

# Traffic-Aware Backscatter Communications in Wireless-Powered Heterogeneous Networks

Sung Hoon Kim and Dong In Kim

Department of Electrical and Computer Engineering  
Sungkyunkwan University (SKKU), Suwon, Korea

**Abstract**—With the emerging Internet-of-Things services, massive machine-to-machine (M2M) communication will be deployed on top of human-to-human (H2H) communication in the near future. Due to the coexistence of M2M and H2H communications, the performance of M2M (i.e., secondary) network depends largely on the H2H (i.e., primary) network. In this paper, we propose ambient backscatter communication for the M2M network which exploits the energy (signal) sources of the H2H network, referring to traffic applications and popularity. In order to maximize the harvesting and transmission opportunities offered by varying traffic sources of the H2H network, we adopt a Bayesian nonparametric (BNP) learning algorithm to classify traffic applications (patterns) for secondary user (SU). We then analyze the performance of SU using the stochastic geometrical approach, based on a criterion for optimal traffic pattern selection. Results are presented to validate the performance of the proposed BNP classification algorithm and the criterion, as well as the impact of traffic sources and popularity.

**Keywords**—Wireless-powered heterogeneous networks (WPHet-Nets), ambient backscatter, Ginibre point process, traffic patterns classification, Bayesian nonparametric identification.

## I. INTRODUCTION

Due to the growing interest in Internet-of-Things (IoT) services, smart devices such as implanted sensors, flexible epidermal devices, RFID tags, and wearable devices are interconnected for machine-to-machine (M2M) communication. Unlike human-to-human (H2H) communication devices, M2M communication devices will be deployed densely in the IoT network because these are in small form factor with intermittent and low-rate communication capability [1]. Therefore, the coexistence of massive M2M and H2H communications is becoming a critical issue for realizing the IoT network. To address this issue, the radio frequency (RF) energy harvesting and backscatter communication are considered promising techniques for enabling low-power massive M2M communication on top of H2H communication.

Recently the RF energy harvesting is gaining growing interest to utilize ambient or dedicated RF signals as the energy (signal) sources for low-power M2M communication devices. The proposals such as [2], [3] and [4] have studied wireless-powered communication networks (WPCNs) where devices first harvest energy from the RF signals in downlink, and then use the harvested energy for transmitting their collected information in uplink [4]. Namely, the well-known “harvest-then-transmit (HTT)” protocol was proposed in [4] for WPCNs.

However, traditional active-radio based RF communications cannot fully support WPCNs because it requires high circuit power, which is not appropriate for the low-power devices. To tackle this problem, researchers have adopted backscatter communication.

Backscatter communication, classified as bistatic scatter (BS) and ambient backscatter (AB), is low-power low-cost communication technique, and it is recognized as a key enabler for battery-free communication [5]. The AB and BS communications were developed for passive communications by utilizing ambient and dedicated RF signals, respectively, as the only source of energy during absorbing state while transmitting information by simply choosing between the absorbing and reflecting states via antenna impedance switching [6]. Especially, to realize the AB communication, there is no need to install additional infrastructure in the network in contrast to the BS communication, and it can be more cost effective than the BS communication.

Hardware prototypes for tag-to-tag AB communication were first developed in [7]. The transmitter utilizes ambient TV signal for transmission, and the receiver averages out its received signal for information decoding due to which low data rate ( $\sim 10$ kbps) and short-range communication (1m) can only be supported. To increase the data rate as well as for an increased range, authors in [8] developed multiple receive antennas ( $\mu$ mo) which can support up to 1Mbps data rate, and novel coding ( $\mu$ code) which can increase the operational range up to 30m. Authors in [9] proposed the inter-technology backscatter so called *interscatter*. The interscatter device reflects Bluetooth signal for transmitting its data which is transformed to Wi-Fi and ZigBee-compatible signals. In addition, the proposed epidermal prototypes show 2-11 Mbps data rate.

