) f_k'(x_{m_k})}\right) = 0,$$

$$\tag{47c}$$

where $f_k'(x_{m_k}) \triangleq \exp\big(\frac{\rho}{1-\rho} I_k'(x_{m_k})\big)$ and the equality holds if and only if (46) converges. Therefore, we have $I(x_\mathcal{K}) \geq I'(x_\mathcal{K})$. Since mutual information is bounded above, we conclude the sequence (23) is non-decreasing and convergent in mutual information. Next, we prove any converging point of sequence (23), denoted as $P_k^\star(x_{m_k})$, fulfills KKT conditions (22). To see this, let

$$D_k^{(r)}(x_{m_k}) \triangleq \frac{P_k^{(r+1)}(x_{m_k})}{P_k^{(r)}(x_{m_k})} = \frac{f_k^{(r)}(x_{m_k})}{\sum_{m_k'} P_k^{(r)}(x_{m_k'}) f_k^{(r)}(x_{m_k'})}. \tag{48}$$

As sequence (23) is convergent, any state with $P_k^\star(x_{m_k}) > 0$ need to satisfy $D_k^\star(x_{m_k}) \triangleq \lim_{r \to \infty} D_k^{(r)}(x_{m_k}) = 1$, namely $I_k^\star(x_{m_k}) = \frac{1-\rho}{\rho} \log \sum_{m_k'} P_k^\star(x_{m_k'}) f_k^\star(x_{m_k'})$, which is reminiscent of (43a) and (22a). That is, given $P_k^{(0)}(x_{m_k}) > 0$, any converging point with $P_k^\star(x_{m_k}) > 0$ must satisfy (22a). On the other hand, we assume $P_k^\star(x_{m_k})$ does not satisfy (22b), such that for any state with $P_k^\star(x_{m_k}) = 0$, $I_k^\star(x_{m_k}) > I^\star(x_\mathcal{K}) = \sum_{m_k'} P_k^\star(x_{m_k'}) I_k^\star(x_{m_k'})$, where the equality inherits from (19). Since the exponential function is monotonically increasing, we have $f_k^\star(x_{m_k}) > \sum_{m_k'} P_k^\star(x_{m_k'}) f_k^\star(x_{m_k'})$ and $D_k^\star(x_{m_k}) > 1$. Considering $P_k^{(0)}(x_{m_k}) > 0$ and $P_k^\star(x_{m_k}) = 0$, it contradicts with $P_k^{(r)}(x_{m_k}) = P_k^{(0)}(x_{m_k}) \prod_{n=1}^{r} D_k^{(n)}(x_{m_k})$. That is, given $P_k^{(0)}(x_{m_k}) > 0$, any converging point with $P_k^\star(x_{m_k}) = 0$ must satisfy (22b). The proof is completed.

REFERENCES

[1] C. Boyer and S. Roy, "Backscatter communication and RFID: Coding, energy, and MIMO analysis," *IEEE Transactions on Communications*, vol. 62, pp. 770–785, Mar 2014.

[2] D. Dobkin, *The RF in RFID: Passive UHF RFID in Practice*. Newnes, Nov 2012.

[3] J. Landt, "The history of RFID," *IEEE Potentials*, vol. 24, pp. 8–11, Oct 2005.

[4] G. Vannucci, A. Bletsas, and D. Leigh, "A software-defined radio system for backscatter sensor networks," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 2170–2179, Jun 2008.

[5] S. D. Assimonis, S. N. Daskalakis, and A. Bletsas, "Sensitive and efficient RF harvesting supply for batteryless backscatter sensor networks," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, pp. 1327–1338, Apr 2016.

[6] V. Liu, A. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. R. Smith, "Ambient backscatter: Wireless communication out of thin air," *ACM SIGCOMM Computer Communication Review*, vol. 43, pp. 39–50, Sep 2013.

[7] G. Yang, Q. Zhang, and Y.-C. Liang, "Cooperative ambient backscatter communications for green internet-of-things," *IEEE Internet of Things Journal*, vol. 5, pp. 1116–1130, Apr 2018.

[8] Y.-C. Liang, Q. Zhang, E. G. Larsson, and G. Y. Li, "Symbiotic radio: Cognitive backscattering communications for future wireless networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, pp. 1242–1255, Dec 2020.

[9] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Transactions on Communications*, vol. 69, pp. 3313–3351, May 2021.

[10] H. Guo, Y.-C. Liang, R. Long, and Q. Zhang, "Cooperative ambient backscatter system: A symbiotic radio paradigm for passive IoT," *IEEE Wireless Communications Letters*, vol. 8, pp. 1191–1194, Aug 2019.

