Metascatter: Unifying Symbiotic Radio and Intelligent Reflecting Surface

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

***Abstract*—Backscatter allows wireless nodes to harvest energy from, modulate information over, and reconfigure propagation of surrounding radio waves. We uniquely introduce the concept of Metascatter, which adapts the input probability distribution of a finite-state backscatter device based on Channel State Information (CSI), to simultaneously act as parasitic source of Symbiotic Radio (SR) that conveys sensor data, and reflecting element of Intelligent Reflecting Surface (IRS) that assists legacy transmission. Using shared spectrum, energy, and infrastructures, fully cognitive networks can be built over Metascatter, and we consider a fundamental scenario where a multi-antenna Access Point (AP) and multiple Metascatter simultaneously transmit to a single-antenna user. Since tags involve load switching and the primary and backscatter messages are multiplied (instead of superposed) in the received signal, it is assumed the user employ joint energy detection on all tags, then model their reflection pattern within the equivalent CSI for primary decoding. We characterize the achievable primary-(sum-)backscatter rate region through a joint optimization of transmit beamforming vector at the AP, input probability distribution at the tags, and energy detection regions at the user. A suboptimal solution is provided by the Block Coordinate Descent (BCD) method, where the beamforming, input distribution, and decision thresholds are obtained by Successive Convex Approximation (SCA), Karush–Kuhn–Tucker (KKT), and Dynamic Programming (DP), respectively. Simulation results demonstrate Metascatter can exploit the additional backscatter path to transmit and assist.**

I. Introduction

BACKSCATTER is recently re-innovated as a promising approach to support low-power communications and control wireless propagation environments. The concept of Ambient Backscatter Communications (AmBC) was first proposed in [1], where each tag harvests energy from and modulates information over ambient Radio Frequency (RF) signals by periodically switching between different reflection patterns. It enables battery-free communication between fully passive nodes, but the backscatter signal strength can be relatively weak and the decoding is subject to strong direct-link interference. To solve this, [2] exploited the repeating structure of Orthogonal Frequency-Division Multiplexing (OFDM) symbol for interference cancellation based on backscatter channel strength. Cooperative primary-backscatter decoding using co-located receiver, known as cooperative AmBC, was proposed in [3]. 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 with ambient OFDM carrier. When transmit cooperation is also available, cooperative AmBC is more commonly known as Symbiotic Radio (SR) [4]. The authors of [5] classified SR into commensal, parasitic, and competitive modes based on the relative link priority, then derived their instantaneous achievable rate and optimal power allocation. The outage probability for those modes was later studied in [6]. On the other hand, it was concluded in [7] that if the backscatter symbol period is much longer than the primary symbol period, then the non-coherent achievable rate of primary decoding would approach that of the coherent case. The authors also considered the transmit precoder design for weighted sum-rate maximization and transmit power minimization problems. [8] explored how the number of transmit antennas, receive antennas, and symbol period ratio asymptotically influence the ergodic rate of primary and backscatter links. With multi-antenna at all ends, [9] optimized the transmit precoder to maximize the backscatter achievable rate under primary rate constraint. However, those paper only considered single tag scenario and the optimal tag multiple access scheme remains unknown. [10] proposed a Non-Orthogonal Multiple Access (NOMA)-based SR and studied the receive combiner design when the receiver decodes in the order of equivalent channel strength. A Time-Division Multiple Access (TDMA)-based SR with energy harvesting constraints was also presented in [11], and energy efficiency maximization was considered with respect to transmit power, reflection efficiency, and time allocation design. In [12], random code-assisted backscatter multiple access was combined with SR, and the authors evaluated the asymptotic Signal-to-Interference-plus-Noise Ratio (SINR) performance using random matrix theory.

On the other hand, Intelligent Reflecting Surface (IRS) has recently emerged as a promising technique to enhance the energy efficiency. In practice, an IRS consists of multiple individual sub-wavelength reflecting elements to adjust the amplitude and phase of the incoming signal (i.e., passive beamforming). Different from relay, backscatter and frequency-selective surface [13], IRS assists the primary transmission using passive components with much lower energy consumption and thermal noise, but is limited to frequency-dependent reflection. [14] proposed an IRS-assisted Multiple-Input Single-Output (MISO) system and jointly optimized the precoder and the phase shifts to minimize the transmit power. The active and passive beamforming problem was then extended to the discrete phase shift case [15] and the multi-user case [16]. In [17], the authors investigated the impact of non-zero resistance on the reflection pattern and emphasized the coupling between reflection amplitude and phase shift in practice. To estimate the cascaded link without RF-chains at the IRS, practical protocols were developed based on element-

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).

wise on/off switching [18], training sequence and reflection pattern design [19], [20], and compressed sensing [21]. IRS was also added to single-tag and multi-tag SR systems to improve backscatter efficiency [22], [23]. Recently, the idea of IRS-empowered SR was proposed in [24], [25], where the reflection pattern can be decoupled as the multiplication of whole phase shift matrix and one backscatter symbol. The idea was combined with spatial modulation to divide the IRS into reflection and information and elements, and the error performance of non-coherent backscatter detection was studied accordingly [26].

To the best of our knowledge, most existing SR systems [3], [5]–[10] assumed Gaussian codebook at the tags and employed SIC from primary to backscatter link. It thus requires non-coherent primary encoding at the transmitter and involves re-encoding, precoding, and subtraction operations that may not be available to legacy receivers. Besides, the primary and backscatter symbols in SR are mixed by multiplication coding instead of superposition coding, and the backscatter symbol period can be much longer than the primary symbol period due to physical constraints. Inspired by above, we propose a novel Metascatter network where multiple battery-free tags ride over a conventional point-to-point system and adapt their input probability distribution to simultaneously act as backscatter tags of SR and reflecting elements of IRS. To fully utilize the reflection pattern and signal structure, we also present a novel decoding strategy with minor modifications on existing receivers. The contributions of this paper is summarized as follows.

First, Metascatter adapts the input probability distribution of a finite-state backscatter device based on the primary and backscatter Channel State Information (CSI). It generalizes the backscatter tags of SR and reflecting elements of IRS to improve the primary-backscatter tradeoff in a flexible manner. When the primary link is absolutely prioritized, the tag input probability boils down to 1 at the state that maximizes the equivalent channel strength and 0 at other states, which is essentially the phase shift selection of an IRS element. When we only care about the backscatter performance, the network boils down to a multi-tag AmBC and the input design can be simplified accordingly.

Second, we propose a novel receiving strategy where the backscatter symbol of all tags are first recovered by energy-based decoding, then modeled within equivalent CSI for primary decoding. When the ratio of backscatter symbol period over primary symbol period is large enough, the randomness of primary source can be averaged out and the performance of non-coherent backscatter decoding can be greatly improved. Since the primary and backscatter symbols are mixed by multiplication coding, the backscatter decoding can be viewed as part of primary channel training, which avoids non-coherent encoding at the transmitter and SIC at the receiver.

