Metascatter: Unifying Symbiotic Radio and Intelligent Reflecting Surface

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

***Abstract*—Backscatter nodes can harvest energy from, modulate information over, and reconfigure propagation of ambient signals. We uniquely introduce *Metascatter* that adapts the input distribution of a finite-state passive backscatter node based on Channel State Information (CSI), to *simultaneously* encode self message and assist legacy transmission while powered by surrounding waves. Compared to existing Intelligent Reflecting Surface (IRS)-empowered Symbiotic Radio (SR) that encodes and precodes independently using advanced architecture (e.g., overlay or parallel), Metascatter softly bridges and generalizes parasitic source of SR and reflecting element of IRS via smart input design. This not only reduces hardware complexity and optimization cost, but also enables a flexible tradeoff between primary (legacy) and backscatter links using shared spectrum, energy, and infrastructures. Moreover, we consider a specific scenario where a multi-antenna Access Point (AP) serves a single-antenna user surrounded by multiple Metascatters. For simplicity, it is assumed the user first jointly decodes all backscatter messages from accumulated energy, then models reflection patterns and backscatter paths within equivalent channel for primary decoding. We characterize the achievable primary-(total-)backscatter rate region by optimizing the input distribution at Metascatters, the active beamforming at the AP, and the decision regions at the user. A suboptimal Block Coordinate Descent (BCD) algorithm is proposed, where the Karush–Kuhn–Tucker (KKT) input distribution is evaluated in closed form, the active beamforming is updated by Projected Gradient Descent (PGD), and the suboptimal convex decision regions are refined by Dynamic Programming (DP). Simulation results demonstrate Metascatters can exploit additional propagation paths to transmit and assist via input design.**

I. Introduction

BACKSCATTER is recently re-innovated as a promising approach to support low-power communications and control wireless propagation environments. Ambient Backscatter Communications (AmBC) that enables battery-free communication between interactive nodes was first introduced and prototyped in [1], where devices harvest energy from and modulate information over ambient Radio-Frequency (RF) signals by switching between reflecting and absorbing states. To combat the strong direct-link interference of AmBC, [2] exploited the repeating structure of Orthogonal Frequency-Division Multiplexing (OFDM) symbol and proposed a multi-antenna detector that only requires the backscatter channel strength. Cooperative AmBC that decodes primary and backscatter links using co-located receiver was proposed in [3], where the authors evaluated the error performance of Maximum-Likelihood (ML), linear, and Successive Interference Cancellation (SIC) detectors for flat fading channels and proposed a low-complexity ML detector for frequency-selective fading channels. The concept was then generalized to SR in [4] to "exploit the benefits and address the drawbacks" of Cognitive Radio (CR) and AmBC. The authors of [5] classified SR into commensal, parasitic, and competitive types based on link priority, and derived their instantaneous rates and optimal power allocations. The corresponding outage probabilities were also studied in [6]. Besides, [7] concluded that if the backscatter symbol period is sufficiently long, then the non-coherent primary rate would approach its coherent counterpart. The authors thus proposed to decode the primary link under backscatter uncertainty, then perform SIC and decode the backscatter link. [8] also explored the asymptotic impact of transmit/receive antenna and backscatter symbol period on the ergodic rate of primary and backscatter links. For a Multiple-Input Multiple-Output (MIMO) SR system with a multi-antenna backscatter node, [9] proposed a beamforming design to maximize the backscatter rate while guaranteeing the primary performance. However, those paper only considered one backscatter node and backscatter multiple access remains an open issue. In [10], a Non-Orthogonal Multiple Access (NOMA)-based SR was proposed and receive combining was investigated when SIC order follows equivalent channel strength. A Time-Division Multiple Access (TDMA)-based SR with energy harvesting constraints was also presented in [11], where transmit power, reflection efficiency, and time allocation were jointly optimized to maximize energy efficiency. To reduce coordination between passive nodes, [12] proposed a random code-assisted multiple access for SR and evaluated the asymptotic Signal-to-Interference-plus-Noise Ratio (SINR) using random matrix theory.

On the other hand, IRS alters propagation environment for legacy networks using numerous sub-wavelength passive elements with adaptive amplitudes and/or phases. Extensive research has been devoted to optimizing the phase shifts for the whole channel block to improve communication, sensing, and power performances [13]–[18]. Besides, [19] proposed a dynamic passive beamforming that further adjusts IRS over fine-grained time slots for OFDM systems to balance beamforming gain and multiuser diversity. The idea was further applied to Wireless Powered Communication Network (WPCN) to accommodate the downlink Wireless Power Transfer (WPT) phase and uplink Wireless Information Transfer (WIT) phase of the "harvest-then-transmit" protocol [20], [21]. For multiuser WPT, [22] also reported that dynamic beamforming can mimic multi-beam reflection in a time-division manner and reduce the phase extraction loss. Although dynamic passive beamforming artificially creates temporal diversity for flexible channel

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).

reconfiguration and resource allocation, it demands additional computational cost and control overhead, and the tradeoff deserves further attention especially for large IRS. In [23], [24], IRS was also introduced to single- and multi-node SR systems to reduce the total transmit power. When joint transmitter-IRS encoding is possible, [25] proved using IRS only for passive beamforming is generally suboptimal in terms of achievable rate for finite input constellations. Recently, IRS-empowered SR was introduced in [26]–[28] where independent passive beamforming and backscatter encoding were combined using advanced architectures. The authors of [26], [27] proposed the IRS to modulate binary message over the whole phase shift matrix and the receiver to decodes from (non-coherent) primary to (coherent) backscatter link. In contrast, [28] divided the IRS into reflection and information elements, and evaluated the error performance of non-coherent backscatter detection.

To the best of our knowledge, all existing SR literatures assumed backscatter modulation employs either Gaussian codebook [5]–[11] or finite equiprobable inputs [3], [12], [26]–[29]. The former is impractical for passive nodes with constrained number of states, while the latter does not fully exploit CSI to boost achievable backscatter rate. Besides, most relevant designs [3], [5]–[12], [26], [27], [29] were built over ideal ML or SIC receiver. However, the advantage of SIC is questionable because 1) it requires non-coherent primary encoding at the transmitter and re-encoding, precoding, and subtraction at the receiver, 2) the primary and backscatter symbols are mixed by multiplication instead of superposition, and 3) the backscatter symbol period is typically much longer due to physical constraints. Motivated by those, we propose the concept of Metascatter, which adapts the input distribution of a passive backscatter node to generalize backscatter sources of SR and reflecting elements of IRS. The contributions of this paper is summarized as follows.

First, Metascatter adapts the input probability distribution of a finite-state passive backscatter device based on primary and (cascaded) backscatter CSI to unify and generalize parasitic source of SR and reflecting elements of IRS. The reflection pattern over time is no longer fully random or deterministic, but can be flexibly distributed to balance backscatter encoding and passive beamforming. For single-user scenario, when primary link is absolutely prioritized, the distribution falls on one state and Metascatter boils down to conventional IRS. When only considering backscatter performance, the distribution involves the highest entropy and Metascatter is essentially an AmBC node.

Second, we consider an application scenario where multiple Metascatters ride over a point-to-point transmission, exploiting additional propagation paths to simultaneously transmit and assist. To fully accommodate backscatter characteristics, we also propose a novel receiving strategy that first jointly decodes all Metascatter from accumulated energy, then models their reflection patterns and backscatter paths within equivalent channel for primary decoding. Since backscatter message is modulated over primary signal, backscatter decoding is indeed part of primary channel training, and there is no need for operation-intensive SIC at the receiver.

Third, we evaluate the achievable primary-(total-)backscatter rate region by optimizing the input distribution at Metascatters, the active beamforming at the AP, and the decision regions at the user. Since the original problem is highly non-convex, we consider a suboptimal BCD algorithm where the KKT input distribution is numerically evaluated by limit of sequences, the active beamforming is sequentially updated by PGD accelerated by backtracking line search, and the decision regions are first restricted to convex, then refined by DP and Shor-Moran-Aggarwal-Wilber-Klawe (SMAWK) algorithm.

