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 external electromagnetic waves. Reconfigurable Intelligent Surface (RIS) adapts its phase shift response to alter channel strength in specific directions. In this paper, we show how those two seemingly different technologies (and their derivatives) can be unified into one architecture called RIScatter. RIScatter consists of dispersed or co-located scatter nodes, whose reflection states are adapted to partially modulate their information and partially engineer the wireless channel. The key is to render the probability distribution of reflection states as a joint function of the information source, Channel State Information (CSI), and relative priority of coexisting links. This enables RIScatter to softly bridge BackCom and RIS; reduce to either under specific setup; or evolve in a mixed form for heterogeneous traffic control and universal hardware design. We also propose a low-complexity Successive Interference Cancellation (SIC)-free receiver that exploits the properties of RIScatter. For a single-user multi-node network, we characterize the achievable primary-(total-)backscatter rate region by optimizing the input distribution at scatter nodes, the active beamforming at the Access Point (AP), and the energy decision regions at the user. Simulations demonstrate RIScatter nodes can recycle surrounding radios for backscatter modulation and passive beamforming.**

Index Terms—**Backscatter communication, reconfigurable intelligent surface, active-passive coexisting network, input distribution design, SIC-free receiver.**

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 superimposes 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 shaper* 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]. The achievable rate and optimal input distribution for binary-input AmBC are investigated in [12], but its impact on the primary link is omitted. In [13], the authors analyzed the energy efficiency and achievable rate region for an AmBC-aided multi-user downlink Non-Orthogonal Multiple Access (NOMA) system. However, they assumed equal symbol duration and perfect synchronization for the coexisting 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 *double fading*;
3) The spreading factor (i.e., backscatter symbol duration over primary) is usually large[1].

The second property motivated [7], [8], [10], [11], [15]–[22] to view SR as a multiplicative NOMA and perform Successive Interference Cancellation (SIC) from primary to backscatter link. During primary decoding, the backscatter signal can be

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 [14], accounting for a typical spreading factor between 10 and 10^3.
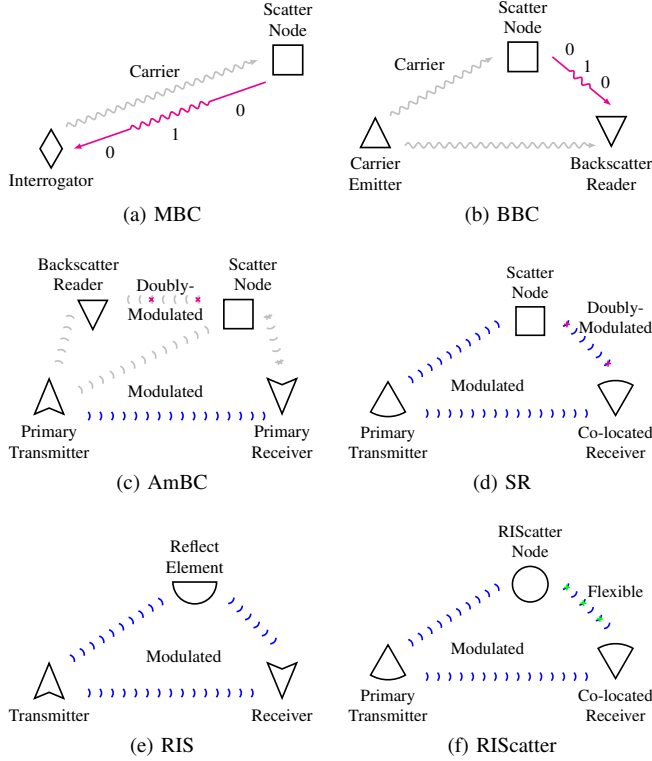
Fig. 1. Illustration of scattering applications. The blue flow(s) constitutes the primary link while the magenta/green flow denotes the backscatter link.

modelled as channel uncertainty or multiplicative interference, when the spreading factor is large or small. Decoding each backscatter symbol also requires multiple SIC followed by a Maximal Ratio Combining (MRC) over primary blocks, which is operation-intensive and CSI-sensitive. Under those assumptions, the achievable rate region of cell-free SR is characterized in [22]. When the spreading factor is sufficiently large, the primary achievable rate under semi-coherent detection[2] asymptotically approaches its coherent counterpart such that both links are approximately interference-free [15]. However, this assumption severely limits the backscatter throughput.

On the other hand, static RIS that employs fixed reflection pattern per channel block has been extensively studied in wireless communication, sensing, and power literature [23]–[28]. Dynamic RIS performs time sharing between different phase shifts and introduces artificial channel diversity within each channel block. The idea was first proposed to fine-tune the Orthogonal Frequency-Division Multiplexing (OFDM) resource blocks [29], then extended to the downlink power and uplink information phases of Wireless Powered Communication Network (WPCN) [30]–[32]. However, dynamic RIS carries no information because the reflection state at a specific time is known to the receiver. RIS can also be used as an information source, and prototypes have been developed for Phase Shift Keying (PSK) [33] and Quadrature Amplitude Modulation (QAM) [34]. From an information-theoretic perspective, the authors of [35] 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 rate-suboptimal. This inspired [36]–[45] to combine passive beamforming and backscatter modulation in the overall RIS design. In particular, *symbol level precoding* maps the information symbols to the optimized RIS coefficient sets [36], [37], *overlay modulation* superposes the information symbols over a common auxiliary matrix [38]–[41], *spatial modulation* switches between the reflection coefficient sets that maximize SNR at different receive antennas [42]–[44], and *index modulation* employs dedicated reflection elements (resp. information elements) for passive beamforming (resp. backscatter modulation) [45]. Those RIS-based backscatter modulation schemes incur advanced hardware architecture and high optimization complexity. In contrast, [46] exploits commodity RFID tags, powered and controlled by a software-defined radio reader at a different frequency, to perform passive beamforming (but no backscatter modulation) towards a legacy user. Most relevant literature considers either Gaussian codebook [10], [11], [15]–[19], [40] that is impractical for low-power nodes, or finite equiprobable inputs [7], [8], [20], [21], [36]–[39], [41]–[45] 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 to simultaneously modulate their information and engineer the wireless channel. A co-located receiver cooperatively decodes both coexisting links. Each reflection state is simultaneously an information codeword and a passive beamforming codeword. The reflection pattern over time is semi-random and guided by the input probability assigned to each state. This probability distribution is carefully designed to incorporate the backscatter information, CSI, and Quality of Service (QoS)[3]. 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. Table I compares RIScatter to BackCom and RIS. However, two major challenges for RIScatter are the receiver design 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 low-complexity SIC-free receiver. It semi-coherently decodes the weak backscatter signal using an energy detector, re-encodes for the exact reflection pattern, then coherently decodes the primary link. Thanks to double modulation, backscatter detection can be viewed as part of channel training, and the impact of backscatter modulation can be modelled as dynamic passive beamforming afterwards. The proposed receiver may be built over legacy receivers with minor hardware upgrade, as it only requires one additional energy comparison and re-encoding per backscatter symbol (instead of primary symbol). The energy detector can also be tailored for arbitrary

---

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

[3]QoS refers to the relative priority of the primary link.

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

input distribution and spreading factor to increase backscatter throughput. This is the first paper to propose a SIC-free cooperative receiver for active-passive coexisting networks.

Third, we address the second challenge in 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 the converging point of a sequence, the active beamforming is optimized by Projected Gradient Ascent (PGA), and the decision regions are refined by state-of-the-art sequential quantizer designs for Discrete Memoryless Thresholding Channel (DMTC). Uniquely, our optimization problem takes into account the CSI, QoS, and backscatter constellation, and the resulting input distribution is applicable to other 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. $\mathrm{diag}(\cdot)$ returns a square matrix with the input vector on its main diagonal and zeros elsewhere. $\mathrm{card}(\cdot)$ returns the cardinality of a set. $\log(\cdot)$ 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 or reflecting are often manipulated by passive antennas or programmable metamaterial [47]. The former receives the impinging signals and reradiates some back to the space, while the latter reflects at the space-cell boundary and mainly applies a phase shift. In the scattered signal, the structural mode component depends on the scatterer geometry and material. Its impact is usually modelled as part of environment multipath [8], [48], or simply a baseband Direct Current (DC) offset when the impinging signal is a Continuous Waveform (CW) [1]. On the other hand, the antenna mode component depends on the impedance mismatch and is widely exploited in scattering applications. For an antenna (resp. metamaterial) scatterer with M reflection states, the reflection coefficient at state $m \in \mathcal{M} \triangleq \{1, ..., M\}$ is [47], [49]

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

where Z_m is the antenna load (resp. metamaterial cell) impedance at state m, and Z is the antenna input (resp. medium characteristic) impedance. Specifically,

- BackCom: The scatterer is an information source with random reflection pattern over time. The reflection coefficient is used merely as an information codeword [50]

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

where $\alpha_m \in \mathbb{I}$ is the amplitude scattering ratio at state m, and c_m is the corresponding constellation point.

