RIScatter: Unifying Backscatter Communications, Symbiotic Radio, and Reconfigurable Intelligent Surface

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

***Abstract*—Backscatter Communications (BackCom) nodes harvest energy from and modulate information over an external carrier by manipulating the magnitude, phase, and/or frequency of the scattered signal. Symbiotic Radio (SR) incorporates a passive scatter node into existing radio networks to create additional propagation path and ride its own information towards the cooperative receiver. Reconfigurable Intelligent Surface (RIS) is a programmable reflector array that adapts the phase shift response to enhance or suppress signal strength in specific directions. In this paper, we show how those three seemingly different technologies can be unified to leverage their benefits simultaneously into a single architecture called RIScatter. RIScatter is a new paradigm for future wireless networks and consists of multiple dispersed or co-located passive scatter nodes, whose reflection states can be adapted to partially engineer the wireless channel and partially modulate information onto the scattered wave. This contrasts with BackCom/SR (resp. RIS) where states are exclusively a function of information symbols (resp. Channel State Information (CSI)). The key principle in RIScatter is to render the probability distribution of reflection states as a joint function of the CSI and input information source. This enables RISscatter to softly bridge, generalize, and outperform BackCom, SR, RIS; boil down to any of those under specific reflection states; or evolve in a mixed form for universal hardware design and heterogeneous traffic control. For a typical setup, we characterize the achievable primary-(total-)backscatter rate region by optimizing input distribution at the nodes, active beamforming at the Access Point (AP), and backscatter detector at the user. Simulation results demonstrate RIScatter nodes can exploit the additional propagation paths to smoothly transition between backscatter modulation and passive beamforming via smart input distribution design, and the proposed receiver significantly reduces the decoding complexity and effectively accommodates the double modulation and signal difference in scatter-based cooperative networks.**

I. Introduction

A. Fundamentals

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 Communications (BackCom) separates 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, while the backscatter reader can be either co-located or separated with the carrier emitter, shown as Monostatic Backscatter Communications (MBC) and Bistatic Backscatter Communications (BBC) in Fig. 1(a) and 1(b), respectively. Its applications such as Radio-Frequency Identification (RFID) [2], [3] and passive sensor network [4], [5] have been 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 an Ambient Backscatter Communications (AmBC) system in Fig. 1(c), where battery-free nodes recycle ambient legacy signals (e.g., radio, television and Wi-Fi) to harvest energy and establish connections in between. It eliminates the need of dedicated power supply, carrier emitter and frequency spectrum, but is subject to the strong direct-link interference. In [7], the authors proposed a cooperative AmBC system where both primary (legacy) and backscatter links are decoded by the same receiver under various detection schemes. The concept of cooperative AmBC was further refined as Symbiotic Radio (SR) in Fig. 1(d) that cognitively incorporates AmBC with existing systems [8]. In a SR system, the active transmitter generates RF wave carrying primary information, the passive node enhances the radio propagation and superimposes its own message on the scattered signal, and the cooperative receiver jointly or sequentially decodes both links. The applications above employ scatter node as information sources, and the active primary link (if exists) is subject to the influence of backscatter randomness. On the other hand, Reconfigurable Intelligent Surface (RIS) is a smart planar reflector that consists of numerous low-power and low-cost elements with adjustable amplitude and phase responses. The reflection pattern is deterministic over time, which can be optimized and coordinated before the transmission. RIS recycles and redistributes surrounding RF waves to customize wireless propagation environment for signal enhancement, interference suppression, scattering enrichment and non-line-of-sight bypassing [9]. A comparison of different scattering applications is summarized in Table I.

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. Related Works

Similar to Cognitive Radio (CR), the coexistence of primary and backscatter links in SR can be classified into commensal, parasitic, and competitive relationships, whose instantaneous rates, power schemes, and outage probabilities were acquired in [10], [11]. To evaluate the performance of cooperative receivers,

TABLE I
COMPARISON OF SCATTERING APPLICATIONS

	MBC	BBC	AmBC	SR	RIS	RIScatter
Coexisting systems	1	1	2 (competitive)	2 (collaborative)	1	2 (collaborative)
Scatterer contribution	Backscatter modulation	Backscatter modulation	Backscatter modulation and primary interference	Backscatter modulation and primary multipath	Passive beamforming	Backscatter modulation and passive beamforming
Cooperative transmissions	—	—	No (individual transmitters)	Active beamforming	—	Active beamforming and smart reflection
Cooperative reception	—	—	No (separated receivers)	Joint detection or SIC	—	Backscatter detection as primary channel training
Primary detection	—	—	Semi-coherent	Coherent or semi-coherent	Coherent	Coherent
Backscatter detection	Coherent	Coherent	Semi-coherent or noncoherent	Coherent	—	Semi-coherent
Reflection state distribution	Equiprobable (line-coded)	Equiprobable (line-coded)	Equiprobable (line-coded)	Equiprobable or Gaussian	Degenerate (CSI-adaptive)	Flexible (CSI- and traffic-adaptive)
Load-switching frequency	Fast or slow	Fast or slow	Fast or slow	Slow	Slow (dynamic) or quasi-static (blockwise)	Fast or slow

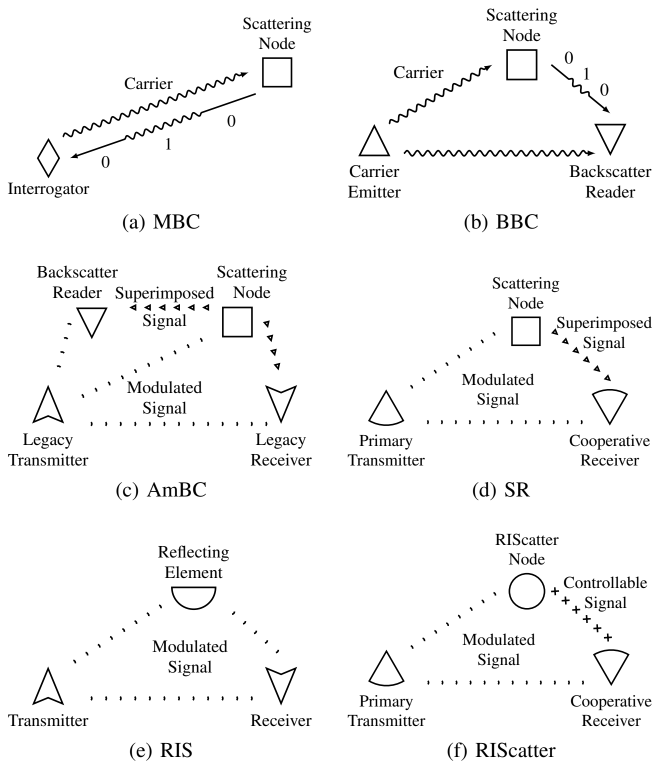


Fig. 1. Illustration of scattering applications.

[7] derived the bit error rates of Maximum-Likelihood (ML) and SIC detectors for flat fading channels, and proposed a low-complexity detector for frequency-selective fading channels. However, it only considered the case where the primary and backscatter symbols of the same period are perfectly aligned in time. Joint ML decoding can achieve the best error performance but comes with prohibitive computational complexity especially for sources with high-order constellation [7], [8], [12]. One special property of scatter-based cooperative networks is the backscatter signal strength is significantly weaker than primary due to the double fading effect. It motivated [7], [8], [10]–[18] to view SR as a multiplicative Non-Orthogonal Multiple Access (NOMA) and perform sequential primary-backscatter decoding based on SIC. During primary decoding, the randomness from backscatter modulation can be modelled as either interference or channel uncertainty, depending on the relationship between the primary sampling rate and backscatter switching speed. If the former is much higher (i.e., commensal SR), the average primary achievable rate under noncoherent detection would asymptotically approach its coherent counterpart [13], and both links may be decoded in an interference-free manner. However, the assumption of very large backscatter-over-primary symbol period ratio may not hold in practice [1], and such a SIC-based sequential decoding requires re-encoding, precoding, and subtraction at each primary symbol block with a time-domain Maximal Ratio Combining (MRC), which can be operation-intensive and CSI-sensitive. Another open issue for BackCom and SR is efficient node multiple access. [16] proposed a NOMA-based SR where the SIC order depends on the backscatter channel strength, and the performance deteriorates fast as the number of nodes increases. Time-Division Multiple Access (TDMA)-based SR was also considered in [17] where each node transmits information during its dedicated slot and harvests energy during others. It enables adaptive transmission time and reflection ratio optimization but requires regular coordination and incurs high coordination cost. [20] controls the load-switching speed to shift the scattered signal to the desired frequency band. This enables backscatter Frequency-Division

[1]For example, Wi-Fi sampling rate is $20\,\mathrm{MHz}$ while RFID load switching speed varies between 100s of kHz to 10s of MHz [19], corresponding to a typical symbol period ratio between 1 and 100.

Multiple Access (FDMA) at the cost of extra bandwidth and higher node power consumption. To reduce coordination burden for passive nodes, [18] proposed a random code-assisted multiple access for SR and evaluated the asymptotic Signal-to-Interference-plus-Noise Ratio (SINR) using random matrix theory. However, this code-domain solution suffers from the near-far problem and imperfect synchronization. On the other hand, conventional RIS design with fixed reflection coefficients during each channel block has been extensively studied in communication, sensing, and power literatures [21]–[26]. Dynamic RIS, which employs independent reflection patterns during 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 resource blocks for Orthogonal Frequency-Division Multiplexing (OFDM) systems, then extended to the downlink power and uplink information phases of Wireless Powered Communication Network (WPCN) [28]–[30]. It creates artificial channel diversity and enables flexible resource allocation, but misses the opportunity to encode its own message. RIS can also be used a transmitter when placed in the near field of a carrier emitter, and prototypes for Phase Shift Keying (PSK) and Quadrature Amplitude Modulation (QAM) have been implemented in [31], [32]. From an information-theoretic perspective, [33] reported using RIS as a passive beamformer to maximize the 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 multiplication coding with SIC decoding (i.e., SIC-based SR) can outperform pure passive beamforming at high SNR. It inspired [34]–[43] to combine passive beamforming and backscatter modulation in the overall reflection pattern. In particular, *symbol level precoding* maps backscatter symbols to optimized RIS coefficient sets [34], [35], *overlay modulation* superposes information-bearing symbols over a common auxiliary matrix [36]–[39], *spatial modulation* switches between reflection coefficient sets that maximize SNR at different receive antennas [40]–[42], and *index modulation* employs dedicated reflection elements for passive beamforming and dedicated information elements for backscatter modulation [43]. Those RIS-empowered BackCom/SR designs involve advanced hardware architecture and high optimization complexity. Besides, all relevant literatures considered either Gaussian codebook [10], [11], [13]–[17], [38] or finite equiprobable inputs [7], [8], [12], [18], [34]–[37], [39]–[43] for backscatter information sources. The former is impractical for passive backscatter devices while the latter does not fully exploit CSI and signal characteristics.