[11] H. Ding, D. B. da Costa, and J. Ge, "Outage analysis for cooperative ambient backscatter systems," *IEEE Wireless Communications Letters*, vol. 9, pp. 601–605, May 2020.

[12] R. Long, Y.-C. Liang, H. Guo, G. Yang, and R. Zhang, "Symbiotic radio: A new communication paradigm for passive internet of things," *IEEE Internet of Things Journal*, vol. 7, pp. 1350–1363, Feb 2020.

[13] S. Zhou, W. Xu, K. Wang, C. Pan, M.-S. Alouini, and A. Nallanathan, "Ergodic rate analysis of cooperative ambient backscatter communication," *IEEE Wireless Communications Letters*, vol. 8, pp. 1679–1682, Dec 2019.

[14] T. Wu, M. Jiang, Q. Zhang, Q. Li, and J. Qin, "Beamforming design in multiple-input-multiple-output symbiotic radio backscatter systems," *IEEE Communications Letters*, vol. 25, pp. 1949–1953, Jun 2021.

[15] J. Xu, Z. Dai, and Y. Zeng, "Enabling full mutualism for symbiotic radio with massive backscatter devices," *arXiv:2106.05789*, Jun 2021.

[16] Z. Yang and Y. Zhang, "Optimal SWIPT in RIS-aided MIMO networks," *IEEE Access*, vol. 9, pp. 112 552–112 560, 2021.

[17] S. Han, Y.-C. Liang, and G. Sun, "The design and optimization of random code assisted multi-BD symbiotic radio system," *IEEE Transactions on Wireless Communications*, vol. 20, pp. 5159–5170, Aug 2021.

[18] Q. Zhang, Y.-C. Liang, H.-C. Yang, and H. V. Poor, "Mutualistic mechanism in symbiotic radios: When can the primary and secondary transmissions be mutually beneficial?" *IEEE Transactions on Wireless Communications*, vol. 1276, pp. 1–1, 2022.

[19] R. Torres, R. Correia, N. Carvalho, S. N. Daskalakis, G. Goussetis, Y. Ding, A. Georgiadis, A. Eid, J. Hester, and M. M. Tentzeris, "Backscatter communications," *IEEE Journal of Microwaves*, vol. 1, pp. 864–878, Oct 2021.

[20] G. Vougioukas and A. Bletsas, "Switching frequency techniques for universal ambient backscatter networking," *IEEE Journal on Selected Areas in Communications*, vol. 37, pp. 464–477, Feb 2019.

[21] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," vol. 18. IEEE, Dec 2018, pp. 1–6.

[22] S. Zhang and R. Zhang, "Capacity characterization for intelligent reflecting surface aided MIMO communication," *IEEE Journal on Selected Areas in Communications*, vol. 38, pp. 1823–1838, Aug 2020.

[23] S. Lin, B. Zheng, F. Chen, and R. Zhang, "Intelligent reflecting surface-aided spectrum sensing for cognitive radio," *IEEE Wireless Communications Letters*, vol. 11, pp. 928–932, May 2022.

[24] Y. Liu, Y. Zhang, X. Zhao, S. Geng, P. Qin, and Z. Zhou, "Dynamic-controlled RIS assisted multi-user MISO downlink system: Joint beamforming design," *IEEE Transactions on Green Communications and Networking*, vol. 6, pp. 1069–1081, Jun 2022.

[25] Z. Feng, B. Clerckx, and Y. Zhao, "Waveform and beamforming design for intelligent reflecting surface aided wireless power transfer: Single-user and multi-user solutions," *IEEE Transactions on Wireless Communications*, 2022.

[26] Y. Zhao, B. Clerckx, and Z. Feng, "IRS-aided SWIPT: Joint waveform, active and passive beamforming design under nonlinear harvester model," *IEEE Transactions on Communications*, vol. 70, pp. 1345–1359, 2022.

[27] Y. Yang, S. Zhang, and R. Zhang, "IRS-enhanced OFDMA: Joint resource allocation and passive beamforming optimization," *IEEE Wireless Communications Letters*, vol. 9, pp. 760–764, Jun 2020.

[28] Q. Wu, X. Zhou, and R. Schober, "IRS-assisted wireless powered NOMA: Do we really need different phase shifts in DL and UL?" *IEEE Wireless Communications Letters*, vol. 10, pp. 1493–1497, Jul 2021.

[29] Q. Wu, X. Zhou, W. Chen, J. Li, and X. Zhang, "IRS-aided WPCNs: A new optimization framework for dynamic IRS beamforming," *IEEE Transactions on Wireless Communications*, pp. 1–1, Dec 2021.