Notations: Scalars, vectors and matrices are respectively denoted by italic, bold lower-case, and bold upper-case letters. j represents the imaginary unit. $\mathbb{R}_+^{x\times y}$ and $\mathbb{C}^{x\times y}$ respectively denote the space of real nonnegative and complex $x \times y$ matrices. $\Delta^n = \left\{(p_0,\ldots,p_n) \in \mathbb{R}_+^{n+1} | \sum_{i=0}^n t_i = 1\right\}$ denotes the standard n-simplex. $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^+$, $|\cdot|$, $\|\cdot\|$ respectively represent the conjugate, transpose, conjugate transpose, ramp function, absolute value, and Euclidean norm. The distribution of a CSCG random vector with mean $\mathbf{0}$ and covariance $\boldsymbol{\Sigma}$ is denoted by $\mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$. $\sim$ means "distributed as". $(\cdot)^{(i)}$ represents the i-th iterated value and $(\cdot)^\star$ represents the end solution.

II. Backscatter Principles

Passive backscatter nodes harvest energy from and modulate information over surrounding RF signals. As shown in Fig. 1(a), a typical passive node consists of a scattering antenna, an energy harvester, an integrated receiver, a load-switching modulator, and on-chip components (e.g., micro-controller and sensors) [31]. Its equivalent circuit is presented in Fig. 1(b). When illuminated, the node absorbs a portion of the impinging wave for information decoding and/or energy harvesting [32], and backscatters the remaining as *structural* and *antenna* components. The former consistently contributes to environment multipath and can be modelled by channel estimation [33], while the latter depends on antenna-load impedance mismatch and can be used for backscatter modulation [34] and/or channel reconfiguration [35]. Fig. 1(c) illustrates the scatter model of a node with M states, where the reflection coefficient of state $m \in \mathcal{M} \triangleq \{1,\ldots,M\}$ is defined as[1]

$$\Gamma_m = \frac{Z_m - Z_\mathrm{A}^*}{Z_m + Z_\mathrm{A}}, \tag{1}$$

where Z_m is the load impedance at state m and Z_A is the antenna input impedance.

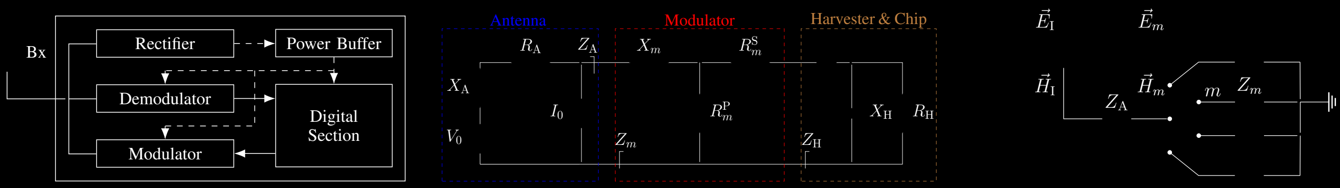
A. SR: Backscatter Modulation

Backscatter sources encode self message by *random reflection states variation.* For M-ary Quadrature Amplitude Modulation (QAM), reflection coefficient Γ_m maps to the corresponding *complex constellation point* c_m by [36]

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

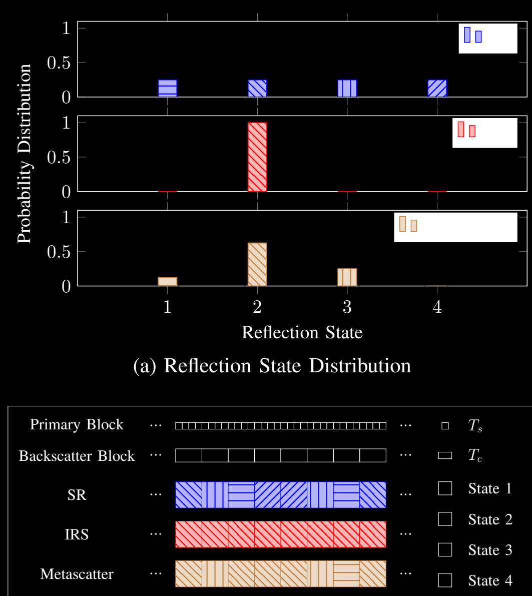
where $0 \leq \alpha \leq 1$ is the amplitude reflect ratio that controls the harvest-scatter tradeoff at the direction of interest.

[1]It corresponds to a linear backscatter model where the reflection coefficient is irrelevant to incident electromagnetic field strength.



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

Fig. 1. Block diagram, equivalent circuit, and scatter model of a passive backscatter node. The solid and dashed vectors represent signal and energy flows. The backscatter 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 $\vec{H}_{\mathrm{I}}$ [30].



(a) Reflection State Distribution

(b) Time Slot Structure

Fig. 2. Reflection state distribution and time block structure of SR, IRS, and Metascatter. T_s and T_c respectively denote the primary and backscatter symbol period. Within channel coherence time, Metascatter semi-randomly selects reflection state for each backscatter block, with guidance from input probability distribution.

B. IRS: Channel Reconfiguration

IRS elements assist legacy transmission by *deterministic phase shifts selection* based on relevant CSI. For a reflecting element with M candidate states, reflection coefficient Γ_m relates to the corresponding *phase shift* θ_m by [13]

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

where $0 \leq \beta_m \leq 1$ is overall amplitude reflect ratio of state m.

C. Metascatter: Bridge and Generalization

Metascatters simultaneously transmit and assist by *adaptive input distribution design* based on primary and (cascaded) backscatter CSI. Instead of using fully random or deterministic reflection pattern over time, as shown in Fig. 2, Metascatter semi-randomly selects reflection state for each backscatter block, with *guidance of input probability* $P(\Gamma_m)$ for state m. In other words, it flexibly controls input distribution of candidate reflection states to balance backscatter encoding and passive beamforming. SR and IRS can be regarded as extreme cases of Metascatter, where node input distribution boils down to uniform and deterministic, respectively.

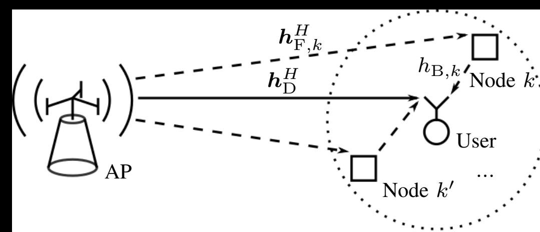


Fig. 3. A Metascatter-enabled single-user multi-node network.

Remark 1. *Compared to conventional IRS literatures that optimize phase shifts under unit-module constraint, Metascatter starts from predefined reflection coefficients and designs their input distribution under sum-probability constraint to achieve flexible primary-backscatter tradeoff.*

III. Metascatter-Enabled Network

A. System Model

As shown in Fig. 3, we propose a Metascatter-enabled single-user multi-node network where two coexisting systems share spectrum, energy and infrastructures. The primary point-to-point transmission from a Q-antenna AP to a single-antenna user is assisted by K nearby single-antenna Metascatters. In the secondary backscatter Multiple Access Channel (MAC) system, the AP serves as the carrier emitter, K nearby single-antenna Metascatters modulate information over reradiated RF signals, and the user decodes their messages. For simplicity, we consider a quasi-static block fading model where channels remain constant within coherence interval while vary independently between consecutive blocks. Due to physical constraints on load switching, we assume the backscatter symbol period is $N \gg 1$ times longer than primary and consider integer N without loss of generality. We also assume the direct channel and all cascaded channels can be estimated and fed back to the AP.[2] Besides, we omit the signal reflected by two or more times [40] and ignore the time difference of arrival from different paths [5].