Third, we characterize the achievable primary-(sum-)backscatter rate region by iteratively optimizing the tag input distribution, the backscatter decision threshold, and the transmit precoder. For the input design, we propose a numerical method to evaluate the Karush–Kuhn–Tucker (KKT) solutions by limits of sequences. For the threshold design, we derive the minimum number of decision thresholds to maximize the total backscatter rate, and obtain the optimal thresholds by Dynamic Programming (DP) accelerated by Shor-Moran-Aggarwal-Wilber-Klawe (SMAWK) algorithm. However, the optimal precoder design can be non-trivial and we may end up with a low-complexity suboptimal solution.

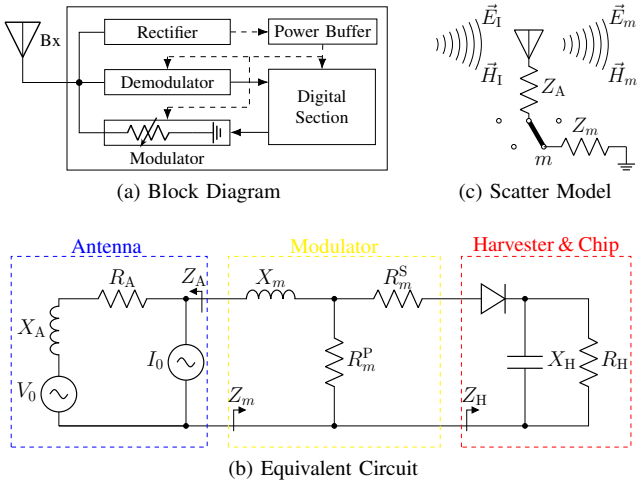


Fig. 1. Block diagram, equivalent circuit, and scatter model of a typical passive tag. The solid and dashed vectors represent signal and energy flows. The backscatter antenna behaves as a constant power source, where the voltage V_0 and current I_0 are introduced by incident electric field $\vec{E}_{\rm I}$ and magnetic field $\vec{H}_{\rm I}$ [27].

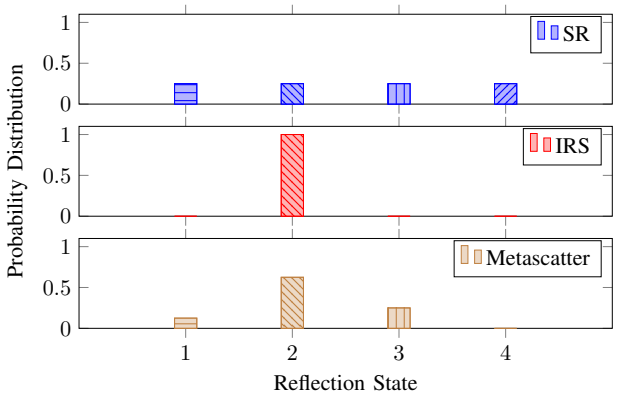
II. Backscatter Principles

Passive backscatter devices (a.k.a., tags) harvest energy from and modulate information over surrounding RF signal. As shown in Fig. 1(a), a typical passive tag 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) [28]. Its equivalent circuit is also given by Fig. 1(b). When illuminated, the tag absorbs a portion of the impinging wave and scatters the rest back to the space. The absorbed signal can be used for source decoding and energy harvesting [29], while the scattered signal can be decomposed into the *structural* and *antenna* mode components [30]. The structural mode scattering applies to general objects and is related to the antenna structure, shape and material, which contributes to environment multipath and can be covered by channel estimation [31]. On the other hand, the antenna mode scattering models the antenna radiation and depends on the antenna-load impedance mismatch, which can be used for backscatter modulation [32] or channel reconfiguration [33]. Fig. 1(c) illustrates the scatter model. For a tag with M candidate states, the reflection coefficient of state $m \in \mathcal{M} \triangleq \{1,...,M\}$ is defined as [1]t corresponds to a linear backscatter model where the reflection coefficient is irrelevant to incident electromagnetic field strength.

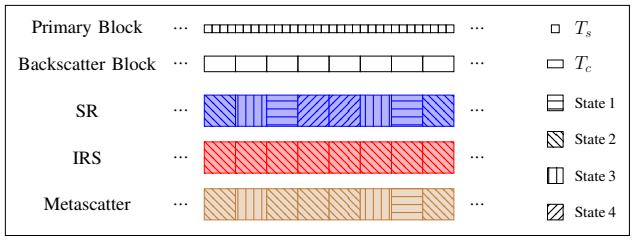
$$\Gamma_m = \frac{Z_m - Z_{\rm A}^*}{Z_m + Z_{\rm A}}, \tag{1}$$

where Z_m is the load impedance at state m and $Z_{\rm A}$ is the antenna input impedance.

[1]I



(a) Reflection State Distribution



(b) Time Block Structure

Fig. 2. Reflection state distribution and time block structure of SR, IRS, and Metascatter. T_s and T_c denote the primary and backscatter symbol length, respectively. Based on CSI, Metascatter adapts the reflection state distribution to unify data transmission and channel reconfiguration.

A. Backscatter Modulation

In backscatter modulation, the tag *periodically varies* the reflection coefficient to embed its own information. As such, there exists a mapping from the alphabet to the set. When the CSI and reflection alphabet $\{\Gamma_1,...,\Gamma_M\}$ are available at the reader, it can decode the tag message from the observed scattered signal.

For M-ary Quadrature Amplitude Modulation (QAM), the complex constellation point c_m maps to the corresponding reflection coefficient by [34]

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

where $\alpha \in [0,1]$ models the amplitude scatter ratio at the direction of interest.

B. Channel Reconfiguration

2 For passive tags, $\alpha \ll 1$ is commonly assumed as the majority of the incident wave can be harvested to support tag modules [34].

Remark 1. *By choosing proper load impedance, it is possible to implement any signal constellation or think beyond modulation. For example, perfect matching enables maximum power transfer to the tag, while perfect mismatch leads to complete reflection as desired by ideal IRS. It inspires Metascatter to design a smart reflection and unify wireless power transfer, backscatter modulation, and passive beamforming.*

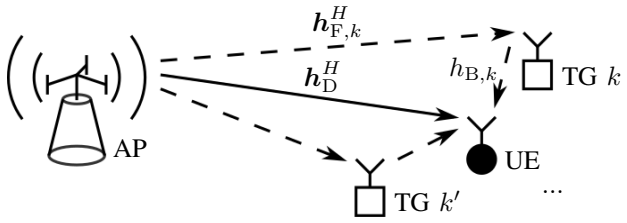


Fig. 3. The proposed Metascatter system.