There have been many works such as [1], [5], [6], [10], [11], [12] which studied backscatter communication for battery-free massive M2M communication. In [11], AB communication was introduced for overlay/underlay RF-powered cognitive radio networks to overcome the range discrimination of HTT protocol. In addition, a multiple access scheme for AB assisted WPCN was analyzed in [6]. To ensure both uniform coverage and rate distribution for WPCNs, [5] proposed hybrid of AB and BS for wireless-powered heterogeneous networks (WPHetNets). Especially, dual mode operation was proposed which utilizes AB and BS as the secondary access on top of the primary HTT protocol. The dual mode operation was optimized by maximizing the overall throughput of WPHetNets.

Authors in [12] have performed a comprehensive survey which highlights the state-of-art researches and open issues about AB communication.

Meanwhile, authors in [10] and [1] invoke stochastic geometry to analyze backscatter networks. For this, wireless-powered backscatter communication networks were modelled in [10] that power beacons (PBs) and transmitting nodes are Poisson point process (PPP) and Poisson cluster process (PCP) distributed in the network, respectively. The coverage and network capacities were analyzed and optimized with regard to the duty cycle, reflection coefficient, and density of PBs. In order to analyze the impact of environment factors such as the distribution, spatial density, and transmission load of ambient transmitters, authors in [1] modelled ambient transmitters by  $\alpha$ -Ginibre point process ( $\alpha$ -GPP) which is kind of a repulsive point process. For flexible adaption to various environments, two mode selection protocols were designed, termed power threshold-based and SNR threshold-based protocols. The impact of environment factors and the validity of the two protocols were investigated through stochastic geometrical approach.

The analysis made in [1] may not be sufficient enough to reflect the entire features of H2H network since only popularity was considered. In fact, traffic sources of the H2H network should also be taken into account as an important feature as well as the popularity because channel busy/idle distributions which influence the performance of M2M network depends heavily on specific traffic applications. In our earlier work [13], we applied the traffic classifications [14], [15] for the network model considered in [1]. For this, we adopted a Bayesian nonparametric (BNP) learning algorithm to classify traffic applications to maximize the harvesting/transmission opportunities for AB communication where PUs were distributed according to PPP. In this paper, we extend our earlier work [13] by applying  $\alpha$ -GPP which enables to analyze a general popularity and includes PPP as a special case. This will allow in-depth analysis of AB communication for the M2M network coexisting with the H2H network.

The rest of this paper is organized as follows. Section II describes the network and traffic model for the H2H network. Section III describes how to classify traffic applications by using the BNP learning algorithm. In section IV, introduction of  $\alpha$ -GPP and optimal/suboptimal traffic pattern selection criterion determined by stochastic geometrical approach are given. Section V presents numerical and simulation results to show the validity of the BNP learning algorithm and traffic pattern selection criterion as well as the impact of traffic sources and popularity. Concluding remarks are given in Section VI.

## II. SYSTEM MODEL

We consider a pair of secondary users (SUs)<sup>1</sup> employing low-power AB communication in WPHetNets where the primary user (PU) network coexists, comprising of WiFi access points (APs) or cellular base stations and primary users (PUs). SUs are assumed to be passive tags which are not equipped

<sup>1</sup>To focus on the impact of traffic applications (patterns), in this paper we simply assume there is one pair of SUs.