- RIS: The scatterer is a channel shaper with deterministic reflection pattern over time. The reflection coefficient is used merely as a passive beamforming codeword [23]

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

where θ_m is the phase shift at state m.[4]

---

[4]Most papers assume $\alpha_m = \alpha$, with $\alpha \ll 1$ for BackCom and $\alpha = 1$ for RIS.

(a) Input Distribution

(b) Reflection Pattern

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 corresponds to state 2. The spreading factor is 4 for RIScatter and 8 for AmBC/SR. BackCom and RIS can be viewed as extreme cases of RIScatter, where the input distribution boils down to uniform and degenerate, respectively.

RIScatter generalizes BackCom and RIS from a probabilistic perspective. Each reflection coefficient simultaneously acts as an information codeword and a passive beamforming codeword. As shown in Fig. 2, the reflection pattern of each RIScatter node over time is semi-random and guided by the input probability assigned to each state. This probability distribution is carefully designed to incorporate the backscatter information, CSI, and QoS, in order to strike a balance between backscatter modulation and passive beamforming.

Remark 1. *Unlike dynamic RIS that simply performs a time sharing between reflection states, RIScatter conveys additional information by randomizing the reflection pattern over time while still guaranteeing the probability of occurrence of each state. Upon successful backscatter detection, the impact of RIScatter nodes on the primary link can be modelled as dynamic passive beamforming.*

RIScatter nodes can be implemented, for example, by adding an integrated receiver[5] [51] and adaptive encoder [52] to off-the-shelf passive RFID tags. The block diagram, equivalent circuit, and scatter model are illustrated in Fig. 3.

B. System Model

As shown in Fig. 4, we consider a RIScatter network where a Q-antenna AP serves a single-antenna user and K nearby dispersed or co-located RIScatter nodes. Without loss of generality, we assume 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[6] enhanced by RIScatter nodes. In the backscatter multiple access system, the AP acts as a carrier emitter, the RIScatter nodes modulate over the scattered signal, and the user jointly decodes all nodes.[7] For simplicity, we consider a quasi-static block fading model and focus on a specific channel block where the CSI remains constant. Denote the AP-user direct channel as $\boldsymbol{h}_\mathrm{D}^\mathsf{H} \in \mathbb{C}^{1 \times Q}$, the AP-node $k \in \mathcal{K} \triangleq \{1,...,K\}$ forward channel as $\boldsymbol{h}_{\mathrm{F},k}^\mathsf{H} \in \mathbb{C}^{1 \times Q}$, the node k-user backward channel as $h_{\mathrm{B},k}$, and the cascaded AP-node k-user channel as $\boldsymbol{h}_{\mathrm{C},k}^\mathsf{H} \triangleq h_{\mathrm{B},k} \boldsymbol{h}_{\mathrm{F},k}^\mathsf{H} \in \mathbb{C}^{1 \times Q}$. We assume the direct and cascaded CSI are available at the AP and user[8].

Let $\alpha_k \in \mathbb{I}$ be the amplitude scattering ratio of node k, $x_k \in \mathcal{X} \triangleq \{c_1,...,c_M\}$ be the (coded) backscatter symbol of node k, and $x_\mathcal{K} \triangleq (x_1,...,x_K)$ be the backscatter symbol tuple of all nodes. Due to double modulation, the composite channel is a function of backscatter symbol tuple[9]

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

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

where $\boldsymbol{\alpha} \triangleq [\alpha_1,...,\alpha_K]^\mathsf{T} \in \mathbb{I}^K$, $\boldsymbol{x} \triangleq [x_1,...,x_K]^\mathsf{H} \in \mathcal{X}^K$, and $\boldsymbol{H}_\mathrm{C} \triangleq [\boldsymbol{h}_{\mathrm{C},1},...,\boldsymbol{h}_{\mathrm{C},K}]^\mathsf{H} \in \mathbb{C}^{K \times Q}$. Without loss of generality, we assume the spreading factor N is a positive integer. Within one backscatter block, the signal received by the user at primary block $n \in \mathcal{N} \triangleq \{1,...,N\}$ 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 $\|\boldsymbol{w}\|^2 \leq P$, P is the maximum average transmit power, $s \sim \mathcal{CN}(0,1)$ is the primary symbol, and $v \sim \mathcal{CN}(0,\sigma_v^2)$ is the Additive White Gaussian Noise (AWGN) with variance σ_v^2.

Let $m_k \in \mathcal{M} \triangleq \{1,...,M\}$ be the reflection state index of node k, and $m_\mathcal{K} \triangleq (m_1,...,m_K)$ be the state index tuple of all nodes. The backscatter symbol x_k (resp. symbol tuple $x_\mathcal{K}$) is a random variable that takes value x_{m_k} (resp. value tuple $x_{m_\mathcal{K}}$) with probability $p(x_{m_k})$ (resp. $p(x_{m_\mathcal{K}})$).

Remark 2. *Dispersed RIScatter nodes encode independently such that*

$$p(x_{m_\mathcal{K}}) = \prod_k p(x_{m_k}). \tag{6}$$

When multiple nodes are co-located, they can jointly encode by designing the joint probability $p(x_{m_\mathcal{K}})$ directly.

Let $z = \sum_n |y[n]|^2$ be the receive energy per backscatter block. When $x_{m_\mathcal{K}}$ is transmitted, the receive signal y follows CSCG distribution $\mathcal{CN}(0,\sigma_{m_\mathcal{K}}^2)$ with variance

$$\sigma_{m_\mathcal{K}}^2 = |\boldsymbol{h}^\mathsf{H}(x_{m_\mathcal{K}}) \boldsymbol{w}|^2 + \sigma_v^2, \tag{7}$$

---

[5]The aim is to coordinate the node with the active source and to acquire the optimized input distribution, instead of decoding the primary information dedicated for the user. The node receiver can be implemented using simple circuits or even integrated with the rectifier [51] to reduce cost and complexity.

[6]It is assumed the primary symbol duration is much longer than multipath delay spread (i.e., no inter-symbol interference).

[7]It is assumed the signal going through two or more RIScatter nodes is too weak to be received by the user.

[8]The cascaded CSI can be estimated by sequential [54]–[56] or parallel [57] approaches for dispersed nodes, or group-based [58] or hierarchical [59] approaches for co-located nodes. The impact of channel estimation error will be investigated in Section IV.

[9](4a) and (4b) are often used in BackCom and RIS literature, respectively.

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 $\vec{E}_{\text{I}}$ and magnetic field $\vec{H}_{\text{I}}$ [53].
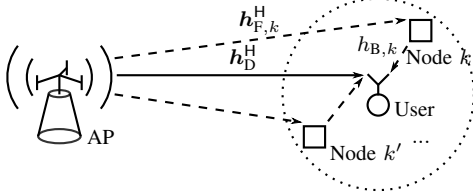

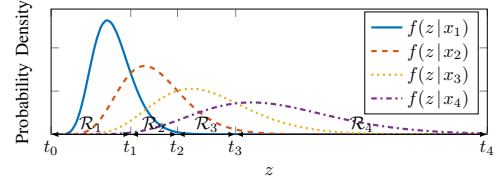Fig. 4. A single-user multi-node RIScatter network.

Fig. 5. PDF of the receive energy per backscatter block conditioned on different reflection state.

and z follows Gamma distribution with conditional Probability Density Function (PDF)

$$f(z|x_{m_{\mathcal{K}}}) = \frac{z^{N-1}\exp(-z/\sigma_{m_{\mathcal{K}}}^2)}{\sigma_{m_{\mathcal{K}}}^{2N}(N-1)!}. \tag{8}$$

Remark 3. *We have assumed Gaussian codebook for the primary source and finite support for the backscatter nodes, since they are relatively practical and widely adopted in relevant literatures. The proposed framework is extendable to non-Gaussian primary source, and the conditional PDF* (8) *can be approximated using Central Limit Theorem (CLT) for large N.*

The user first jointly decodes the backscatter message of all nodes using a low-complexity energy detector.[10] The energy detector formulates a DMTC of size $M^K \times M^K$.

