RIScatter: Unifying Backscatter Communication and Reconfigurable Intelligent Surface

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

***Abstract*—Backscatter Communication (BackCom) nodes harvest energy from and modulate information over an external electromagnetic wave. Reconfigurable Intelligent Surface (RIS) adapts its phase shift response to enhance or attenuate channel strength in specific directions. In this paper, we show how those two seemingly different technologies (and their derivatives) can be unified into a single architecture called RIScatter. RIScatter consists of multiple dispersed or co-located scatter nodes, whose reflection states are adapted to partially modulate their own information and partially engineer the wireless channel. The key principle is to render the probability distribution of reflection states as a joint function of the information source, Channel State Information (CSI), and Quality of Service (QoS) of coexisting active primary and passive backscatter links. This enables RIScatter to softly bridge BackCom and RIS; boil down to either under specific input distribution; or evolve in a mixed form for heterogeneous traffic control and universal hardware design. To reap the benefits of RIScatter, we also propose a co-located Successive Interference Cancellation (SIC)-free receiver that semi-coherently decodes the backscatter information, recovers the reflection pattern, and coherently decodes the primary link. For a single-user multi-node RIScatter network, we characterize the achievable primary-(total-)backscatter rate region by designing the input distribution at scatter nodes, the active beamforming at the Access Point (AP), and the energy decision regions at the user. Simulation results demonstrate RIScatter nodes can exploit the scattered paths to smoothly shift between backscatter modulation and passive beamforming.**

***Index Terms*—Input distribution design, symbol-level precoding, index modulation, active-passive coexisting network, SIC-free receiver, energy detection, backscatter communication, reconfigurable intelligent surface, ambient backscatter communication, symbiotic radio.**

I. Introduction

FUTURE wireless network is envisioned to provide high throughput, uniform coverage, pervasive connectivity, heterogeneous control, and cognitive intelligence for trillions of low-power devices. Backscatter Communication (BackCom) separates a 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]. The receiver (reader) can be either co-located or separated with the carrier emitter, known as Monostatic BackCom (MBC) and Bistatic BackCom (BBC) in Fig. 1(a) and 1(b), respectively. Relevant applications such as Radio-Frequency Identification (RFID) [2], [3] and passive sensor network [4], [5] have been extensively researched, standardized, and commercialized to embrace the Internet of Everything (IoE). However, conventional backscatter nodes only respond when externally inquired by a nearby reader. Ambient Backscatter Communication (AmBC) in Fig. 1(c) was proposed a decade ago where battery-free nodes recycle ambient signals (e.g., radio, television and Wi-Fi) to harvest energy and establish connections [6]. It does not require dedicated power source, carrier emitter, or frequency spectrum, but the backscatter decoding is subject to the strong interference from the primary (legacy) link. To tackle this, cooperative AmBC [7] employs a co-located receiver to decode both coexisting links and the concept was further refined as Symbiotic Radio (SR) in Fig. 1(d) [8]. Specifically, the active transmitter generates RF wave carrying primary information, the passive node creates a rich-scattering environment and rides its own information, and the co-located receiver cooperatively decodes both links. In those BackCom applications, the scatter node is considered as an *information source* and the reflection pattern depends exclusively on the information symbol. On the other hand, Reconfigurable Intelligent Surface (RIS) in Fig. 1(e) is a smart signal reflector with numerous passive elements of adjustable phase shifts. It customizes the wireless environment for signal enhancement, interference suppression, scattering enrichment, and/or non-line-of-sight bypassing [9]. Each RIS element is considered as a *channel adaptor* and the reflection pattern depends exclusively on the Channel State Information (CSI).

As a special case of Cognitive Radio (CR), active and passive transmissions coexist and interplay in AmBC and SR. Such a coexistence is classified into commensal (overlay), parasitic (underlay), and competitive (interfering) paradigms, and their achievable rate and outage performance were investigated in [10], [11]. For the co-located cooperative receiver, the Bit Error Rate (BER) performance of Maximum-Likelihood (ML), linear, and SIC detectors are derived over flat fading channels [7]. However, the work assumed equal symbol duration and perfect synchronization for primary and backscatter links. Importantly, active-passive coexisting networks have three special and important properties:

1) Primary and backscatter symbols are superimposed by *double modulation* (i.e., multiplication coding);
2) Backscatter signal strength is much weaker than primary due to the *double fading* effect;
3) The spreading factor (i.e., backscatter symbol duration over primary) is usually large[1].

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

[1]The load-switching interval of low-power backscatter modulators is usually 0.1 to 10 μs [12], accounting for a typical spreading factor between 10 and 10^3.

TABLE I
COMPARISON OF SCATTERING APPLICATIONS

	MBC/BBC	AmBC	SR (large spreading factor)	RIS	RIScatter
Information link(s)	Backscatter	Coexisting	Coexisting	Primary	Coexisting
Primary signal on backscatter decoding	Carrier	Multiplicative interference	Spreading code	—	Energy uncertainty
Backscatter signal on primary decoding	—	Multiplicative interference	CSI uncertainty	Passive beamforming	Dynamic passive beamforming
Cooperative devices	—	No	Primary transmitter and co-located receiver	—	Primary transmitter, scatter nodes, and co-located receiver
Sequential decoding	—	No	Primary-to-backscatter, SIC and MRC	—	Backscatter-to-primary, no SIC/MRC
Reflection pattern depends on	Information source	Information source	Information source	CSI	Information source, CSI, and QoS
Reflection state distribution	Equiprobable	Equiprobable	Equiprobable or Gaussian	Degenerate	Flexible
Load-switching speed	Fast	Slow	Slow	Quasi-static	Arbitrary

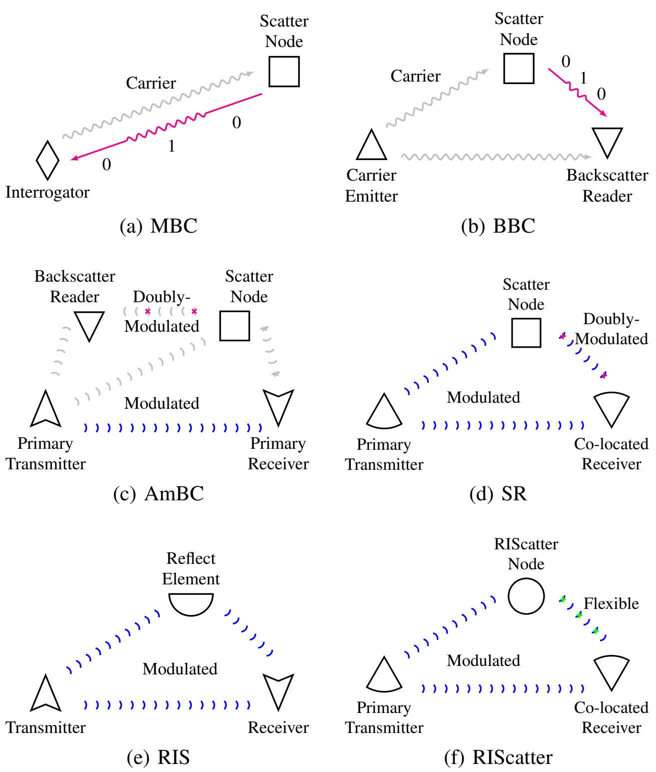


Fig. 1. Illustration of scattering applications. The blue flow(s) constitutes the primary link while the magenta/green flow denotes the backscatter link.

The second property motivated [7], [8], [10], [11], [13]–[19] to view SR as a multiplicative Non-Orthogonal Multiple Access (NOMA) and perform SIC from primary to backscatter link. During primary decoding, the backscatter signal can be modelled as channel uncertainty or multiplicative interference when the spreading factor is large or small, respectively. Decoding each backscatter symbol also requires multiple SIC followed by a MRC and is operation-intensive and CSI-sensitive. When the spreading factor is sufficiently large, the primary achievable rate under semi-coherent detection[2] asymptotically approaches its coherent counterpart and both links are decoded interference-free [13]. However, this severely limits the backscatter throughput and requires numerous SIC per backscatter symbol.