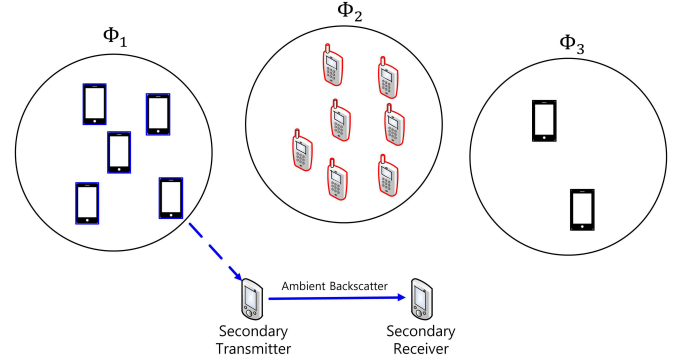


Fig. 1. An illustration of the proposed network model, where three traffic applications are utilized (i.e.,  $K = 3$ ) and  $\Phi_1$  is selected for AB communication.

with the battery for communication and may not collect any information about the PU network. In this situation, SUs attempt to select an optimal traffic application offered by PUs in order to maximize the harvesting/transmission opportunities for AB communication.

To enable this function, SUs are required to learn the characteristics of the traffic sources in the PU network, such as traffic applications and popularity. Since SUs may obtain these information based only on their observations, an unsupervised BNP learning algorithm is appropriate to extract such information. Through the BNP learning algorithm, SUs can not only classify traffic applications (i.e., patterns), but also obtain key network parameters which help to improve the performance of SUs. As for the performance metrics, energy outage probability and coverage probability will be analyzed by using the stochastic geometrical approach [16]. We will then formulate an optimal traffic pattern selection criterion. To begin with, the PU network and its traffic sources are modelled in the sequel.

### A. Network Model

In this paper, we consider a WPHetNet in which PUs deliver various traffic applications (patterns) in the PU network, as illustrated in Fig. 1. The whole PU network  $\Phi$  is modeled by  $\alpha$ -GPP (to be described later) which is defined as  $\Phi = \{X, \zeta, T, \alpha\}$ .  $X = \{x_i | i = 1, 2, \dots\}$  represents the locations of PUs where  $x_i \in \mathbb{R}^2$ ,  $\zeta$  denotes the spatial density,  $T = \{T_i | i = 1, 2, \dots\}$  the traffic indicator for the  $i$ th PU where  $T_i \in \{1, 2, \dots, k, \dots\}$ , and  $\alpha \in [-1, 0)$  the repulsion factor to measure the correlation among the spatial points. Then we can define the  $k$ th traffic pattern  $\Phi_k = \{X^k, \zeta_k, p_b^k\}$  where  $X^k = \{x_i | T_i = k\}$  denotes the location of PUs delivering the  $k$ th traffic pattern,  $p_b^k$  indicates busy period statistics for the  $k$ th traffic pattern.  $\zeta_k$  is the spatial density of the  $k$ th traffic application being defined as  $\zeta_k = l_k \zeta$ , where  $l_k = \Pr[T_i = k]$  is portion of the  $k$ th traffic application generated in the network. SU transmitter (ST) is assumed located at the origin of  $\mathbb{R}^2$  plane and SU receiver (SR) near ST.

### B. Traffic Model

We introduce the packet/energy arrival based traffic model for which ST examines the packet header of PUs to identify the packet arrivals. We assume that ST identifies a total of  $N$  PUs deployed in the PU network  $\Phi$  with various traffic applications. Traffic applications show unique behavior in terms of their features [14]. Here, to properly classify traffic patterns, we consider the following three features: packet length ( $p_{l,r}^{(n)}$ ), packet interarrival time ( $p_{i,r}^{(n)}$ ), and variance in packet length ( $\Delta_{W_r}^{(n)}$ ), where  $n \in \{1, 2, \dots, N\}$  and  $r \in \{1, 2, \dots, R\}$  denote the PU and observation indexes, respectively. We define  $P_{len}^{(n)}$ ,  $P_{inter}^{(n)}$ , and  $\Delta^{(n)}$  as packet length vector, packet interarrival time vector, and variance in packet length vector for the  $n$ th PU. Then, they can be defined as  $P_{len}^{(n)} = [p_{l,1}^{(n)}, p_{l,2}^{(n)}, \dots, p_{l,R}^{(n)}]$ ,  $P_{inter}^{(n)} = [p_{i,1}^{(n)}, p_{i,2}^{(n)}, \dots, p_{i,R}^{(n)}]$ , and  $\Delta^{(n)} = [\Delta_{W_1}^{(n)}, \Delta_{W_2}^{(n)}, \dots, \Delta_{W_R}^{(n)}]$ , respectively. Here,  $\Delta_{W_r}^{(n)}$  denotes the temporal variance of packet length in a window  $W_r$  of size  $r$ , spanning over the  $[1, 2, \dots, r]$ th observations. We denote  $y_r$  as the feature space vector with the  $r$ th training feature point, which is defined by