C. Contributions

As presented in Fig. 1(f), we propose RIScatter as a novel scattering protocol that generalizes BackCom, SR, and RIS. The contributions of this paper are summarized as follows.

First, we propose RIScatter nodes that adapt the reflection state distribution of passive scatter devices based on CSI and link weights, in order to flexibly transition between backscatter modulation and passive beamforming. Scattering sources of BackCom/SR and reflecting elements of RIS can be regarded as special cases with uniform and degenerate input distributions, respectively. On the contrary, adaptive RIScatter encoding boils down to deterministic RIS when primary link is prioritized, and achieves higher backscatter rate than conventional line coding when backscatter link is prioritized. Multiple RIScatter nodes can be either co-located to enable joint distribution design for improved total backscatter rate, or dispersed to guarantee uniformly good performance for both links.

Second, we propose a practical receiver that accommodates the double modulation and signal difference in scatter-based cooperative networks. Since the primary and backscatter messages are superimposed by multiplication coding and the backscatter symbol is typically longer than primary, the scattered signals from RIScatter nodes can be treated as multipath components during primary decoding, and the primary message can be viewed as a spreading code during backscatter decoding. Conventional sequential primary-backscatter decoder eliminates primary interference by SIC at each primary block, while our sequential backscatter-primary detector semi-coherently decodes RIScatter nodes from the received energy, re-encodes to recover exact reflection patterns, and models the deterministic multipath in the primary equivalent channel as dynamic passive beamforming at each backscatter block. It enables backscatter modulation and dynamic passive beamforming at much lower operational complexity, and is suitable for scatter nodes with various load switching speeds.

Third, we consider a scenario where multiple RIScatter nodes ride over an active point-to-point Multiple-Input Single-Output (MISO) transmission to perform backscatter modulation and passive beamforming towards a nearby user using shared spectrum, energy, and infrastructures. We provide primary and total backscatter rate analyses and characterize the achievable rate region by optimizing input distribution at RIScatter nodes, active beamforming at the Access Point (AP), and backscatter decision regions at the user. Since the original problem is highly non-convex, we decouple it into individual subproblems and 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 Channel (DMC). This is the first paper to reveal the importance of input distribution and decision region designs in relevant literatures.

Fourth, we provide numerical results to demonstrate the benefits of RIScatter and proposed algorithms. We conclude: 1) adaptive reflection state distribution design can flexibly transition between backscatter modulation and passive beamforming; 2) when primary link is prioritized, input distribution becomes degenerate and RIScatter nodes coincide with discrete RIS; 3) when backscatter link is prioritized, adaptive RIScatter encoding achieves higher backscatter rate than conventional line coding with equiprobable inputs; 4) co-located RIScatter nodes can further leverage total backscatter rate by joint encoding; 5) the proposed receiver provides comparable backscatter detection performance than SIC-based SR while significantly reduces the encoding and precoding (and avoids subtraction) costs;

6) it also supports fast-switching nodes and allows higher backscatter rate per unit time; 7) PGD active beamformer enlarges achievable rate region by boosting the receive SNR and/or widening the energy gap under different reflection states; 8) distribution-adaptive backscatter detectors provide higher total 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 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. $(\cdot)^{*}$, $(\cdot)^{T}$, $(\cdot)^{H}$, $|\cdot|$, 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 terminal solutions, 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 are usually realized by a *variable-load antenna* or *programmable metamaterial* and described by a unified signal model [44]. A typical antenna-based scatterer 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 remainder. 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 reflects residual and avoids leakage, an inner circuit board layer that adjusts the amplitude and phase responses of patches, and an integrated microcontroller/FPGA that coordinates with the network and controls the circuit [45]. Ideally, it reflects the incident waves at the air-metamaterial boundary without receiving them and mainly applies a phase shift on the reflected wave. In practice, both kinds of scatterers should have finite reflection states and non-zero reflection loss, and the scatter-absorb tradeoff depends on impedance matching. For a scatter node with M reflection states, the reflection coefficient at state $m\in\mathcal{M}\triangleq\{1,...,M\}$ is

$$\mathit{\Gamma}_m=\frac{Z_m-Z^*}{Z_m+Z}, \tag{1}$$

where Z_m is the antenna load (resp. metamaterial unit) impedance at state m and Z is the antenna input (resp. medium characteristic) impedance. Both backscatter modulation and passive beamforming are realized by manipulating the reflection coefficient. Specifically, BackCom, AmBC and SR employ scatter nodes as information sources that encode by frequently switching between available states. For M-ary QAM, reflection coefficient $\mathit{\Gamma}_m$ maps to the corresponding *complex constellation point* c_m by

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

where $\alpha\in\mathbb{I}$ is the amplitude scattering ratio at the direction of interest.

A. RIS: Channel Reconfiguration

RIS typically involves a planar antenna array or metasurface of numerous metamaterial units as a passive channel reconfigurator. By *adaptively choosing* the reflection pattern, each reflecting element alters the phase of scattered signal for constructive/destructive superposition at the receiver. For each unit with M candidate states, reflection coefficient $\mathit{\Gamma}_m$ maps to the corresponding *phase shift* θ_m by

$$\mathit{\Gamma}_m=\beta_m\exp(j\theta_m), \tag{3}$$

where $\beta_m\in\mathbb{I}$ is the overall amplitude scattering ratio of state m. Lossless RIS assumes $\beta_m=1$ to maximize the available scattering power.

B. Opportunities

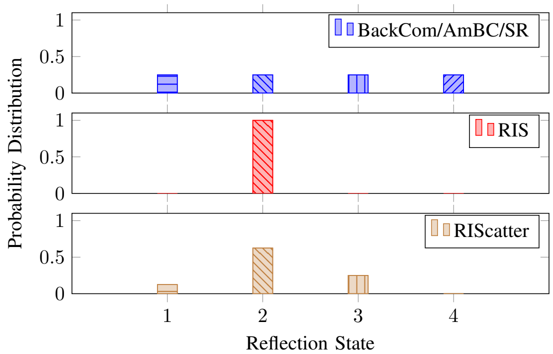
We have seen conventional scattering applications as BackCom, SR and RIS using the reflection pattern (**??**), (**??**) for either backscatter modulation (2) or passive beamforming (3). A natural but critical question is, for a given scatterer with pre-determined reflection pattern set, *can we softly bridge backscatter modulation and passive beamforming by appropriate reflection design?* This question is answered in Section III.

III. RIScatter

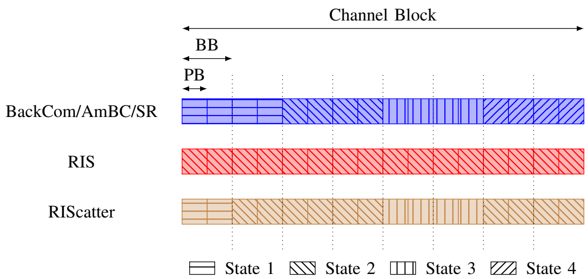
Using shared spectrum, energy and infrastructures, RIScatter is a novel passive communication/assistance protocol that generalizes BackCom, SR and RIS from the perspective of input distribution. In particular, the potentially dispersed passive RIScatter nodes ride over an active legacy transmission in a flexible and mutualistic manner, while the practical RIScatter receiver cooperatively decodes both primary and backscatter links under the benefits of backscatter modulation and passive beamforming.

A. RIScatter Nodes

Based on link priority and CSI, RIScatter nodes adapt the *state input probability distribution* of antenna-based scatterers to unify backscatter modulation and passive beamforming. Conventional line-coded BackCom/SR nodes (e.g., EPC Gen 2 UHF RFID using FM0) and deterministic RIS elements can be viewed as its extreme cases, where the input distribution boils down to equiprobable and degenerate, respectively. As illustrated in Fig. 2, instead of always using fully random or deterministic reflection pattern over time, each RIScatter node can *flexibly adjust* its input distribution to balance information encoding and channel reconfiguration. That is, it semi-randomly select the reflection state per backscatter symbol block under the guidance of input probability $P(\mathit{\Gamma}_m)$ for state m. Such an adaptive backscatter channel coding coincides with RIS when primary link is absolutely prioritized, and outperforms the conventional equiprobable line-coded BackCom/SR when backscatter link is absolutely prioritized.



(a) Input Distribution



(b) Time Structure

Fig. 2. Input probability distribution and time block structure of BackCom, SR, RIS, and RIScatter. T_s and T_c respectively denote the primary and backscatter symbol period. Within channel coherence time, RIScatter nodes semi-randomly select reflection state for each backscatter symbol block under the guidance of input probability distribution.

Besides, the dispersed characteristics of RIScatter nodes (and legacy users) avoids the development optimization and allows uniformly good performance for both links.

As shown in Fig. 3(a), RIScatter nodes can be implemented over readily available passive backscatter devices with energy harvester and information decoder. Its equivalent circuit and scatter model are presented in Fig. 3(b) and 3(c). When activated by primary transmitter, each node harvests a proportion of the impinging wave for its own operation, decode the embedded coordination information for input distribution control, then applies a potentially random phase shift on the reradiated signal due to backscatter modulation. Energy harvesting and information decoding at passive RIScatter nodes can be realized using conventional power-splitting or time-sharing Simultaneous Wireless Information and Power Transfer (SWIPT) protocols [46] or integrated energy-information receiver with pulse position modulated signal [47]. Besides, relevant CSI can be acquired using either traditional sequential methods [48]–[50] or the state-of-the-art parallel approach [51]. In the following context, we assume perfect primary-total-backscatter coordination and perfect channel estimation for all links.