Remark 4. *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 regions (consist of non-adjacent partitions) are generally required and the optimal number of thresholds is unknown [61], [62]. Next, we restrict the energy detector to convex deterministic decision regions and consider sequential threshold design.*

Let $L = M^K$ be the number of decision regions. Sort $\{\sigma_{m_{\mathcal{K}}}^2\}$ in ascending order and denote the result sequence as $\sigma_1^2,...,\sigma_L^2$. With sequential thresholding, the decision region of backscatter symbol tuple $l \in \mathcal{L} \triangleq \{1,...,L\}$ is[11]

$$\mathcal{R}_l \triangleq [t_{l-1}, t_l), \quad 0 \leq t_{l-1} \leq t_l, \tag{9}$$

[10]The reliability of the energy detector is improved by the adaptive input distribution and thresholding design. With high-order modulation or large number of scatter nodes, the reliability can be further enhanced by increasing the spreading factor or using error correction codes with low code-rate. In practice, users can decode backscatter nodes ranging from a few to hundreds of meters in the presence of noise and interference, and the backscatter throughput can reach few Kbps to tens of Mbps [60].

[11]$m_{\mathcal{K}}$ and l are one-to-one mapped. Both notations are used interchangeably in the following context.

where t_l is the decision threshold between hypotheses x_l and x_{l+1}. An example is shown in Fig. 5.

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

$$q(\hat{x}_{m'_{\mathcal{K}}}|x_{m_{\mathcal{K}}}) = \int_{\mathcal{R}_{m'_{\mathcal{K}}}} f(z|x_{m_{\mathcal{K}}})dz. \tag{10}$$

The backscatter mutual information is

$$I_{\text{B}}(x_{\mathcal{K}};\hat{x}_{\mathcal{K}}) = \sum_{m_{\mathcal{K}}} p(x_{m_{\mathcal{K}}})I_{\text{B}}(x_{m_{\mathcal{K}}};\hat{x}_{\mathcal{K}}), \tag{11}$$

where $I_{\text{B}}(x_{m_{\mathcal{K}}};\hat{x}_{\mathcal{K}})$ is the backscatter information function

$$I_{\text{B}}(x_{m_{\mathcal{K}}};\hat{x}_{\mathcal{K}}) \triangleq \sum_{m'_{\mathcal{K}}} q(\hat{x}_{m'_{\mathcal{K}}}|x_{m_{\mathcal{K}}})\log\frac{q(\hat{x}_{m'_{\mathcal{K}}}|x_{m_{\mathcal{K}}})}{p(\hat{x}_{m'_{\mathcal{K}}})}. \tag{12}$$

Once the backscatter information is successfully decoded, the user re-encodes to recover the reflection pattern, constructs the composite channel by (4), then coherently decodes the primary link. The primary mutual information is

$$I_{\text{P}}(s;y|x_{\mathcal{K}}) = \sum_{m_{\mathcal{K}}} p(x_{m_{\mathcal{K}}})I_{\text{P}}(s;y|x_{m_{\mathcal{K}}}), \tag{13}$$

where $I_{\text{P}}(s;y|x_{m_{\mathcal{K}}})$ is the primary information function

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

III. RATE-REGION CHARACTERIZATION

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

$$I(x_{\mathcal{K}}) \triangleq \rho I_{\text{P}}(s;y|x_{\mathcal{K}}) + (1-\rho)I_{\text{B}}(x_{\mathcal{K}};\hat{x}_{\mathcal{K}}), \tag{15}$$

$$I(x_{m_{\mathcal{K}}}) \triangleq \rho I_{\text{P}}(s;y|x_{m_{\mathcal{K}}}) + (1-\rho)I_{\text{B}}(x_{m_{\mathcal{K}}};\hat{x}_{\mathcal{K}}), \tag{16}$$

where $\rho \in \mathbb{I}$ is the QoS. To obtain the achievable primary-(total-)backscatter rate region, we consider the weighted sum mutual information maximization problem with independent encoding at all nodes[12]

$$\max_{\{p_k\}_{k\in\mathcal{K}}, w, t} I(x_{\mathcal{K}}) \tag{17a}$$

$$\text{s.t.} \quad \mathbf{1}^{\mathsf{T}} p_k = 1, \quad \forall k, \tag{17b}$$

$$p_k \geq 0, \quad \forall k, \tag{17c}$$

$$\|w\|^2 \leq P, \tag{17d}$$

$$t_{l-1} \leq t_l, \quad \forall l, \tag{17e}$$

$$t \geq 0, \tag{17f}$$

where $p_k = [p(x_{1_k}),...,p(x_{M_k})]^{\mathsf{T}} \in \mathbb{I}^M$ is the input distribution of node k. Problem (17) generalizes BackCom by allowing CSI- and QoS-adaptive input distribution and decision region design. On the other hand, it also relaxes the feasible domain of discrete RIS phase shift selection problems from the vertices of M-dimensional probability simplex to the simplex itself. The original problem is highly non-convex and we propose a BCD algorithm that iteratively updates $\{p_k\}_{k\in\mathcal{K}}$, w and t.

A. Input Distribution

For any given w and t, we can formulate a DMMAC by (10) and simplify (17) to

$$\max_{\{p_k\}_{k\in\mathcal{K}}} I(x_{\mathcal{K}}) \tag{18a}$$

$$\text{s.t.} \quad (17b),(17c), \tag{18b}$$

which involves the product term (6) and is generally non-convex (unless $K = 1$). Following [63], we first recast the KKT conditions to their equivalent forms, then propose a numerical solution that guarantees those conditions on the converging point of a sequence.

Proposition 1. *The KKT optimality conditions for problem (18) are equivalent to, $\forall k, m_k$,*

$$I_k^\star(x_{m_k}) = I^\star(x_{\mathcal{K}}), \quad \text{if } p^\star(x_{m_k}) > 0, \tag{19a}$$

$$I_k^\star(x_{m_k}) \leq I^\star(x_{\mathcal{K}}), \quad \text{if } p^\star(x_{m_k}) = 0, \tag{19b}$$

where $I_k(x_{m_k})$ is the weighted sum marginal information

$$I_k(x_{m_k}) \triangleq \sum_{m_{\mathcal{K}\setminus\{k\}}} p(x_{m_{\mathcal{K}\setminus\{k\}}}) I(x_{m_{\mathcal{K}}}). \tag{20}$$

Proof. Please refer to Appendix A. □

For each RIScatter node, (19a) suggests each probable state should produce the same marginal information (averaged over all states of other nodes), while (19b) suggests any state with potentially less marginal information should not be used.

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

$$p^{(r+1)}(x_{m_k}) = \frac{p^{(r)}(x_{m_k})\exp\left(\frac{\rho}{1-\rho}I_k^{(r)}(x_{m_k})\right)}{\sum_{m'_k} p^{(r)}(x_{m'_k})\exp\left(\frac{\rho}{1-\rho}I_k^{(r)}(x_{m'_k})\right)}, \tag{21}$$

[12]Joint encoding over multiple nodes can be viewed as its special case with an augmented backscatter source.

Algorithm 1: Input Distribution Evaluation by a Sequence

Input: K, N, $h_{\mathrm{D}}^{\mathsf{H}}$, H_{C}, α, $\mathcal{X}$, σ_v^2, ρ, w, t, ϵ
Output: $\{p_k^\star\}_{k\in\mathcal{K}}$
1: Set $h^{\mathsf{H}}(x_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}$ by (4)
2: $\sigma_{m_{\mathcal{K}}}^2$, $\forall m_{\mathcal{K}}$ by (7)
3: $f(z|x_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}$ by (8)
4: $q(\hat{x}_{m'_{\mathcal{K}}}|x_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}, m'_{\mathcal{K}}$ by (10)
5: Initialize $r \leftarrow 0$
6:   $p_k^{(0)} > 0$, $\forall k$
7: Get $p^{(r)}(x_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}$ by (6)
8: $I^{(r)}(x_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}$ by (12), (14), (16)
9: $I_k^{(r)}(x_{m_k})$, $\forall k, m_k$ by (20)
10: $I^{(r)}(x_{\mathcal{K}})$ by (11), (13), (15)
11: **Repeat**
12: Update $r \leftarrow r+1$
13:   $p_k^{(r)}$, $\forall k$ by (21)
14: Redo step 7–10
15: **Until** $I^{(r)}(x_{\mathcal{K}}) - I^{(r-1)}(x_{\mathcal{K}}) \leq \epsilon$

---

where r is the iteration index.