$$y_r = [p_{l,r}^{(1)}, p_{i,r}^{(1)}, \Delta_{W_r}^{(1)} | \dots | p_{l,r}^{(N)}, p_{i,r}^{(N)}, \Delta_{W_r}^{(N)}], \quad (1)$$

$$\mathbf{Y} = [y_1^T | \dots | y_R^T]. \quad (2)$$

Here, the matrix  $\mathbf{Y}$  represents an observation matrix as the set of feature points.

### III. TRAFFIC CLASSIFICATION

We assume that the set of data follows a specific generative model. As for the generative model, we adopt the finite Gaussian-mixture model (FGMM) which is frequently used to model an arbitrary multi-modal probability density function (pdf) and the infinite Gaussian-mixture model (IFGMM) to handle the case when the number of traffic applications is unknown. First, we define the general Gaussian-mixture model (GMM). Let  $\{z_r\}_{r=1}^R$  be the traffic assignment indicators for all observations, then the general GMM is given by

$$P(y_r | \Theta, \vec{\pi}) = \sum_{k=1}^K \pi_k p_k(y_r | \theta_k) \quad (3)$$

where  $p_k(x_r | \theta_k)$  is the Gaussian pdf for the  $k$ th cluster (traffic pattern) with multivariate Gaussian parameter  $\theta_k = \{\vec{\mu}_k, \Sigma_k\}$ .  $\vec{\mu}_k = \{\mu_l^k, \mu_i^k, \mu_\Delta^k\}$  denotes the set of mean values where  $\mu_l^k, \mu_i^k, \mu_\Delta^k$  denote the means of packet length, packet interarrival time, and variance in packet length, respectively.  $\Sigma_k$  denotes the covariance for cluster  $k$ , and  $\Theta = \{\theta_k\}_{k=1}^K$  is the collection of all cluster parameters.  $\vec{\pi} = \{\pi_k\}_{k=1}^K$  is the collection of the Gaussian-mixture weights where  $\pi_k = \Pr(z_r = k)$  representing the prior probability that the feature point was generated from the  $k$ th traffic pattern.

#### A. Finite Gaussian-Mixture Model (FGMM)

To complete the above modeling, it is required to find the model parameters  $\Theta$  and their prior probability vector  $\vec{\pi}$  for the

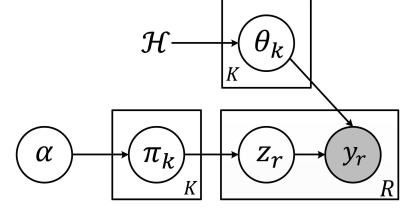


Fig. 2. The graphical illustration of the proposed FGMM. As  $K \rightarrow \infty$ , it becomes the IFGMM. Circles and boxes represent random variables and repetition of the random variables, respectively. The links show the dependency among random variables.

FGMM. According to Bayes' rule, the posterior distribution of the model  $\mathcal{M}$  given an observation matrix  $\mathbf{Y}$  is

$$P(\mathcal{M} | \mathbf{Y}) \propto P(\mathbf{Y} | \mathcal{M}) P(\mathcal{M}). \quad (4)$$

The FGMM model is then defined below.