B. RIScatter Receiver

Since each passive RIScatter node directly modulates its own message over the legacy signal, it involves a *double modulation* where the primary and backscatter symbols are superimposed by *multiplication coding* instead of superposition coding. Besides, the backscatter symbol period is typically longer than primary due to the load switching speed constraint. Those facts imply *backscatter detection* under primary uncertainty can be viewed as part of *primary channel training*. Hence, we propose a practical RIScatter receiver that semi-coherently decodes RIScatter nodes from the received energy per backscatter block, re-encodes for their reflection patterns, retrieves primary equivalent channel as if dynamic passive beamforming, then coherently decodes the primary link under the enhanced multipath. As illustrated in Fig. 4, the total received energy per backscatter block is a random variable that follows different distributions conditioned on different input state hypotheses. Compared with conventional joint ML and SIC from stronger primary link to weaker backscatter links, RIScatter receiver not only preserves the benefits of backscatter modulation and passive beamforming, but also enjoys much lower computational and operational complexities.

C. System Model

To demonstrate the advantages of RIScatter, we consider a single-user multi-node MISO RIScatter system as shown in Fig. 5. In the primary active point-to-point system, a Q-antenna AP transmits to a single-antenna user assisted by K nearby single-antenna RIScatter nodes with M reflection states. In the secondary backscatter Multiple Access Channel (MAC) system, the AP serves as the carrier emitter, the distributed RIScatter nodes modulate information over reradiated RF signals, and the user decodes their messages. For simplicity, we consider a quasi-static block fading model where channels remain constant within coherence block and vary independently between consecutive blocks. Due to the physical constraints of load switching speed, we assume the backscatter over primary symbol period ratio is an integer $N \gg 1$. We also omit the signal reflected by two or more times and ignore the propagation time difference of different paths.

Denote the AP-user direct channel as $\boldsymbol{h}_{\mathrm{D}}^H \in \mathbb{C}^{1\times Q}$, the AP-node $k \in \mathcal{K} \triangleq \{1,\ldots,K\}$ forward channel as $\boldsymbol{h}_{\mathrm{F},k}^H \in \mathbb{C}^{1\times Q}$, the node k-user backward channel as $h_{\mathrm{B},k}$, and the cascaded channel via tag k as $\boldsymbol{h}_{\mathrm{C},k}^H \triangleq h_{\mathrm{B},k}\boldsymbol{h}_{\mathrm{F},k}^H \in \mathbb{C}^{1\times Q}$. Let $x_{\mathcal{K}} \triangleq (x_1,\ldots,x_K)$ be the backscatter symbol tuple of all RIScatter nodes. Without loss of generality, we consider one specific backscatter block (i.e., N primary blocks) in the following context. Due to double modulation, the primary equivalent channel is a function of *coded* backscatter symbols[2]

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

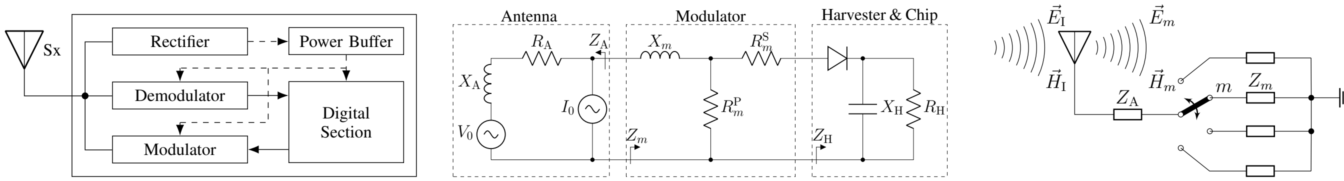
$$= \boldsymbol{h}_{\mathrm{D}}^H + \boldsymbol{x}^H \operatorname{diag}(\boldsymbol{\alpha}) \boldsymbol{H}_{\mathrm{C}}, \tag{4b}$$

where $\alpha_k \in \mathbb{I}$ is the amplitude scattering ratio of RIScatter node k, $x_k \in \mathcal{X} \triangleq \{c_1,\ldots,c_M\}$ is the *coded* backscatter symbol of node k, $\boldsymbol{\alpha} \triangleq [\alpha_1,\ldots,\alpha_K]^T \in \mathbb{I}^K$, $\boldsymbol{x} \triangleq [x_1,\ldots,x_K]^H \in \mathcal{X}^K$, and $\boldsymbol{H}_{\mathrm{C}} \triangleq [\boldsymbol{h}_{\mathrm{C},1},\ldots,\boldsymbol{h}_{\mathrm{C},K}]^H \in \mathbb{C}^{K\times Q}$. The signal received by the user at primary block $n \in \mathcal{N} \triangleq \{1,\ldots,N\}$ is

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

where $s \sim \mathcal{CN}(0,1)$ is the primary symbol, $v \sim \mathcal{CN}(0,\sigma_v^2)$ is the Additive White Gaussian Noise (AWGN), and $\boldsymbol{w} \in \mathbb{C}^Q$

[2]Expression (4a) is often used in BackCom literatures while expression (4b) is commonly observed in RIS literatures.



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

Fig. 3. Block diagram, equivalent circuit, and scatter model of a RIScatter node. The solid and dashed vectors represent signal and energy flows. The scattering 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 $\vec{H}_\mathrm{I}$ [52].

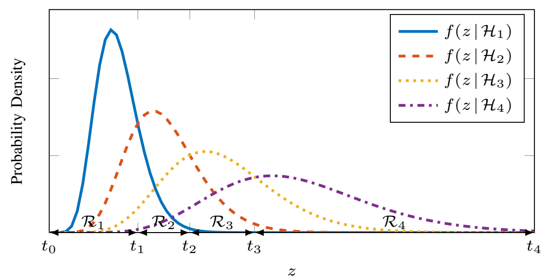


Fig. 4. PDF of total received energy per backscatter block, conditioned on different input state hypotheses. t and $\mathcal{R}$ denote the energy decision thresholds and regions, respectively.

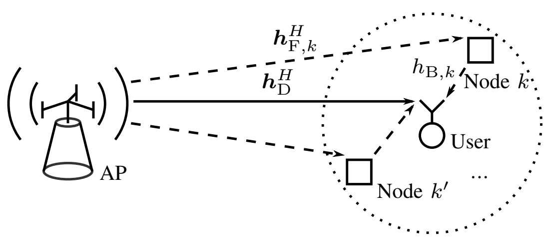


Fig. 5. A single-user multi-node RIScatter system.

is the active beamforming vector subject to average transmit power constraint $\|\boldsymbol{w}\|^2 \leq P$. Denote $m_k \in \mathcal{M} \triangleq \{1,\dots,M\}$ as the reflection state index of node k, and $m_\mathcal{K} \triangleq (m_1,\dots,m_K)$ as the state index tuple of all nodes. Conditioned on backscatter index tuple $m_\mathcal{K}$, the received signal at each primary block follows CSCG distribution $\mathcal{CN}(0,\sigma^2_{m_\mathcal{K}})$, where

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

is the received variance and $x_{m_\mathcal{K}}$ is the backscatter symbol tuple associated with $m_\mathcal{K}$. For node k, let x_{m_k} be the backscatter symbol associated with index m_k.[3] Besides, we define the total received energy per backscatter block as $z = \sum_{n\in\mathcal{N}} \left|y[n]\right|^2$. Since z is the sum of N independent and identically distributed (i.i.d.) exponential variables, its PDF conditioned on $m_\mathcal{K}$ follows Gamma distribution

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

where $\mathcal{H}_{m_\mathcal{K}}$ denotes hypothesis $m_\mathcal{K}$. As illustrated in Fig. 4, the RIScatter receiver divides the received energy space into disjoint decision regions associated with those hypotheses. For example, if the total received energy during a backscatter block falls within region $\mathcal{R}_{m_\mathcal{K}}$, then the detector output would be $x_{m_\mathcal{K}}$.

[3]Please note x_k and $x_\mathcal{K}$ are random variable and tuple, while x_{m_k} and $x_{m_\mathcal{K}}$ are their instances of index m_k and $m_\mathcal{K}$.

Remark 1. *Interestingly, the capacity-achieving decision design for Discrete Memoryless Thresholding Channel (DMTC) remains under-investigated, and some attempts were made for a single source with binary inputs [53], [54]. For non-binary inputs with arbitrary distribution, the optimal decision region for each letter can be non-convex (i.e., with non-adjacent partitions) and the optimal number of thresholds is still unknown [55], [56].*

Like most existing literatures, we limit the scope of this paper to convex decision regions and consider sequential decision thresholds design therein. For the ease of notations, we define a general bijective function from backscatter index tuple $m_\mathcal{K}$ to integer $l \in \mathcal{L} \triangleq \{1,\dots,L \triangleq M^K\}$, where $\{\sigma_l^2\}_{l\in\mathcal{L}}$ are sorted in an ascending order. Both notations are used interchangeably in the following context. As such, the convex decision region of backscatter tuple l can be written as

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

Once the decision threshold vector $\boldsymbol{t} \triangleq [t_0,\dots,t_L]^T \in \mathbb{R}_+^{(L+1)}$ is determined, we can formulate a Discrete Memoryless Thresholding Multiple Access Channel (DMTMAC) with transition probability from input $x_{m_\mathcal{K}}$ to output $\hat{x}_{m'_\mathcal{K}}$

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

then perform backscatter channel coding on top of it.

D. Information Theory

Denote the input probability distribution vector of node k as $\boldsymbol{p}_k \triangleq [P_k(c_1),\dots,P_k(c_M)]^T \in \mathbb{I}^M$, and the probability of state m_k as $P_k(x_{m_k})$. With independent encoding at all nodes, the probability of backscatter symbol tuple $x_{m_\mathcal{K}}$ is

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

Similar to [57], we define the backscatter information function between input symbol tuple instance $x_{m_\mathcal{K}}$ and output symbol tuple variable $\hat{x}_\mathcal{K}$ 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 channel output tuple $\hat{x}_{m'_{\mathcal{K}}}$. We also define the backscatter marginal information of letter x_{m_k} of node k as

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

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

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

Once backscatter messages of all nodes are successfully decoded, we can re-code for backscatter symbol tuple $x_{\mathcal{K}}$, recover their reflection patterns by (2), and retrieve the primary equivalent channel by (4). Moreover, we define the primary information function conditioned on backscatter symbol tuple $x_{m_{\mathcal{K}}}$ as

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

the primary marginal information conditioned on letter x_{m_k} of node k as

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

and the primary ergodic mutual information as

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

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

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

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

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

where $\rho \in \mathbb{I}$ is the relative priority of the primary link.