Denote the AP-user direct channel as $\boldsymbol{h}_{\mathrm{D}}^H \in \mathbb{C}^{1\times Q}$, the AP-node $k \in \mathcal{K} \triangleq \{1,\ldots,K\}$ forward channel as $\boldsymbol{h}_{\mathrm{F},k}^H \in \mathbb{C}^{1\times Q}$, and the node k-user backward channel as $h_{\mathrm{B},k}$. Also, define the cascaded channel of tag k as $\boldsymbol{h}_{\mathrm{C},k}^H \triangleq h_{\mathrm{B},k}\boldsymbol{h}_{\mathrm{F},k}^H \in \mathbb{C}^{1\times Q}$, and $\boldsymbol{H}_{\mathrm{C}} \triangleq [\boldsymbol{h}_{\mathrm{C},1},\ldots,\boldsymbol{h}_{\mathrm{C},K}]^H \in \mathbb{C}^{K\times Q}$. Let $x_{\mathcal{K}} \triangleq (x_1,\ldots,x_K)$ be

[2]Due to the lack of RF chains at the passive tag, accurate and efficient CSI acquisition at the AP can be challenging. One possibility is the AP sends training pilots, the tags respond in predefined manners, and the user performs least-square estimation with feedbacks [37]–[39].

the backscatter symbol tuple of all Metascatters. Consider the signal model during one backscatter block (i.e., N primary blocks). Under perfect synchronization, the equivalent primary channel is a function of backscatter symbols

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

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

where α_k is the amplitude reflect ratio of Metascatter k, $\boldsymbol{\alpha} \triangleq [\alpha_1, \ldots, \alpha_K]^T \in \mathbb{R}_{+}^{K\times 1}$, $x_k \in \mathcal{X} \triangleq \{c_1, \ldots, c_M\}$ is the backscatter symbol of Metascatter k, and $\boldsymbol{x} \triangleq [x_1, \ldots, x_K]^H \in \mathbb{C}^{K\times 1}$. The signal received by the user at primary block $n \in \mathcal{N} \triangleq \{1,\ldots,N\}$ is

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

where $s \sim \mathcal{CN}(0,1)$ is the primary symbol, $v \sim \mathcal{CN}(0,\sigma_v^2)$ is the Additive White Gaussian Noise (AWGN), and $\boldsymbol{w} \in \mathbb{C}^{Q\times 1}$ is the active beamforming vector with average power constraint $\|\boldsymbol{w}\|^2 \le P$.

Remark 2. *Metascatter involves a symbiotic MAC where the primary and backscatter symbols of different duration are mixed by Multiplication Coding (MC) instead of Superposition Coding (SC). For each node, the reflection coefficient not only encodes the backscatter message, but also influences the equivalent primary channel* (4). *To accommodate such signal characteristics, novel receiving strategy apart from SIC is desired to better utilize the reflection pattern and boost the primary-backscatter tradeoff.*

B. Receiving Strategy

We propose a Metascatter receiver where the backscatter symbols of all Metascatters are first jointly and semi-coherently detected using total received energy per backscatter block, then modeled within equivalent channel (4) as dynamic passive beamforming. Compared with ML and SIC, Metascatter receiver allows practical and low-complexity node multiple access with minor adjustment over legacy equipments.

At a specific backscatter block, denote $m_k \in \mathcal{M}$ as the state index of Metascatter k, and let $m_{\mathcal{K}} \triangleq (m_1, \ldots, m_K)$ be the state index tuple of all Metascatters. Conditioned on $m_{\mathcal{K}}$, the received signal at primary block n is subject to the variation of $s[n]$ and $v[n]$, distributed as $y[n] \sim \mathcal{CN}(0, \sigma_{m_{\mathcal{K}}}^2)$ with

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

where x_{m_k} and $x_{m_{\mathcal{K}}}$ are the symbol and symbol tuple associated with state m_k and state tuple $m_{\mathcal{K}}$, respectively.[3] Also, denote the total received energy within backscatter block as $z = \sum_{n=1}^{N} |y[n]|^2$. As the sum of N independent and identically distributed (i.i.d.) exponential variables, the Probability Density Function (PDF) of z conditioned on $m_{\mathcal{K}}$ follows Erlang distribution

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

[3] x_k and $x_{\mathcal{K}}$ are random variables, while x_{m_k} and $x_{m_{\mathcal{K}}}$ are their instances indexed by m_k and $m_{\mathcal{K}}$.

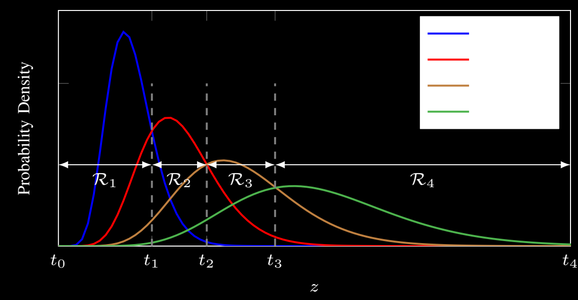


Fig. 4. PDF of total received energy per backscatter block, conditioned on different input hypothesis. Here, the convex ML decision regions are generally rate-suboptimal except for equiprobable inputs.

where $\mathcal{H}_{m_{\mathcal{K}}}$ denotes hypothesis $m_{\mathcal{K}}$. To accommodate backscatter characteristics and reduce decoding complexity, we consider a joint semi-coherent detection for all Metascatters over accumulated energy z. Once disjoint energy decision regions are determined, we can construct a Discrete Thresholding Multiple Access Channel (DTMAC) and formulate the transition probability from input $x_{m_{\mathcal{K}}}$ to output $\hat{x}_{m'_{\mathcal{K}}}$ as

$$P(\hat{x}_{m'_{\mathcal{K}}} | x_{m_{\mathcal{K}}}) = \int_{\mathcal{R}_{m'_{\mathcal{K}}}} f(z|\mathcal{H}_{m_{\mathcal{K}}}) \mathrm{d}z, \tag{8}$$

where $\mathcal{R}_{m'_{\mathcal{K}}}$ is the decision region of hypothesis $\mathcal{H}_{m'_{\mathcal{K}}}$. An example of ML energy decision is illustrated in Fig. 4.

Remark 3. *The rate-optimal thresholding channel design remains under-investigated, and some attempts were made for single source with binary inputs in [41], [42]. For non-binary inputs with general distribution, the optimal decision region for each letter can be non-convex (i.e., with non-adjacent partitions) and the optimal number of thresholds is still unknown.*

In the following context, we restrict all decision regions to convex and optimize decision thresholds accordingly. That is, for any bijective mapping $f: m_{\mathcal{K}} \to \mathcal{L} \triangleq \{1,\ldots,M^K\}$, the decision region of letter $l \in \mathcal{L}$ is defined as $\mathcal{R}_l \triangleq [t_{l-1}, t_l)$, where $t_{l-1} < t_l$. We also define the decision threshold vector as $\boldsymbol{t} \triangleq [t_0,\ldots,t_L]^T \in \mathbb{R}_{+}^{(L+1)\times 1}$.

C. Achievable Rates

Denote the input probability of state m_k of Metascatter node k as $P_k(x_{m_k})$, and define the input probability distribution vector of node k as $\boldsymbol{p}_k \triangleq [P_k(c_1),\ldots,P_k(c_M)]^T \in \mathbb{R}^{M\times 1}$. With independent encoding at all nodes, the probability of backscatter symbol tuple $x_{m_{\mathcal{K}}}$ is $P_{\mathcal{K}}(x_{m_{\mathcal{K}}}) = \prod_{k\in\mathcal{K}} P_k(x_{m_k})$. Similar to [43], we define the backscatter information function between input symbol tuple instance $x_{m_{\mathcal{K}}}$ and output symbol tuple $\hat{x}_{\mathcal{K}}$ as

$$I^{\mathrm{B}}(x_{m_{\mathcal{K}}}; \hat{x}_{\mathcal{K}}) \triangleq \sum_{m'_{\mathcal{K}}} P(\hat{x}_{m'_{\mathcal{K}}} | x_{m_{\mathcal{K}}}) \log \frac{P(\hat{x}_{m'_{\mathcal{K}}} | x_{m_{\mathcal{K}}})}{P(\hat{x}_{m'_{\mathcal{K}}})}, \tag{9}$$

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

$$I_k^{\mathrm{B}}(x_{m_k}; \hat{x}_{\mathcal{K}}) \triangleq \sum_{m_{\mathcal{K}\setminus\{k\}}} P_{\mathcal{K}\setminus\{k\}}(x_{m_{\mathcal{K}\setminus\{k\}}}) I^{\mathrm{B}}(x_{m_{\mathcal{K}}}; \hat{x}_{\mathcal{K}}), \tag{10}$$