[30] M. Hua and Q. Wu, "Joint dynamic passive beamforming and resource allocation for IRS-aided full-duplex WPCN," *IEEE Transactions on Wireless Communications*, pp. 1–1, Dec 2021.

[31] W. Tang, J. Y. Dai, M. Chen, X. Li, Q. Cheng, S. Jin, K.-K. Wong, and T. J. Cui, "Programmable metasurface-based RF chain-free 8PSK wireless transmitter," *Electronics Letters*, vol. 55, pp. 417–420, Apr 2019.

[32] J. Y. Dai, W. Tang, L. X. Yang, X. Li, M. Z. Chen, J. C. Ke, Q. Cheng, S. Jin, and T. J. Cui, "Realization of multi-modulation schemes for wireless communication by time-domain digital coding metasurface," *IEEE Transactions on Antennas and Propagation*, vol. 68, pp. 1618–1627, Mar 2020.

[33] R. Karasik, O. Simeone, M. D. Renzo, and S. S. Shitz, "Beyond max-SNR: Joint encoding for reconfigurable intelligent surfaces," vol. 2020-June.   IEEE, Jun 2020, pp. 2965–2970.

[34] R. Liu, H. Li, M. Li, and Q. Liu, "Symbol-level precoding design for intelligent reflecting surface assisted multi-user MIMO systems." IEEE, Oct 2019, pp. 1–6.

[35] A. Bereyhi, V. Jamali, R. R. Muller, A. M. Tulino, G. Fischer, and R. Schober, "A single-RF architecture for multiuser massive MIMO via reflecting surfaces."   IEEE, May 2020, pp. 8688–8692.

[36] X. Xu, Y.-C. Liang, G. Yang, and L. Zhao, "Reconfigurable intelligent surface empowered symbiotic radio over broadcasting signals," vol. 2020-Janua.   IEEE, Dec 2020, pp. 1–6.

[37] Q. Zhang, Y.-C. Liang, and H. V. Poor, "Reconfigurable intelligent surface assisted MIMO symbiotic radio networks," *IEEE Transactions on Communications*, vol. 69, pp. 4832–4846, Jul 2021.

[38] J. Hu, Y. C. Liang, and Y. Pei, "Reconfigurable intelligent surface enhanced multi-user MISO symbiotic radio system," *IEEE Transactions on Communications*, vol. 69, pp. 2359–2371, Apr 2021.

[39] M. Hua, Q. Wu, L. Yang, R. Schober, and H. V. Poor, "A novel wireless communication paradigm for intelligent reflecting surface based symbiotic radio systems," *IEEE Transactions on Signal Processing*, vol. 70, pp. 550–565, Apr 2022.

[40] E. Basar, "Reconfigurable intelligent surface-based index modulation: A new beyond MIMO paradigm for 6G," *IEEE Transactions on Communications*, vol. 68, pp. 3187–3196, May 2020.

[41] T. Ma, Y. Xiao, X. Lei, P. Yang, X. Lei, and O. A. Dobre, "Large intelligent surface assisted wireless communications with spatial modulation and antenna selection," *IEEE Journal on Selected Areas in Communications*, vol. 38, pp. 2562–2574, Nov 2020.

[42] J. Yuan, M. Wen, Q. Li, E. Basar, G. C. Alexandropoulos, and G. Chen, "Receive quadrature reflecting modulation for RIS-empowered wireless communications," *IEEE Transactions on Vehicular Technology*, vol. 70, pp. 5121–5125, May 2021.

[43] S. Hu, C. Liu, Z. Wei, Y. Cai, D. W. K. Ng, and J. Yuan, "Beamforming design for intelligent reflecting surface-enhanced symbiotic radio systems," *arxiv:2110.10316*, Oct 2021.

[44] Y. C. Liang, Q. Zhang, J. Wang, R. Long, H. Zhou, and G. Yang, "Backscatter communication assisted by reconfigurable intelligent surfaces," *Proceedings of the IEEE*, 2022.

[45] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Communications Magazine*, vol. 58, pp. 106–112, 1 2020.

[46] R. Hansen, "Relationships between antennas as scatterers and as radiators," *Proceedings of the IEEE*, vol. 77, pp. 659–662, May 1989.

[47] S. J. Thomas, E. Wheeler, J. Teizer, and M. S. Reynolds, "Quadrature amplitude modulated backscatter in passive and semipassive UHF RFID systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, pp. 1175–1182, Apr 2012.