Proof. Please refer to Appendix B. □

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

Remark 5. *The insufficiency of the KKT conditions for problem (18) implies that the proposed method may not converge to the global-optimal solution. However, simulation results in Section IV will show that the average performance gap is indistinguishable at a moderate K.*

B. Active Beamforming

For any given $\{p_k\}_{k\in\mathcal{K}}$ and t, problem (17) reduces to

$$\max_{w} I(x_{\mathcal{K}}) \tag{22a}$$

$$\text{s.t.} \quad (17d), \tag{22b}$$

which is still non-convex due to the integration and entropy terms. Note the DMMAC $q(x_l|x_{m_{\mathcal{K}}})$ depends on the variance of accumulated receive energy $\sigma_{m_{\mathcal{K}}}^2$, which is a function of w. Plugging (8) into (10), we have

$$q(x_l|x_{m_{\mathcal{K}}}) = \frac{\int_{t_{l-1}/\sigma_{m_{\mathcal{K}}}^2}^{t_l/\sigma_{m_{\mathcal{K}}}^2} z^{N-1}\exp(-z)dz}{(N-1)!} \tag{23}$$

$$= Q\left(N, \frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}}^2}, \frac{t_l}{\sigma_{m_{\mathcal{K}}}^2}\right), \tag{24}$$

where $Q(N, b_1, b_2) \triangleq \int_{b_1}^{b_2} z^{N-1}\exp(-z)/(N-1)!\, dz$ is the regularized incomplete Gamma function. Its series expansion is given by [64, Theorem 3]

$$Q\left(N, \frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}}^2}, \frac{t_l}{\sigma_{m_{\mathcal{K}}}^2}\right) = \exp\left(-\frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}}^2}\right)\sum_{n=0}^{N-1}\frac{\left(\frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}}^2}\right)^n}{n!}$$

$$-\exp\left(-\frac{t_l}{\sigma_{m_{\mathcal{K}}}^2}\right)\sum_{n=0}^{N-1}\frac{\left(\frac{t_l}{\sigma_{m_{\mathcal{K}}}^2}\right)^n}{n!}, \tag{25}$$

whose gradient with respect to $\boldsymbol{w}^*$ is

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

where

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

On top of (25) and (26), we rewrite $I(x_{\mathcal{K}})$ and $\nabla_{\boldsymbol{w}^*}I(x_{\mathcal{K}})$ as (28) and (29) at the end of page 7, respectively. Problem (22) can thus be solved by the PGA method. At iteration $r+1$, the unregulated active beamformer is updated by

$$\bar{\boldsymbol{w}}^{(r+1)}=\boldsymbol{w}^{(r)}+\gamma\nabla_{\boldsymbol{w}^*}I^{(r)}(x_{\mathcal{K}}), \quad (30)$$

where γ is the step size (refinable by backtracking line search [65, Section 9.2]). Then, $\bar{w}$ is projected onto the feasible domain (17d) to retrieve the active beamformer

$$\boldsymbol{w}=\frac{\sqrt{P}\bar{\boldsymbol{w}}}{\max(\sqrt{P},\|\bar{\boldsymbol{w}}\|)}. \quad (31)$$

The PGA active beamforming design is summarized in Algorithm 2.

C. Decision Threshold

For any given $\{\boldsymbol{p}_k\}_{k\in\mathcal{K}}$ and $\boldsymbol{w}$, problem (17) reduces to

$$\max_{\boldsymbol{t}} \quad I(x_{\mathcal{K}}) \quad (32a)$$
$$\text{s.t.} \quad (17e),(17f), \quad (32b)$$

which is still non-convex since $\boldsymbol{t}$ appears on the limits of integration (10). Instead of solving it directly, we constrain the feasible domain from continuous space $\mathbb{R}_+^{L+1}$ to discrete candidates (i.e., fine-grained energy levels) $\mathcal{T}^{L+1}$. As shown in Fig. 6, the decision regions are formulated by grouping adjacent energy bins.

Remark 6. *The design of the energy detector does not affect the primary achievable rate, since the composite channel (4) can always be determined after backscatter decoding and re-encoding. This implies that any thresholding maximizing the total backscatter rate is optimal for problem (32).*

Algorithm 2: Active Beamforming Optimization by PGA

Input: Q, N, $\boldsymbol{h}_{\mathsf{D}}^{\mathsf{H}}$, $\boldsymbol{H}_{\mathsf{C}}$, $\boldsymbol{\alpha}$, $\mathcal{X}$, 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: $\quad p(x_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}$ by (6)
3: Initialize $r\leftarrow 0$
4: $\quad \boldsymbol{w}^{(0)}$, $\|\boldsymbol{w}^{(0)}\|^2\leq P$
5: Get $(\sigma_{m_{\mathcal{K}}}^{(r)})^2$, $\forall m_{\mathcal{K}}$ by (7)
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: $\quad\quad \bar{\boldsymbol{w}}^{(r)}$ by (30)
14: $\quad\quad \boldsymbol{w}^{(r)}$ by (31)
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** $\|\boldsymbol{w}^{(r)}-\boldsymbol{w}^{(r-1)}\|\leq\epsilon$

Remark 7. *In terms of total backscatter rate, the nodes can be viewed as an augmented source, and problem (32) becomes the rate-optimal quantizer design for DMTC.*

Thanks to Remark 6 and 7, problem (32) can be recast as

$$\max_{\boldsymbol{t}\in\mathcal{T}^{L+1}} \quad I_{\mathsf{B}}(x_{\mathcal{K}};\hat{x}_{\mathcal{K}}) \quad (33a)$$
$$\text{s.t.} \quad (17e), \quad (33b)$$

whose global optimal solution has been obtained in recent works. [66] 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 $\mathcal{O}(L^2(\text{card}(\mathcal{T})-L))$. On the other hand, [67] started from the optimality condition for three neighbor thresholds and presented a traverse-then-bisect algorithm with complexity $\mathcal{O}(\text{card}(\mathcal{T})L\log(\text{card}(\mathcal{T})L))$. In Section IV, both schemes will be compared with the

---

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

$$\nabla_{\boldsymbol{w}^*}I(x_{\mathcal{K}})=\sum_{m_{\mathcal{K}}}p(x_{m_{\mathcal{K}}})\Bigg(\rho\frac{\boldsymbol{h}(x_{m_{\mathcal{K}}})\boldsymbol{h}^{\mathsf{H}}(x_{m_{\mathcal{K}}})\boldsymbol{w}}{\sigma_{m_{\mathcal{K}}}^2}+(1-\rho)\sum_l\Bigg(\log\frac{Q\Big(N,\frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}}^2},\frac{t_l}{\sigma_{m_{\mathcal{K}}}^2}\Big)}{\sum_{m_{\mathcal{K}}'}p(x_{m_{\mathcal{K}}'})Q\Big(N,\frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}'}^2},\frac{t_l}{\sigma_{m_{\mathcal{K}}'}^2}\Big)}+1\Bigg)$$
$$\times\nabla_{\boldsymbol{w}^*}Q\Big(N,\frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}}^2},\frac{t_l}{\sigma_{m_{\mathcal{K}}}^2}\Big)-\frac{Q\Big(N,\frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}}^2},\frac{t_l}{\sigma_{m_{\mathcal{K}}}^2}\Big)\sum_{m_{\mathcal{K}}'}p(x_{m_{\mathcal{K}}'})\nabla_{\boldsymbol{w}^*}Q\Big(N,\frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}'}^2},\frac{t_l}{\sigma_{m_{\mathcal{K}}'}^2}\Big)}{\sum_{m_{\mathcal{K}}'}p(x_{m_{\mathcal{K}}'})Q\Big(N,\frac{t_{l-1}}{\sigma_{m_{\mathcal{K}}'}^2},\frac{t_l}{\sigma_{m_{\mathcal{K}}'}^2}\Big)}\Bigg) \quad (29)$$
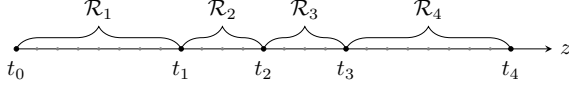
Fig. 6. The thresholds are chosen from fine-grained candidates instead of the continuous space. Each decision region consists of at least one bin.