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

$$I^{\mathrm{B}}(x_{\mathcal{K}};\hat{x}_{\mathcal{K}}) = \sum_{m_{\mathcal{K}}} P_{\mathcal{K}}(x_{m_{\mathcal{K}}}) I^{\mathrm{B}}(x_{m_{\mathcal{K}}};\hat{x}_{\mathcal{K}}). \tag{11}$$

Once backscatter symbols are successfully decoded, we can eliminate modulation uncertainty and retrieve equivalent primary channel by (4). We define the primary information function conditioned on backscatter symbol tuple $x_{m_{\mathcal{K}}}$ as

$$I^{\mathrm{P}}(s;y|x_{m_{\mathcal{K}}}) \triangleq \log\left(1 + \frac{|\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2}{\sigma_v^2}\right), \tag{12}$$

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

$$I_k^{\mathrm{P}}(s;y|x_{m_k}) \triangleq \sum_{m_{\mathcal{K}\setminus\{k\}}} P_{\mathcal{K}\setminus\{k\}}(x_{m_{\mathcal{K}\setminus\{k\}}}) I^{\mathrm{P}}(s;y|x_{m_{\mathcal{K}}}), \tag{13}$$

and the primary ergodic mutual information as

$$I^{\mathrm{P}}(s;y|x_{\mathcal{K}}) = \sum_{m_{\mathcal{K}}} P_{\mathcal{K}}(x_{m_{\mathcal{K}}}) I^{\mathrm{P}}(s;y|x_{m_{\mathcal{K}}}). \tag{14}$$

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

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

$$I_k(x_{m_k}) \triangleq \rho I_k^{\mathrm{P}}(s;y|x_{m_k}) + (1-\rho) I_k^{\mathrm{B}}(x_{m_k};\hat{x}_{\mathcal{K}}), \tag{16}$$

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

where $0 \le \rho \le 1$ is the relative priority of the primary link.

IV. Input Distribution, Active Beamforming, and Decision Threshold, Design

To characterize the achievable primary-(total-)backscatter rate region of the proposed Metascatter-enabled network, we aim to maximize the weighted sum mutual information with respect to node input distributions $\{\boldsymbol{p}_k\}_{k\in\mathcal{K}}$, active beamforming vector $\boldsymbol{w}$, and decision threshold vector $\boldsymbol{t}$ as

$$\max_{\{\boldsymbol{p}_k\}_{k\in\mathcal{K}},\boldsymbol{w},\boldsymbol{t}} \quad I(x_{\mathcal{K}}) \tag{18a}$$

$$\text{s.t.} \quad \sum\nolimits_{m_k} P_k(x_{m_k}) = 1, \quad \forall k, \tag{18b}$$

$$P_k(x_{m_k}) \ge 0, \quad \forall k, m_k, \tag{18c}$$

$$\|\boldsymbol{w}\|^2 \le P. \tag{18d}$$

Problem (18) is highly non-convex, and we propose a BCD algorithm that iteratively updates $\{\boldsymbol{p}_k\}_{k\in\mathcal{K}}$, $\boldsymbol{w}$ and $\boldsymbol{t}$ until convergence.

A. Input Distribution

For any given $\boldsymbol{w}$ and $\boldsymbol{t}$, we can construct the equivalent DTMAC by (8) and simplify (18) to

$$\max_{\{\boldsymbol{p}_k\}_{k\in\mathcal{K}}} \quad I(x_{\mathcal{K}}) \tag{19a}$$

$$\text{s.t.} \quad \text{(18b),(18c)}, \tag{19b}$$

which involves coupled term $\prod_{k\in\mathcal{K}} P_k(x_{m_k})$ and is non-convex when $K > 1$. Next, we propose a numerical method that evaluate the KKT input distribution by limit of sequences.

Algorithm 1: Numerical Evaluation of KKT Input Distribution

Input: K, N, $\boldsymbol{h}_{\mathrm{D}}^H$, $\boldsymbol{H}_{\mathrm{C}}$, $\boldsymbol{\alpha}$, $\mathcal{X}$, σ_v^2, ρ, $\boldsymbol{w}$, $\boldsymbol{t}$, ϵ
Output: $\{\boldsymbol{p}_k^\star\}_{k\in\mathcal{K}}$
Set $\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}$ by (4)
    $\sigma_{m_{\mathcal{K}}}^2$, $\forall m_{\mathcal{K}}$ by (6)
    $f(z|\mathcal{H}_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}$ by (7)
    $P(\hat{x}_{m'_{\mathcal{K}}}|x_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}, m'_{\mathcal{K}}$ by (8)
Initialize $r \leftarrow 0$
    $\boldsymbol{p}_k^{(0)} > \boldsymbol{0}$, $\forall k$
    $I^{(0)}(x_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}$ by (9), (12), (15)
    $I_k^{(0)}(x_{m_k})$, $\forall k, m_k$ by (10), (13), (16)
    $I^{(0)}(x_{\mathcal{K}})$ by (11), (14), (17)
Repeat
    Update $r \leftarrow r+1$
        $\boldsymbol{p}_k^{(r)}$, $\forall k$ by (21)
        $I^{(r)}(x_{m_{\mathcal{K}}})$, $\forall m_{\mathcal{K}}$ by (9), (12), (15)
        $I_k^{(r)}(x_{m_k})$, $\forall k, m_k$ by (10), (13), (16)
        $I^{(r)}(x_{\mathcal{K}})$ by (11), (14), (17)
Until $I^{(r)}(x_{\mathcal{K}}) - I^{(r-1)}(x_{\mathcal{K}}) \le \epsilon$

Remark 4. *As pointed out in [44], KKT conditions are generally necessary but insufficient for total rate maximization in discrete memoryless MAC. Therefore, KKT solutions may end up being saddle points of problem* (19).

Following [43], we first recast KKT conditions to their equivalent form for problem (19), then propose an iterative method that guarantees input distribution satisfying above conditions on convergence.

Proposition 1. *The KKT optimality conditions for problem* (19) *are equivalent to,* $\forall k, m_k$,

$$I_k^\star(x_{m_k}) = I^\star(x_{\mathcal{K}}), \quad P_k^\star(x_{m_k}) > 0, \tag{20a}$$

$$I_k^\star(x_{m_k}) \le I^\star(x_{\mathcal{K}}), \quad P_k^\star(x_{m_k}) = 0. \tag{20b}$$

Proof. Please refer to Appendix A. □

For each node, (20a) suggests each probable state should produce the same marginal information (averaged over all states of other nodes), while (20b) implies any state with potentially less marginal information should not be used.

Proposition 2. *The KKT input probability of node k of state m_k is given by the converging point of the sequence*

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

where r is the iteration index and $\boldsymbol{p}_k^{(0)} > \boldsymbol{0}$, $\forall k$.

Proof. Please refer to Appendix B. □

At each iteration, the input distribution of node k is evaluated over updated input distribution of node 1 to $k-1$, together with previous input distribution of node $k+1$ to K. The KKT input distribution design is summarized in Algorithm TODO.

B. Active Beamforming

For any given $\{\boldsymbol{p}_k\}_{k\in\mathcal{K}}$ and $\boldsymbol{t}$, problem (18) reduces to

$$\max_{\boldsymbol{w}} \quad I(x_{\mathcal{K}}) \tag{22a}$$
$$\text{s.t.} \quad (18\text{d}), \tag{22b}$$

which is still non-convex due to integration and entropy terms. To see this, we explicitly write (22a) as (23) at the bottom of page 6, where

$$Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)=\frac{\int_{t_{l-1}/\sigma^2_{m_{\mathcal{K}}}}^{t_l/\sigma^2_{m_{\mathcal{K}}}} z^{N-1}\exp(-z)\mathrm{d}z}{(N-1)!} \tag{24}$$

is the regularized incomplete Gamma function that explicitly substitutes the DTMAC transition probability (8). Its series representation is given by [45, Theorem 3]

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

We further write the gradient of (25) and (23) w.r.t. $\boldsymbol{w}^*$ as (26) and TODO at the end of page 6 and TODO, respectively.