IV. Rate-Region Characterization

We notice the primary ergodic mutual information (16) relates to input distribution and active beamforming, while the total backscatter mutual information (13) depends explicitly on input distribution and implicitly on active beamforming and decision thresholds via backscatter DMTMAC (9). To characterize the achievable primary-total-backscatter rate region of the proposed RIScatter system, we aim to maximize the weighted sum mutual information with respect to input distribution $\{\boldsymbol{p}_k\}_{k\in\mathcal{K}}$, active beamforming $\boldsymbol{w}$, and decision thresholds $\boldsymbol{t}$ as

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

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

$$\boldsymbol{p}_k \geq \boldsymbol{0}, \quad \forall k, \quad (20c)$$

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

$$t_{l-1} \leq t_l, \quad \forall l, \quad (20e)$$

$$\boldsymbol{t} \geq \boldsymbol{0}. \quad (20f)$$

Problem (20) introduces adaptive backscatter channel coding over conventional BackCom/SR designs. On the other hand, it also generalizes the discrete RIS phase shift selection problem by 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

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

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

which involves the coupled term $\prod_{k\in\mathcal{K}} P_k(x_{m_k})$ and is non-convex when $K > 1$. Following [57], we first recast KKT conditions of problem (21) to their equivalent forms, then propose a numerical method that guarantees those conditions on convergence of sequences.

Remark 2. *As demonstrated in [58], KKT conditions are generally necessary but insufficient for total rate maximization problems. Although KKT solutions to problem* (21) *may end up being saddle points, we will later show by simulation their average performance can be reasonably close to optimal for a moderate K.*

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

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

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

Proof. Please refer to Appendix A. □

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 $\{\boldsymbol{p}_k^{(0)}\}_{k\in\mathcal{K}}$, the KKT input probability of node k at state m_k is given by the converging point of the sequence*

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

where r is the iteration index.

Proof. Please refer to Appendix B. □

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

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

Input: K, N, $\boldsymbol{h}_{\mathrm{D}}^{H}$, $\boldsymbol{H}_{\mathrm{C}}$, $\boldsymbol{\alpha}$, $\mathcal{X}$, σ_v^2, ρ, $\boldsymbol{w}$, $\boldsymbol{t}$, ε
Output: $\{\boldsymbol{p}_k^{\star}\}_{k\in\mathcal{K}}$
1: Set $\boldsymbol{h}_{\mathrm{E}}^{H}(x_{m_\mathcal{K}})$, $\forall m_\mathcal{K}$ by (4)
2: $\quad\sigma_{m_\mathcal{K}}^2$, $\forall m_\mathcal{K}$ by (6)
3: $\quad f(z|\mathcal{H}_{m_\mathcal{K}})$, $\forall m_\mathcal{K}$ by (7)
4: $\quad P(\hat{x}_{m'_\mathcal{K}}|x_{m_\mathcal{K}})$, $\forall m_\mathcal{K}, m'_\mathcal{K}$ by (9)
5: Initialize $r\leftarrow 0$
6: $\quad\quad \boldsymbol{p}_k^{(0)} > \boldsymbol{0}$, $\forall k$
7: Get $P_\mathcal{K}^{(r)}(x_{m_\mathcal{K}})$, $\forall m_\mathcal{K}$ by (10)
8: $\quad I^{(r)}(x_{m_\mathcal{K}})$, $\forall m_\mathcal{K}$ by (11), (14), (17)
9: $\quad I_k^{(r)}(x_{m_k})$, $\forall k, m_k$ by (12), (15), (18)
10: $\quad I^{(r)}(x_\mathcal{K})$ by (13), (16), (19)
11: **Repeat**
12: $\quad$ Update $r\leftarrow r+1$
13: $\quad\quad \boldsymbol{p}_k^{(r)}$, $\forall k$ by (23)
14: $\quad$ Redo step 7–10
15: **Until** $I^{(r)}(x_\mathcal{K}) - I^{(r-1)}(x_\mathcal{K}) \le \varepsilon$

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 [59, Theorem 3]

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

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

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

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

$$g_{m_\mathcal{K}}(t_l) = t_l \exp\Big(-\frac{t_l}{\sigma_{m_\mathcal{K}}^2}\Big) \Bigg(-1 + \sum_{n=1}^{N-1} \frac{\Big(n - \frac{t_l}{\sigma_{m_\mathcal{K}}^2}\Big)\Big(\frac{t_l}{\sigma_{m_\mathcal{K}}^2}\Big)^{n-1}}{n!}\Bigg). \tag{27}$$

On top of (25) and (26), we explicitly express the objective function (24a) and its gradient as (28) and (29) at the end of page 9, respectively. Those allows problem (24) to be solved by the PGD method, where any unregulated beamforming vector $\bar{\boldsymbol{w}}$ can be projected onto the feasible domain of average transmit power constraint (20d) by

$$\boldsymbol{w} = \sqrt{P} \frac{\bar{\boldsymbol{w}}}{\max\big(\sqrt{P}, \|\bar{\boldsymbol{w}}\|\big)}. \tag{30}$$

Algorithm 2: Iterative Active Beamforming Optimization by PGD with BLS

Input: $Q, N, \boldsymbol{h}_{\mathrm{D}}^{H}, \boldsymbol{H}_{\mathrm{C}}, \boldsymbol{\alpha}, \mathcal{X}, P, \sigma_v^2, \rho, \{\boldsymbol{p}_k\}_{k\in\mathcal{K}}, \boldsymbol{t}, \alpha, \beta, \gamma, \varepsilon$
Output: $\boldsymbol{w}^{\star}$
1: Set $\boldsymbol{h}_{\mathrm{E}}^{H}(x_{m_\mathcal{K}})$, $\forall m_\mathcal{K}$ by (4)
2: $\quad P_\mathcal{K}(x_{m_\mathcal{K}})$, $\forall m_\mathcal{K}$ by (10)
3: Initialize $r\leftarrow 0$
4: $\quad\quad \boldsymbol{w}^{(0)}$, $\|\boldsymbol{w}^{(0)}\|^2 \le P$
5: Get $(\sigma_{m_\mathcal{K}}^{(r)})^2$, $\forall m_\mathcal{K}$ by (6)
6: $\quad Q^{(r)}\big(N, \frac{t_{l-1}}{\sigma_{m_\mathcal{K}}^2}, \frac{t_l}{\sigma_{m_\mathcal{K}}^2}\big)$, $\forall m_\mathcal{K}, l$ by (25)
7: $\quad I^{(r)}(x_\mathcal{K})$ by (28)
8: $\quad \nabla_{\boldsymbol{w}^*} Q^{(r)}\big(N, \frac{t_{l-1}}{\sigma_{m_\mathcal{K}}^2}, \frac{t_l}{\sigma_{m_\mathcal{K}}^2}\big)$, $\forall m_\mathcal{K}, l$ by (26)
9: $\quad \nabla_{\boldsymbol{w}^*} I^{(r)}(x_\mathcal{K})$ by (29)
10: **Repeat**
11: $\quad$ Update $r\leftarrow r+1$
12: $\quad\quad \gamma^{(r)} \leftarrow \gamma$
13: $\quad\quad \bar{\boldsymbol{w}}^{(r)} \leftarrow \boldsymbol{w}^{(r-1)} + \gamma \nabla_{\boldsymbol{w}^*} I^{(r-1)}(x_\mathcal{K})$
14: $\quad\quad \boldsymbol{w}^{(r)}$ by (30)
15: $\quad$ Redo step 5–7
16: $\quad$ **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: $\quad\quad$ Set $\gamma^{(r)} \leftarrow \beta\gamma^{(r)}$
18: $\quad\quad$ Redo step 13–15
19: $\quad$ **End While**
20: $\quad$ Redo step 8, 9
21: **Until** $\|\boldsymbol{w}^{(r)} - \boldsymbol{w}^{(r-1)}\| \le \varepsilon$

The PGD active beamforming optimization with step size determined by BLS is summarized in Algorithm 2.

C. Decision Threshold

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

$$\max_{\boldsymbol{t}} \quad I(x_\mathcal{K}) \tag{31a}$$

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

which is still non-convex because variable $\boldsymbol{t}$ appears on the limits of integration (9). Fortunately, we can further simplify problem (31) as a point-to-point rate-optimal quantizer optimization for a discrete-input continuous-output memoryless channel, thanks to Remark 3 and 4.

Remark 3. *Upon successful backscatter decoding, RIScatter receiver can always re-encode node messages to recover their reflection patterns at each primary block. Therefore, backscatter decision design has no impact on the primary achievable rate, and any thresholding scheme maximizing the total backscatter mutual information* (13) *is also optimal for problem* (31).

Remark 4. *In terms of total backscatter rate, the potentially dispersed nodes with given input distribution can be viewed as an equivalent source with augmented alphabet of backscatter symbol tuples* $\{x_{m_\mathcal{K}}\}$. *As such, the DMTMAC* (9) *is essentially a DMTC, and problem* (31) *reduces to the rate-optimal quantization design for a discrete-input continuous-output memoryless channel.*

Next, we constrain the feasible domain of problem (31) 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 the extra analog-to-digital conversion step, we can group adjacent high-resolution energy bins for backscatter decision regions. Thus, problem (31) is recast as

$$\max_{\boldsymbol{t}\in\mathcal{T}^{L+1}} \quad I_{\mathrm{B}}(x_{\mathcal{K}};\hat{x}_{\mathcal{K}}) \tag{32a}$$

$$\text{s.t.} \quad (20\text{e}), \tag{32b}$$

which can be solved by existing rate-optimal sequential quantizer designs for DMC. To obtain global optimal solution, [60] 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 [61] 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 thresholding [62]

$$t_l^{\mathrm{ML}} = N\frac{\sigma_{l-1}^2\sigma_l^2}{\sigma_{l-1}^2-\sigma_l^2}\log\frac{\sigma_{l-1}^2}{\sigma_l^2}, \quad l\in\mathcal{L}\setminus\{L\}, \tag{33}$$

which is generally suboptimal for problem (31) 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 design for the RIScatter system in Fig. 5. 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 $2\,\mathrm{m}$. The AP is with an average transmit power budget $P=36\mathrm{dBm}$, all nodes employs M-QAM with amplitude scattering ratio $\alpha=0.5$, and the user is with average noise power $\sigma_v^2=-40\,\mathrm{dBm}$. For all channels involved, we consider a distance-dependent path loss model

$$L(d) = L_0\left(\frac{d_0}{d}\right)^{\gamma}, \tag{34}$$

together with a Rician fading model

$$\boldsymbol{H} = \sqrt{\frac{\kappa}{1+\kappa}}\bar{\boldsymbol{H}} + \sqrt{\frac{1}{1+\kappa}}\tilde{\boldsymbol{H}}, \tag{35}$$

where d is the transmission distance, $L_0=-30\mathrm{dB}$ is the reference path loss at distance $d_0=1\mathrm{m}$, κ is the Rician K-factor, $\bar{\boldsymbol{H}}$ is the deterministic line-of-sight component with entries of unit magnitude, and $\tilde{\boldsymbol{H}}$ is the Rayleigh fading component with entries in standard i.i.d. CSCG distribution. We choose $\gamma_{\mathrm{D}}=2.6$, $\gamma_{\mathrm{F}}=2.4$, $\gamma_{\mathrm{B}}=2$, and $\kappa_{\mathrm{D}}=\kappa_{\mathrm{F}}=\kappa_{\mathrm{B}}=5$ for direct, forward and backward links. The finite 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 level $1-\varepsilon$, where $b=9$ and $\varepsilon=1\times10^{-3}$. All average achievable rate points/regions are evaluated over 1000 realizations, and the parameters remain fixed unless otherwise specified.