On the other hand, static RIS design with fixed reflection pattern per channel block has been extensively studied in wireless communication, sensing, and power literature [20]–[25]. Dynamic RIS performs time sharing between different phase shifts and introduces artificial channel diversity within each channel block. It was first proposed to fine-tune the Orthogonal Frequency-Division Multiplexing (OFDM) resource blocks [26] then extended to the downlink power and uplink information phases of Wireless Powered Communication Network (WPCN) [27]–[29]. However, dynamic RIS carries *no* additional information since the reflection state at each time slot is known to the receiver. RIS can also be used as an information source and prototypes have been developed for Phase Shift Keying (PSK) [30] and Quadrature Amplitude Modulation (QAM) [31]. From an information-theoretic perspective, the authors of [32] reported that joint transmitter-RIS encoding achieves the capacity of RIS-aided finite-input channel and using RIS as a naive passive beamformer to maximize the receive Signal-to-Noise Ratio (SNR) is generally suboptimal. This inspired RIS-empowered BackCom [33]–[42] to combine passive beamforming and backscatter modulation in the overall reflection pattern. In particular, *symbol level precoding* maps the information symbols to the optimized RIS coefficient sets [33], [34], *overlay modulation* superposes the information symbols over a common auxiliary matrix [35]–[38], *spatial modulation* switches between the reflection coefficient sets that maximize SNR at different receive antennas [39]–[41], and *index modulation* employs dedicated reflection elements (resp. information elements) for passive beamforming (resp. backscatter modulation) [42]. However, those joint designs incur advanced hard-

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

ware architecture and high optimization complexity. Relevant literature also considers either Gaussian codebook [10], [11], [13]–[17], [37] that is impractical for low-power nodes or finite equiprobable inputs [7], [8], [18], [19], [33]–[36], [38]–[42] that does not fully exploit the CSI and properties of active-passive coexisting networks. Those problems are addressed in this paper and the contributions are summarized below.

First, we propose RIScatter as a novel protocol that unifies BackCom and RIS by adaptive reflection state (backscatter input) distribution design. The concept is shown in Fig. 1(f) where one or more RIScatter nodes ride over an active transmission and simultaneously modulate their information and engineer the wireless channel. A co-located receiver cooperatively decodes both coexisting links and each reflection state simultaneously acts as information and passive beamforming codewords. The reflection pattern of each node is semi-randomly chosen with the guidance of input distribution as a joint function of the information source, CSI, and QoS. Such an adaptive channel coding boils down to the degenerate distribution of RIS when the primary link is prioritized, and outperforms the uniform distribution of BackCom (by accounting the CSI) when the backscatter link is prioritized. Joint and independent encoding are available when multiple RIScatter nodes are co-located and dispersed, respectively. Table I compares RIScatter to BackCom and RIS. However, two major challenges for RIScatter are the practical receiver and input distribution design. This is the first paper to unify BackCom and RIS from the perspective of input distribution.

Second, we address the first challenge and propose a practical receiver that semi-coherently decodes the backscatter information, recovers the reflection pattern, and coherently decodes the primary link. We consider backscatter energy detection and formulate a Discrete Memoryless Multiple Access Channel (DMMAC) over disjoint decision regions. Thanks to the double modulation and symbol-level passive precoding, the semi-coherent backscatter decoding is essentially part of channel training. Once the backscatter information is successfully decoded, its contribution can be modelled within composite channel as dynamic passive beamforming (rather than re-precoded and cancelled). It requires only one energy comparison and re-encoding per backscatter symbol and is much efficient than conventional decoding based on SIC and MRC. It is also suitable for arbitrary input distribution and spreading factor, which significantly improves backscatter throughput. This is the first paper to consider practical backscatter-primary decoding scheme to exploit backscatter modulation as dynamic passive beamforming at the co-located receiver.

Third, we address the second challenge and consider a single-user multi-node Multiple-Input Single-Output (MISO) scenario. We characterize the achievable primary-(total-)backscatter rate region by optimizing the input distribution at RIScatter nodes, the active beamforming at the Access Point (AP), and the energy decision regions at the user under different QoS. A Block Coordinate Descent (BCD) algorithm is proposed where the Karush-Kuhn-Tucker (KKT) input distribution is numerically evaluated by limit of sequences, the active beamforming is optimized by Projected Gradient Descent (PGD), and the decision regions are refined by existing sequential quantizer for Discrete Memoryless Thresholding Channel (DMTC). Uniquely, we consider CSI, QoS, and backscatter constellation in the optimization and the input distribution result is applicable to general detection schemes. This is also the first paper to reveal the importance of backscatter input distribution and decision region designs in active-passive coexisting networks.

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. $\operatorname{diag}(\cdot)$ returns a square matrix with the input vector on its main diagonal and zeros elsewhere. $\operatorname{card}(\cdot)$ returns the cardinality of a set. $\log(\cdot)$ denotes logarithm of base e. $(\cdot)^*$, $(\cdot)^\mathsf{T}$, $(\cdot)^\mathsf{H}$, $|\cdot|$, and $\|\cdot\|$ denote the conjugate, transpose, conjugate transpose (Hermitian), absolute value, and Euclidean norm operators, respectively. $(\cdot)^{(r)}$ and $(\cdot)^\star$ denote the r-th iterated and optimal results, respectively. The distribution of a Circularly Symmetric Complex Gaussian (CSCG) random variable with zero mean and variance σ^2 is denoted by $\mathcal{CN}(0,\sigma^2)$, and $\sim$ means "distributed as".

II. RISCATTER

A. Principles

RF wave scattering/reflecting are often manipulated by passive antennas or programmable metamaterial [43]. The former first receives the impinging signals then reradiates some back to the space and dissipates the remaining [2]. In contrast, the latter reflects the wave at the space-metamaterial boundary without receiving them and mainly applies a phase shift. The scattered signal consists of a structural mode component that consistently contributes to environment multipath (usually modelled within CSI) and an antenna mode component that depends on the impedance mismatch (mainly used for backscatter modulation and/or passive beamforming) [44]. For an antenna (resp. metamaterial) scatterer 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.

BackCom treats scatterers as information sources that *randomly* choosing the reflection state based on information symbol. For M-ary QAM, constellation point c_m maps to reflection coefficient $\mathit{\Gamma}_m$ by [45]

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

where $\alpha\in\mathbb{I}$ is the common amplitude scattering ratio at the direction of interest. In contrast, RIS treats scatterers as channel adaptors that *deterministically* choosing the reflection state based on CSI. For a RIS element with M available states, phase shift θ_m maps to reflection coefficient $\mathit{\Gamma}_m$ by [20]

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

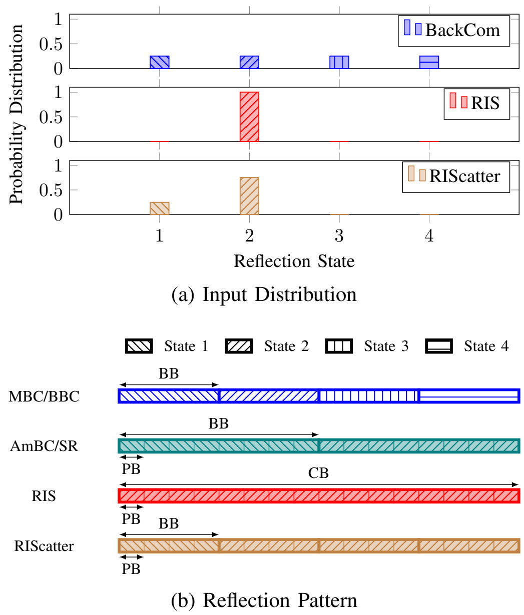


Fig. 2. Input distribution and reflection pattern of scattering applications. "PB", "BB", and "CB" refer to primary symbol block, backscatter symbol block, and channel block, respectively. Shadowing means presence of primary link. In this example, the optimal passive beamformer is state 2 and its input probability for RIScatter is 0.75. The spreading factor is 4 for RIScatter and 8 for AmBC/SR.

where $\beta_m \in \mathbb{I}$ is the overall amplitude scattering ratio at state m.[3] For a scatter node with given load impedance set, BackCom employs $\mathit{\Gamma}_m$ only as an information codeword while RIS employs $\mathit{\Gamma}_m$ only as a passive beamforming codeword. RIScatter generalizes both from a probabilistic perspective where each node *semi-randomly* chooses the reflection state with the guidance of probability $P(\mathit{\Gamma}_m)$. The input distribution is flexibly designed as a joint function of the information source, CSI, and QoS of both links. BackCom and (static) RIS can be regarded as its extreme cases where the input distribution boils down to uniform and degenerate, respectively. Fig. 2 compares the input distribution and reflection pattern of RIScatter to BackCom and RIS.