C. Decision Threshold

For a given tag input distribution $\{\boldsymbol{p}_k\}_{k\in\mathcal{K}}$, we can formulate an equivalent information source with augmented alphabet of tag input combination. This equivalent source transmits at the total backscatter rate, and the input distribution is given by $P_{\mathcal{K}}(x_{m_{\mathcal{K}}})=\prod_{k\in\mathcal{K}}P_k(x_{m_k})$, $\forall k\in\mathcal{K}$ and $\forall m_k\in\mathcal{M}$.

Remark 5. *Since the equivalent source transmits at the total backscatter rate, the DTMAC* (8) *is essentially a point-to-point Discrete Memoryless Thresholding Channel (DMTC) and we can simplify the decision threshold design accordingly.*

Interestingly, the optimal threshold design to maximize the mutual information for a general DMTC with a fixed number of output letters remains an open issue. The reason is that each decision region may contain more than one disjoint partitions (i.e., non-convex) and the number of thresholds are unknown. Fortunately, for the proposed energy detection, we proved that the DMTC capacity can be achieved using only convex decision regions. This conclusion is summarized below.

Proposition 3. *For a discrete-input continuous-output channel in Erlang distribution* (7)*, if the DMTC is constructed for detection (i.e., same input/output alphabet) and L input letters are with non-zero probability, then it is possible to achieve the DMTC capacity by L non-empty convex decision regions defined by $L+1$ distinct decision thresholds.*

Proof. Please refer to Appendix C. ☐

Once the optimal number of decision threshold is determined, we can first discretize the output energy level into numerous bins, then obtain the optimal decision regions that maximizes

$$I(x_{\mathcal{K}})=\sum_{m_{\mathcal{K}}}P_{\mathcal{K}}(x_{m_{\mathcal{K}}})\left(\rho\log\Big(1+\frac{|\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2}{\sigma_v^2}\Big)+(1-\rho)\sum_l Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)\log\frac{Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)}{\sum_{m'_{\mathcal{K}}}P_{\mathcal{K}}(x_{m'_{\mathcal{K}}})Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m'_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m'_{\mathcal{K}}}}\Big)}\right) \tag{23}$$

$$\begin{aligned}\nabla_{\boldsymbol{w}^*}Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)=&\frac{\boldsymbol{h}_{\mathrm{E}}(x_{m_{\mathcal{K}}})\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}}{\big(|\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2+\sigma_v^2\big)^2}\\&\times\left(t_l\exp\Big(-\frac{t_l}{|\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2+\sigma_v^2}\Big)\left(-1+\sum_{n=1}^{N-1}\frac{n\Big(\frac{t_l}{|\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2+\sigma_v^2}\Big)^{n-1}-\Big(\frac{t_l}{|\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2+\sigma_v^2}\Big)^n}{n!}\right)\right.\\&\left.-t_{l-1}\exp\Big(-\frac{t_{l-1}}{|\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2+\sigma_v^2}\Big)\left(-1+\sum_{n=1}^{N-1}\frac{n\Big(\frac{t_{l-1}}{|\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2+\sigma_v^2}\Big)^{n-1}-\Big(\frac{t_{l-1}}{|\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2+\sigma_v^2}\Big)^n}{n!}\right)\right)\end{aligned} \tag{26}$$

$$\begin{aligned}\nabla_{\boldsymbol{w}^*}I(x_{\mathcal{K}})=\sum_{m_{\mathcal{K}}}P_{\mathcal{K}}(x_{m_{\mathcal{K}}})&\left(\rho\frac{\boldsymbol{h}_{\mathrm{E}}(x_{m_{\mathcal{K}}})\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}}{|\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2+\sigma_v^2}+(1-\rho)\sum_l\left(\log\frac{Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)}{\sum_{m'_{\mathcal{K}}}P_{\mathcal{K}}(x_{m'_{\mathcal{K}}})Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m'_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m'_{\mathcal{K}}}}\Big)}+1\right)\right.\\&\left.\times\nabla_{\boldsymbol{w}^*}Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)-\frac{Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m_{\mathcal{K}}}}\Big)\sum_{m'_{\mathcal{K}}}P_{\mathcal{K}}(x_{m'_{\mathcal{K}}})\nabla_{\boldsymbol{w}^*}Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m'_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m'_{\mathcal{K}}}}\Big)}{\sum_{m'_{\mathcal{K}}}P_{\mathcal{K}}(x_{m'_{\mathcal{K}}})Q\Big(N,\frac{t_{l-1}}{\sigma^2_{m'_{\mathcal{K}}}},\frac{t_l}{\sigma^2_{m'_{\mathcal{K}}}}\Big)}\right)\end{aligned} \tag{27}$$

TABLE I
PARAMETERS IN SIMULATION

Transmit antenna Q	1
Tags K	2
States M	2
Reflect ratio α	0.5
Duration ratio N	10
Noise power σ_v^2	1
Discretization bins	256

the total backscatter rate by DP accelerated by SMAWK algorithm [46].

> We may need to tailor threshold design for Metascatter as threshold design has potential impact on primary achievable rate, as implied by some simulation results.

V. Preliminary Results

Table I shows the parameters used in simulation. We assume all links are in standard Circularly Symmetric Complex Gaussian (CSCG) distribution and evaluated the rate regions on two instances. For the input design, "Cooperation" assumes full transmit cooperation at all tags (i.e., joint encoding), "Exhaustion" runs exhaustive search on all possible input distributions, "KKT" is proposed KKT input design (21), and "Marginalization" marginalizes the joint input array by "Cooperation" to obtain independent tag input distribution. For the threshold design, "SMAWK" refers to the DP-based quantization proposed in [46], "Bisection" sequentially optimizes each threshold by bisection [47], and "ML" is the ML detector that requires no knowledge of input distribution.

Figs. 5 and 6 show two typical scenarios where joint encoding can be helpful and unnecessary, respectively. Although the proposed KKT input design converges to the optimal solutions in both examples, it may be trapped at saddle points under poor initialization, especially when the number of tags increases.

We believed threshold design has no impact on the primary achievable rate, because primary decoding acts on $\boldsymbol{y}$ while thresholding acts on z. As such, the threshold that maximizes the total backscatter rate should also maximize the weighted sum rate. Interestingly, this may not be the case because the SMAWK and Bisection threshold designs, both maximizing the total backscatter rate, can be outperformed by the ML when it comes to weighted sum primary-(total-)backscatter rate. It may result from a precision issue, but inspires further research on threshold design for Metascatter.

The parameters remain fixed unless specified.

Appendix

A. *Proof of Proposition 1*

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

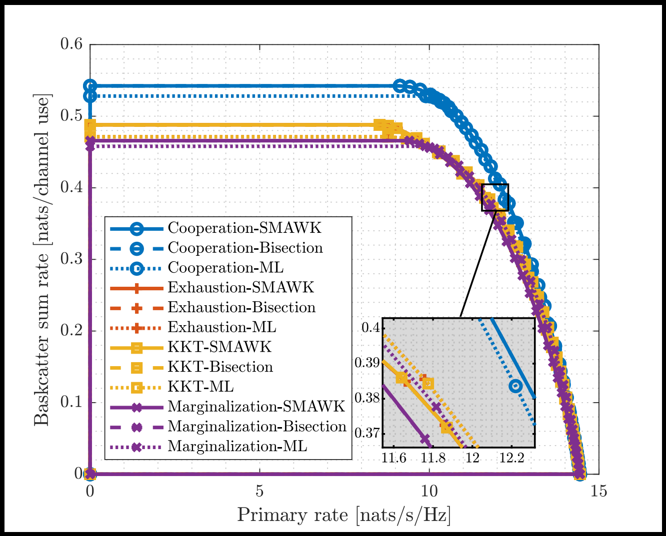


Fig. 5. Achievable rate regions by input-threshold design: Case I.