A. Evaluation of Proposed Algorithms

1) Initialization: To characterize the achievable rate region, we progressively obtain all boundary points by successively increasing the primary priority ρ and solving problem (20). At $\rho=0$ where backscatter performance is absolutely prioritized, we initialize Algorithm 1 and 2 by uniform input distribution and Maximum Ratio Transmission (MRT) towards sum cascaded channel $\sum_k \boldsymbol{h}_{\mathrm{C},k}^H$, 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 (on the first call within BCD) and BCD algorithms in each iteration. For $K=8$ and $M=2$, Algorithm 1 typically takes around 100 fast iterations over closed-form expression (23) to converge to the KKT input distribution. For $Q=4$, around 10 iterations are required for Algorithm 2 to converge, where the gradient needs to be computed by (29) and the step size needs to be refined by BLS. In comparison, the BCD algorithm typically requires at most 5 iterations to converge. At the following rate points (not presented), the convergence of all three algorithms become much faster thanks to the progressive initialization. Overall, we conclude the proposed algorithms are able to converge within moderate iterations and provide significant rate benefits for the RIScatter system.

$$I(x_{\mathcal{K}}) = \sum_{m_{\mathcal{K}}} P_{\mathcal{K}}(x_{m_{\mathcal{K}}})\left(\rho\log\Big(1+\frac{|\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2}{\sigma_v^2}\Big) + (1-\rho)\sum_l Q\Big(N,\frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}}^2},\frac{t_l}{\sigma_{m_{\mathcal{K}}}^2}\Big)\log\frac{Q\Big(N,\frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}}^2},\frac{t_l}{\sigma_{m_{\mathcal{K}}}^2}\Big)}{\sum_{m'_{\mathcal{K}}} P_{\mathcal{K}}(x_{m'_{\mathcal{K}}})Q\Big(N,\frac{t_{l-1}}{\sigma_{m'_{\mathcal{K}}}^2},\frac{t_l}{\sigma_{m'_{\mathcal{K}}}^2}\Big)}\right) \tag{28}$$

$$\begin{aligned}\nabla_{\boldsymbol{w}^*} I(x_{\mathcal{K}}) = \sum_{m_{\mathcal{K}}} P_{\mathcal{K}}(x_{m_{\mathcal{K}}})\Bigg(&\rho\frac{\boldsymbol{h}_{\mathrm{E}}(x_{m_{\mathcal{K}}})\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}}{\sigma_{m_{\mathcal{K}}}^2} + (1-\rho)\sum_l\Bigg(\log\frac{Q\Big(N,\frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}}^2},\frac{t_l}{\sigma_{m_{\mathcal{K}}}^2}\Big)}{\sum_{m'_{\mathcal{K}}} P_{\mathcal{K}}(x_{m'_{\mathcal{K}}})Q\Big(N,\frac{t_{l-1}}{\sigma_{m'_{\mathcal{K}}}^2},\frac{t_l}{\sigma_{m'_{\mathcal{K}}}^2}\Big)}+1\Bigg)\\ &\times\nabla_{\boldsymbol{w}^*} Q\Big(N,\frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}}^2},\frac{t_l}{\sigma_{m_{\mathcal{K}}}^2}\Big) - \frac{Q\Big(N,\frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}}^2},\frac{t_l}{\sigma_{m_{\mathcal{K}}}^2}\Big)\sum_{m'_{\mathcal{K}}} P_{\mathcal{K}}(x_{m'_{\mathcal{K}}})\nabla_{\boldsymbol{w}^*} Q\Big(N,\frac{t_{l-1}}{\sigma_{m'_{\mathcal{K}}}^2},\frac{t_l}{\sigma_{m'_{\mathcal{K}}}^2}\Big)}{\sum_{m'_{\mathcal{K}}} P_{\mathcal{K}}(x_{m'_{\mathcal{K}}})Q\Big(N,\frac{t_{l-1}}{\sigma_{m'_{\mathcal{K}}}^2},\frac{t_l}{\sigma_{m'_{\mathcal{K}}}^2}\Big)}\Bigg)\end{aligned} \tag{29}$$

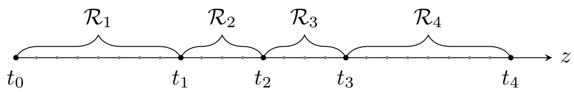


Fig. 6. The decision thresholds are selected from fine-grained discrete energy levels instead of continuous space, and each decision region consists of at least one neighbor energy bins.

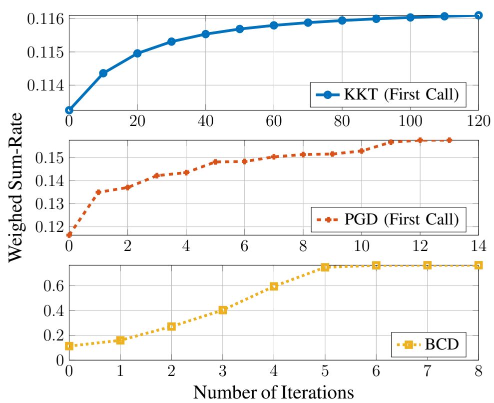


Fig. 7. Convergence behavior of the proposed algorithms for $\rho = 0$.

B. Comparison of Scattering Applications

On top of the setup in Fig. 5, we compare the achievable rate pairs of RIScatter and these scattering applications:

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

$$y^{\text{BBC}}[n] = \Big(\boldsymbol{h}_{\text{D}}^H + \sum_k \alpha_k \boldsymbol{h}_{\text{C},k}^H x_k\Big)\boldsymbol{w} + v[n], \tag{36}$$

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

$$f^{\text{BBC}}(z|\mathcal{H}_{m_{\mathcal{K}}}) = \frac{(z - \mu_{m_{\mathcal{K}}}^{\text{BBC}})^{N-1} \exp\big(-(z - \mu_{m_{\mathcal{K}}}^{\text{BBC}})/\sigma_v^2\big)}{\sigma_v^{2N}(N-1)!}, \tag{37}$$

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

$$t_l^{\text{BBC}} = \frac{\mu_{l-1}^{\text{BBC}} \exp\big((\mu_{l-1}^{\text{BBC}} - \mu_l^{\text{BBC}})/\sigma_v^2(N-1)\big) - \mu_l^{\text{BBC}}}{\exp\big((\mu_{l-1}^{\text{BBC}} - \mu_l^{\text{BBC}})/\sigma_v^2(N-1)\big) - 1}. \tag{38}$$

- *AmBC:* The user decodes both links independently and semi-coherently by treating the other as interference or uncertainty. Hence, the primary achievable rate can be approximated by[4]

$$I_{\text{P}}^{\text{AmBC}}(s;y) \approx \log\Big(1 + \frac{|\boldsymbol{h}_{\text{D}}^H \boldsymbol{w}|^2}{\sum_k |\alpha_k \boldsymbol{h}_{\text{C},k}^H \boldsymbol{w}|^2 + \sigma_v^2}\Big), \tag{39}$$

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

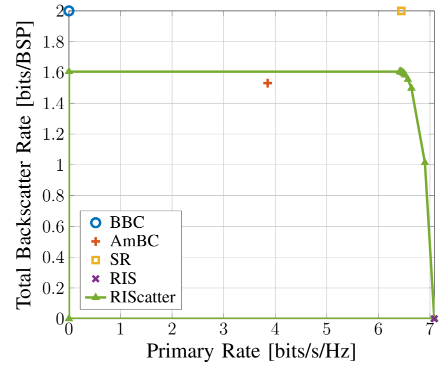


Fig. 8. Typical achievable rate points of scattering applications for $Q = 1$, $K = 1$, $M = 4$ and $N = 1000$.

- *SR:* For a sufficiently large N, the primary ergodic rate under semi-coherent detection asymptotically approaches (16) with uniform input distribution [13]. When $s[n]$ is successfully decoded and the direct interference $\boldsymbol{h}_{\text{D}}^H \boldsymbol{w} s[n]$ is perfectly cancelled, the intermediate signal for backscatter detection is

$$\hat{y}^{\text{SR}}[n] = \sum_k \alpha_k \boldsymbol{h}_{\text{C},k}^H x_k \boldsymbol{w} s[n] + v[n], \tag{40}$$

which only involves noise uncertainty under hypothesis $\mathcal{H}_{m_{\mathcal{K}}}$. By MRC over N primary blocks, the total achievable for nodes with equiprobable inputs is [63]

$$I_{\text{B}}(x_{\mathcal{K}}; \hat{y}_{\text{SR}}) = K \log M - \frac{\epsilon}{M^K}, \tag{41}$$

where $\epsilon \triangleq \sum_{m_{\mathcal{K}}} \mathbb{E}_{\hat{v}} \log \sum_{m'_{\mathcal{K}}} \exp(-|x_{m_{\mathcal{K}}} - x_{m'_{\mathcal{K}}} + \hat{v}|^2/2\sigma^2)$ and $\hat{v} \sim \mathcal{CN}(0, \sigma_v^2/N)$. For a sufficiently large N, ϵ is negligible and the total backscatter rate approaches $K \log M$.
- *RIS:* Since the backscatter symbol tuple $x_{\mathcal{K}}$ becomes deterministic, the total backscatter rate is zero and the primary achievable rate boils down to a special case of (16)

$$I_{\text{P}}^{\text{RIS}}(s;y|x_{\mathcal{K}}) = I_{\text{P}}(s;y|x_{m_{\mathcal{K}}^\star}) = \log\Big(1 + \frac{|\boldsymbol{h}_{\text{E}}^H(x_{m_{\mathcal{K}}^\star})\boldsymbol{w}|^2}{\sigma_v^2}\Big), \tag{42}$$

where $m_{\mathcal{K}}^\star = \arg\max_{m_{\mathcal{K}}} I_{\text{P}}(s;y|x_{m_{\mathcal{K}}})$.