Remark 1. *RIScatter is different from conventional dynamic passive beamforming since the reflection pattern conveys additional information. The latter only involves time sharing between different states and the actual order has no impact on the (ergodic) primary performance. For RIScatter, different state combinations correspond to different messages and the impact reduces to dynamic passive beamforming upon successful backscatter decoding.*

RIScatter nodes can be implemented, for example, by adding an integrated receiver [46] and adaptive encoder [47] to off-the-shelf passive RFID tags. The block diagram, equivalent circuit, and scatter model are illustrated in Fig. 3. To exploit the benefits of RIScatter, we also propose a practical co-located receiver that semi-coherently decodes the backscatter information, recovers the reflection pattern, and coherently decodes the primary link. The detail will be covered in Section II-B.

[3]Most existing RIS literature assumes lossless reflection $\beta_m = 1$, $\forall m$.

B. System Model

As shown in Fig. 4, we consider a RIScatter system where a Q-antenna AP serves a single-antenna user and K nearby dispersed or co-located RIScatter nodes. All nodes have M available reflection states. In the primary point-to-point system, the AP transmits information to the user over a multipath channel enhanced by RIScatter nodes. In the backscatter multiple access system, the AP acts as carrier emitter, the RIScatter nodes modulate over scattered signal, and the user jointly decodes all node messages. For simplicity, we consider a quasi-static block fading model and focus on a specific block where the CSI remains constant. It is assumed the signal reflected by two or more times is negligible, the primary symbol duration is much longer than multipath delay spread (no inter-symbol interference), and the spreading factor N is a positive integer. Denote the AP-user direct channel as $\boldsymbol{h}_{\mathrm{D}}^{\mathsf{H}} \in \mathbb{C}^{1\times Q}$, the AP-node $k \in \mathcal{K} \triangleq \{1,\ldots,K\}$ forward channel as $\boldsymbol{h}_{\mathrm{F},k}^{\mathsf{H}} \in \mathbb{C}^{1\times Q}$, the node k-user backward channel as $h_{\mathrm{B},k}$, and the cascaded AP-node k-user channel as $\boldsymbol{h}_{\mathrm{C},k}^{\mathsf{H}} \triangleq h_{\mathrm{B},k}\boldsymbol{h}_{\mathrm{F},k}^{\mathsf{H}} \in \mathbb{C}^{1\times Q}$. We assume perfect direct and cascaded CSI are available at the AP and the user[4]. Let $x_k \in \mathcal{X} \triangleq \{c_1,\ldots,c_M\}$ be the coded backscatter symbol of node k and $x_{\mathcal{K}} \triangleq (x_1,\ldots,x_K)$ be the backscatter symbol tuple of all nodes. Due to double modulation, the composite channel is a function of backscatter symbol tuple[5]

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

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

where $\alpha_k \in \mathbb{I}$ is the common amplitude scattering ratio of node k, $\boldsymbol{\alpha} \triangleq [\alpha_1,\ldots,\alpha_K]^{\mathsf{T}} \in \mathbb{I}^K$, $\boldsymbol{x} \triangleq [x_1,\ldots,x_K]^{\mathsf{H}} \in \mathcal{X}^K$, and $\boldsymbol{H}_{\mathrm{C}} \triangleq [\boldsymbol{h}_{\mathrm{C},1},\ldots,\boldsymbol{h}_{\mathrm{C},K}]^{\mathsf{H}} \in \mathbb{C}^{K\times Q}$. Next, we consider the signal model within one backscatter block. The signal received by the user at primary block $n \in \mathcal{N} \triangleq \{1,\ldots,N\}$ is

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

where $\boldsymbol{w} \in \mathbb{C}^Q$ is the active beamformer satisfying average transmit power constraint $\|\boldsymbol{w}\|^2 \le P$, $s \sim \mathcal{CN}(0,1)$ is the primary symbol, and $v \sim \mathcal{CN}(0,\sigma_v^2)$ is the Additive White Gaussian Noise (AWGN) with average power σ_v^2. Let $m_k \in \mathcal{M} \triangleq \{1,\ldots,M\}$ be the reflection state index of node k and $m_{\mathcal{K}} \triangleq (m_1,\ldots,m_K)$ be the state index tuple of all nodes. The backscatter symbol tuple $x_{\mathcal{K}}$ is a random variable that takes value $x_{m_{\mathcal{K}}}$ when state tuple $m_{\mathcal{K}}$ is used. Conditioned on hypothesis $\mathcal{H}_{m_{\mathcal{K}}}$, the receive signal per primary block follows CSCG distribution $\mathcal{CN}(0,\sigma_{m_{\mathcal{K}}}^2)$, where

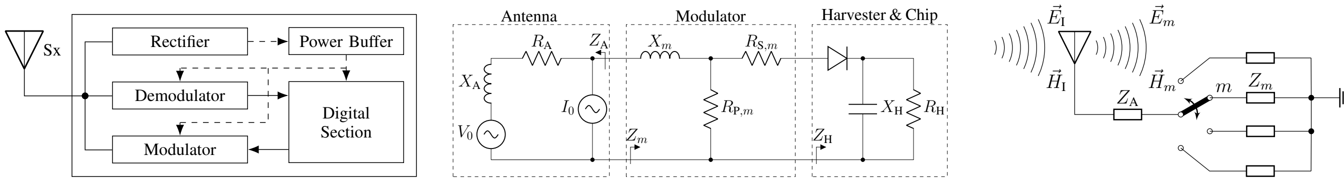
$$\sigma_{m_{\mathcal{K}}}^2 = |\boldsymbol{h}^{\mathsf{H}}(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2 + \sigma_v^2 \tag{6}$$

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

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

[4]The cascaded CSI can be estimated by sequential [49]–[51] or parallel [52] approaches for dispersed nodes, or group-based [53] or hierarchical [54] approaches for co-located nodes.

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



(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 scatter antenna behaves as a constant power source, where the voltage V_0 and current I_0 are introduced by incident electric field $\vec{E}_{\mathrm{I}}$ and magnetic field $\vec{H}_{\mathrm{I}}$ [48].

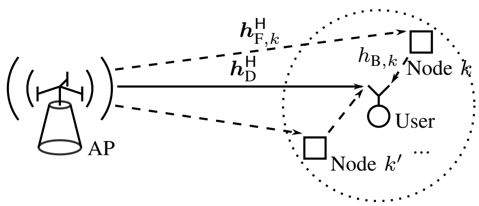


Fig. 4. A single-user multi-node RIScatter system.

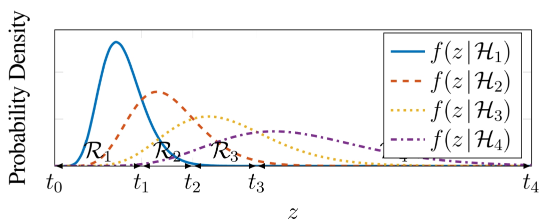


Fig. 5. PDF of accumulated energy per backscatter block conditioned on different reflection state hypotheses.

The user first jointly decodes all backscatter messages under primary interference by performing energy detection over disjoint decision regions. For the ease of notation, we map the state index tuple $m_{\mathcal{K}}$ to $l \in \mathcal{L} \triangleq \{1,\ldots,L\}$, where $L \triangleq M^K$ and $\sigma_1^2,\ldots,\sigma_L^2$ forms an increasing sequence. Both notations are used interchangeably in the following context. The decision region of backscatter symbol tuple l is

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

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

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

over which input distribution of all RIScatter nodes can be designed. An example of energy detection is illustrated in Fig. 5.