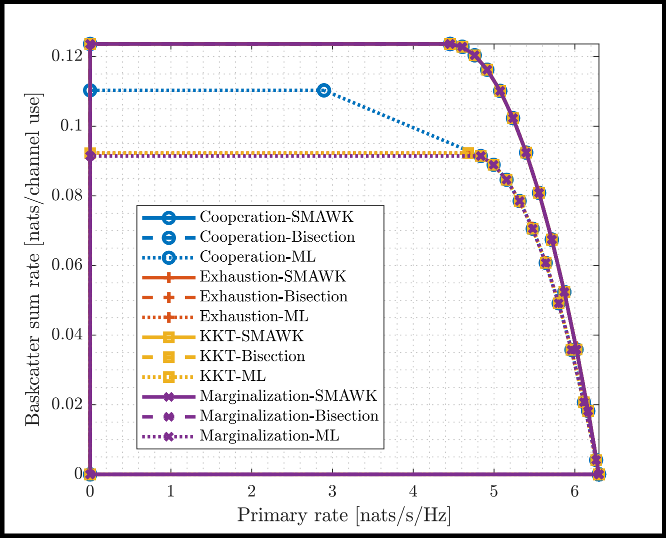


Fig. 6. Achievable rate regions by input-threshold design: Case II.

Lagrangian function of problem (19) is

$$L = -I(x_{\mathcal{K}}) + \sum_k \nu_k \left(\sum_{m_k\in\mathcal{M}} P_k(x_{m_k}) - 1 \right) - \sum_k \sum_{m_k} \lambda_{k,m_k} P_k(x_{m_k}), \tag{28}$$

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

$$-\nabla_{P_k^\star(x_{m_k})} I^\star(x_{\mathcal{K}}) + \nu_k^\star - \lambda_{k,m_k}^\star = 0, \tag{29a}$$

$$\lambda_{k,m_k}^\star = 0, \quad P_k^\star(x_{m_k}) > 0, \tag{29b}$$

$$\lambda_{k,m_k}^\star \geq 0, \quad P_k^\star(x_{m_k}) = 0. \tag{29c}$$

The directional derivative can be explicitly expressed as

$$\nabla_{P_k^\star(x_{m_k})} I^\star(x_{\mathcal{K}}) = I_k^\star(x_{m_k}) - (1-\rho). \tag{30}$$

Combining (29) and (30), we have

$$I_k^\star(x_{m_k})=\nu_k^\star+(1-\rho),\quad P_k^\star(x_{m_k})>0, \tag{31a}$$
$$I_k^\star(x_{m_k})\leq\nu_k^\star+(1-\rho),\quad P_k^\star(x_{m_k})=0, \tag{31b}$$

which suggests

$$\sum_{m_k}P_k^\star(x_{m_k})I_k^\star(x_{m_k})=\nu_k^\star+(1-\rho). \tag{32}$$

On the other hand, by definition of weighted sum marginal information (16),

$$\sum_{m_k}P_k^\star(x_{m_k})I_k^\star(x_{m_k})=I^\star(x_\mathcal{K}), \tag{33}$$

where the right-hand side is irrelevant to k. (31), (32), and (33) together complete the proof. □

B. Proof of Proposition 2

We first prove sequence (21) 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 $k\in\mathcal{K}$ at state $m_k\in\mathcal{M}$, and define an intermediate function $J(P_\mathcal{K}(x_{m_\mathcal{K}}),P'_\mathcal{K}(x_{m_\mathcal{K}}))$ as (34) at the end of page 8. It is straightforward to verify $J(P_\mathcal{K}(x_{m_\mathcal{K}}),P_\mathcal{K}(x_{m_\mathcal{K}}))=I(x_\mathcal{K})$ and $J(P_\mathcal{K}(x_{m_\mathcal{K}}),P'_\mathcal{K}(x_{m_\mathcal{K}}))$ is a concave function for a fixed $P'_\mathcal{K}(x_{m_\mathcal{K}})$. By choosing $\nabla_{P_k^\star(x_{m_k})}J(P_\mathcal{K}(x_{m_\mathcal{K}}),P'_\mathcal{K}(x_{m_\mathcal{K}}))=0$, we have

$$S'_k(x_{m_k})-S'_k(x_{i_k})+(1-\rho)\log\frac{P_k(x_{i_k})}{P_k^\star(x_{m_k})}=0, \tag{35}$$

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

$$S'_k(x_{m_k})\triangleq I'_k(x_{m_k})+(1-\rho)\sum_{m_{\mathcal{K}\setminus\{k\}}}P_{\mathcal{K}\setminus\{k\}}(x_{m_{\mathcal{K}\setminus\{k\}}})\times\sum_{m'_\mathcal{K}}P(\hat{x}_{m'_\mathcal{K}}|x_{m_\mathcal{K}})\log P'_\mathcal{K}(x_{m_\mathcal{K}}). \tag{36}$$

Evidently, $\forall m_k\neq i_k$, (35) boils down to

$$P_k^\star(x_{m_k})=\frac{P'_k(x_{m_k})\exp\left(\frac{\rho}{1-\rho}I'_k(x_{m_k})\right)}{\sum_{m'_k}P'_k(x_{m'_k})\exp\left(\frac{\rho}{1-\rho}I'_k(x_{m'_k})\right)}. \tag{37}$$

We also notice $P_k(x_{i_k})=1-\sum_{m_k\neq i_k}P_k^\star(x_{m_k})$ has exactly the same expression as (37). Therefore, the result is irrelevant to the choice of reference state, and (37) 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 (37) ensures

$$J(P_\mathcal{K}(x_{m_\mathcal{K}}),P'_\mathcal{K}(x_{m_\mathcal{K}}))\geq I'(x_\mathcal{K}). \tag{38}$$

On the other hand, it also guarantees

$$\Delta\triangleq I(x_\mathcal{K})-J(P_\mathcal{K}(x_{m_\mathcal{K}}),P'_\mathcal{K}(x_{m_\mathcal{K}})) \tag{39a}$$
$$=(1-\rho)\sum_{m_k}\frac{P'_k(x_{m_k})f'_k(x_{m_k})}{\sum_{m'_k}P'_k(x_{m'_k})f'_k(x_{m'_k})}\sum_{m''_\mathcal{K}}P(\hat{x}_{m''_\mathcal{K}}|x_{m_k})\times\log\frac{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_\mathcal{K}}|x_{m'_k})f'_k(x_{m_k})}{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_\mathcal{K}}|x_{m'_k})f'_k(x_{m'_k})} \tag{39b}$$
$$\geq(1-\rho)\sum_{m_k}\frac{P'_k(x_{m_k})f'_k(x_{m_k})}{\sum_{m'_k}P'_k(x_{m'_k})f'_k(x_{m'_k})}\sum_{m''_\mathcal{K}}P(\hat{x}_{m''_\mathcal{K}}|x_{m_k})\times\left(1-\frac{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_\mathcal{K}}|x_{m'_k})f'_k(x_{m'_k})}{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_\mathcal{K}}|x_{m'_k})f'_k(x_{m_k})}\right) \tag{39c}$$
$$=0, \tag{39d}$$

where $f'_k(x_{m_k})\triangleq\exp\left(\frac{\rho}{1-\rho}I'_k(x_{m_k})\right)$ and the equality holds if and only if (37) converges. (38) and (39) 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 in mutual information.