III. Metascatter

A. System Model

As shown in Fig. 3, we propose a single-user (UE) multi-tag (TG) symbiotic radio network where the RF signal generated by the Q-antenna Access Point (AP) is shared by two coexisting systems. In the primary point-to-point system, the AP transmits information to the single-antenna user. In the secondary backscatter system, the AP acts as the carrier emitter, the K nearby single-antenna tags modulate their information over the reradiated RF signals, and the user receives and decodes the backscatter messages. Denote the AP-UE direct channel as $\boldsymbol{h}_{\mathrm{D}}^H \in \mathbb{C}^{1\times Q}$, the AP-TG $k \in \mathcal{K} \triangleq \{1,...,K\}$ forward channel as $\boldsymbol{h}_{\mathrm{F},k}^H \in \mathbb{C}^{1\times Q}$, the TG k-UE backward channel as $h_{\mathrm{B},k}$, and define the cascaded channel of tag k as $\boldsymbol{h}_{\mathrm{C},k}^H \triangleq h_{\mathrm{B},k}\boldsymbol{h}_{\mathrm{F},k}^H \in \mathbb{C}^{1\times Q}$. For simplicity, we consider a quasi-static block fading model where the channel response remain constant within each coherence interval and vary independently between consecutive blocks. Since backscatter modulation involves load switching, tags typically transmit at much lower rate, and we assume the backscatter symbol length is $N \in \mathbb{Z}_{++}$ times the primary symbol length. We also assume the direct channel and all cascaded channels can be estimated and fed back to the AP. [2]ue to the lack of RF chains at the passive tag, accurate and efficient CSI acquisition at the AP can be challenging. One possible approach is that the AP sends pre-defined pilots, the tags respond in well-designed manners, and the user performs least-square estimation with feedbacks [35]–[37].

Consider the signal model during one backscatter symbol period. We assume the primary symbol follows standard Circularly Symmetric Complex Gaussian (CSCG) distribution and the backscatter symbol of all tags employs M-QAM. Under perfect synchronization, the signal received by the user at primary symbol block $n \in \mathcal{N} \triangleq \{1,...,N\}$ is [3]e omit the signal reflected by two or more times [16] and assume the time difference of arrival from different paths are negligible [5].

$$y[n] = \left(\boldsymbol{h}_{\mathrm{D}}^H + \sum_{k\in\mathcal{K}} \sqrt{\alpha_k}\boldsymbol{h}_{\mathrm{C},k}^H x_k\right)\boldsymbol{w}s[n] + w[n], \tag{3}$$

where α_k and x_k denote respectively the harvest-backscatter efficiency and backscatter symbol of tag k, $s[n]$ and $w[n] \sim \mathcal{CN}(0,\sigma_w^2)$ denote respectively the primary symbol and Additive White Gaussian Noise (AWGN) at block n, and $\boldsymbol{w} \in \mathbb{C}^{Q\times 1}$ is the transmit precoder satisfying power constraint $\|\boldsymbol{w}\|^2 \le P$. For the ease of notation, we

[2]D
[3]W

define $x_{\mathcal{K}} \triangleq \{x_k : k \in \mathcal{K}\}$ as the tag input combination, and $\boldsymbol{y} \triangleq [y[1],...,y[N]]^T \in \mathbb{C}^{N\times 1}$ as the received signal per backscatter symbol period. For primary transmission, the equivalent channel is subject to backscatter modulation uncertainty as

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

Remark 2. *Metascatter involves an unconventional Multiple Access Channel (MAC) where the primary and backscatter symbols of different duration are mixed by Multiplication Coding (MC) instead of Superposition Coding (SC). For each tag, the reflection coefficient not only encodes the backscatter message but also influences the equivalent primary channel* (4)*. As such, novel receiving strategy other than SIC should be tailored to signal characteristics to unveil how reflection pattern potentially influences the primary-backscatter tradeoff.*

B. Receiving Strategy

We propose a novel receiving strategy where the backscatter symbol of all tags are first jointly identified by non-coherent energy detection, then modeled within equivalent channel (4) as if passive beamforming. Compared with conventional schemes as joint decoding and SIC, the Metascatter receiver may not achieve as high data rate for one tag, but it avoids non-coherent primary encoding and enables tag multiple access in a practical and low-complexity manner. Therefore, we believe it can be compatible to and readily implemented over legacy point-to-point systems.

Define $m_k \in \mathcal{M}$ as the modulation index of tag k, and $m_{\mathcal{K}} \triangleq \{m_k : k \in \mathcal{K}\}$ as the modulation index combination of all tags. Recall that x_k and $x_{\mathcal{K}}$ are random variables representing the backscatter symbol of tag k and the symbol combination of all tags, respectively. Their instances are denoted by modulation index (combination) as x_{m_k} and $x_{m_{\mathcal{K}}}$. Since any backscatter combination $x_{m_{\mathcal{K}}}$ remains unchanged over N primary symbol duration, the received signal at symbol block n is only subject to the variation of $s[n]$ and $w[n]$, and follows zero-mean CSCG distribution with variance

$$\sigma_{m_{\mathcal{K}}}^2 = \Big| \underbrace{\big(\boldsymbol{h}_{\mathrm{D}}^H + \sum_{k\in\mathcal{K}} \sqrt{\alpha_k}\boldsymbol{h}_{\mathrm{C},k}^H x_{m_k}\big)}_{\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})} \boldsymbol{w} \Big|^2 + \sigma_w^2. \tag{5}$$

Besides, we denote the total received energy as $z \triangleq \|\boldsymbol{y}\|^2$ and the hypothesis of modulation index combination $m_{\mathcal{K}}$ as $\mathcal{H}_{m_{\mathcal{K}}}$. Since z is the sum of N independent and identically distributed (i.i.d.) exponential variables, its conditional Probability Density Function (PDF) under $\mathcal{H}_{m_{\mathcal{K}}}$ follows Erlang distribution with shape N and scale $\sigma_{m_{\mathcal{K}}}^2$ as

$$f(z\,|\,\mathcal{H}_{m_{\mathcal{K}}}) = \frac{z^{N-1}e^{-z/\sigma_{m_{\mathcal{K}}}^2}}{\sigma_{m_{\mathcal{K}}}^{2N}(N-1)!}. \tag{6}$$

Energy-based backscatter detection essentially formulates a discrete-input continuous-output memoryless MAC. To further reduce decoding complexity, we divide the continuous output z into disjoint decision regions to construct a Discrete Memoryless Thresholding Multiple Access Channel (DMTMAC), where the transition probability from tag input combination $x_{m_{\mathcal{K}}}$ to detector output combination $\hat{x}_{m'_{\mathcal{K}}}$ is

$$P(\hat{x}_{m'_{\mathcal{K}}}\,|\,x_{m_{\mathcal{K}}}) = \int_{\mathcal{R}_{m'_{\mathcal{K}}}} f(z\,|\,\mathcal{H}_{m_{\mathcal{K}}})\mathrm{d}z, \tag{7}$$

where $\mathcal{R}_{m'_{\mathcal{K}}}$ is the decision region over z for hypothesis $\mathcal{H}_{m'_{\mathcal{K}}}$.