Remark 2. *The capacity-achieving decision region design for DMTC with non-binary inputs in arbitrary distribution remains an open issue. It was proved deterministic detectors can be rate-optimal, but non-convex decision regions (comprise non-adjacent partitions) are generally required and the optimal number of thresholds remains unknown [55], [56]. Hence, we limit the energy detector to convex deterministic decision regions and consider sequential threshold design.*

After successful backscatter decoding, the user can re-encode for exact reflection pattern and obtain the composite channel by (4), then coherently decode the primary link. Such an SIC-free sequential decoding enables backscatter modulation and dynamic passive beamforming by only one energy comparison and re-encoding per backscatter symbol. It also supports arbitrary input distribution and spreading factor for higher backscatter throughput.

C. Achievable Rates

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

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

Following [57], we define the backscatter information function between input value tuple $x_{m_{\mathcal{K}}}$ and output variable tuple $\hat{x}_{\mathcal{K}}$ as

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

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

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

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

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

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

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

the primary marginal information of x_{m_k} is

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

and the average primary mutual information is

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

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

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

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

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

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

III. Rate-Region Characterization

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

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

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

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

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

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

$$\boldsymbol{t} \geq \mathbf{0}. \tag{20f}$$

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

A. Input Distribution

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

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

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

which is convex when $K=1$ or joint encoding[6] over $K>1$ co-located nodes. When the nodes are dispersed, problem (21) involves product term $\prod_{k\in\mathcal{K}} P_k(x_{m_k})$ and is non-convex. Following [57], we first recast the KKT conditions to their equivalent forms then propose a numerical method that guarantees those conditions on convergence of sequences.

Remark 3. *As demonstrated in [58], KKT conditions are generally necessary but insufficient for total rate maximization of DMMAC. We will later show by simulation that, for a moderate K, the average achievable rate regions of KKT and global-optimal input distributions completely coincide with each other.*

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

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

---

Input: K, N, $\boldsymbol{h}_{\mathrm{D}}^{\mathsf{H}}$, $\boldsymbol{H}_{\mathrm{C}}$, $\boldsymbol{\alpha}$, $\mathcal{X}$, σ_v^2, ρ, $\boldsymbol{w}$, $\boldsymbol{t}$, ϵ
Output: $\{\boldsymbol{p}_k^\star\}_{k\in\mathcal{K}}$
1: Set $\boldsymbol{h}^{\mathsf{H}}(x_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}$ 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: $\quad \boldsymbol{p}_k^{(0)} > \mathbf{0}$, $\forall k$
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: $\quad \boldsymbol{p}_k^{(r)}$, $\forall k$ by (23)
14: $\quad$ Redo step 7–10
15: **Until** $I^{(r)}(x_{\mathcal{K}}) - I^{(r-1)}(x_{\mathcal{K}}) \leq \epsilon$

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

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

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

Proof. Please refer to Appendix A. □

For each RIScatter 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*

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

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.

B. Active Beamforming

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

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

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

which is still non-convex due to the integration and entropy terms. To tackle this, we rewrite the DMMAC transition probability (9) as a regularized incomplete Gamma function in the series representation [59, Theorem 3]

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

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

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

where $g_{m_\mathcal{K}}(t_{l-1},t_l)\triangleq g_{m_\mathcal{K}}(t_l)-g_{m_\mathcal{K}}(t_{l-1})$ and

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

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

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

The PGD active beamforming optimization with step size γ refined by Backtracking Line Search (BLS) [60, Section 9.2] is summarized in Algorithm 2.

Algorithm 2: Iterative Active Beamforming Optimization by PGD with BLS

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

1: Set $\boldsymbol{h}^{\mathsf{H}}(x_{m_\mathcal{K}})$, $\forall m_\mathcal{K}$ by (4)
2: $P_\mathcal{K}(x_{m_\mathcal{K}})$, $\forall m_\mathcal{K}$ by (10)
3: Initialize $r\leftarrow 0$
4: $\boldsymbol{w}^{(0)}$, $\|\boldsymbol{w}^{(0)}\|^2\leq P$
5: Get $(\sigma^{(r)}_{m_\mathcal{K}})^2$, $\forall m_\mathcal{K}$ by (6)
6: $Q^{(r)}\big(N,\frac{t_{l-1}}{\sigma^2_{m_\mathcal{K}}},\frac{t_l}{\sigma^2_{m_\mathcal{K}}}\big)$, $\forall m_\mathcal{K},l$ by (25)
7: $I^{(r)}(x_\mathcal{K})$ by (28)
8: $\nabla_{\boldsymbol{w}^*}Q^{(r)}\big(N,\frac{t_{l-1}}{\sigma^2_{m_\mathcal{K}}},\frac{t_l}{\sigma^2_{m_\mathcal{K}}}\big)$, $\forall m_\mathcal{K},l$ by (26)
9: $\nabla_{\boldsymbol{w}^*}I^{(r)}(x_\mathcal{K})$ by (29)
10: **Repeat**
11: Update $r\leftarrow r+1$
12: $\gamma^{(r)}\leftarrow\gamma$
13: $\bar{\boldsymbol{w}}^{(r)}\leftarrow\boldsymbol{w}^{(r-1)}+\gamma\nabla_{\boldsymbol{w}^*}I^{(r-1)}(x_\mathcal{K})$
14: $\boldsymbol{w}^{(r)}$ by (30)
15: Redo step 5–7
16: **While** $I^{(r)}(x_\mathcal{K}) < I^{(r-1)}(x_\mathcal{K}) + \alpha\gamma\|\nabla_{\boldsymbol{w}^*}I^{(r-1)}(x_\mathcal{K})\|^2$
17: Set $\gamma^{(r)}\leftarrow\beta\gamma^{(r)}$
18: Redo step 13–15
19: **End While**
20: Redo step 8, 9
21: **Until** $\|\boldsymbol{w}^{(r)}-\boldsymbol{w}^{(r-1)}\|\leq\epsilon$

C. Decision Threshold

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

$$\max_{\boldsymbol{t}} \quad I(x_\mathcal{K}) \tag{31a}$$
$$\text{s.t.} \quad \text{(20e),(20f)}, \tag{31b}$$

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 design for a Discrete-input Continuous-output Memoryless Channel (DCMC), thanks to Remark 4 and 5.

Remark 4. *Energy detector design has no impact on the primary achievable rate since the composite channel* (4) *can always be reconstructed after backscatter decoding and re-encoding. It suggests any thresholding scheme maximizing the total backscatter rate* (13) *is also optimal for problem* (31).

Remark 5. *In terms of total backscatter rate, the potentially dispersed nodes with known input distribution can be viewed as an equivalent source with tuple codewords. As such, the DMMAC* (9) *becomes a DMTC and problem* (31) *reduces to the rate-optimal quantization design for a DCMC.*

Next, we constrain the feasible domain of problem (31) from continuous space $\mathbb{R}_+^{L+1}$ to finite candidate set (i.e., fine-grained

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

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

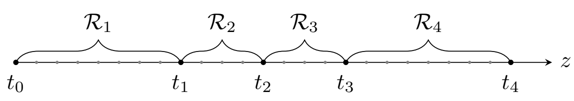


Fig. 6. The decision thresholds are selected from fine-grained discrete energy levels instead of continuous space. Each decision region consists of at least one neighbor energy bins.

discrete energy levels) $\mathcal{T}^{L+1}$. As shown in Fig. 6, by introducing this extra analog-to-digital conversion, we can group adjacent high-resolution energy bins to construct backscatter decision regions. Thus, problem (31) can be recast as

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

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

which can be solved by existing rate-optimal sequential quantizer designs for DMTC. To obtain global optimal solution, [61] 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 $\mathcal{O}\big(L^2(\text{card}(\mathcal{T})-L)\big)$, while [62] started from the optimality condition for three neighbor thresholds and presented a traverse-then-bisect algorithm with complexity $\mathcal{O}\big(\text{card}(\mathcal{T})L\log(\text{card}(\mathcal{T})L)\big)$. In Section IV, both schemes will be compared with the ML scheme [63]

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

which is generally suboptimal for problem (31) unless all RIScatter nodes are with equiprobable inputs.