*Definition 1:*

$$\begin{aligned} z_r | \vec{\pi} &\sim \text{Multinomial}(\cdot | \vec{\pi}), \\ y_r | z_r = k; \Theta &\sim \mathcal{N}(\cdot | \theta_k). \end{aligned} \quad (5)$$

The first term of Definition 1 stands for the probability of choosing a specific collection of  $K$  clusters from an infinite number of clusters with repetitions and the probabilities of each choice given by  $\vec{\pi}$ . To define priors on the model parameters, we perform the maximum a posteriori (MAP) estimation with Markov chain Monte Carlo (MCMC) method. To do this, we use Dirichlet distribution for  $\vec{\pi}$  and Normal times Inverse Wishart for the Normal parameters  $\Theta$  which are the conjugate priors to the multinomial and multivariate normal distributions, respectively. The conjugate priors play an important role in performing the marginalization steps in estimating the posterior distribution of the model [14]. Finally, we can define the generative model below.

*Definition 2:* The generative model under Bayesian setting for the FGMM can be defined as

$$\begin{aligned} \vec{\pi} | \alpha_o &\sim \text{Dirichlet}\left(\cdot \mid \frac{\alpha_o}{K}, \dots, \frac{\alpha_o}{K}\right), \\ \Sigma_k &\sim \text{Inverse-Wishart}_{\nu_0}(\Lambda_0^{-1}), \\ \vec{\mu}_k &\sim \mathcal{N}(\vec{\mu}_0, \Sigma_k / \kappa_0) \end{aligned} \quad (6)$$

where the parameters  $\alpha_o, \Lambda_0^{-1}, \nu_0, \vec{\mu}_0, \kappa_0$  are hyperparameters. The Dirichlet prior  $\alpha_o$  encodes our prior knowledge about the number of traffic applications. Another parameters are the hyperparameters for the Inverse-Wishart  $\mathcal{H} = \{\Lambda_0^{-1}, \nu_0, \vec{\mu}_0, \kappa_0\}$  which encode our prior belief about the traffic application variability [14].

By exploiting the dependence among random variables in Fig. 2, the joint distribution of the data and model parameters can be expressed as

$$\begin{aligned} \Pr(\mathbf{Y}, \mathbf{Z}, \Theta, \vec{\pi}, \alpha_o; \mathcal{H}) &= \left( \prod_{k=1}^K \Pr(\theta_k; \mathcal{H}) \right) \\ &\times \left( \prod_{r=1}^R \Pr(y_r | z_r, \theta_{z_r}) \Pr(z_r | \vec{\pi}) \right) \Pr(\vec{\pi} | \alpha_o) \Pr(\alpha_o). \end{aligned} \quad (7)$$

Then, Bayes' rule gives the posterior probability conditioned by the observed data and hyperparameters, which can be expressed as

$$\Pr(Z, \Theta, \vec{\pi}, \alpha_o | \mathbf{Y}; \mathcal{H}) \propto \Pr(\mathbf{Y} | Z, \Theta) \Pr(\Theta; \mathcal{H}) \times \left( \prod_{r=1}^R \Pr(z_r | \vec{\pi}) \right) \Pr(\vec{\pi} | \alpha_o) \Pr(\alpha_o) \quad (8)$$

where  $\Pr(\mathbf{Y} | Z, \Theta) = \prod_{r=1}^R \Pr(y_r | z_r, \theta_{z_r})$  and  $\Pr(\Theta; \mathcal{H}) = \prod_{k=1}^K \Pr(\theta_k; \mathcal{H})$ . In (8), we express the conditional probability with proportionality sign, as the marginal probability of the data under the model cannot be evaluated analytically. To obtain a discrete representation of the posterior by sampling from the unnormalized pdf in (8), we can invoke the MCMC method for inference.

The FGMM is appropriate when the number of traffic patterns  $K$  is *a priori* known. In practice, it is hard to know how many distinct patterns are generated by PUs. To this end, it is necessary to apply the IFGMM to handle an infinite number of traffic applications (i.e.,  $K \rightarrow \infty$ ).