Maximum-Likelihood (ML) scheme [68]

$$t_l^{\mathrm{ML}} = N \frac{\sigma_{l-1}^2 \sigma_l^2}{\sigma_{l-1}^2 - \sigma_l^2} \log \frac{\sigma_{l-1}^2}{\sigma_l^2}, \quad l \in \mathcal{L} \setminus \{L\}, \tag{34}$$

which is suboptimal for problem (32) unless all nodes are with equiprobable inputs.

IV. SIMULATION RESULTS

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

$$L(d) = L_0 \left(\frac{d_0}{d} \right)^\gamma, \tag{35}$$

together with a Rician fading model

$$\boldsymbol{H} = \sqrt{\frac{\kappa}{1+\kappa}} \bar{\boldsymbol{H}} + \sqrt{\frac{1}{1+\kappa}} \tilde{\boldsymbol{H}}, \tag{36}$$

where d is the transmission distance, $L_0 = -30\,\mathrm{dB}$ is the reference path loss at $d_0 = 1\mathrm{m}$, 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 independent and identically distributed (i.i.d.) CSCG entries. 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 decision threshold domain $\mathcal{T}$ is obtained by b-bit uniform discretization over the critical interval defined by the $1 - \varepsilon$ confidence bounds of edge hypotheses (i.e., lower bound of x_1 and upper bound of x_L). We set $b = 9$ and $\varepsilon = 10^{-3}$. All achievable rate regions are averaged over 10^3 channel realizations.

A. Evaluation of Proposed Algorithms

1) Initialization: To characterize the achievable rate region, we progressively obtain all boundary points by successively increasing ρ and solving problem (17). 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 $\sum_k \boldsymbol{h}_{\mathrm{C},k}^{\mathsf{H}}$, respectively. At the following points, both algorithms are initialized by the solutions at the previous point.

2) Convergence: The BCD algorithm is convergent for problem (17) since the input distribution and active beamforming subproblems converge and the thresholding subproblem attains global optimality. In company with BCD, we also plotted the convergence results of KKT and PGA algorithms in Fig. 7 to show how much performance is gained by solving each
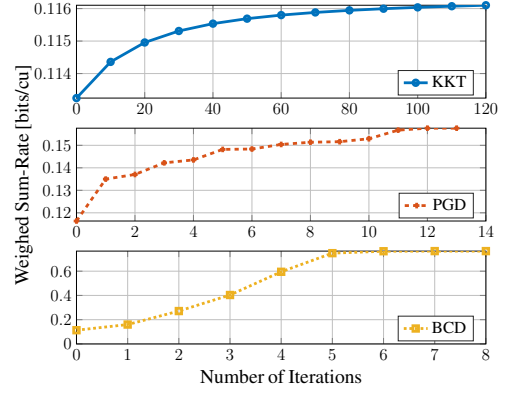

Fig. 7. Typical convergence curves at $\rho = 0$ for $Q = 4$, $K = 8$, $M = 2$, $N = 20$, $\sigma_v^2 = -40\mathrm{dBm}$ and $r = 2\mathrm{m}$.

subproblem. It is observed that Algorithm 1 and 2 take around 100 and 10 iterations to converge, respectively. Overall, the BCD algorithm requires at most 5 iterations to converge. As ρ increases (not presented here), the convergence of all three algorithms are much faster thanks to the progressive initialization.

B. Comparison of Scattering Applications

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

- *Legacy:* Legacy transmission without scatter nodes.
- *BBC:* The primary signal is CW and the receive signal is

$$y^{\mathrm{BBC}}[n] = \left(\boldsymbol{h}_{\mathrm{D}}^{\mathsf{H}} + \sum_k \alpha_k \boldsymbol{h}_{\mathrm{C},k}^{\mathsf{H}} x_k \right) \boldsymbol{w} + v[n]. \tag{37}$$

  The total backscatter rate approaches $K \log M$ when N is sufficiently large.
- *AmBC:* The user decodes each link independently and semi-coherently while treating the other as interference. The primary achievable rate is approximately[13]

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

  while the total backscatter rate follows (11) with uniform input distribution.
- *SR:* For a sufficiently large N, the average primary rate under semi-coherent detection asymptotically approaches (13) with uniform input distribution [15]. 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

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

  The total backscatter rate approaches $K \log M$.

---

[13]The scattered component is treated as interference with average power $\mathbb{E}\{\sum_k \alpha_k \boldsymbol{h}_{\mathrm{C},k}^{\mathsf{H}} x_k \boldsymbol{w} s[n]\} = \sum_k |\alpha_k \boldsymbol{h}_{\mathrm{C},k}^{\mathsf{H}} \boldsymbol{w}|^2$ [15].
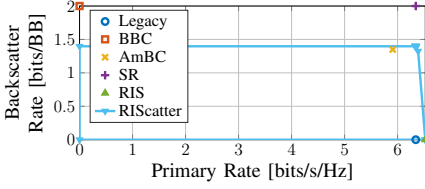
Fig. 8. Typical achievable rate region/points of scattering applications for $Q=1$, $K=1$, $M=4$, $N=10^3$, $\sigma_v^2=-40$dBm and $r=2$m.

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.

- *RIS:* The reflection pattern is deterministic and the total backscatter rate is zero. The primary achievable rate is a special case of (13)

$$I_{\mathrm{P}}^{\mathrm{RIS}}(s;y|x_{\mathcal{K}}) = I_{\mathrm{P}}(s;y|x_{m_{\mathcal{K}}^\star}) = \log\Big(1 + \frac{|\boldsymbol{h}^{\mathrm{H}}(x_{m_{\mathcal{K}}^\star})\boldsymbol{w}|^2}{\sigma_v^2}\Big), \tag{40}$$

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 those strategies. *First,* we observe BBC and SR achieve the best backscatter performance thanks to the coherent decoding. For SR, this comes with the cost of N re-encoding, precoding, subtraction together with a time-domain MRC per backscatter symbol. Since SR requires a very large N to guarantee the primary rate, the signal processing cost at the receiver can be prohibitive, and the backscatter *throughput* can be severely constrained. *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 around 25 dB weaker than the direct channel. Here, RIS ensures a constructive superposition of the 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 stable multipath, and the SR point should move vertically 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 by using the static reflection pattern that maximizes the primary SNR all the time. On the other hand, its maximum backscatter achievable rate is higher than that of AmBC. This is because the adaptive channel coding of RIScatter outperforms the equiprobable inputs of AmBC, especially at the low backscatter SNR caused by double fading. When multiple antenna is available at the AP, active beamforming can be optimized for RIScatter nodes and the advantage over AmBC should be more prominent.

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 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
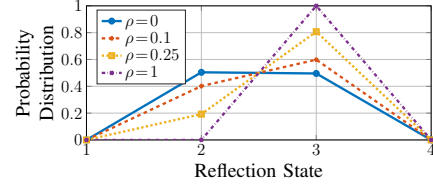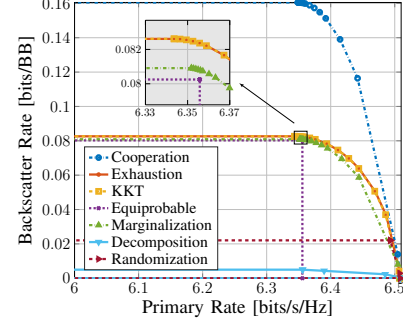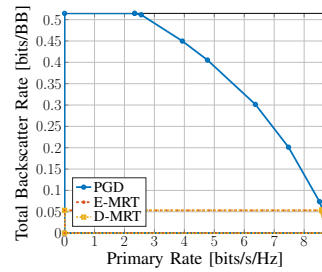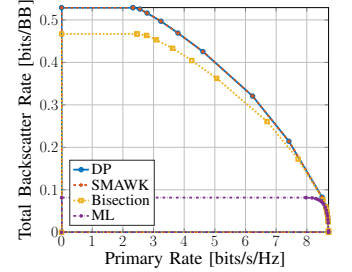

(a) Input Distribution, $Q=1$

(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.

on the other two. This is inline with Shannon's observation that binary antipodal inputs is good enough for channel capacity at low SNR [69]. When the scattered signal is relatively weak, 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, while the middles cannot provide enough energy diversity and end up unused. At $\rho = 1$ where the primary link is prioritized, the optimal input distribution is $[0,0,1,0]^{\mathsf{T}}$ since state 3 provides higher primary SNR than other states. That is, the reflection pattern becomes deterministic and the RIScatter node boils down to a static discrete RIS element. Increasing ρ from 0 to 1 creates a smooth transition from backscatter modulation to passive beamforming, suggesting 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 (18):

- *Cooperation:* Joint encoding using a joint probability array $p(x_{m_{\mathcal{K}}})$ with M^K entries by Algorithm 1;

- *Exhaustion:* Exhaustive search over the M-dimensional probability simplex with resolution $\Delta p = 10^{-2}$;
- *KKT:* Solution by Algorithm 1;
- *Equiprobable:* Uniform input distribution.