C. Achievable Rates

Denote the input probability distribution vector of tag k as $\boldsymbol{p}_k \triangleq [P_k(x_1),...,P_k(x_M)]^T \in \mathbb{R}^{M\times 1}$, where $P_k(x_{m_k})$ is the probability at state m_k. Consider independent encoding at all tags such that the probability of input combination $x_{m_{\mathcal{K}}}$ is $P_{\mathcal{K}}(x_{m_{\mathcal{K}}}) = \prod_{k\in\mathcal{K}} P_k(x_{m_k})$. The backscatter information function of input combination $x_{m_{\mathcal{K}}}$ is defined as

$$I_{\mathrm{B}}(x_{m_{\mathcal{K}}};\hat{x}_{\mathcal{K}}) \triangleq \sum_{m'_{\mathcal{K}}} P(\hat{x}_{m'_{\mathcal{K}}}\,|\,x_{m_{\mathcal{K}}}) \log \frac{P(\hat{x}_{m'_{\mathcal{K}}}\,|\,x_{m_{\mathcal{K}}})}{P(\hat{x}_{m'_{\mathcal{K}}})}, \tag{8}$$

where $P(\hat{x}_{m'_{\mathcal{K}}}) = \sum_{m_{\mathcal{K}}} P_{\mathcal{K}}(x_{m_{\mathcal{K}}}) P(\hat{x}_{m'_{\mathcal{K}}}\,|\,x_{m_{\mathcal{K}}})$. Besides, the backscatter marginal information function associated with letter x_{m_k} of tag k is

$$I_{\mathrm{B},k}(x_{m_k};\hat{x}_{\mathcal{K}}) \triangleq \sum_{m_{\mathcal{K}\setminus\{k\}}} P_{\mathcal{K}\setminus\{k\}}(x_{m_{\mathcal{K}\setminus\{k\}}}) I_{\mathrm{B}}(x_{m_{\mathcal{K}}};\hat{x}_{\mathcal{K}}), \tag{9}$$

where $P_{\mathcal{K}\setminus\{k\}}(x_{m_{\mathcal{K}\setminus\{k\}}}) = \prod_{q\in\mathcal{K}\setminus\{k\}} P_q(x_{m_q})$. Therefore, the backscatter mutual information is

$$I_{\mathrm{B}}(x_{\mathcal{K}};\hat{x}_{\mathcal{K}}) = \sum_{m_{\mathcal{K}}} P_{\mathcal{K}}(x_{m_{\mathcal{K}}}) \sum_{m'_{\mathcal{K}}} P(\hat{x}_{m'_{\mathcal{K}}}\,|\,x_{m_{\mathcal{K}}}) \log \frac{P(\hat{x}_{m'_{\mathcal{K}}}\,|\,x_{m_{\mathcal{K}}})}{P(\hat{x}_{m'_{\mathcal{K}}})}. \tag{10}$$

Once the tag input combination is successfully decoded, the backscatter uncertainty can be eliminated and the equivalent channel for primary transmission can be updated by (4). Similar to above, the primary information function of input combination $x_{m_{\mathcal{K}}}$ is

$$I_{\mathrm{P}}(x_{m_{\mathcal{K}}};\boldsymbol{y}) \triangleq N\log_2\left(1 + \frac{|\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2}{\sigma_w^2}\right), \tag{11}$$

the primary marginal information function associated with letter x_{m_k} of tag k is

$$I_{\mathrm{P},k}(x_{m_k};\boldsymbol{y}) \triangleq \sum_{m_{\mathcal{K}\setminus\{k\}}} P_{\mathcal{K}\setminus\{k\}}(x_{m_{\mathcal{K}\setminus\{k\}}}) I_{\mathrm{P}}(x_{m_{\mathcal{K}}};\boldsymbol{y}), \tag{12}$$

and the primary (ergodic) mutual information is

$$I_{\mathrm{P}}(x_{\mathcal{K}};\boldsymbol{y}) = \sum_{m_{\mathcal{K}}} P_{\mathcal{K}}(x_{m_{\mathcal{K}}}) N\log_2\left(1 + \frac{|\boldsymbol{h}_{\mathrm{E}}^H(x_{m_{\mathcal{K}}})\boldsymbol{w}|^2}{\sigma_w^2}\right). \tag{13}$$

Therefore, the weighted sum information function, marginal information function, and mutual information of primary and backscatter links are respectively given by

$$I(x_{m_{\mathcal{K}}};\hat{x}_{\mathcal{K}},\boldsymbol{y}) \triangleq \rho I_{\mathrm{P}}(x_{m_{\mathcal{K}}};\boldsymbol{y}) + (1-\rho) I_{\mathrm{B}}(x_{m_{\mathcal{K}}};\hat{x}_{\mathcal{K}}), \tag{14}$$

$$I_k(x_{m_k};\hat{x}_{\mathcal{K}},\boldsymbol{y}) \triangleq \rho I_{\mathrm{P},k}(x_{m_k};\boldsymbol{y}) + (1-\rho) I_{\mathrm{B},k}(x_{m_k};\hat{x}_{\mathcal{K}}), \tag{15}$$

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

where $\rho \in [0,1]$ represents the priority of the primary link.

IV. Input, Threshold, and Precoder Design

To characterize the achievable primary-(sum-)backscatter rate region of Metascatter, we aim to maximize the weighted sum mutual information with respect to tag input probability distribution $\{\boldsymbol{p}_k\}$, decision threshold vector $\boldsymbol{t}$, and transmit precoder $\boldsymbol{w}$

$$\max_{\{\boldsymbol{p}_k\},\boldsymbol{t},\boldsymbol{w}} \quad I(x_{\mathcal{K}};\hat{x}_{\mathcal{K}},\boldsymbol{y}) \tag{17a}$$

$$\text{s.t.} \quad \sum\nolimits_{m_k} P_k(x_{m_k})=1, \quad \forall k\in\mathcal{K}, \tag{17b}$$

$$P_k(x_{m_k})\geq 0, \quad \forall k\in\mathcal{K},\ \forall m_k\in\mathcal{M}, \tag{17c}$$

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

Since problem (17) is not jointly convex, we propose a Block Coordinate Descent (BCD) algorithm that iteratively updates $\{\boldsymbol{p}_k\}$, $\boldsymbol{t}$ and $\boldsymbol{w}$ until convergence.