Fig. 8 presents typical achieve rate points of different scattering applications for $Q = 1$, $K = 1$, $M = 4$ and $N = 1000$. First, we observe both BBC and SR can achieve the highest backscatter rate with coherent detection over the receive energy and intermediate signal, respectively. This is because repetition coding with MRC over N primary blocks effectively increases the backscatter SNR by N times, and the information loss due to the signal-to-energy conversion on top may be insignificant. However, the asymptotic primary and backscatter rate analysis only provides upper bounds for SR. Using a very large N not only reduces the system throughput[5] but also requires the operation-intensive SIC over each primary block. When the duration of primary and backscatter symbol are comparable, the randomly scattered signals should be modelled as interference

[4] To provide a preliminary benchmark, we approximate the unknown signal components from finite-input backscatter sources by independent interference from Gaussian sources.

[5] For example, Wi-Fi sampling rate is 20 MHz while RFID load switching speed varies between 100s of kHz to 10s of MHz [19], corresponding to a typical symbol period ratio N between 1 and 100.

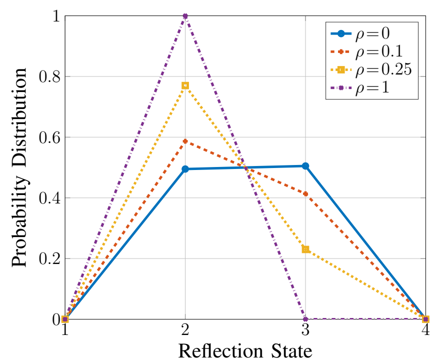


Fig. 9. Typical reflection state distribution under different weights for $Q=1$, $K=1$, $M=4$ and $N=20$.

to primary link rather than quasi-static multipath components, and the achievable rate point of SR will move towards the origin. On the other hand, RIScatter provides a lower backscatter rate than BBC and SR due to semi-coherent detection, but preserves the primary rate benefit for arbitrary N.

C. Input Distribution under Different Weights

The objective of this study is to demonstrate each RIScatter node can adjust its input distribution based on CSI and link priority to balance backscatter modulation and passive beamforming. For this aim, we evaluate the KKT input distribution by Algorithm 1 for a single node under different weights in Fig. 9. We observe that as ρ increases from 0 to 1, the KKT input distribution gradually progresses from slightly skewed to completely biased. For $\rho = 0$ where backscatter performance is absolutely prioritized, the CSI-adaptive channel coding at the RIScatter node provides higher backscatter rate than conventional BackCom and SR nodes with equiprobable inputs, and some states may end up unused since they cannot produce enough energy difference over the existing alphabet. On the other hand, for $\rho=1$ where primary performance is absolutely prioritized, the input probability is 1 at the state that maximizes the primary equivalent channel strength, and 0 at the others. In this case, the reflection pattern over time is fully deterministic, and the RIScatter node indeed boils down to a RIS element with discrete phase shifts. Those observations meet our s that RIScatter nodes include scattering sources of BackCom/SR and reflecting elements of RIS as special cases, and generalize backscatter modulation and passive beamforming from the perspective of input distribution design. For the BCD algorithm, we choose the KKT input distribution by Algorithm 1, the PGD active beamforming by Algorithm 2, and the SMAWK decision threshold by [60] as reference.

D. Rate Region by Different Schemes

The average achievable primary-total-backscatter rate regions of various input distribution, active beamforming, and decision threshold schemes are explored in this study.

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

- *Exhaustion:* Exhaustive grid search over the M-dimensional probability simplex with resolution $\Delta p=1\times 10^{-3}$;
- *KKT:* Numerical KKT solution evaluation by Algorithm 1;
- *Cooperation:* Numerical symbol tuple distribution (i.e., K-dimensional joint probability array) evaluation by the Blahut-Arimoto algorithm [64], [65].

Since *Cooperation* involves full transmit cooperation of all nodes, to support independent backscatter encoding, we also propose the following methods to recover individual input distributions from the joint probability array:

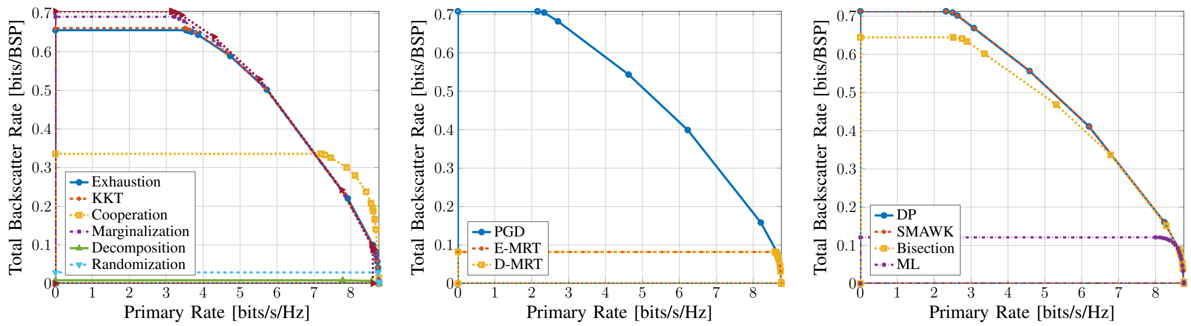
- *Marginalization:* Marginal probability distributions;
- *Decomposition:* Normalized rank-1 Canonical Polyadic (CP) decomposition tensors;
- *Randomization:* Random search guided with correlation matrix [66].

Fig. 10(a) shows the achievable rate regions for those input distribution schemes. We observe that *Cooperation* provides the outer bound for all achievable rate regions, which suggests transmit cooperation between RIScatter nodes can be generally helpful. However, it requires joint encoding with arbitrarily dependent codewords, and the extra coordination burden may be unaffordable for a large K. We also notice the average performance of *KKT* input distribution design by Algorithm 1 is extremely close to that of *Exhaustion*, and the loss is indistinguishable for $K=2$. This verifies our claim in Remark 2 that the average performance of KKT solutions can be reasonably good for a moderate K. Despite *Randomization* yields a similar result, the computational complexity is much higher (loops over $K+1$ linear programming problems) and the performance is backed by numerous randomly generated instances. In contrast, the low-complexity *Marginalization* returns a comparable result and the rate loss is insignificant when K is small. Finally, *Decomposition* suffers from much lower average total backscatter rate, as the rank-1 CP approximation can be inaccurate for some channel-drop realizations. Those observations demonstrate the high performance and low complexity of the proposed KKT input distribution design, and reveal the potential of joint encoding at RIScatter nodes.

2) Active Beamforming: We compare the following active beamforming schemes for problem (24):

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

Fig. 10(b) plots the achievable rate regions for those active beamforming schemes. In the low-ρ regime, we observe the proposed *PGD* design by Algorithm 2 outperforms *E-MRT* and *D-MRT* in terms of backscatter rate. This is because backscatter decoding relies on the *difference* of received energy expectations under different backscatter symbol tuples. Such an energy diversity is not fully exploited by simply maximizing the direct/ergodic equivalent SNR, but can be properly controlled by *PGD* for enhanced detection performance. As ρ further increases, the primary equivalent SNR outweighs the backscatter energy difference in the weighted sum-rate expression (28), and the input distribution is gradually biased



(a) Input Distribution (b) Active Beamforming (c) Decision Threshold

Fig. 10. Average primary-total-backscatter rate regions for different input distribution, active beamforming, and decision threshold schemes. "BSP" means backscatter symbol period.

to the state that roughly align direct and scattered components. Therefore, the advantages of distribution-adaptive schemes *PGD* and *E-MRT* become marginal, the ergodic equivalent and direct channels are closer in direction, and the bounds of all schemes approach each other. For the extreme case $\rho = 1$, both distribution-adaptive schemes boil down to MRT towards the deterministic primary equivalent channel, as considered in RIS literatures. Those observations prove the proposed *PGD* active beamforming design can balance the primary equivalent SNR and backscatter energy diversity for enlarged rate regions.

3) Decision Threshold: We compare the following decision thresholds for problem (32):

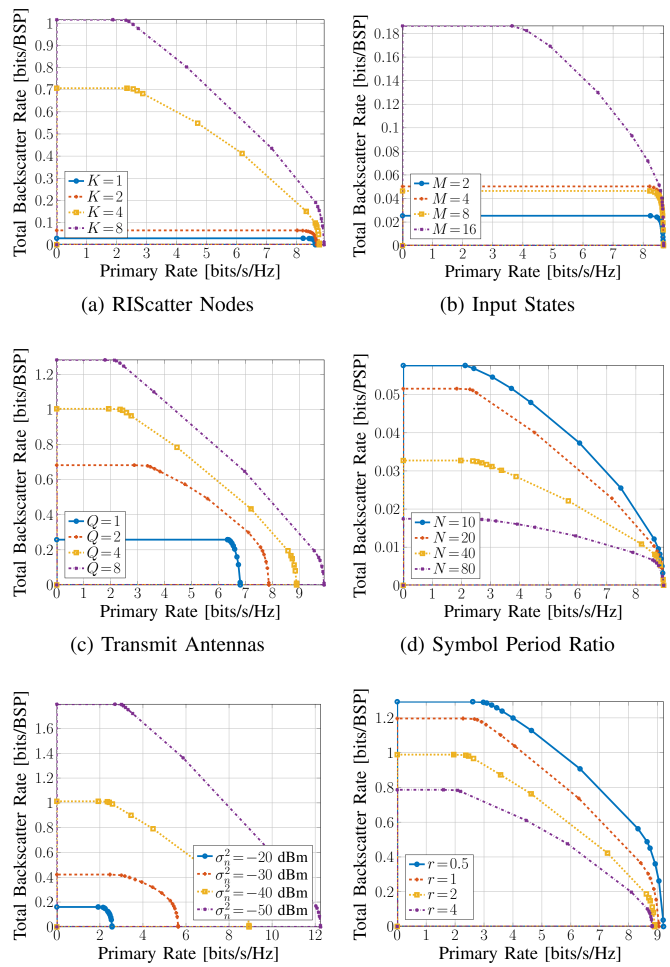
- *DP:* Benchmark DP method for sequential quantizer [60];
- *SMAWK: DP* accelerated by the SMAWK algorithm [60];
- *Bisection:* The traverse-then-bisect algorithm [61];
- *ML:* Maximum likelihood detector (33) [62].