We also consider these independent distribution recovery methods from the joint probability array:

- *Marginalization:* Marginal probability distributions;
- *Decomposition:* Normalized rank-1 Canonical Polyadic (CP) decomposed tensors by `Tensor Toolbox` [70];
- *Randomization:* Gaussian randomization with the guidance of correlation matrix [71].

Fig. 10(a) shows their average achievable rate regions. Cooperation achieves the outer bound of all schemes, but joint encoding over passive devices may incur additional hardware cost. Besides, the average rate performance of Exhaustion and KKT completely coincide with each other when $K = 2$. This confirms Remark 5 that the KKT input distribution can be good enough when K is moderate. Equiprobable experiences minor backscatter and major primary rate losses without exploiting CSI and QoS, and those gaps should be larger when M or K increases. Marginalization provides a close performance to KKT, but Randomization and Decomposition fail our expectations for most channel realizations. Those observations emphasize the importance of 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 (22):

- *PGA:* Solution by Algorithm 2;
- *E-MRT:* MRT towards the ergodic composite channel $\sum_{m_{\mathcal{K}}} p(x_{m_{\mathcal{K}}}) \boldsymbol{h}^{\mathsf{H}}(x_{m_{\mathcal{K}}})$;
- *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 PGA beamformer significantly outperforms both MRT schemes in terms of total backscatter rate. This is because the semi-coherent backscatter decoding relies on the relative energy difference under different backscatter symbol tuples. Such an energy diversity is enhanced by PGA that effectively exploits backscatter constellation and input distribution knowledge rather than simply maximizes the channel strength. As ρ increases, the primary SNR outweighs the backscatter energy difference in (28), and PGA beamformer approaches E-MRT. At $\rho = 1$, both PGA and E-MRT boil down to MRT towards the strongest composite channel. The difference between E-MRT and D-MRT is insignificant when RIScatter nodes are dispersed. Those observations confirm that the proposed PGA active beamforming design can exploit the CSI, QoS, and backscatter constellation to enlarge the achievable rate region.

3) Decision Threshold: We evaluate the following decision threshold strategies for problem (33):

- *DP:* Benchmark DP method for sequential quantizer [66];
- *SMAWK:* DP accelerated by the SMAWK algorithm [66];
- *Bisection:* The traverse-then-bisect algorithm [67];
- *ML:* Maximum likelihood detector (34) [68].

Fig. 10(c) reveals the average achievable rate region for those strategies. The distribution-aware schemes DP, SMAWK

(a) RIScatter Nodes

(b) Reflection States

(c) Transmit Antennas

(d) Spreading Factor

(e) Average Noise Power

(f) Imperfect CSI

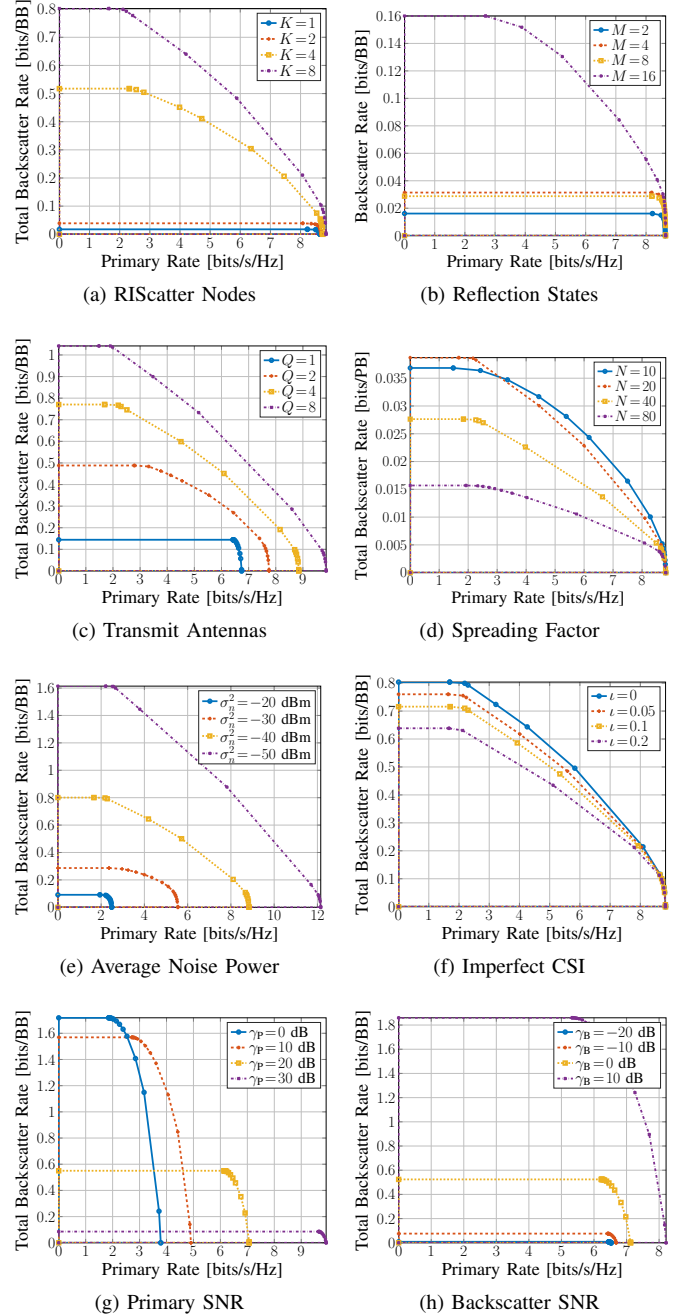(g) Primary SNR

(h) Backscatter SNR

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

and Bisection ensure higher total backscatter rate than ML. This is because the total backscatter rate (11) is a function of both input distribution and decision regions, and the rate-optimal thresholding depends heavily on the 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, unless otherwise specified.

1) Number of Nodes: Fig. 11(a) reveals how the number RIScatter nodes K influences 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 SNR but also influences the relative energy difference that other nodes can make. 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 passive array gain of K^2.

2) Number of States: Fig. 11(b) shows the relationship between the available reflection states (i.e., QAM order) M and the achievable rate region when $K=1$. We notice that increasing the reflection states has a marginal effect on the primary rate but significantly improves the backscatter rate. This is because the maximum amplitude normalized-QAM (2) involves more weak reflection points as M increases. It enhances the receive energy diversity but cannot provide enough phase shift resolution with maximum reflection.

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 become available and the channel diversity can be better exploited to improve the primary-backscatter tradeoff. It emphasizes the importance of multi-antenna RIScatter systems and demonstrate the effectiveness of the proposed PGA design.

4) Spreading factor: Fig. 11(d) shows how the spreading factor N affects the achievable rate region.[14] Using a very large N (as in the case of SR) can severely constrain the backscatter throughput, since the gain in energy certainty (by the law of large numbers) cannot withstand the loss in the gross rate. As $N \to \infty$, RIScatter nodes boil down to static RIS elements and the total backscatter rate approaches 0. On the other hand, when N is too small, the DMMAC (10) becomes unreliable and energy detection is error-prone. It explains the observation that $N=10$ provides lower backscatter throughput than $N=20$. Therefore, we conclude the spreading factor N should be carefully designed over multiple factors (e.g., primary and backscatter SNR, data rate requirements, load switching speed 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. Note that the noise influences both primary and backscatter SNR. When σ_v^2 relatively high, one can choose a larger N to improve the SNR of energy detection.