[48] J. Kim and B. Clerckx, "Wireless information and power transfer for IoT: Pulse position modulation, integrated receiver, and experimental validation," *IEEE Internet of Things Journal*, vol. 9, pp. 12 378–12 394, Jul 2022.

[49] X. He, W. Jiang, M. Cheng, X. Zhou, P. Yang, and B. Kurkoski, "GuardRider: Reliable WiFi backscatter using reed-solomon codes with QoS guarantee."   IEEE, Jun 2020, pp. 1–10.

[50] Y. Huang, A. Alieldin, and C. Song, "Equivalent circuits and analysis of a generalized antenna system," *IEEE Antennas and Propagation Magazine*, vol. 63, pp. 53–62, Apr 2021.

[51] D. Bharadia, K. R. Joshi, M. Kotaru, and S. Katti, "BackFi: High throughput WiFi backscatter," vol. 45.   ACM, Aug 2015, pp. 283–296.

[52] G. Yang, C. K. Ho, and Y. L. Guan, "Multi-antenna wireless energy transfer for backscatter communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 33, pp. 2974–2987, Dec 2015.

[53] H. Guo, Q. Zhang, S. Xiao, and Y.-C. Liang, "Exploiting multiple antennas for cognitive ambient backscatter communication," *IEEE Internet of Things Journal*, vol. 6, pp. 765–775, Feb 2019.

[54] M. Jin, Y. He, C. Jiang, and Y. Liu, "Parallel backscatter: Channel estimation and beyond," *IEEE/ACM Transactions on Networking*, vol. 29, pp. 1128–1140, Jun 2021.

[55] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," *IEEE Wireless Communications Letters*, vol. 9, pp. 518–522, Apr 2020.

[56] C. You, B. Zheng, and R. Zhang, "Intelligent reflecting surface with discrete phase shifts: Channel estimation and passive beamforming."   IEEE, Jun 2020, pp. 1–6.

[57] J. Qian, F. Gao, G. Wang, S. Jin, and H. Zhu, "Semi-coherent detection and performance analysis for ambient backscatter system," *IEEE Transactions on Communications*, vol. 65, pp. 5266–5279, Dec 2017.

[58] T. Nguyen, Y.-J. Chu, and T. Nguyen, "On the capacities of discrete memoryless thresholding channels," vol. 2018-June.   IEEE, Jun 2018, pp. 1–5.

[59] T. Nguyen and T. Nguyen, "Optimal quantizer structure for maximizing mutual information under constraints," *IEEE Transactions on Communications*, vol. 69, pp. 7406–7413, Nov 2021.

[60] M. Rezaeian and A. Grant, "Computation of total capacity for discrete memoryless multiple-access channels," *IEEE Transactions on Information Theory*, vol. 50, pp. 2779–2784, Nov 2004.

[61] J. Buhler and G. Wunder, "A note on capacity computation for the discrete multiple access channel," *IEEE Transactions on Information Theory*, vol. 57, pp. 1906–1910, Apr 2011.

[62] G. J. O. Jameson, "The incomplete gamma functions," *The Mathematical Gazette*, vol. 100, pp. 298–306, Jul 2016.

[63] S. Boyd and L. Vandenberghe, *Convex Optimization*.   Cambridge University Press, Mar 2004.

[64] X. He, K. Cai, W. Song, and Z. Mei, "Dynamic programming for sequential deterministic quantization of discrete memoryless channels," *IEEE Transactions on Communications*, vol. 69, pp. 3638–3651, Jun 2021.

[65] T. Nguyen and T. Nguyen, "On thresholding quantizer design for mutual information maximization: Optimal structures and algorithms," vol. 2020-May.   IEEE, May 2020, pp. 1–5.

[66] J. Qian, A. N. Parks, J. R. Smith, F. Gao, and S. Jin, "IoT communications with M-PSK modulated ambient backscatter: Algorithm, analysis, and implementation," *IEEE Internet of Things Journal*, vol. 6, pp. 844–855, Feb 2019.

[67] S. X. Ng and L. Hanzo, "On the MIMO channel capacity of multidimensional signal sets," *IEEE Transactions on Vehicular Technology*, vol. 55, pp. 528–536, 3 2006.

[68] B. W. Bader and T. G. Kolda, "Tensor toolbox for MATLAB," https://www.tensortoolbox.org/, Sep 2022.

[69] E. Calvo, D. P. Palomar, J. R. Fonollosa, and J. Vidal, "On the computation of the capacity region of the discrete MAC," *IEEE Transactions on Communications*, vol. 58, pp. 3512–3525, Dec 2010.

[70] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, pp. 5394–5409, Nov 2019.