Fig. 10(c) presents the achievable rate region for those decision threshold schemes. We observe that the distribution-adaptive schemes *DP*, *SMAWK* and *Bisection* achieve higher total backscatter rate than the non-adaptive *ML*. This is because the total backscatter mutual information (13) is a function of both input distribution and decision regions, and the rate-optimal threshold design depends heavily on input distribution. For example, the backscatter symbol tuples with zero input probability should never be detected and use empty decision regions, in order to increase the success detection chance of other tuples. Those observations highlight the importance of joint input distribution and decision threshold design.

E. Rate Region under Different Configurations

In this study, we investigate the impact of system configuration on the average achievable rate regions.

1) Number of Nodes: Fig. 11(a) reveals how the number of two-state RIScatter nodes K influence the primary-backscatter tradeoff. We observe the total backscatter rate roughly scales proportionally with the number of nodes, and the decrease of individual rate is unobvious. This demonstrates the proposed low-complexity semi-coherent energy detector performs reasonably well for a moderate K. On the other hand, we observe adding RIScatter nodes can also boost the achievable primary rate. The reason is the nodes can exploit the



(a) RIScatter Nodes (b) Input States

(c) Transmit Antennas (d) Symbol Period Ratio

(e) Average Noise Power (f) Disk Radius

Fig. 11. Average primary-total-backscatter rate regions for different system configurations. "PSP" means primary symbol period.

additional propagation path for a constructive primary signal superposition that enjoys a "squared law" similar to RIS.

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 is roughly determined, using denser constellation points may not create enough phase resolution and energy diversity for primary and backscatter links. Another possible reason is the amplitude normalization step in the reflection-constellation mapping (2) grants less scattered (and more harvestable) power for the inner points, which are thus less frequently employed. It motivates the use of numerous elementary RIScatter nodes, instead of high-order transponders or high-resolution metasurfaces.

3) Number of Transmit Antennas: Fig. 11(c) illustrates the impact of transmit antennas Q on the average performance. We observe larger Q produces greater achievable rate regions. Thanks to the PGD active beamforming, the expectations of the received power under different backscatter symbol tuples can be flexibly controlled to satisfy the link priority requirements. For example, beamforming towards the equivalent channel can produce higher primary SNR, while beamforming towards the cascaded channels can enhance the relative energy difference for better backscatter detection. Those observations emphasize the importance of multi-antenna RIScatter systems and demonstrate the advantages of the proposed PGD design.

4) Symbol Period Ratio: Fig. 11(d) presents how symbol period ratio N affects the achievable rate region. We notice using a small N can effectively boost the total backscatter rate per primary symbol, and the conventional SR literatures assuming sufficiently large N is generally inefficient. However, it not only requires a frequent change of reflection states that consumes more power at the passive nodes, but also involves more detection and re-encoding operations at the user to maintain the primary rate. When N becomes sufficiently large, the total backscatter rate approaches 0, and RIScatter nodes boil down to RIS elements with fixed reflection pattern during whole channel block. Therefore, we conclude N should be properly tweaked based on the rate requirements and physical properties of the passive nodes.

5) Average Noise Power: Fig. 11(e) depicts the impact of average noise power σ_v^2 on average rate regions. Despite the proposed low-complexity energy detection is suitable for a wide range of noise levels, there exists a backscatter performance upper bound due to its semi-coherent property. When σ_v^2 relatively high, we can choose a longer backscatter symbol period (i.e., larger N) to maintain the backscatter SNR for better detection performance.

6) Coverage Disk Radius: Fig. 11(f) shows the relationship between RIScatter cover disk radius and achievable rate regions. We observe both primary and backscatter performance are enhanced when nodes are dropped closer to the user. This is because the product path loss (a.k.a. double fading) model for finite-size scatterers is less severe for such near-far setups. Thanks to the dispersed characteristics of RIScatter nodes and legacy users, each node can simply scatter and be decoded by the closest user, thus avoids the development optimization and allows uniformly good performance for both links.

VI. Conclusion

This paper introduced RIScatter as a low-power transmit-assist protocol that generalizes BackCom, SR and RIS by smart input distribution and practical receiver design. Starting from scattering principles, we showed how RIScatter nodes include scattering source of BackCom/SR and reflecting 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 low-complexity RIScatter receiver that preserves the benefits of backscatter modulation and passive beamforming. 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 demonstrated the effectiveness of the proposed algorithms, but also emphasized the importance of adaptive input distribution and cooperative decoding on both primary and backscatter subsystems.

One possible direction is to consider backscatter detection over the received signal domain rather than energy domain. Learning-based classification approaches can be prominsing in such cases. Another interesting question is how to design RIScatter node and receiver in a multi-user system. If one node can be decoded by multiple users, its input distribution may be further adjusted to mimic the multi-beam gain of dynamic passive beamforming [67].

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 Lagrangian function of problem (21) is

$$-I(x_{\mathcal{K}})+\sum_k \nu_k\left(\sum_{m_k} P_k(x_{m_k})-1\right)-\sum_k\sum_{m_k}\lambda_{k,m_k}P_k(x_{m_k}) \tag{43}$$

and the KKT conditions are, $\forall k,m_k$,

$$-\nabla_{P_k^\star(x_{m_k})} I^\star(x_{\mathcal{K}})+\nu_k^\star-\lambda_{k,m_k}^\star=0, \tag{44a}$$

$$\lambda_{k,m_k}^\star=0, \quad P_k^\star(x_{m_k})>0, \tag{44b}$$

$$\lambda_{k,m_k}^\star\geq 0, \quad P_k^\star(x_{m_k})=0, \tag{44c}$$

where directional derivative is explicitly written as

$$\nabla_{P_k^\star(x_{m_k})} I^\star(x_{\mathcal{K}})=I_k^\star(x_{m_k})-(1-\rho). \tag{45}$$

Combining (44) and (45), we have

$$I_k^\star(x_{m_k})=\nu_k^\star+(1-\rho), \quad P_k^\star(x_{m_k})>0, \tag{46a}$$
$$I_k^\star(x_{m_k})\leq\nu_k^\star+(1-\rho), \quad P_k^\star(x_{m_k})=0, \tag{46b}$$

such that

$$\sum_{m_k} P_k^\star(x_{m_k}) I_k^\star(x_{m_k})=\nu_k^\star+(1-\rho). \tag{47}$$

On the other hand, by definition (18) we have

$$\sum_{m_k} P_k^\star(x_{m_k}) I_k^\star(x_{m_k})=I^\star(x_{\mathcal{K}}), \tag{48}$$

where the right-hand side is irrelevant to k. (46), (47), and (48) together 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 $J\big(P_{\mathcal{K}}(x_{m_{\mathcal{K}}}), P'_{\mathcal{K}}(x_{m_{\mathcal{K}}})\big)$ as (49) at the end of page 14. 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 $\nabla_{P^\star_k(x_{m_k})} J\big(P_{\mathcal{K}}(x_{m_{\mathcal{K}}}),P'_{\mathcal{K}}(x_{m_{\mathcal{K}}})\big)=0$, we have

$$S'_k(x_{m_k}) - S'_k(x_{i_k}) + (1-\rho)\log\frac{P_k(x_{i_k})}{P^\star_k(x_{m_k})} = 0, \tag{50}$$

where $i_k \neq m_k$ is the reference state and

$$S'_k(x_{m_k}) \triangleq I'_k(x_{m_k}) + (1-\rho)\sum_{m_{\mathcal{K}\setminus\{k\}}} P_{\mathcal{K}\setminus\{k\}}(x_{m_{\mathcal{K}\setminus\{k\}}}) \times \sum_{m'_{\mathcal{K}}} P(\hat{x}_{m'_{\mathcal{K}}}|x_{m_{\mathcal{K}}}) \log P'_{\mathcal{K}}(x_{m_{\mathcal{K}}}). \tag{51}$$

Evidently, $\forall m_k \neq i_k$, (50) boils down to

$$P^\star_k(x_{m_k}) = \frac{P'_k(x_{m_k})\exp\Big(\frac{\rho}{1-\rho} I'_k(x_{m_k})\Big)}{\sum_{m'_k} P'_k(x_{m'_k})\exp\Big(\frac{\rho}{1-\rho} I'_k(x_{m'_k})\Big)}. \tag{52}$$

Since $P_k(x_{i_k}) = 1 - \sum_{m_k\neq i_k} P^\star_k(x_{m_k})$ has exactly the same form as (52), the choice of reference state i_k does not matter and (52) 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 (52) ensures

$$J\big(P_{\mathcal{K}}(x_{m_{\mathcal{K}}}),P'_{\mathcal{K}}(x_{m_{\mathcal{K}}})\big) \geq I'(x_{\mathcal{K}}). \tag{53}$$

On the other hand, we also have

$$\begin{aligned}
\varDelta &\triangleq I(x_{\mathcal{K}}) - J\big(P_{\mathcal{K}}(x_{m_{\mathcal{K}}}),P'_{\mathcal{K}}(x_{m_{\mathcal{K}}})\big) && \text{(54a)}\\
&= (1-\rho)\sum_{m_k}\frac{P'_k(x_{m_k})f'_k(x_{m_k})}{\sum_{m'_k}P'_k(x_{m'_k})f'_k(x_{m'_k})}\sum_{m''_{\mathcal{K}}}P(\hat{x}_{m''_{\mathcal{K}}}|x_{m_k}) \\
&\quad \times \log\frac{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_{\mathcal{K}}}|x_{m'_k})f'_k(x_{m_k})}{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_{\mathcal{K}}}|x_{m'_k})f'_k(x_{m'_k})} && \text{(54b)}\\
&\geq (1-\rho)\sum_{m_k}\frac{P'_k(x_{m_k})f'_k(x_{m_k})}{\sum_{m'_k}P'_k(x_{m'_k})f'_k(x_{m'_k})}\sum_{m''_{\mathcal{K}}}P(\hat{x}_{m''_{\mathcal{K}}}|x_{m_k}) \\
&\quad \times \left(1 - \frac{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_{\mathcal{K}}}|x_{m'_k})f'_k(x_{m'_k})}{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_{\mathcal{K}}}|x_{m'_k})f'_k(x_{m_k})}\right) && \text{(54c)}\\
&= 0, && \text{(54d)}
\end{aligned}$$