A. Input Probability Distribution

For any fixed decision boundary $\boldsymbol{t}$ and transmit precoder $\boldsymbol{w}$, the equivalent DMTMAC can be formulated by (7) and problem (17) boils down to

$$\max_{\{\boldsymbol{p}_k\}} \quad I(x_{\mathcal{K}};\hat{x}_{\mathcal{K}},\boldsymbol{y}) \tag{18a}$$

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

which involves the product term $\prod_{k\in\mathcal{K}} P_k(x_{m_k})$ and is non-convex when $K>1$. Interestingly, the sum-rate optimal input design for general discrete memoryless MAC remains an open problem, and we instead propose a KKT solution to problem (18). [4]s pointed out in [38], KKT conditions are only necessary for sum-rate optimality and these solutions may end up being saddle points.

Proposition 1. *The KKT optimality conditions for problem* (18) *are equivalent to,* $\forall k\in\mathcal{K}$ *and* $\forall m_k\in\mathcal{M}$*,*

$$I_k^\star(x_{m_k};\hat{x}_{\mathcal{K}},\boldsymbol{y})=I^\star(x_{\mathcal{K}};\hat{x}_{\mathcal{K}},\boldsymbol{y}), \quad P_k^\star(x_{m_k})>0, \tag{19a}$$

$$I_k^\star(x_{m_k};\hat{x}_{\mathcal{K}},\boldsymbol{y})\leq I^\star(x_{\mathcal{K}};\hat{x}_{\mathcal{K}},\boldsymbol{y}), \quad P_k^\star(x_{m_k})=0. \tag{19b}$$

Proof. Please refer to Appendix A. □

We notice (19a) means each probable state of each tag should produce the same marginal information (averaged over all states of other tags), while (19b) implies any state of any tag with potentially less marginal information than above should not be used. Next, we generalize the Blahut-Arimoto algorithm [39], [40] and propose a numerical evaluation of KKT points by limits of sequences.

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

$$P_k^{(r+1)}(x_{m_k})=\frac{P_k^{(r)}(x_{m_k})\exp\Big(\frac{\rho}{1-\rho}I_k^{(r)}(x_{m_k};\hat{x}_{\mathcal{K}},\boldsymbol{y})\Big)}{\sum_{m_k'}P_k^{(r)}(x_{m_k'})\exp\Big(\frac{\rho}{1-\rho}I_k^{(r)}(x_{m_k'};\hat{x}_{\mathcal{K}},\boldsymbol{y})\Big)}, \tag{20}$$

where r *is the iteration index and* $\boldsymbol{p}_k^{(0)}>\boldsymbol{0}$*,* $\forall k\in\mathcal{K}$*. At each iteration, the input distribution of tag* k *is evaluated based on the updated input distribution of tags* 1 *to* $k-1$*.*

Proof. Please refer to Appendix B. □

B. Decision Threshold

For a given tag input distribution $\{\boldsymbol{p}_k\}$, 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 3. *Since the equivalent source transmits at the total backscatter rate, the DMTMAC* (7) *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* (6)*, 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 the total backscatter rate by DP accelerated by SMAWK algorithm [41].

> 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.

C. Precoder

> The optimal precoder design is highly non-convex that involves integration, entropy term, and variable on exponential. We may propose a suboptimal precoder based on linear combination of equivalent and cascaded channels, or consider single transmit antenna instead.

[4]A

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_w^2	1
Discretization bins	256

$$I(x_i;\hat{x}_j)=\sum_{j\in\mathcal{I}}\int_{t_{j-1,j}}^{t_{j,j+1}}\frac{z^{N-1}\exp\left(-\frac{z}{\mathrm{tr}(H_{\mathrm{E},i}W)+\sigma_w^2}\right)}{\left(\mathrm{tr}(H_{\mathrm{E},i}W)+\sigma_w^2\right)^N(N-1)!}\mathrm{d}z \times\log\frac{\int_{t_{j-1,j}}^{t_{j,j+1}}\frac{z^{N-1}\exp\left(-\frac{z}{\mathrm{tr}(H_{\mathrm{E},i}W)+\sigma_w^2}\right)}{\left(\mathrm{tr}(H_{\mathrm{E},i}W)+\sigma_w^2\right)^N(N-1)!}\mathrm{d}z}{\sum_{i'\in\mathcal{I}}\int_{t_{j-1,j}}^{t_{j,j+1}}\frac{z^{N-1}\exp\left(-\frac{z}{\mathrm{tr}(H_{\mathrm{E},i'}W)+\sigma_w^2}\right)}{\left(\mathrm{tr}(H_{\mathrm{E},i'}W)+\sigma_w^2\right)^N(N-1)!}\mathrm{d}z}, \tag{21}$$

V. Preliminary Results

Table I shows the parameters used in simulation. We assume all links are in standard 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 (20), 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 [41], "Bisection" sequentially optimizes each threshold by bisection [42], and "ML" is the ML detector that requires no knowledge of input distribution.

Figs. 4 and 5 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-(sum-)backscatter rate. It may result from a precision issue, but inspires further research on threshold design for Metascatter.

Appendix

A. Proof of Proposition 1

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

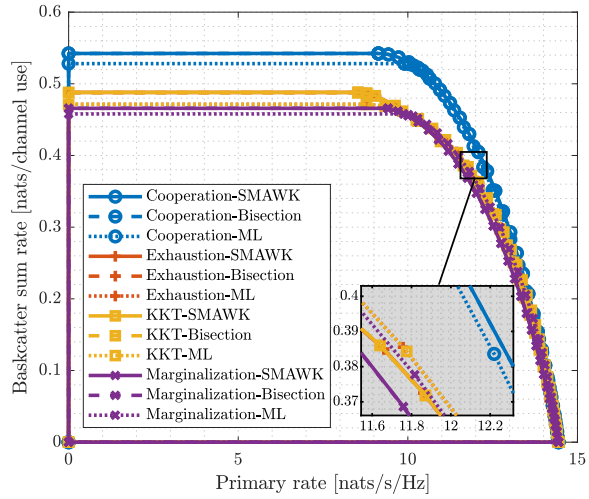


Fig. 4. Achievable rate regions by input-threshold design: Case I.