Next, we prove any converging point of sequence (21), denoted as $P_k^\star(x_{m_k})$, fulfills KKT conditions (20). To see this, define

$$D_k^{(r)}(x_{m_k})\triangleq\frac{P_k^{(r+1)}(x_{m_k})}{P_k^{(r)}(x_{m_k})}=\frac{f_k^{(r)}(x_{m_k})}{\sum_{m'_k}P_k^{(r)}(x_{m'_k})f_k^{(r)}(x_{m'_k})}. \tag{40}$$

As sequence (21) 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

$$I_k^\star(x_{m_k})=\frac{1-\rho}{\rho}\log\sum_{m'_k}P_k^\star(x_{m'_k})f_k^\star(x_{m'_k}), \tag{41}$$

which is reminiscent of (31a) and (20a). That is to say, given $P_k^{(0)}(x_{m_k})>0$, any converging point with $P_k^\star(x_{m_k})>0$ must satisfy (20a). On the other hand, we assume $P_k^\star(x_{m_k})$ does not satisfy (20b), such that for any state with $P_k^\star(x_{m_k})=0$,

$$I_k^\star(x_{m_k})>I^\star(x_\mathcal{K})=\sum_{m'_k}P_k^\star(x_{m'_k})I_k^\star(x_{m'_k}), \tag{42}$$

where the equality inherits from (17). Since exponential function is monotonically increasing, we have $f_k^\star(x_{m_k})>\sum_{m'_k}P_k^\star(x_{m'_k})f_k^\star(x_{m'_k})$ and $D_k^\star(x_{m_k})>1$. Considering $P_k^{(0)}(x_{m_k})>0$ and $P_k^\star(x_{m_k})=0$, it contradicts with

$$P_k^{(r)}(x_{m_k})=P_k^{(0)}(x_{m_k})\prod_{n=1}^{r}D_k^{(n)}(x_{m_k}). \tag{43}$$

Therefore, given $P_k^{(0)}(x_{m_k})>0$, any converging point with $P_k^\star(x_{m_k})=0$ must satisfy (20b). This completes the proof. □

$$J(P_\mathcal{K}(x_{m_\mathcal{K}}),P'_\mathcal{K}(x_{m_\mathcal{K}}))\triangleq\rho\sum_{m_\mathcal{K}}P_\mathcal{K}(x_{m_\mathcal{K}})\log\left(1+\frac{|\boldsymbol{h}_\mathrm{E}^H(x_{m_\mathcal{K}})\boldsymbol{w}|^2}{\sigma_v^2}\right)+(1-\rho)\sum_{m_\mathcal{K}}P_\mathcal{K}(x_{m_\mathcal{K}})\sum_{m'_\mathcal{K}}P(\hat{x}_{m'_\mathcal{K}}|x_{m_\mathcal{K}})\log\frac{P(\hat{x}_{m'_\mathcal{K}}|x_{m_\mathcal{K}})P'_\mathcal{K}(x_{m_\mathcal{K}})}{P'(\hat{x}_{m'_\mathcal{K}})P_\mathcal{K}(x_{m_\mathcal{K}})}. \tag{34}$$

C. Proof of Proposition 3

Since L input letters are with non-zero probability and $x \to z \to \hat{x}$ formulates a Markov chain, we need L non-empty decision regions and at least $L+1$ distinct thresholds (including 0 and ∞) to minimize the distortion between source and decision. On the other hand, the optimal decision regions are apparently empty for those unused letters.

Suppose the optimal number of thresholds is $S+1$ where $S \geq L$. Let $\boldsymbol{t} \triangleq [t_0,...,t_S]^T \in \mathbb{R}_+^{(S+1)\times 1}$ be the optimal threshold vector where $t_{s-1} < t_s$, $\forall s \in \mathcal{S} \triangleq \{1,...,S\}$. Since the optimal decision region for any letter may consist of multiple partitions, without loss of generality, we assume the mapping from threshold vector to decision region $l' \in \mathcal{L} \triangleq \{1,...,L\}$ is given by $\mathcal{R}_{l'} = \bigcup_{s \equiv l' \pmod{L}} [t_{s-1}, t_s)$. [4]he proof holds for any valid mapping from threshold vector to decision regions, and we consider this specific case for the ease of presentation. The threshold optimization problem is

$$\max_{\boldsymbol{t}} \quad I_{\mathrm{B}}(x;\hat{x}) \tag{44a}$$

$$\text{s.t.} \quad t_{s-1} < t_s, \quad \forall s \in \mathcal{S}. \tag{44b}$$

Problem (44) is intricate due to the strict inequality constraint (44b). Following [48], we first relax it to the convex counterpart, then discard the solutions that violate any original constraint. The Lagrangian function for the relaxed problem is

$$L = -I_{\mathrm{B}}(x;\hat{x}) + \sum_{s \in \mathcal{S}} \mu_s (t_{s-1} - t_s), \tag{45}$$

where μ_s is the Lagrange multiplier associated with the non-strict version of (44b). The KKT conditions on the optimal primal and dual solutions are, $\forall s \in \mathcal{S}$,

$$-\nabla_{t_s^\star} I_{\mathrm{B}}^\star(x;\hat{x}) + \mu_{s-1}^\star - \mu_s^\star = 0, \tag{46a}$$

$$\mu_s^\star \geq 0, \tag{46b}$$

$$\mu_s^\star (t_{s-1}^\star - t_s^\star) = 0. \tag{46c}$$

Due to the strict inequality constraint (44b), conditions (46b) and (46c) together imply $\mu_s^\star = 0$, $\forall s \in \mathcal{S}$. Besides, it is trivial to conclude $t_0^\star = 0$ for energy-based detection. As such, the necessary optimality conditions for problem (44), $\forall s \in \mathcal{S}$,

$$\nabla_{t_s^\star} I_{\mathrm{B}}^\star(x;\hat{x}) = 0, \tag{47}$$

which can be explicitly written as, $\forall s \equiv l' \pmod{L}$,

$$\sum_l P(x_l) \frac{(t_s^\star)^{N-1} \exp\{-t_s^\star/\sigma_l^2\}}{\sigma_l^{2N}(N-1)!} \log \frac{P(x_l|\hat{x}_{l'+1})}{P(x_l|\hat{x}_{l'})} = 0, \tag{48}$$

According to [46], the optimal backward channel quantizer is convex and separates each pair of posterior distribution by a hyperplane. It implies, for a given output letter l', the sequence $\{\log P(x_l|\hat{x}_{l'+1})/P(x_l|\hat{x}_{l'})\}_{l \in \mathcal{L}}$ changes sign exactly once. We notice the left-hand side of (48) is a generalized Dirichlet polynomial, and by Descartes' rule of signs [49], has at most one positive solution. In other words, starting from $t_0^\star$, each optimal decision region requires at most one additional distinct threshold, and we have $S \leq L$. Therefore, we conclude $S = L$ and the proof is completed. □

[4]T

References

[1] V. Liu, A. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. R. Smith, "Ambient backscatter: Wireless communication out of thin air," *ACM SIGCOMM Computer Communication Review*, vol. 43, pp. 39–50, 9 2013. [Online]. Available: https://dl.acm.org/doi/10.1145/2534169.2486015

[2] G. Yang, Y.-C. C. Liang, R. Zhang, and Y. Pei, "Modulation in the air: Backscatter communication over ambient ofdm carrier," *IEEE Transactions on Communications*, vol. 66, pp. 1219–1233, 3 2018. [Online]. Available: http://ieeexplore.ieee.org/document/8103807/

[3] G. Yang, Q. Zhang, and Y.-C. Liang, "Cooperative ambient backscatter communications for green internet-of-things," *IEEE Internet of Things Journal*, vol. 5, pp. 1116–1130, 4 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8274950/

[4] Y.-C. Liang, Q. Zhang, E. G. Larsson, and G. Y. Li, "Symbiotic radio: Cognitive backscattering communications for future wireless networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, pp. 1242–1255, 12 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9193946/

[5] H. Guo, Y.-C. Liang, R. Long, and Q. Zhang, "Cooperative ambient backscatter system: A symbiotic radio paradigm for passive iot," *IEEE Wireless Communications Letters*, vol. 8, pp. 1191–1194, 8 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8692391/

[6] H. Ding, D. B. da Costa, and J. Ge, "Outage analysis for cooperative ambient backscatter systems," *IEEE Wireless Communications Letters*, vol. 9, pp. 601–605, 5 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8941106/

[7] R. Long, Y.-C. Liang, H. Guo, G. Yang, and R. Zhang, "Symbiotic radio: A new communication paradigm for passive internet of things," *IEEE Internet of Things Journal*, vol. 7, pp. 1350–1363, 2 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8907447/

[8] S. Zhou, W. Xu, K. Wang, C. Pan, M.-S. Alouini, and A. Nallanathan, "Ergodic rate analysis of cooperative ambient backscatter communication," *IEEE Wireless Communications Letters*, vol. 8, pp. 1679–1682, 12 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8807353/

[9] T. Wu, M. Jiang, Q. Zhang, Q. Li, and J. Qin, "Beamforming design in multiple-input-multiple-output symbiotic radio backscatter systems," *IEEE Communications Letters*, vol. 25, pp. 1949–1953, 6 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9358202/

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

[11] H. Yang, Y. Ye, K. Liang, and X. Chu, "Energy efficiency maximization for symbiotic radio networks with multiple backscatter devices," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1431–1444, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9461158/

[12] S. Han, Y.-C. Liang, and G. Sun, "The design and optimization of random code assisted multi-bd symbiotic radio system," *IEEE Transactions on Wireless Communications*, vol. 20, pp. 5159–5170, 8 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9382925/

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

[14] S. Zhang and R. Zhang, "Capacity characterization for intelligent reflecting surface aided mimo communication," *IEEE Journal on Selected Areas in Communications*, vol. 38, pp. 1823–1838, 8 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9110912/

[15] S. Lin, B. Zheng, F. Chen, and R. Zhang, "Intelligent reflecting surface-aided spectrum sensing for cognitive radio," *IEEE Wireless Communications Letters*, vol. 11, pp. 928–932, 5 2022.