IV. Simulation Results

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

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

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

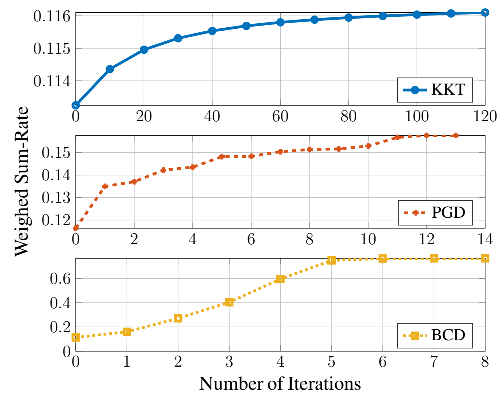


Fig. 7. Typical convergence curves at $\rho = 0$ for $Q = 4$, $K = 8$, $M = 2$, $N = 20$, $\sigma_v^2 = -40\text{dBm}$ and $r = 2\text{m}$.

A. Evaluation of Proposed Algorithms

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

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

B. Comparison of Scattering Applications

On top of the setup in Fig. 4, we consider RIScatter and the following benchmark applications:

- *Legacy:* Active transmission without scatterers.
- *BBC:* The primary symbol becomes deterministic $s[n] = 1$ and the receive signal becomes

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

  The total backscatter rate of coherent decoding approaches $K\log M$ when N is sufficiently large.

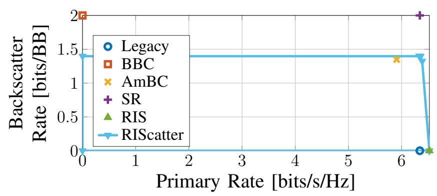


Fig. 8. Typical achievable rate region/points of scattering applications for $Q=1$, $K=1$, $M=4$, $N=10^3$, $\sigma_v^2=-40$dBm and $r=2$m.

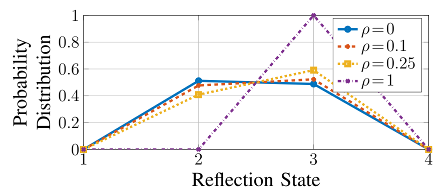


Fig. 9. Typical RIScatter reflection state distribution at different ρ for $Q=1$, $K=1$, $M=4$, $N=20$, $\sigma_v^2=-40$dBm and $r=2$m.

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

$$I_{\mathrm{P}}^{\mathrm{AmBC}}(s;y) \approx \log\left(1+\frac{|\boldsymbol{h}_{\mathrm{D}}^{\mathsf{H}}\boldsymbol{w}|^2}{\sum_k |\alpha_k \boldsymbol{h}_{\mathrm{C},k}^{\mathsf{H}}\boldsymbol{w}|^2+\sigma_v^2}\right), \tag{37}$$

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

$$\hat{y}^{\mathrm{SR}}[n] = \sum_k \alpha_k \boldsymbol{h}_{\mathrm{C},k}^{\mathsf{H}} x_k \boldsymbol{w} s[n] + v[n]. \tag{38}$$

The total backscatter rate of coherent decoding approaches $K\log M$ when N is sufficiently large.
- *RIS:* The reflection pattern is deterministic and the total backscatter rate is zero. The primary achievable rate is a special case of (16)

$$I_{\mathrm{P}}^{\mathrm{RIS}}(s;y|x_{\mathcal{K}}) = I_{\mathrm{P}}(s;y|x_{m_{\mathcal{K}}^\star}) = \log\left(1+\frac{|\boldsymbol{h}^{\mathsf{H}}(x_{m_{\mathcal{K}}^\star})\boldsymbol{w}|^2}{\sigma_v^2}\right), \tag{39}$$

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

Fig. 8 compares the typical achievable rate region/points of RIScatter and the strategies above. *First,* we observe BBC and SR achieve the best backscatter performance thanks to coherent decoding. For SR, this comes with the cost of N re-encoding, precoding, subtraction together with a time-domain MRC per backscatter symbol. *Second,* the average primary rate slightly decreases/increases in the presence of a AmBC/RIS node and the benefit of SR is not obvious. This is because the cascaded channel is usually orders of magnitude weaker than the direct channel. RIS always ensures constructive superposition of direct and scattered components while SR only creates a quasi-static rich-scattering environment that marginally enhances the average primary rate. When N is moderate, the randomly scattered signals should be modelled as interference rather than multipath components and the SR point should move towards the AmBC point. *Third,* RIScatter enables a flexible primary-backscatter tradeoff with adaptive input distribution design. In terms of maximum primary achievable rate, RIScatter coincides with RIS and outperforms the others using deterministic reflection pattern. On the other hand, for a large N, the maximum backscatter achievable rate of RIScatter is lower than BBC and SR but higher than AmBC. This is because both RIScatter and AmBC employ semi-coherent energy detection, and the adaptive channel coding of RIScatter provides better performance than equiprobable inputs of AmBC especially at the low SNR of double fading. When multiple antenna is available at the AP, active beamforming can be optimized for RIScatter nodes and the backscatter rate gain over AmBC can be even larger.

C. Input Distribution under Different QoS

The objective is to demonstrate RIScatter nodes can leverage CSI- and QoS-adaptive input distribution design to balance backscatter modulation and passive beamforming. For one RIScatter node with $M = 4$, we evaluate the KKT input distribution[8] at different QoS and present the result in Fig. 9. At $\rho=0$ where the backscatter performance is prioritized, the optimal input distribution is 0 on two states and nearly uniform on the other two. This is inline with Shannon's observation that binary antipodal inputs can be good enough for channel capacity at the low SNR [64]. Due to the weak scattered signal, the conditional energy PDF under different hypotheses can be closely spaced as in Fig. 5. The extreme states producing the lowest/highest energy are always assigned with non-zero probability and the middle ones may not provide enough energy difference and thus end up unused. At $\rho=1$ where the primary performance is prioritized, the optimal input distribution is 1 at the state that maximizes the primary SNR and 0 at the others. That is, the reflection pattern becomes deterministic and the RIScatter node boils down to a discrete RIS element. Increasing ρ from 0 to 1 provides a smooth transition from backscatter modulation to passive beamforming and proves RIScatter unifies BackCom and RIS from a probabilistic perspective.

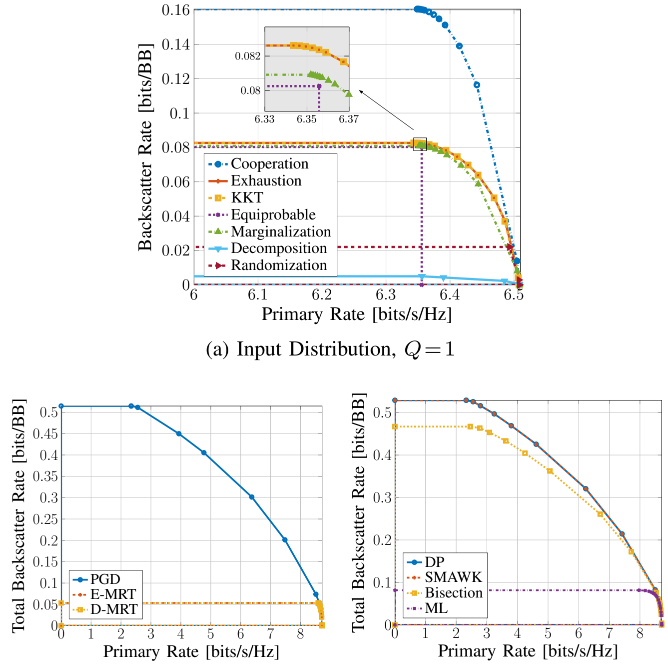
D. Rate Region by Different Schemes

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

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

[7] To provide a preliminary benchmark, we regard the scattered component $\sum_k \alpha_k \boldsymbol{h}_{\mathrm{C},k}^{\mathsf{H}} x_k \boldsymbol{w} s[n]$ as AWGN during primary decoding [13].

[8] Problem (21) is convex when $K=1$ and the KKT solution is global optimal.



(a) Input Distribution, $Q=1$

(b) Active Beamforming, $Q=4$

(c) Decision Threshold, $Q=4$

Fig. 10. Average primary-total-backscatter rate regions by different input distribution, active beamforming, and decision threshold schemes for $K=2$, $M=4$, $N=20$, $\sigma_v^2=-40$dBm and $r=2$m.