### B. Infinite Gaussian-Mixture Model (IFGMM)

By marginalizing out  $\vec{\pi}$  as  $K \rightarrow \infty$ , we can derive the IFGMM to apply for the unknown number of traffic patterns from the FGMM,<sup>2</sup> which can be expressed as [17]

$$\begin{aligned} \Pr(Z | \alpha_o) &= \int \prod_{r=1}^R \Pr(z_r | \vec{\pi}) \Pr(\vec{\pi} | \alpha_o) d\vec{\pi} \\ &= \frac{\prod_{k=1}^K \Gamma(m_k + \frac{\alpha_o}{K})}{\Gamma(\frac{\alpha_o}{K})^K} \frac{\Gamma(\alpha_o)}{\Gamma(R + \alpha_o)} \\ &\stackrel{(i)}{=} \frac{K!}{(K - K_+)!} \frac{\prod_{k=1}^K \Gamma(m_k + \frac{\alpha_o}{K})}{\Gamma(\frac{\alpha_o}{K})^K} \frac{\Gamma(\alpha_o)}{\Gamma(R + \alpha_o)} \end{aligned} \quad (9)$$

where  $\Gamma(a) = \int_0^\infty x^{a-1} \exp(-x) dx$  is the Gamma function and  $m_k = \sum_{r=1}^R I(z_r = k)$  is the number of data points belonging to class  $k$  for the indicator function  $I(\cdot)$ . In the above, (9) is the joint probability of a single labelling of all observations. Because the structure of the labelling is not changed by permuting the labels allocated to sets of observations, a model that expresses the probability of partitions of the data is preferred rather than a specific labelling for simplicity. Thus, the term  $\frac{K!}{(K - K_+)!}$  is multiplied at (i), which is the number of different ways where we can apply  $K$  labels to a single partitioning of the data with  $K_+ < K$  bins. Next, if we take the limit as  $K \rightarrow \infty$ , (9) turns out to be

$$\Pr(Z | \alpha_o) = \alpha_o^{K_+} \left[ \prod_{k=1}^{K_+} (m_k - 1)! \right] \frac{\Gamma(\alpha_o)}{\Gamma(R + \alpha_o)}. \quad (10)$$

Given (8) and (10), we can carry out the model estimation through sampling from the posterior. Toward this, an expression for the conditional distribution of a single class label

given the values of all others is required for the MCMC based sampling (i.e., the collapsed Gibbs sampling). The conditional distribution can be derived as

$$\Pr(z_r = k | Z_{-r}, \alpha_o) = \begin{cases} \frac{m_k}{r-1+\alpha_o}, & \text{if } k \leq K_+ \\ \frac{\alpha_o}{r-1+\alpha_o}, & \text{if } k > K_+ \end{cases} \quad (11)$$

where  $Z_{-r} = Z/z_r$  is the set of all other indicators except for  $z_r$ . This generative process is called Chinese restaurant process (CRP) [17]. The process performs as the following procedure. The class indicators  $\{z_r\}_{r=1}^R$  are first generated by the CRP, which will result in some classes  $K$ . Then, specific observations are generated from each class of Normal densities whose parameters are drawn independently from the multivariate Normal-Inverse-Wishart prior. [14]

### C. Inference

To estimate the posterior distribution for the IFGMM, we can use two inference methods: one is a sampling-based MCMC method which is also known as the collapsed Gibbs sampling, and the other is a variational inference [18]. There exists the trade-off between classification accuracy and convergence speed for the two methods. The former has higher classification accuracy while suffering lower convergence speed because of the sampling based approach. To the contrary, the latter has lower accuracy but higher speed because it simplifies a complicate optimization problem with an approximation. In this paper, we resort to the collapsed Gibbs sampling method for traffic classification.