6) Imperfect CSI: Due to the lack of RF chains at RIScatter nodes, fast and accurate acquisition of the cascaded CSI can

be challenging, especially when the backscatter SNR is weak or the number of nodes is large. We consider an imperfect CSI model, where the cascaded channel of node k is estimated as

$$\hat{\boldsymbol{h}}_{\mathrm{C},k} = \boldsymbol{h}_{\mathrm{C},k} + \tilde{\boldsymbol{h}}_{\mathrm{C},k}, \tag{41}$$

$\tilde{\boldsymbol{h}}_{\mathrm{C},k}$ is the estimation error with entries following i.i.d. CSCG distribution $\mathcal{CN}(0, \iota \Lambda_{\mathrm{C}})$, ι is the relative estimation error, and Λ_{C} is the cascaded path loss. In Fig. 11(f), it is observed that the channel estimation error mainly affects the backscatter rate. When ι increases from 0 to 0.2, the maximum total backscatter rate decreases by $20\,\%$. This is because the energy detector is sensitive to the DMMAC (10) and thus the estimation of the cascaded channel. On the other hand, a small estimation error may be insufficient to change the optimal passive beamforming state and the primary rate is almost unchanged.

7) Primary SNR: The backscatter SNR is fixed to $0\,\text{dB}$ with $Q=1$ in this study. Interestingly, Fig. 11(g) shows that increasing the primary SNR can improve the primary rate but degrade the backscatter rate. The reason is that the relative strength of the scattered signal compared to the direct signal is weakened, such that the nodes cannot make enough difference to the energy detector. This, together with Fig. 10(b) and 11(a), emphasizes the importance of balancing the primary and backscatter SNR in the design of active-passive coexisting networks.

8) Backscatter SNR: The primary SNR is fixed to $20\,\text{dB}$ with $Q=1$ in this study. Fig. 11(h) shows that the primary and backscatter rates are both improved when the backscatter SNR increases. This motivates one to use high-efficiency or semi-passive RIScatter nodes to improve the overall performance. In a multi-user RIScatter network, 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 novel scatter protocol that bridges backscatter modulation and passive beamforming. Starting from scattering principles, we showed how RIScatter nodes generalize information nodes of BackCom and reflect elements of RIS, how they can be built over existing passive scatter devices, and how they simultaneously encode self information and assist legacy transmission. We also proposed a practical SIC-free receiver that exploits the properties of active-passive coexisting networks to benefit both subsystems. The achievable primary-total-backscatter rate region was then studied for a single-user multi-node RIScatter network, where the input distribution, active beamforming, and decision thresholds are iteratively updated. Numerical results validated the proposed algorithms and emphasized the importance of adaptive input distribution and cooperative receiver design.

To support massive RIScatter networks with large number of nodes and states, one possible future direction is to consider backscatter detection over the received signal domain rather than energy domain, where multi-antenna [72] and learning-based approaches can be promising. Another interesting question is how to design RIScatter nodes and receivers in a multi-user system to fully exploit the dynamic passive beamforming that naturally origins from backscatter modulation.

---

[14]Here, the unit of total backscatter rate is bits per primary block to show backscatter throughput.

APPENDIX

A. Proof of Proposition 1

Denote the Lagrange multipliers associated with (17b) and (17c) as $\{\nu_k\}_{k\in\mathcal{K}}$ and $\{\lambda_{m_k}\}_{k\in\mathcal{K},m_k\in\mathcal{M}}$, respectively. The Lagrangian function of problem (18) is

$$L(p,\nu,\lambda) = I(x_\mathcal{K}) + \sum_k \nu_k\Big(\sum_{m_k}p(x_{m_k})-1\Big) + \sum_{k,m_k}\lambda_{m_k}p(x_{m_k}),$$
(42)

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

$$I_k^\star(x_{m_k}) - (1-\rho) + \nu_k^\star + \lambda_{m_k}^\star = 0, \tag{43a}$$
$$\lambda_{m_k}^\star = 0, \quad \text{if } p^\star(x_{m_k})>0, \tag{43b}$$
$$\lambda_{m_k}^\star \geq 0, \quad \text{if } p^\star(x_{m_k})=0. \tag{43c}$$

Plugging (43b) and (43c) into (43a) yields

$$I_k^\star(x_{m_k}) = 1-\rho-\nu_k^\star, \quad \text{if } p^\star(x_{m_k})>0, \tag{44a}$$
$$I_k^\star(x_{m_k}) \leq 1-\rho-\nu_k^\star, \quad \text{if } p^\star(x_{m_k})=0, \tag{44b}$$

such that

$$\sum_{m_k}p^\star(x_{m_k})I_k^\star(x_{m_k}) = 1-\rho-\nu_k^\star. \tag{45}$$

On the other hand, by definition (20) we have

$$\sum_{m_k}p^\star(x_{m_k})I_k^\star(x_{m_k}) = I^\star(x_\mathcal{K}), \tag{46}$$

where the right-hand side is irrelevant to k. (44), (45), and (46) together complete the proof.

B. Proof of Proposition 2

We first prove sequence (21) is non-decreasing in weighted sum mutual information. Let $p(x_{m_\mathcal{K}}) = \prod_{q\in\mathcal{K}}p(x_{m_q})$ and $p'(x_{m_\mathcal{K}}) = p'(x_{m_k})\prod_{q\in\mathcal{K}\setminus\{k\}}p(x_{m_q})$ be two distributions potentially different at x_{m_k}, and $J\big(p(x_{m_\mathcal{K}}),p'(x_{m_\mathcal{K}})\big)$ be a joint function defined in (47) at the end of page 12. It is straightforward to verify $J\big(p(x_{m_\mathcal{K}}),p(x_{m_\mathcal{K}})\big) = I(x_\mathcal{K})$ and $J\big(p(x_{m_\mathcal{K}}),p'(x_{m_\mathcal{K}})\big)$ is a concave function for a given $p'(x_{m_k})$. Setting $\nabla_{p(x_{m_k})}J\big(p(x_{m_\mathcal{K}}),p'(x_{m_\mathcal{K}})\big)=0$ yields

$$S_k'(x_{m_k}) - S_k'(x_{i_k}) + (1-\rho)\log\frac{p(x_{i_k})}{p^\star(x_{m_k})} = 0, \tag{48}$$

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

$$S_k'(x_{m_k}) \triangleq I_k'(x_{m_k}) + (1-\rho)\sum_{m_{\mathcal{K}\setminus\{k\}}}p(x_{m_{\mathcal{K}\setminus\{k\}}})$$
$$\times \sum_{m_\mathcal{K}'}q(\hat{x}_{m_\mathcal{K}'}|x_{m_\mathcal{K}})\log p'(x_{m_k}). \tag{49}$$

Evidently, $\forall m_k\neq i_k$, (48) boils down to

$$p^\star(x_{m_k}) = \frac{p'(x_{m_k})\exp\Big(\frac{\rho}{1-\rho}I_k'(x_{m_k})\Big)}{\sum_{m_k'}p'(x_{m_k'})\exp\Big(\frac{\rho}{1-\rho}I_k'(x_{m_k'})\Big)}. \tag{50}$$

Since $p(x_{i_k}) = 1-\sum_{m_k\neq i_k}p^\star(x_{m_k})$ has exactly the same form as (50), the choice of reference does not matter and (50) is optimal $\forall m_k\in\mathcal{M}$. That is, for a fixed $p'(x_{m_\mathcal{K}})$, (50) ensures

$$J\big(p(x_{m_\mathcal{K}}),p'(x_{m_\mathcal{K}})\big) \geq I'(x_\mathcal{K}). \tag{51}$$