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

- *Marginalization:* Marginal probability distributions;
- *Decomposition:* Normalized rank-1 Canonical Polyadic (CP) decomposition tensors by `Tensor Toolbox` [65];
- *Randomization:* Gaussian randomization with the guidance of correlation matrix [66].

Fig. 10(a) shows the average achievable rate regions for those designs. We observe Cooperation provides the outer bound and joint encoding can be beneficial. The average rate performance of Exhaustion and KKT completely coincide with each other, demonstrating KKT input distribution is favorable for a moderate K as mentioned in Remark 3. Equiprobable experiences minor backscatter and major primary rate losses without exploiting CSI and QoS. Both rate gaps should be larger when M and/or K increase. For the recovery methods, the simple Marginalization provides a close performance to KKT, but Randomization and Decomposition fail our expectations for most channel realizations. Those observations emphasize the importance of (joint) adaptive RIScatter encoding and demonstrate the advantage of the proposed KKT input distribution design.

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

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

Fig. 10(b) presents the average achievable rate regions for those schemes. In the low-ρ regime, the proposed PGD beamformer significantly outperforms both MRT schemes in terms of total backscatter rate. This is because the semi-coherent backscatter decoding relies on the *relative difference* of accumulated energy under different symbol tuples. Such an energy diversity is enhanced by PGD that effectively exploits backscatter constellation and input distribution knowledge rather than simply maximizing the direct/ergodic SNR. As ρ increases, the primary equivalent SNR outweighs the backscatter energy difference in (28) and PGD beamformer becomes closer to both MRT schemes. At $\rho=1$, both PGD and E-MRT boil down to MRT towards the composite channel. The difference between E-MRT and D-MRT is insignificant when RIScatter nodes are dispersed. Those observations prove the proposed PGD active beamforming design can exploit the CSI, QoS, and backscatter constellation to balance the primary equivalent SNR and backscatter energy difference and enlarge the achievable rate region.

3) Decision Threshold: We evaluate the following decision threshold strategies for problem (32):

- *DP:* Benchmark DP method for sequential quantizer [61];
- *SMAWK: DP* accelerated by the SMAWK algorithm [61];
- *Bisection:* The traverse-then-bisect algorithm [62];
- *ML:* Maximum likelihood detector (33) [63].

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

E. Rate Region under Different Configurations

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

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

2) Number of States: Fig. 11(b) shows the relationship between available reflection states (i.e., QAM order) M and achievable rate regions. We notice increasing the reflection states has a marginal effect on both primary and total backscatter rates. This is because once the scope of reflection

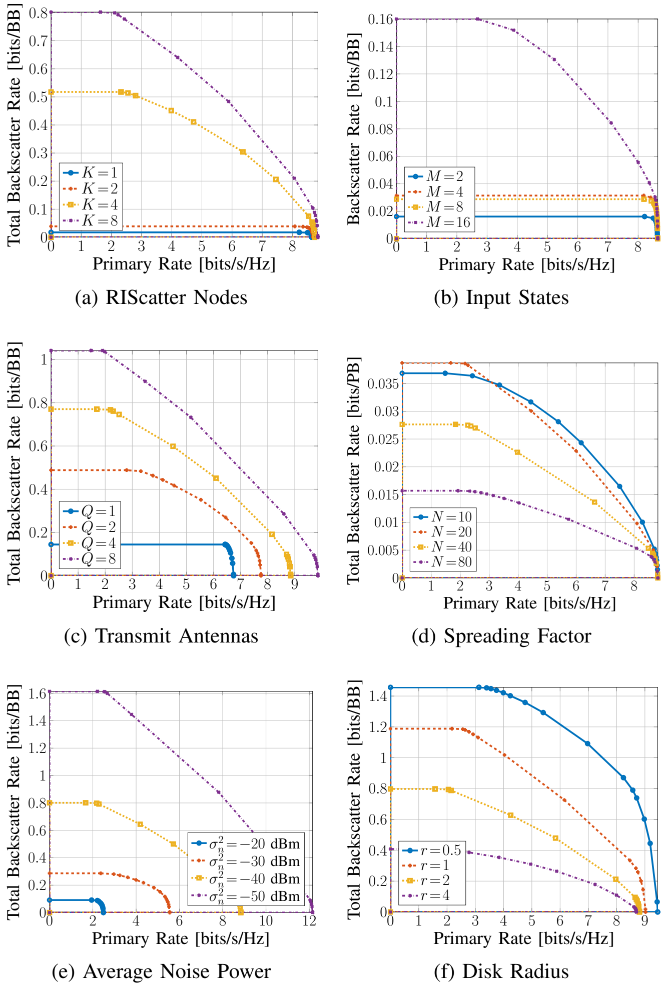


Fig. 11. Average primary-total-backscatter rate regions for different system configurations.

coefficient is determined, using denser constellation points only creates marginal phase resolution and energy diversity for primary and backscatter links. Due to the maximum amplitude normalization in (2), the average rate region of 8-QAM (points on rectangle border) is smaller than that of 4-QAM (points on square vertices), and the 4 points on the rectangle edges are less frequently used. On the other hand, 16-QAM provides significant higher backscatter rate benefit than primary. It motivates one to use high-resolution metamaterial units as RIScatter nodes.

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

4) Spreading factor: Fig. 11(d) presents how the spreading factor N affects the achievable rate region. Please notice here the unit of total backscatter rate is bits per primary block to indicate throughput. Apparently, using a very large N can severely constrain the backscatter throughput, since the gain in energy certainty (by law of large numbers) cannot withstand the loss in gross rate. As $N \rightarrow \infty$, RIScatter nodes boil down to RIS elements with fixed reflection pattern during whole channel block and the total backscatter rate approaches 0. On the other hand, when N is too small, the equivalent DMMAC (9) has very high error probabilities and the energy detection is thus unreliable. It explains the observation that $N = 10$ provides lower backscatter throughput than $N = 20$. Therefore, we conclude the design of spreading factor N in RIScatter systems should accommodate multiple factors (e.g., data rate requirements, load switching constraints at the nodes, and signal processing capability at the user).

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

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

V. Conclusion

This paper introduced RIScatter as a low-power scatter protocol that simultaneously provides backscatter modulation and passive beamforming by smart input distribution design. Starting from scattering principles, we showed how RIScatter nodes generalize information nodes of BackCom and reflect elements of RIS as special cases, how they can be built over existing passive scatter devices, and how they simultaneously encode self information and assist legacy transmission. We also propose a practical SIC-free receiver that exploits the properties of active-passive coexisting networks to maintain primary benefit and improve backscatter throughput. The achievable primary-total-backscatter rate region is then studied for a single-user multi-node RIScatter system, where the input distribution, active beamforming, and decision thresholds are iteratively updated. Numerical results not only validated the proposed algorithms but also emphasized the importance of adaptive input distribution and cooperative decoding for both primary and backscatter links.

One possible future direction is to consider backscatter detection over the received signal domain rather than energy domain, where multi-antenna [67] and learning-based approaches can be promising. Another interesting question is how to design RIScatter node and receiver in a multi-user system to fully exploit the dynamic passive beamforming that naturally origins from backscatter modulation.