Note that the Normal-Inverse-Wishart prior was chosen for multivariate normal distribution, and hence we can marginalize out these parameters as

$$\begin{aligned} \Pr(Z | \mathbf{Y}; \mathcal{H}) &= \int d\Theta \Pr(Z, \Theta | \mathbf{Y}; \mathcal{H}) \\ &\propto \Pr(Z; \mathcal{H}) \int d\Theta \Pr(\mathbf{Y} | Z, \Theta; \mathcal{H}) \Pr(\Theta; \mathcal{H}). \end{aligned} \quad (12)$$

Collapsed Gibbs sampler is applied to draw samples from the conditional distributions of the variables such that it approximates the joint distribution over time. In the collapsed Gibbs sampling, the sampler state is composed of  $Z$  and  $\alpha_o$ . The updates for the labels of the class are performed as follows:

$$\begin{aligned} \Pr(z_r = k | Z_{-r}, \mathbf{Y}, \alpha_o; \mathcal{H}) \\ \propto \Pr(\mathbf{y}_r | \mathbf{Y}_{-r}^{(k)}; \mathcal{H}) \Pr(z_r = k | Z_{-r}, \alpha_o) \end{aligned} \quad (13)$$

where  $\mathbf{Y}_{-r}^{(k)}$  is the set of observations currently allocated to cluster  $k$  except  $\mathbf{y}_r$ .

By the choice of conjugate prior, the first term of the right-hand side (RHS) in (13) is the multivariate Student-t distribution [17], which is given as

$$\mathbf{y}_r | \mathbf{Y}_{-r}^{(k)}; \mathcal{H} \sim \mathbf{t}_{(\nu_r - D + 1)} \left\{ \vec{\mu}_r, \frac{\Lambda_r(\kappa_r + 1)}{\kappa_r(\nu_r - D + 1)} \right\}. \quad (14)$$

<sup>2</sup>The terms dependent on  $K$  in (8) are  $\Pr(z_r | \vec{\pi})$  and  $\Pr(\vec{\pi} | \alpha_o)$ . In addition, the Dirichlet prior is conjugate to the discrete multinomial likelihood.

The hyperparameters for the distribution in (14) are defined as

$$\begin{aligned}\vec{\mu}_r &= \frac{\kappa_0}{\kappa_0 + R} \vec{\mu}_0 + \frac{R}{\kappa_0 + R} \bar{\mathbf{y}}, \\ \kappa_r &= \kappa_0 + R, \quad \nu_r = \nu_0 + R, \\ \mathbf{\Lambda}_r &= \mathbf{\Lambda}_0 + \mathbf{S} + \frac{\kappa_0 R}{\kappa_0 + R} (\bar{\mathbf{y}} - \vec{\mu}_0)(\bar{\mathbf{y}} - \vec{\mu}_0)^T\end{aligned}$$

where  $\bar{\mathbf{y}}$  is the sample mean and  $\mathbf{S} = \sum_{r=1}^R (\mathbf{y}_r - \bar{\mathbf{y}})(\mathbf{y}_r - \bar{\mathbf{y}})^T$  is the scatter matrix of the evidence.  $D$  is the dimension of  $\mathbf{y}_r$  which implies the observed feature space (i.e.,  $D = 3$  in this paper). The subscript  $(\nu_r - D + 1)$  is the degrees of freedom of the multivariate Student-t distribution.

So far we have computed the probability of allocating the observed feature data to the existing cluster. We next consider the probability of creating a new cluster, which is expressed as

$$\begin{aligned}\Pr(z_r > k | Z_{-r}, \mathbf{Y}, \alpha_o; \mathcal{H}) \\ \propto \Pr(\mathbf{y}_r; \mathcal{H}) \Pr(z_r > k | Z_{-r}, \alpha_o).\end{aligned}\quad (15)$$

In (15),  $\Pr(\mathbf{y}_r; \mathcal{H})$  has the same form as  $\Pr(\mathbf{y}_r | \mathbf{Y}_{-r}^{(k)}; \mathcal{H})$  in (13), which follows the multivariate Student-t distribution. However, if there is no other observation, the original hyperparameters are used for computation in (14).