On the other hand, we notice

$$I(x_\mathcal{K}) - J\big(p(x_{m_\mathcal{K}}),p'(x_{m_\mathcal{K}})\big)$$
$$= (1-\rho)\sum_{m_k}\frac{p'(x_{m_k})f_k'(x_{m_k})}{\sum_{m_k'}p'(x_{m_k'})f_k'(x_{m_k'})}\sum_{m_\mathcal{K}''}q(\hat{x}_{m_\mathcal{K}''}|x_{m_k})$$
$$\times\log\frac{\sum_{m_k'}p'(x_{m_k'})q(\hat{x}_{m_\mathcal{K}''}|x_{m_k'})f_k'(x_{m_k})}{\sum_{m_k'}p'(x_{m_k'})q(\hat{x}_{m_\mathcal{K}''}|x_{m_k'})f_k'(x_{m_k'})} \tag{52a}$$
$$\geq (1-\rho)\sum_{m_k}\frac{p'(x_{m_k})f_k'(x_{m_k})}{\sum_{m_k'}p'(x_{m_k'})f_k'(x_{m_k'})}\sum_{m_\mathcal{K}''}q(\hat{x}_{m_\mathcal{K}''}|x_{m_k})$$
$$\times\left(1 - \frac{\sum_{m_k'}p'(x_{m_k'})q(\hat{x}_{m_\mathcal{K}''}|x_{m_k'})f_k'(x_{m_k'})}{\sum_{m_k'}p'(x_{m_k'})q(\hat{x}_{m_\mathcal{K}''}|x_{m_k'})f_k'(x_{m_k})}\right) \tag{52b}$$
$$= (1-\rho)\left(1 - \sum_{m_k}\frac{p'(x_{m_k})\cancel{f_k'(x_{m_k})}}{\sum_{m_k'}p'(x_{m_k'})f_k'(x_{m_k'})}\sum_{m_\mathcal{K}''}q(\hat{x}_{m_\mathcal{K}''}|x_{m_k})\right.$$
$$\left.\times\frac{\sum_{m_k'}p'(x_{m_k'})q(\hat{x}_{m_\mathcal{K}''}|x_{m_k'})f_k'(x_{m_k'})}{\sum_{m_k'}p'(x_{m_k'})q(\hat{x}_{m_\mathcal{K}''}|x_{m_k'})\cancel{f_k'(x_{m_k})}}\right) \tag{52c}$$
$$= (1-\rho)\left(1 - \sum_{m_\mathcal{K}''}\frac{\cancel{\sum_{m_k}p'(x_{m_k})q(\hat{x}_{m_\mathcal{K}''}|x_{m_k})}}{\sum_{m_k'}p'(x_{m_k'})f_k'(x_{m_k'})}\right.$$
$$\left.\times\frac{\sum_{m_k'}p'(x_{m_k'})q(\hat{x}_{m_\mathcal{K}''}|x_{m_k'})f_k'(x_{m_k'})}{\cancel{\sum_{m_k'}p'(x_{m_k'})q(\hat{x}_{m_\mathcal{K}''}|x_{m_k'})}}\right) \tag{52d}$$
$$= (1-\rho)\left(1 - \frac{\sum_{m_k'}p'(x_{m_k'})f_k'(x_{m_k'})\sum_{m_\mathcal{K}''}q(\hat{x}_{m_\mathcal{K}''}|x_{m_k'})}{\sum_{m_k'}p'(x_{m_k'})f_k'(x_{m_k'})}\right)$$
$$= 0, \tag{52e}$$

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 $p(x_{m_\mathcal{K}})$ and $p'(x_{m_\mathcal{K}})$ equals (i.e., (50) converges). (51) and (52) together imply $I(x_\mathcal{K}) \geq I'(x_\mathcal{K})$. Since mutual information is bounded above, we conclude the sequence (21) is non-decreasing and convergent.

Next, we prove any converging point of sequence (21), denoted as $p^\star(x_{m_k})$, fulfills the KKT conditions (19). Let

$$D^{(r)}(x_{m_k}) \triangleq \frac{p^{(r+1)}(x_{m_k})}{p^{(r)}(x_{m_k})} = \frac{f_k^{(r)}(x_{m_k})}{\sum_{m_k'}p^{(r)}(x_{m_k'})f_k^{(r)}(x_{m_k'})}. \tag{53}$$

As sequence (21) is convergent, any state with $p^\star(x_{m_k})>0$ need to satisfy $D^\star(x_{m_k}) \triangleq \lim_{r\to\infty}D^{(r)}(x_{m_k})=1$, namely

$$I_k^\star(x_{m_k}) = \frac{1-\rho}{\rho}\log\sum_{m_k'}p^\star(x_{m_k'})\exp\Big(\frac{\rho}{1-\rho}I_k^\star(x_{m_k'})\Big). \tag{54}$$

$$J\big(p(x_{m_\mathcal{K}}),p'(x_{m_\mathcal{K}})\big) \triangleq \sum_{m_\mathcal{K}}p(x_{m_\mathcal{K}})\left(\rho\log\Big(1+\frac{|\mathbf{h}^\mathsf{H}(x_{m_\mathcal{K}})\mathbf{w}|^2}{\sigma_v^2}\Big) + (1-\rho)\sum_{m_\mathcal{K}'}q(\hat{x}_{m_\mathcal{K}'}|x_{m_\mathcal{K}})\log\frac{q(\hat{x}_{m_\mathcal{K}'}|x_{m_\mathcal{K}})p'(x_{m_\mathcal{K}})}{p'(\hat{x}_{m_\mathcal{K}'})p(x_{m_\mathcal{K}})}\right). \tag{47}$$

The right-hand side is a constant for node k and implies (44a). That is, any converging point with nonzero probability must satisfy (19a). On the other hand, we assume $p^\star(x_{m_k}) = 0$ does not satisfy (19b), namely

$$I_k^\star(x_{m_k}) > I^\star(x_\mathcal{K}) = \sum_{m_k'} p^\star(x_{m_k'}) I_k^\star(x_{m_k'}), \qquad (55)$$

Since the exponential function is monotonically increasing, (55) implies $f_k^\star(x_{m_k}) > \sum_{m_k'} p^\star(x_{m_k'}) f_k^\star(x_{m_k'})$ and $D^\star(x_{m_k}) > 1$. It contradicts with

$$p^{(r)}(x_{m_k}) = p^{(0)}(x_{m_k}) \prod_{n=1}^{r} D^{(n)}(x_{m_k}), \qquad (56)$$

since the left-hand side is zero while all terms on the right-hand side are strictly positive. The proof is completed.

ACKNOWLEDGEMENT

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 SIG-COMM 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] 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, Sep 2019.

[13] H. E. Hassani, A. Savard, E. V. Belmega, and R. C. de Lamare, "Multi-user downlink NOMA systems aided by ambient backscattering: Achievable rate regions and energy-efficiency maximization," *IEEE Transactions on Green Communications and Networking*, pp. 1–1, 2023.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

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

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

[20] 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.

[21] 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.

[22] Z. Dai, R. Li, J. Xu, Y. Zeng, and S. Jin, "Rate-region characterization and channel estimation for cell-free symbiotic radio communications," *IEEE Transactions on Communications*, vol. 71, pp. 674–687, Feb 2023.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] 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.

[41] 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.

[42] 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.

[43] 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.

[44] 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.

[45] 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." IEEE, 5 2022, pp. 2651–2657.

[46] I. Vardakis, G. Kotridis, S. Peppas, K. Skyvalakis, G. Vougioukas, and A. Bletsas, "Intelligently wireless batteryless RF-powered reconfigurable surface: Theory, implementation & limitations," *IEEE Transactions on Wireless Communications*, vol. 22, pp. 3942–3954, Jun 2023.

[47] 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.

[48] S. J. Thomas and M. S. Reynolds, "A 96 mbit/sec, 15.5 pj/bit 16-qam modulator for uhf backscatter communication." IEEE, Apr 2012, pp. 185–190.

[49] N. V. Huynh, D. T. Hoang, X. Lu, D. Niyato, P. Wang, and D. I. Kim, "Ambient backscatter communications: A contemporary survey," *IEEE Communications Surveys & Tutorials*, vol. 20, pp. 2889–2922, 2018.

[50] 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.

[51] 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.

[52] 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.

[53] 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.

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

[55] 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.

[56] 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.

[57] 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.

[58] 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.

[59] 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.

[60] W. Wu, X. Wang, A. Hawbani, L. Yuan, and W. Gong, "A survey on ambient backscatter communications: Principles, systems, applications, and challenges," Oct 2022.

[61] 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.

[62] 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.

[63] 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.

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

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

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

[67] 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.

[68] 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.

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

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

[71] 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.

[72] 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.