Appendix

A. Proof of Proposition 1

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

Lagrangian function of problem (21) is

$$-I(x_{\mathcal{K}})+\sum_{k}\nu_k\Bigg(\sum_{m_k}P_k(x_{m_k})-1\Bigg)-\sum_{k}\sum_{m_k}\lambda_{k,m_k}P_k(x_{m_k}) \tag{40}$$

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

$$-\nabla_{P_k^\star(x_{m_k})}I^\star(x_{\mathcal{K}})+\nu_k^\star-\lambda_{k,m_k}^\star=0, \tag{41a}$$

$$\lambda_{k,m_k}^\star=0, \quad P_k^\star(x_{m_k})>0, \tag{41b}$$

$$\lambda_{k,m_k}^\star\geq 0, \quad P_k^\star(x_{m_k})=0, \tag{41c}$$

where directional derivative is explicitly written as

$$\nabla_{P_k^\star(x_{m_k})}I^\star(x_{\mathcal{K}})=I_k^\star(x_{m_k})-(1-\rho). \tag{42}$$

Combining (41) and (42), we have

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

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

such that

$$\sum_{m_k}P_k^\star(x_{m_k})I_k^\star(x_{m_k})=\nu_k^\star+(1-\rho). \tag{44}$$

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

$$\sum_{m_k}P_k^\star(x_{m_k})I_k^\star(x_{m_k})=I^\star(x_{\mathcal{K}}), \tag{45}$$

where the right-hand side is irrelevant to k. (43), (44), and (45) 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 (46) at the end of page 12. It is straightforward to verify $J\big(P_{\mathcal{K}}(x_{m_{\mathcal{K}}}),P_{\mathcal{K}}(x_{m_{\mathcal{K}}})\big)=I(x_{\mathcal{K}})$ and $J\big(P_{\mathcal{K}}(x_{m_{\mathcal{K}}}),P'_{\mathcal{K}}(x_{m_{\mathcal{K}}})\big)$ is a concave function for a fixed $P'_{\mathcal{K}}(x_{m_{\mathcal{K}}})$. Setting $\nabla_{P_k^\star(x_{m_k})}J\big(P_{\mathcal{K}}(x_{m_{\mathcal{K}}}),P'_{\mathcal{K}}(x_{m_{\mathcal{K}}})\big)=0$, we have

$$S'_k(x_{m_k})-S'_k(x_{i_k})+(1-\rho)\log\frac{P_k(x_{i_k})}{P_k^\star(x_{m_k})}=0, \tag{47}$$

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

$$S'_k(x_{m_k})\triangleq I'_k(x_{m_k})+(1-\rho)\sum_{m_{\mathcal{K}\setminus\{k\}}}P_{\mathcal{K}\setminus\{k\}}(x_{m_{\mathcal{K}\setminus\{k\}}}) \times\sum_{m'_{\mathcal{K}}}P(\hat{x}_{m'_{\mathcal{K}}}|x_{m_{\mathcal{K}}})\log P'_{\mathcal{K}}(x_{m_{\mathcal{K}}}). \tag{48}$$

Evidently, $\forall m_k\neq i_k$, (47) boils down to

$$P_k^\star(x_{m_k})=\frac{P'_k(x_{m_k})\exp\Big(\frac{\rho}{1-\rho}I'_k(x_{m_k})\Big)}{\sum_{m'_k}P'_k(x_{m'_k})\exp\Big(\frac{\rho}{1-\rho}I'_k(x_{m'_k})\Big)}. \tag{49}$$

Since $P_k(x_{i_k})=1-\sum_{m_k\neq i_k}P_k^\star(x_{m_k})$ has exactly the same form as (49), the choice of reference state i_k does not matter and (49) 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 (49) ensures

$$J\big(P_{\mathcal{K}}(x_{m_{\mathcal{K}}}),P'_{\mathcal{K}}(x_{m_{\mathcal{K}}})\big)\geq I'(x_{\mathcal{K}}). \tag{50}$$

On the other hand, we also have

$$\begin{aligned}
\Delta &\triangleq I(x_{\mathcal{K}})-J\big(P_{\mathcal{K}}(x_{m_{\mathcal{K}}}),P'_{\mathcal{K}}(x_{m_{\mathcal{K}}})\big) && \text{(51a)}\\
&=(1-\rho)\sum_{m_k}\frac{P'_k(x_{m_k})f'_k(x_{m_k})}{\sum_{m'_k}P'_k(x_{m'_k})f'_k(x_{m'_k})}\sum_{m''_{\mathcal{K}}}P(\hat{x}_{m''_{\mathcal{K}}}|x_{m_k})\\
&\quad\times\log\frac{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_{\mathcal{K}}}|x_{m'_k})f'_k(x_{m_k})}{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_{\mathcal{K}}}|x_{m'_k})f'_k(x_{m'_k})} && \text{(51b)}\\
&\geq(1-\rho)\sum_{m_k}\frac{P'_k(x_{m_k})f'_k(x_{m_k})}{\sum_{m'_k}P'_k(x_{m'_k})f'_k(x_{m'_k})}\sum_{m''_{\mathcal{K}}}P(\hat{x}_{m''_{\mathcal{K}}}|x_{m_k})\\
&\quad\times\left(1-\frac{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_{\mathcal{K}}}|x_{m'_k})f'_k(x_{m'_k})}{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_{\mathcal{K}}}|x_{m'_k})f'_k(x_{m_k})}\right) && \text{(51c)}\\
&=0, && \text{(51d)}
\end{aligned}$$

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 (49) converges. (50) and (51) 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 $P_k^\star(x_{m_k})$, fulfills KKT conditions (22). To see this, let

$$D_k^{(r)}(x_{m_k})\triangleq\frac{P_k^{(r+1)}(x_{m_k})}{P_k^{(r)}(x_{m_k})}=\frac{f_k^{(r)}(x_{m_k})}{\sum_{m'_k}P_k^{(r)}(x_{m'_k})f_k^{(r)}(x_{m'_k})}. \tag{52}$$

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

$$I_k^\star(x_{m_k})=\frac{1-\rho}{\rho}\log\sum_{m'_k}P_k^\star(x_{m'_k})f_k^\star(x_{m'_k}), \tag{53}$$

which is reminiscent of (43a) and (22a). That is, given $P_k^{(0)}(x_{m_k})>0$, any converging point with $P_k^\star(x_{m_k})>0$ must satisfy (22a). On the other hand, we assume $P_k^\star(x_{m_k})$ does not satisfy (22b), such that for any state with $P_k^\star(x_{m_k})=0$,

$$I_k^\star(x_{m_k})>I^\star(x_{\mathcal{K}})=\sum_{m'_k}P_k^\star(x_{m'_k})I_k^\star(x_{m'_k}), \tag{54}$$

where the equality inherits from (19). Since the exponential function is monotonically increasing, we have

---

$$J\big(P_{\mathcal{K}}(x_{m_{\mathcal{K}}}),P'_{\mathcal{K}}(x_{m_{\mathcal{K}}})\big)\triangleq\sum_{m_{\mathcal{K}}}P_{\mathcal{K}}(x_{m_{\mathcal{K}}})\left(\rho\log\Big(1+\frac{|\boldsymbol{h}^{\mathsf{H}}(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2}{\sigma_v^2}\Big)+(1-\rho)\sum_{m'_{\mathcal{K}}}P(\hat{x}_{m'_{\mathcal{K}}}|x_{m_{\mathcal{K}}})\log\frac{P(\hat{x}_{m'_{\mathcal{K}}}|x_{m_{\mathcal{K}}})P'_{\mathcal{K}}(x_{m_{\mathcal{K}}})}{P'_{\mathcal{K}}(\hat{x}_{m'_{\mathcal{K}}})P_{\mathcal{K}}(x_{m_{\mathcal{K}}})}\right). \tag{46}$$

$f_k^\star(x_{m_k})>\sum_{m'_k} P_k^\star(x_{m'_k}) f_k^\star(x_{m'_k})$ and $D_k^\star(x_{m_k})>1$. Considering $P_k^{(0)}(x_{m_k})>0$ and $P_k^\star(x_{m_k})=0$, it contradicts with

$$P_k^{(r)}(x_{m_k})=P_k^{(0)}(x_{m_k})\prod_{n=1}^{r} D_k^{(n)}(x_{m_k}). \tag{55}$$

That is, given $P_k^{(0)}(x_{m_k}) > 0$, any converging point with $P_k^\star(x_{m_k})=0$ must satisfy (22b). The proof is thus completed.

Acknowledgement

The authors would like to thank Prof. Geoffrey Ye Li for his clinical comments on earlier versions of the manuscript.