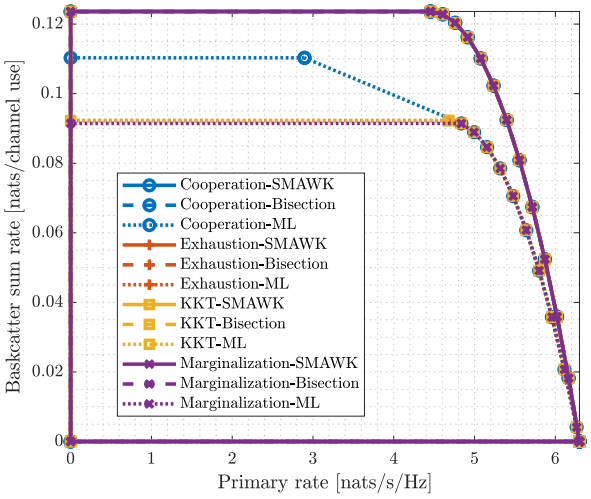


Fig. 5. Achievable rate regions by input-threshold design: Case II.

Lagrangian function of problem (18) is

$$L=-I(x_{\mathcal{K}};\hat{x}_{\mathcal{K}},\boldsymbol{y})+\sum_{k\in\mathcal{K}}\nu_k\left(\sum_{m_k\in\mathcal{M}}P_k(x_{m_k})-1\right) -\sum_{k\in\mathcal{K}}\sum_{m_k\in\mathcal{M}}\lambda_{k,m_k}P_k(x_{m_k}), \tag{22}$$

and the KKT conditions on the optimal primal and dual variables are, $\forall m_k\in\mathcal{M}$ and $\forall k\in\mathcal{K}$,

$$-\nabla_{P_k^\star(x_{m_k})}I^\star(x_{\mathcal{K}};\hat{x}_{\mathcal{K}},\boldsymbol{y})+\nu_k^\star-\lambda_{k,m_k}^\star=0, \tag{23a}$$

$$\lambda_{k,m_k}^\star=0,\quad P_k^\star(x_{m_k})>0, \tag{23b}$$

$$\lambda_{k,m_k}^\star\geq 0,\quad P_k^\star(x_{m_k})=0, \tag{23c}$$

where the directional derivative can be explicitly expressed as

$$\nabla_{P_k^\star(x_{m_k})}I^\star(x_{\mathcal{K}};\hat{x}_{\mathcal{K}},\boldsymbol{y})=I_k^\star(x_{m_k};\hat{x}_{\mathcal{K}},\boldsymbol{y})-(1-\rho). \tag{24}$$

Combining (23) and (24), we have

$$I_k^\star(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})=\nu_k^\star+(1-\rho),\quad P_k^\star(x_{m_k})>0, \tag{25a}$$
$$I_k^\star(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})\le\nu_k^\star+(1-\rho),\quad P_k^\star(x_{m_k})=0, \tag{25b}$$

which suggests

$$\sum_{m_k}P_k^\star(x_{m_k})I_k^\star(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})=\nu_k^\star+(1-\rho). \tag{26}$$

On the other hand, by definition of weighted sum marginal information (15),

$$\sum_{m_k}P_k^\star(x_{m_k})I_k^\star(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})=I^\star(x_\mathcal{K};\hat{x}_\mathcal{K},\boldsymbol{y}), \tag{27}$$

where the right-hand side is irrelevant to k. (25), (26), and (27) together complete the proof. □

B. Proof of Proposition 2

We first prove sequence (20) is non-decreasing in 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 (28). It is straightforward to verify $J\left(P_\mathcal{K}(x_{m_\mathcal{K}}),P_\mathcal{K}(x_{m_\mathcal{K}})\right)=I(x_\mathcal{K};\,\hat{x}_\mathcal{K},\,\boldsymbol{y})$ and $J\left(P_\mathcal{K}(x_{m_\mathcal{K}}),P'_\mathcal{K}(x_{m_\mathcal{K}})\right)$ 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

$$S'_k(x_{m_k})-S'_k(x_{i_k})+(1-\rho)\log\frac{P_k(x_{i_k})}{P_k^\star(x_{m_k})}=0, \tag{29}$$

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

$$S'_k(x_{m_k})\triangleq I'_k(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})+(1-\rho)\sum_{m_{\mathcal{K}\setminus\{k\}}}P_{\mathcal{K}\setminus\{k\}}(x_{m_{\mathcal{K}\setminus\{k\}}})\times\sum_{m'_\mathcal{K}}P(\hat{x}_{m'_\mathcal{K}}\,|\,x_{m_\mathcal{K}})\log P'_\mathcal{K}(x_{m_\mathcal{K}}). \tag{30}$$

Evidently, $\forall m_k\neq i_k$, (29) boils down to

$$P_k^\star(x_{m_k})=\frac{P'_k(x_{m_k})\exp\left(\frac{\rho}{1-\rho}I'_k(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})\right)}{\sum_{m'_k}P'_k(x_{m'_k})\exp\left(\frac{\rho}{1-\rho}I'_k(x_{m'_k};\hat{x}_\mathcal{K},\boldsymbol{y})\right)}. \tag{31}$$

Although it seems $P_k(x_{i_k})=1-\sum_{m_k\neq i_k}P_k^\star(x_{m_k})$ has no optimality guarantee, we can verify that $P_k(x_{i_k})$ has exactly the same form as (31). It implies the selection of reference state does not matter and (31) is indeed optimal $\forall m_k\in\mathcal{M}$. Therefore, for a fixed $P'_\mathcal{K}(x_{m_\mathcal{K}})$, choosing $P_k(x_{m_k})$ by (31) ensures

$$J(P_\mathcal{K}(x_{m_\mathcal{K}}),P'_\mathcal{K}(x_{m_\mathcal{K}}))\ge I'(x_\mathcal{K};\hat{x}_\mathcal{K},\boldsymbol{y}). \tag{32}$$

On the other hand, (31) also guarantees

$$\begin{aligned}\Delta&\triangleq I(x_\mathcal{K};\hat{x}_\mathcal{K},\boldsymbol{y})-J(P_\mathcal{K}(x_{m_\mathcal{K}}),P'_\mathcal{K}(x_{m_\mathcal{K}})) &\text{(33a)}\\
&=(1-\rho)\sum_{m_k}\frac{P'_k(x_{m_k})f'_k(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})}{\sum_{m'_k}P'_k(x_{m'_k})f'_k(x_{m'_k};\hat{x}_\mathcal{K},\boldsymbol{y})}\sum_{m''_\mathcal{K}}P(\hat{x}_{m''_\mathcal{K}}\,|\,x_{m_k})\\
&\quad\times\log\frac{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_\mathcal{K}}\,|\,x_{m'_k})f'_k(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})}{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_\mathcal{K}}\,|\,x_{m'_k})f'_k(x_{m'_k};\hat{x}_\mathcal{K},\boldsymbol{y})} &\text{(33b)}\\
&\ge(1-\rho)\sum_{m_k}\frac{P'_k(x_{m_k})f'_k(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})}{\sum_{m'_k}P'_k(x_{m'_k})f'_k(x_{m'_k};\hat{x}_\mathcal{K},\boldsymbol{y})}\sum_{m''_\mathcal{K}}P(\hat{x}_{m''_\mathcal{K}}\,|\,x_{m_k})\\
&\quad\times\left(1-\frac{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_\mathcal{K}}\,|\,x_{m'_k})f'_k(x_{m'_k};\hat{x}_\mathcal{K},\boldsymbol{y})}{\sum_{m'_k}P'_k(x_{m'_k})P(\hat{x}_{m''_\mathcal{K}}\,|\,x_{m'_k})f'_k(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})}\right) &\text{(33c)}\\
&=0, &\text{(33d)}\end{aligned}$$