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

[17] 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.

[18] 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.

[19] Y. Yang, S. Zhang, and R. Zhang, "Irs-enhanced ofdma: Joint resource allocation and passive beamforming optimization," *IEEE Wireless Communications Letters*, vol. 9, pp. 760–764, 6 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8964457/

[20] Q. Wu, X. Zhou, W. Chen, J. Li, and X. Zhang, "Irs-aided wpcns: A new optimization framework for dynamic irs beamforming," *IEEE Transactions on Wireless Communications*, pp. 1–1, 12 2021.

[21] M. Hua and Q. Wu, "Joint dynamic passive beamforming and resource allocation for irs-aided full-duplex wpcn," *IEEE Transactions on Wireless Communications*, pp. 1–1, 12 2021.

[22] C. Qiu, Q. Wu, M. Hua, X. guan, and Y. Wu, "Achieving multi-beam gain in intelligent reflecting surface assisted wireless energy transfer," *arXiv:2205.08893*, 5 2022. [Online]. Available: http://arxiv.org/abs/2205.08893

[23] H. Chen, G. Yang, and Y.-C. Liang, "Joint active and passive beamforming for reconfigurable intelligent surface enhanced symbiotic radio system," *IEEE Wireless Communications Letters*, vol. 10, pp. 1056–1060, 5 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9345739/

[24] Q. Zhang, Y.-C. Liang, and H. V. Poor, "Reconfigurable intelligent surface assisted mimo symbiotic radio networks," *IEEE Transactions on Communications*, vol. 69, pp. 4832–4846, 7 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9391685/

[25] R. Karasik, O. Simeone, M. D. Renzo, and S. S. Shitz, "Beyond max-snr: Joint encoding for reconfigurable intelligent surfaces," vol. 2020-June. IEEE, 6 2020, pp. 2965–2970. [Online]. Available: https://ieeexplore.ieee.org/document/9174060/

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

[27] M. Hua, Q. Wu, L. Yang, R. Schober, and H. V. Poor, "A novel wireless communication paradigm for intelligent reflecting surface based symbiotic radio systems," *IEEE Transactions on Signal Processing*, vol. 70, pp. 550–565, 4 2022. [Online]. Available: http://arxiv.org/abs/2104.09161https://ieeexplore.ieee.org/document/9652042/

[28] S. Hu, C. Liu, Z. Wei, Y. Cai, D. W. K. Ng, and J. Yuan, "Beamforming design for intelligent reflecting surface-enhanced symbiotic radio systems," *arxiv:2110.10316*, 10 2021. [Online]. Available: http://arxiv.org/abs/2110.10316

[29] Q. Zhang, Y.-C. Liang, H.-C. Yang, and H. V. Poor, "Mutualistic mechanism in symbiotic radios: When can the primary and secondary transmissions be mutually beneficial?" *IEEE Transactions on Wireless Communications*, vol. 1276, pp. 1–1, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9751388/

[30] Y. Huang, A. Alieldin, and C. Song, "Equivalent circuits and analysis of a generalized antenna system," *IEEE Antennas and Propagation Magazine*, vol. 63, pp. 53–62, 4 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9392844/

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

[32] J. Kim and B. Clerckx, "Wireless information and power transfer for iot: Pulse position modulation, integrated receiver, and experimental validation," *arXiv:2104.08404*, pp. 1–15, 4 2021. [Online]. Available: http://arxiv.org/abs/2104.08404

[33] C. Boyer and S. Roy, "Backscatter communication and rfid: Coding, energy, and mimo analysis," *IEEE Transactions on Communications*, vol. 62, pp. 770–785, 3 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6685977/

[34] ——, "Coded qam backscatter modulation for rfid," *IEEE Transactions on Communications*, vol. 60, pp. 1925–1934, 7 2012. [Online]. Available: http://ieeexplore.ieee.org/document/6202629/

[35] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Transactions on Communications*, vol. 69, pp. 3313–3351, 5 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9326394/

[36] 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, 4 2012. [Online]. Available: http://ieeexplore.ieee.org/document/6153042/

[37] D. Bharadia, K. R. Joshi, M. Kotaru, and S. Katti, "Backfi: High throughput wifi backscatter," vol. 45. ACM, 8 2015, pp. 283–296. [Online]. Available: https://dl.acm.org/doi/10.1145/2785956.2787490

[38] G. Yang, C. K. Ho, and Y. L. Guan, "Multi-antenna wireless energy transfer for backscatter communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 33, pp. 2974–2987, 12 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7274644/

[39] H. Guo, Q. Zhang, S. Xiao, and Y.-C. Liang, "Exploiting multiple antennas for cognitive ambient backscatter communication," *IEEE Internet of Things Journal*, vol. 6, pp. 765–775, 2 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8411483/

[40] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, pp. 5394–5409, 11 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8811733/

[41] J. Qian, Y. Zhu, C. He, F. Gao, and S. Jin, "Achievable rate and capacity analysis for ambient backscatter communications," *IEEE Transactions on Communications*, vol. 67, pp. 6299–6310, 9 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8721108/

[42] T. Nguyen and T. Nguyen, "Capacity achieving quantizer design for binary channels," *IEEE Communications Letters*, vol. 25, pp. 759–763, 3 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9247251/

[43] M. Rezaeian and A. Grant, "Computation of total capacity for discrete memoryless multiple-access channels," *IEEE Transactions on Information Theory*, vol. 50, pp. 2779–2784, 11 2004. [Online]. Available: http://ieeexplore.ieee.org/document/1347364/

[44] J. Buhler and G. Wunder, "A note on capacity computation for the discrete multiple access channel," *IEEE Transactions on Information Theory*, vol. 57, pp. 1906–1910, 4 2011. [Online]. Available: https://ieeexplore.ieee.org/document/5730559/

[45] G. J. O. Jameson, "The incomplete gamma functions," *The Mathematical Gazette*, vol. 100, pp. 298–306, 7 2016. [Online]. Available: https://www.cambridge.org/core/product/identifier/S002555721600067X/type/journal_article

[46] X. He, K. Cai, W. Song, and Z. Mei, "Dynamic programming for sequential deterministic quantization of discrete memoryless channels," *IEEE Transactions on Communications*, vol. 69, pp. 3638–3651, 6 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9366549/

[47] T. Nguyen and T. Nguyen, "On thresholding quantizer design for mutual information maximization: Optimal structures and algorithms," vol. 2020-May. IEEE, 5 2020, pp. 1–5. [Online]. Available: https://ieeexplore.ieee.org/document/9128966/

[48] T. D. Nguyen and T. Nguyen, "On binary quantizer for maximizing mutual information," *IEEE Transactions on Communications*, vol. 68, pp. 5435–5445, 9 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9118952/

[49] G. J. O. Jameson, "Counting zeros of generalised polynomials: Descartes' rule of signs and laguerre's extensions," *The Mathematical Gazette*, vol. 90, pp. 223–234, 7 2006. [Online]. Available: https://www.cambridge.org/core/product/identifier/S0025557200179628/type/journal_article