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 (52) converges. (53) and (54) together imply $I(x_{\mathcal{K}}) \geq I'(x_{\mathcal{K}})$. 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 $P^\star_k(x_{m_k})$, fulfills KKT conditions (22). To see this, let

$$D^{(r)}_k(x_{m_k}) \triangleq \frac{P^{(r+1)}_k(x_{m_k})}{P^{(r)}_k(x_{m_k})} = \frac{f^{(r)}_k(x_{m_k})}{\sum_{m'_k}P^{(r)}_k(x_{m'_k})f^{(r)}_k(x_{m'_k})}. \tag{55}$$

As sequence (23) is convergent, any state with $P^\star_k(x_{m_k}) > 0$ need to satisfy $D^\star_k(x_{m_k}) \triangleq \lim_{r\to\infty} D^{(r)}_k(x_{m_k}) = 1$, namely

$$I^\star_k(x_{m_k}) = \frac{1-\rho}{\rho}\log\sum_{m'_k}P^\star_k(x_{m'_k})f^\star_k(x_{m'_k}), \tag{56}$$

which is reminiscent of (46a) and (22a). That is, given $P^{(0)}_k(x_{m_k}) > 0$, any converging point with $P^\star_k(x_{m_k}) > 0$ must satisfy (22a). On the other hand, we assume $P^\star_k(x_{m_k})$ does not satisfy (22b), such that for any state with $P^\star_k(x_{m_k}) = 0$,

$$I^\star_k(x_{m_k}) > I^\star(x_{\mathcal{K}}) = \sum_{m'_k}P^\star_k(x_{m'_k})I^\star_k(x_{m'_k}), \tag{57}$$

where the equality inherits from (19). Since the exponential function is monotonically increasing, we have $f^\star_k(x_{m_k}) > \sum_{m'_k}P^\star_k(x_{m'_k})f^\star_k(x_{m'_k})$ and $D^\star_k(x_{m_k}) > 1$. Considering $P^{(0)}_k(x_{m_k}) > 0$ and $P^\star_k(x_{m_k}) = 0$, it contradicts with

$$P^{(r)}_k(x_{m_k}) = P^{(0)}_k(x_{m_k})\prod_{n=1}^{r}D^{(n)}_k(x_{m_k}). \tag{58}$$

That is, given $P^{(0)}_k(x_{m_k}) > 0$, any converging point with $P^\star_k(x_{m_k}) = 0$ must satisfy (22b). The proof is thus 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, 3 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6685977/

[2] D. Dobkin, *The RF in RFID: Passive UHF RFID in Practice*. Newnes, 11 2012. [Online]. Available: https://www.elsevier.com/books/the-rf-in-rfid/dobkin/978-0-12-394583-9

[3] J. Landt, "The history of rfid," *IEEE Potentials*, vol. 24, pp. 8–11, 10 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, 6 2008. [Online]. Available: http://ieeexplore.ieee.org/document/4543069/

[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, 4 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, 9 2013. [Online]. Available: https://dl.acm.org/doi/10.1145/2534169.2486015

[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, 4 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8274950/

[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, 12 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9193946/

$$J\big(P_{\mathcal{K}}(x_{m_{\mathcal{K}}}),P'_{\mathcal{K}}(x_{m_{\mathcal{K}}})\big) \triangleq \sum_{m_{\mathcal{K}}} P_{\mathcal{K}}(x_{m_{\mathcal{K}}})\left(\rho\log\Big(1+\frac{|\boldsymbol{h}^H_{\mathrm{E}}(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2}{\sigma^2_v}\Big) + (1-\rho)\sum_{m'_{\mathcal{K}}}P(\hat{x}_{m'_{\mathcal{K}}}|x_{m_{\mathcal{K}}})\log\frac{P(\hat{x}_{m'_{\mathcal{K}}}|x_{m_{\mathcal{K}}})P'_{\mathcal{K}}(x_{m_{\mathcal{K}}})}{P'_{\mathcal{K}}(\hat{x}_{m'_{\mathcal{K}}})P_{\mathcal{K}}(x_{m_{\mathcal{K}}})}\right). \tag{49}$$

[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, 5 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9326394/

[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, 8 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8692391/

[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, 5 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8941106/

[12] 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. [Online]. Available: https://ieeexplore.ieee.org/document/9751388/

[13] 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, 2 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8907447/

[14] 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, 12 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8807353/

[15] 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, 6 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9358202/

[16] J. Xu, Z. Dai, and Y. Zeng, "Enabling full mutualism for symbiotic radio with massive backscatter devices," *arXiv:2106.05789*, 6 2021. [Online]. Available: http://arxiv.org/abs/2106.05789

[17] Z. Yang and Y. Zhang, "Optimal swipt in ris-aided mimo networks," *IEEE Access*, vol. 9, pp. 112 552–112 560, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9494365/

[18] 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, 8 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9382925/

[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, 10 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9551877/

[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, 2 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8474355/

[21] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," vol. 18. IEEE, 12 2018, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/8647620/

[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, 8 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9110912/

[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, 5 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, 6 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, 6 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8964457/

[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, 7 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9400380/

[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, 12 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, 12 2021.

[31] W. Tang, J. Y. Dai, M. Chen, X. Li, Q. Cheng, S. Jin, K. Wong, and T. J. Cui, "Programmable metasurface-based rf chain-free 8psk wireless transmitter," *Electronics Letters*, vol. 55, pp. 417–420, 4 2019. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1049/el.2019.0400

[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, 3 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8901437/

[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, 6 2020, pp. 2965–2970. [Online]. Available: https://ieeexplore.ieee.org/document/9174060/

[34] R. Liu, H. Li, M. Li, and Q. Liu, "Symbol-level precoding design for intelligent reflecting surface assisted multi-user mimo systems." IEEE, 10 2019, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/8928065/

[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, 5 2020, pp. 8688–8692. [Online]. Available: https://ieeexplore.ieee.org/document/9052989/

[36] X. Xu, Y.-C. Liang, G. Yang, and L. Zhao, "Reconfigurable intelligent surface empowered symbiotic radio over broadcasting signals," vol. 2020-Janua. IEEE, 12 2020, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/9348236/

[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, 7 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9391685/

[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, 4 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, 4 2022. [Online]. Available: http://arxiv.org/abs/2104.09161https://ieeexplore.ieee.org/document/9652042/

[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, 5 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8981888/

[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, 11 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9133588/

[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, 5 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9405433/

[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*, 10 2021. [Online]. Available: http://arxiv.org/abs/2110.10316

[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. [Online]. Available: https://ieeexplore.ieee.org/document/8910627/

[46] B. Clerckx, R. Zhang, R. Schober, D. W. K. Ng, D. I. Kim, and H. V. Poor, "Fundamentals of wireless information and power transfer: From rf energy harvester models to signal and system designs," *IEEE Journal on Selected Areas in Communications*, vol. 37, pp. 4–33, 1 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8476597/

[47] 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, 7 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9651541/
[48] D. Bharadia, K. R. Joshi, M. Kotaru, and S. Katti, "Backfi: High throughput wifi backscatter," vol. 45. ACM, 8 2015, pp. 283–296. [Online]. Available: https://dl.acm.org/doi/10.1145/2785956.2787490
[49] 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, 12 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7274644/
[50] 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, 2 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8411483/
[51] 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, 6 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9377568/
[52] 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, 4 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9392844/
[53] J. Qian, Y. Zhu, C. He, F. Gao, and S. Jin, "Achievable rate and capacity analysis for ambient backscatter communications," *IEEE Transactions on Communications*, vol. 67, pp. 6299–6310, 9 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8721108/
[54] T. Nguyen and T. Nguyen, "Capacity achieving quantizer design for binary channels," *IEEE Communications Letters*, vol. 25, pp. 759–763, 3 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9247251/
[55] T. Nguyen, Y.-J. Chu, and T. Nguyen, "On the capacities of discrete memoryless thresholding channels," vol. 2018-June. IEEE, 6 2018, pp. 1–5. [Online]. Available: https://ieeexplore.ieee.org/document/8417506/
[56] T. Nguyen and T. Nguyen, "Optimal quantizer structure for maximizing mutual information under constraints," *IEEE Transactions on Communications*, vol. 69, pp. 7406–7413, 11 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9530430/
[57] 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, 11 2004. [Online]. Available: http://ieeexplore.ieee.org/document/1347364/
[58] 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, 4 2011. [Online]. Available: https://ieeexplore.ieee.org/document/5730559/
[59] G. J. O. Jameson, "The incomplete gamma functions," *The Mathematical Gazette*, vol. 100, pp. 298–306, 7 2016. [Online]. Available: https://www.cambridge.org/core/product/identifier/S002555721600067X/type/journal_article
[60] 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, 6 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9366549/
[61] T. Nguyen and T. Nguyen, "On thresholding quantizer design for mutual information maximization: Optimal structures and algorithms," vol. 2020-May. IEEE, 5 2020, pp. 1–5. [Online]. Available: https://ieeexplore.ieee.org/document/9128966/
[62] 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, 2 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8423609/
[63] J. Wang, N. Li, W. Shi, Y. Ma, X. Liang, and X. Dong, "Capacity of 60 ghz wireless communications based on qam," *Journal of Applied Mathematics*, vol. 2014, pp. 1–5, 2014. [Online]. Available: http://www.hindawi.com/journals/jam/2014/815617/
[64] S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Transactions on Information Theory*, vol. 18, pp. 14–20, 1 1972. [Online]. Available: http://ieeexplore.ieee.org/document/1054753/
[65] R. E. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Transactions on Information Theory*, vol. 18, pp. 460–473, 7 1972. [Online]. Available: http://ieeexplore.ieee.org/document/1054855/
[66] 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, 12 2010. [Online]. Available: http://ieeexplore.ieee.org/document/5590317/
[67] C. Qiu, Q. Wu, M. Hua, X. guan, and Y. Wu, "Achieving multi-beam gain in intelligent reflecting surface assisted wireless energy transfer," *arXiv:2205.08893*, 5 2022. [Online]. Available: http://arxiv.org/abs/2205.08893