where $f'_k(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})\triangleq\exp\left(\frac{\rho}{1-\rho}I'_k(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})\right)$ and the equality holds if and only if (31) converges. (32) and (33) together imply $I(x_\mathcal{K};\hat{x}_\mathcal{K},\boldsymbol{y})\ge I'(x_\mathcal{K};\hat{x}_\mathcal{K},\boldsymbol{y})$. Since mutual information is bounded above, we conclude the sequence (20) is non-decreasing and convergent in mutual information.

Next, we prove that any converging point of sequence (20), denoted as $P_k^\star(x_{m_k})$, fulfills KKT conditions (19). To see this, consider $P_k^{(0)}(x_{m_k})>0$ and define

$$D_k^{(r)}(x_{m_k})\triangleq\frac{P_k^{(r+1)}(x_{m_k})}{P_k^{(r)}(x_{m_k})}=\frac{f_k^{(r)}(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})}{\sum_{m'_k}P_k^{(r)}(x_{m'_k})f_k^{(r)}(x_{m'_k};\hat{x}_\mathcal{K},\boldsymbol{y})}. \tag{34}$$

As sequence (20) 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

$$I_k^\star(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})=\frac{1-\rho}{\rho}\log\sum_{m'_k}P_k^\star(x_{m'_k})f_k^\star(x_{m'_k};\hat{x}_\mathcal{K},\boldsymbol{y}), \tag{35}$$

which is reminiscent of (25a) (and hence (19a)). 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 (19a). On the other hand, we assume $P_k^\star(x_{m_k})$ does not satisfy (19b), such that for any state with $P_k^\star(x_{m_k})=0$,

$$I_k^\star(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})>I^\star(x_\mathcal{K};\hat{x}_\mathcal{K},\boldsymbol{y})=\sum_{m'_k}P_k^\star(x_{m'_k})I_k^\star(x_{m'_k};\hat{x}_\mathcal{K},\boldsymbol{y}), \tag{36}$$

where the equality inherits from (16). Since the exponential function is monotonically increasing, we have $f_k^\star(x_{m_k};\hat{x}_\mathcal{K},\boldsymbol{y})>\sum_{m'_k}P_k^\star(x_{m'_k})f_k^\star(x_{m'_k};\hat{x}_\mathcal{K},\boldsymbol{y})$ 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

$$P_k^{(r)}(x_{m_k})=P_k^{(0)}(x_{m_k})\prod_{n=1}^{r}D_k^{(n)}(x_{m_k}). \tag{37}$$

$$\begin{aligned}J(P_\mathcal{K}(x_{m_\mathcal{K}}),P'_\mathcal{K}(x_{m_\mathcal{K}}))&\triangleq\rho\sum_{m_\mathcal{K}}P_\mathcal{K}(x_{m_\mathcal{K}})N\log_2\left(1+\frac{|\boldsymbol{h}_\mathrm{E}^H(x_{m_\mathcal{K}})\boldsymbol{w}|^2}{\sigma_w^2}\right)\\
&\quad+(1-\rho)\sum_{m_\mathcal{K}}P_\mathcal{K}(x_{m_\mathcal{K}})\sum_{m'_\mathcal{K}}P(\hat{x}_{m'_\mathcal{K}}\,|\,x_{m_\mathcal{K}})\log\frac{P(\hat{x}_{m'_\mathcal{K}}\,|\,x_{m_\mathcal{K}})P'_\mathcal{K}(x_{m_\mathcal{K}})}{P'(\hat{x}_{m'_\mathcal{K}})P_\mathcal{K}(x_{m_\mathcal{K}})}.\end{aligned} \tag{28}$$

Therefore, given $P_k^{(0)}(x_{m_k})>0$, any converging point with $P_k^\star(x_{m_k})=0$ must satisfy (19b). This completes the proof. □

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 $\mathcal{R}_{l'}=\bigcup_{s\equiv l'\pmod L}[t_{s-1},t_s)$. [5]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

$$\max_{\boldsymbol{t}} \quad I_{\mathrm{B}}(x;\hat{x}) \tag{38a}$$

$$\text{s.t.} \quad t_{s-1}<t_s, \quad \forall s\in\mathcal{S}. \tag{38b}$$

Problem (38) is intricate due to the strict inequality constraint (38b). Following [43], 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

$$L=-I_{\mathrm{B}}(x;\hat{x})+\sum_{s\in\mathcal{S}}\mu_s(t_{s-1}-t_s), \tag{39}$$

where μ_s is the Lagrange multiplier associated with the non-strict version of (38b). The KKT conditions on the optimal primal and dual solutions are, $\forall s\in\mathcal{S}$,

$$-\nabla_{t_s^\star}I_{\mathrm{B}}^\star(x;\hat{x})+\mu_{s-1}^\star-\mu_s^\star=0, \tag{40a}$$

$$\mu_s^\star\geq 0, \tag{40b}$$

$$\mu_s^\star(t_{s-1}^\star-t_s^\star)=0. \tag{40c}$$

Due to the strict inequality constraint (38b), conditions (40b) and (40c) 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 (38), $\forall s\in\mathcal{S}$,

$$\nabla_{t_s^\star}I_{\mathrm{B}}^\star(x;\hat{x})=0, \tag{41}$$

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

$$\sum_l P(x_l)\frac{(t_s^\star)^{N-1}e^{-t_s^\star/\sigma_l^2}}{\sigma_l^{2N}(N-1)!}\log\frac{P(x_l\,|\,\hat{x}_{l'+1})}{P(x_l\,|\,\hat{x}_{l'})}=0, \tag{42}$$

According to [41], 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 (42) is a generalized Dirichlet polynomial, and by Descartes' rule of signs [44], 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. □

[5]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, no. 4, pp. 39–50, sep 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, no. 3, pp. 1219–1233, mar 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, no. 2, pp. 1116–1130, apr 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, no. 4, pp. 1242–1255, dec 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, no. 4, pp. 1191–1194, aug 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, no. 5, pp. 601–605, may 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, no. 2, pp. 1350–1363, feb 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, no. 6, pp. 1679–1682, dec 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, no. 6, pp. 1949–1953, jun 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*, jun 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, no. March, 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, no. 8, pp. 5159–5170, aug 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9382925/