References

[1] C. Boyer and S. Roy, "Backscatter communication and RFID: Coding, energy, and MIMO analysis," *IEEE Transactions on Communications*, vol. 62, pp. 770–785, Mar 2014.
[2] D. Dobkin, *The RF in RFID: Passive UHF RFID in Practice*. Newnes, Nov 2012.
[3] J. Landt, "The history of RFID," *IEEE Potentials*, vol. 24, pp. 8–11, Oct 2005.
[4] G. Vannucci, A. Bletsas, and D. Leigh, "A software-defined radio system for backscatter sensor networks," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 2170–2179, Jun 2008.
[5] S. D. Assimonis, S. N. Daskalakis, and A. Bletsas, "Sensitive and efficient RF harvesting supply for batteryless backscatter sensor networks," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, pp. 1327–1338, Apr 2016.
[6] V. Liu, A. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. R. Smith, "Ambient backscatter: Wireless communication out of thin air," *ACM SIGCOMM Computer Communication Review*, vol. 43, pp. 39–50, Sep 2013.
[7] G. Yang, Q. Zhang, and Y.-C. Liang, "Cooperative ambient backscatter communications for green internet-of-things," *IEEE Internet of Things Journal*, vol. 5, pp. 1116–1130, Apr 2018.
[8] Y.-C. Liang, Q. Zhang, E. G. Larsson, and G. Y. Li, "Symbiotic radio: Cognitive backscattering communications for future wireless networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, pp. 1242–1255, Dec 2020.
[9] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Transactions on Communications*, vol. 69, pp. 3313–3351, May 2021.
[10] H. Guo, Y.-C. Liang, R. Long, and Q. Zhang, "Cooperative ambient backscatter system: A symbiotic radio paradigm for passive IoT," *IEEE Wireless Communications Letters*, vol. 8, pp. 1191–1194, Aug 2019.
[11] H. Ding, D. B. da Costa, and J. Ge, "Outage analysis for cooperative ambient backscatter systems," *IEEE Wireless Communications Letters*, vol. 9, pp. 601–605, May 2020.
[12] R. Torres, R. Correia, N. Carvalho, S. N. Daskalakis, G. Goussetis, Y. Ding, A. Georgiadis, A. Eid, J. Hester, and M. M. Tentzeris, "Backscatter communications," *IEEE Journal of Microwaves*, vol. 1, pp. 864–878, Oct 2021.
[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, Feb 2020.
[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, Dec 2019.
[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, Jun 2021.
[16] J. Xu, Z. Dai, and Y. Zeng, "Enabling full mutualism for symbiotic radio with massive backscatter devices," *arXiv:2106.05789*, Jun 2021.
[17] Z. Yang and Y. Zhang, "Optimal SWIPT in RIS-aided MIMO networks," *IEEE Access*, vol. 9, pp. 112 552–112 560, 2021.
[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, Aug 2021.
[19] 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.
[20] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," vol. 18. IEEE, Dec 2018, pp. 1–6.
[21] S. Zhang and R. Zhang, "Capacity characterization for intelligent reflecting surface aided MIMO communication," *IEEE Journal on Selected Areas in Communications*, vol. 38, pp. 1823–1838, Aug 2020.
[22] S. Lin, B. Zheng, F. Chen, and R. Zhang, "Intelligent reflecting surface-aided spectrum sensing for cognitive radio," *IEEE Wireless Communications Letters*, vol. 11, pp. 928–932, May 2022.
[23] Y. Liu, Y. Zhang, X. Zhao, S. Geng, P. Qin, and Z. Zhou, "Dynamic-controlled RIS assisted multi-user MISO downlink system: Joint beamforming design," *IEEE Transactions on Green Communications and Networking*, vol. 6, pp. 1069–1081, Jun 2022.
[24] 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.
[25] 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.
[26] Y. Yang, S. Zhang, and R. Zhang, "IRS-enhanced OFDMA: Joint resource allocation and passive beamforming optimization," *IEEE Wireless Communications Letters*, vol. 9, pp. 760–764, Jun 2020.
[27] Q. Wu, X. Zhou, and R. Schober, "IRS-assisted wireless powered NOMA: Do we really need different phase shifts in DL and UL?" *IEEE Wireless Communications Letters*, vol. 10, pp. 1493–1497, Jul 2021.
[28] Q. Wu, X. Zhou, W. Chen, J. Li, and X. Zhang, "IRS-aided WPCNs: A new optimization framework for dynamic IRS beamforming," *IEEE Transactions on Wireless Communications*, pp. 1–1, Dec 2021.
[29] M. Hua and Q. Wu, "Joint dynamic passive beamforming and resource allocation for IRS-aided full-duplex WPCN," *IEEE Transactions on Wireless Communications*, pp. 1–1, Dec 2021.
[30] W. Tang, J. Y. Dai, M. Chen, X. Li, Q. Cheng, S. Jin, K.-K. Wong, and T. J. Cui, "Programmable metasurface-based RF chain-free 8PSK wireless transmitter," *Electronics Letters*, vol. 55, pp. 417–420, Apr 2019.
[31] J. Y. Dai, W. Tang, L. X. Yang, X. Li, M. Z. Chen, J. C. Ke, Q. Cheng, S. Jin, and T. J. Cui, "Realization of multi-modulation schemes for wireless communication by time-domain digital coding metasurface," *IEEE Transactions on Antennas and Propagation*, vol. 68, pp. 1618–1627, Mar 2020.
[32] R. Karasik, O. Simeone, M. D. Renzo, and S. S. Shitz, "Beyond max-SNR: Joint encoding for reconfigurable intelligent surfaces," vol. 2020-June. IEEE, Jun 2020, pp. 2965–2970.
[33] R. Liu, H. Li, M. Li, and Q. Liu, "Symbol-level precoding design for intelligent reflecting surface assisted multi-user MIMO systems." IEEE, Oct 2019, pp. 1–6.
[34] A. Bereyhi, V. Jamali, R. R. Muller, A. M. Tulino, G. Fischer, and R. Schober, "A single-RF architecture for multiuser massive MIMO via reflecting surfaces." IEEE, May 2020, pp. 8688–8692.
[35] X. Xu, Y.-C. Liang, G. Yang, and L. Zhao, "Reconfigurable intelligent surface empowered symbiotic radio over broadcasting signals," vol. 2020-Janua. IEEE, Dec 2020, pp. 1–6.
[36] Q. Zhang, Y.-C. Liang, and H. V. Poor, "Reconfigurable intelligent surface assisted MIMO symbiotic radio networks," *IEEE Transactions on Communications*, vol. 69, pp. 4832–4846, Jul 2021.
[37] J. Hu, Y. C. Liang, and Y. Pei, "Reconfigurable intelligent surface enhanced multi-user MISO symbiotic radio system," *IEEE Transactions on Communications*, vol. 69, pp. 2359–2371, Apr 2021.
[38] M. Hua, Q. Wu, L. Yang, R. Schober, and H. V. Poor, "A novel wireless communication paradigm for intelligent reflecting surface based symbiotic radio systems," *IEEE Transactions on Signal Processing*, vol. 70, pp. 550–565, Apr 2022.
[39] E. Basar, "Reconfigurable intelligent surface-based index modulation: A new beyond MIMO paradigm for 6G," *IEEE Transactions on Communications*, vol. 68, pp. 3187–3196, May 2020.
[40] T. Ma, Y. Xiao, X. Lei, P. Yang, X. Lei, and O. A. Dobre, "Large intelligent surface assisted wireless communications with spatial modulation and antenna selection," *IEEE Journal on Selected Areas in Communications*, vol. 38, pp. 2562–2574, Nov 2020.
[41] J. Yuan, M. Wen, Q. Li, E. Basar, G. C. Alexandropoulos, and G. Chen, "Receive quadrature reflecting modulation for RIS-empowered wireless communications," *IEEE Transactions on Vehicular Technology*, vol. 70, pp. 5121–5125, May 2021.
[42] S. Hu, C. Liu, Z. Wei, Y. Cai, D. W. K. Ng, and J. Yuan, "Beamforming design for intelligent reflecting surface-enhanced symbiotic radio systems," *arxiv:2110.10316*, Oct 2021.

[43] 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

[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, Nov 2004.

[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, Apr 2011.

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

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

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

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

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

[64] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, Jul 1948.

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

[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, Dec 2010.

[67] W. Liu, S. Shen, D. H. K. Tsang, R. K. Mallik, and R. Murch, "An efficient ratio detector for ambient backscatter communication," *arxiv:2210.09920*, Oct 2022.