[13] R. Anwar, L. Mao, and H. Ning, "Frequency Selective Surfaces: A Review," *Applied Sciences*, vol. 8, no. 9, p. 1689, sep 2018. [Online]. Available: http://www.mdpi.com/2076-3417/8/9/1689

[14] Q. Wu and R. Zhang, "Intelligent Reflecting Surface Enhanced Wireless Network: Joint Active and Passive Beamforming Design," in *2018 IEEE Global Communications Conference (GLOBECOM)*, vol. 18, no. 11. IEEE, dec 2018, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/8647620/

[15] ——, "Beamforming Optimization for Intelligent Reflecting Surface With Discrete Phase Shifts," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, may 2019, pp. 7830–7833. [Online]. Available: https://ieeexplore.ieee.org/document/8683145/

[16] ——, "Intelligent Reflecting Surface Enhanced Wireless Network via Joint Active and Passive Beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, nov 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8811733/

[17] S. Abeywickrama, R. Zhang, and C. Yuen, "Intelligent Reflecting Surface: Practical Phase Shift Model and Beamforming Optimization," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*. IEEE, jun 2020, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/9148961/

[18] Q.-U.-A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Intelligent Reflecting Surface Assisted Wireless Communication: Modeling and Channel Estimation," *arXiv:1906.02360*, pp. 1–7, jun 2019. [Online]. Available: http://arxiv.org/abs/1906.02360

[19] C. You, B. Zheng, and R. Zhang, "Intelligent Reflecting Surface With Discrete Phase Shifts: Channel Estimation and Passive Beamforming," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*. IEEE, jun 2020, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/9149292/

[20] J.-M. Kang, "Intelligent Reflecting Surface: Joint Optimal Training Sequence and Refection Pattern," *IEEE Communications Letters*, vol. 24, no. 8, pp. 1784–1788, aug 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9081935/

[21] P. Wang, J. Fang, H. Duan, and H. Li, "Compressed Channel Estimation for Intelligent Reflecting Surface-Assisted Millimeter Wave Systems," *IEEE Signal Processing Letters*, vol. 27, pp. 905–909, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9103231/

[22] 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, no. 5, pp. 1056–1060, may 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9345739/

[23] Q. Zhang, Y.-C. Liang, and H. V. Poor, "Reconfigurable Intelligent Surface Assisted MIMO Symbiotic Radio Networks," *IEEE Transactions on Communications*, vol. 69, no. 7, pp. 4832–4846, jul 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9391685/

[24] X. Xu, Y.-C. Liang, G. Yang, and L. Zhao, "Reconfigurable Intelligent Surface Empowered Symbiotic Radio Over Broadcasting Signals," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, vol. 2020-Janua. IEEE, dec 2020, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/9348236/

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

[26] S. Hu, C. Liu, Z. Wei, Y. Cai, D. W. K. Ng, and J. Yuan, "Beamforming Design for Intelligent Reflecting Surface-Enhanced Symbiotic Radio Systems," *arxiv:2110.10316*, oct 2021. [Online]. Available: http://arxiv.org/abs/2110.10316

[27] Y. Huang, A. Alieldin, and C. Song, "Equivalent Circuits and Analysis of a Generalized Antenna System," *IEEE Antennas and Propagation Magazine*, vol. 63, no. 2, pp. 53–62, apr 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9392844/

[28] Daniel Dobkin, *The RF in RFID: Passive UHF RFID in Practice*. London, U.K.: Newnes, nov 2012. [Online]. Available: https://www.elsevier.com/books/the-rf-in-rfid/dobkin/978-0-12-394583-9

[29] 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, apr 2021. [Online]. Available: http://arxiv.org/abs/2104.08404

[30] R. Hansen, "Relationships Between Antennas as Scatterers and as Radiators," *Proceedings of the IEEE*, vol. 77, no. 5, pp. 659–662, may 1989. [Online]. Available: http://ieeexplore.ieee.org/document/32056/

[31] C. Boyer and S. Roy, "Backscatter Communication and RFID: Coding, Energy, and MIMO Analysis," *IEEE Transactions on Communications*, vol. 62, no. 3, pp. 770–785, mar 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6685977/

[32] ——, "Coded QAM Backscatter Modulation for RFID," *IEEE Transactions on Communications*, vol. 60, no. 7, pp. 1925–1934, jul 2012. [Online]. Available: http://ieeexplore.ieee.org/document/6202629/

[33] 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, no. 5, pp. 3313–3351, may 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9326394/

[34] 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, no. 4, pp. 1175–1182, apr 2012. [Online]. Available: http://ieeexplore.ieee.org/document/6153042/

[35] D. Bharadia, K. R. Joshi, M. Kotaru, and S. Katti, "BackFi: High Throughput WiFi Backscatter," in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, vol. 45, no. 4. New York, NY, USA: ACM, aug 2015, pp. 283–296. [Online]. Available: https://dl.acm.org/doi/10.1145/2785956.2787490

[36] 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, no. 12, pp. 2974–2987, dec 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7274644/

[37] 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, no. 1, pp. 765–775, feb 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8411483/

[38] J. Buhler and G. Wunder, "A Note on Capacity Computation for the Discrete Multiple Access Channel," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1906–1910, apr 2011. [Online]. Available: https://ieeexplore.ieee.org/document/5730559/

[39] S. Arimoto, "An Algorithm for Computing the Capacity of Arbitrary Discrete Memoryless Channels," *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 14–20, jan 1972. [Online]. Available: http://ieeexplore.ieee.org/document/1054753/

[40] R. E. Blahut, "Computation of Channel Capacity and Rate-Distortion Functions," *IEEE Transactions on Information Theory*, vol. 18, no. 4, pp. 460–473, jul 1972. [Online]. Available: http://ieeexplore.ieee.org/document/1054855/

[41] 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, no. 6, pp. 3638–3651, jun 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9366549/

[42] T. Nguyen and T. Nguyen, "On Thresholding Quantizer Design for Mutual Information Maximization: Optimal Structures and Algorithms," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, vol. 2020-May. IEEE, may 2020, pp. 1–5. [Online]. Available: https://ieeexplore.ieee.org/document/9128966/

[43] T. D. Nguyen and T. Nguyen, "On Binary Quantizer For Maximizing Mutual Information," *IEEE Transactions on Communications*, vol. 68, no. 9, pp. 5435–5445, sep 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9118952/

[44] G. J. O. Jameson, "Counting Zeros of Generalised Polynomials: Descartes' Rule of Signs and Laguerre's Extensions," *The Mathematical Gazette*, vol. 90, no. 518, pp. 223–234, jul 2006. [Online]. Available: https://www.cambridge.org/core/product/identifier/S0025557200179628/type/journal_article