Finally, the above collapsed Gibbs sampling process to classify the traffic applications can be described in Algorithm 1 below.

---

**Algorithm 1** Collapsed Gibbs Sampler for IFGMM

---

```

1:  $z_1, \dots, z_R \sim \text{Uniform random integers } (1, \dots, R)$ 
2:  $Z_0 \leftarrow \{z_1, \dots, z_R\}$ 
3:  $K = 0$ 
4: for  $j \leftarrow 1$  to  $M$ , iterate until convergence do
5:    $Z_j \leftarrow Z_{j-1}$ 
6:   for  $i \leftarrow 1$  to  $R$  do
7:      $m_{-i} \leftarrow \sum_{r=1}^R I(z_r = z_i) - 1$ 
8:     if  $m_{-i} = 0$  then
9:        $z_j = z_j - 1$ 
10:       $K = K - 1$ 
11:    end if
12:    Sample  $z_i \sim \Pr(\cdot | Z_{-i}, \mathbf{Y})$  using (13) and (15)
13:    if  $z_i > K$  then
14:       $K = K + 1$ 
15:    end if
16:  end for
17:  Sample  $\alpha_o$  using Gibbs step
18: end for
```

---

#### D. Network Parameter

Using the collapsed Gibbs sampling, we can obtain information about the number of traffic patterns ( $K$ ) and clustered results. Using the results, we can estimate the network parameters such as  $p_b^k$  and  $\zeta_k$ . First, to obtain  $p_b^k$ , ST observes a number of packet arrivals carrying a specific  $k$ th traffic pattern within a time slot  $T_{slot}$ . This can be calculated by summing up all packet interarrival times during the time slot, and then

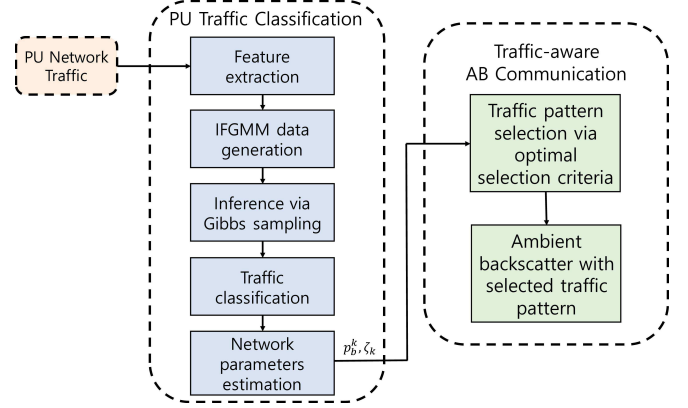


Fig. 3. A flow diagram for traffic applications classification.

normalizing it with the mean value of packet interarrival times of the  $k$ th traffic pattern, namely

$$\lambda^k = \frac{\sum_{\forall t} p_{i,t}^k}{\mu_i^k}. \quad (16)$$

Then, the busy period channel statistics for the  $k$ th traffic pattern can be calculated as

$$p_b^k = \frac{\lambda^k \mu_l^k}{T_{slot}}. \quad (17)$$

Next, the density of the  $k$ th traffic pattern  $\zeta_k$  can easily be estimated, given SU has learned the traffic pattern information, i.e., the number of PUs delivering the  $k$ th traffic pattern.

Finally, an overall framework for Bayesian nonparametric traffic patterns classification and traffic-aware backscatter communication is described in Fig. 3.

#### IV. TRAFFIC SELECTION CRITERION

After classifying traffic patterns, ST determines an optimal traffic pattern being utilized to maximize harvesting and AB communication opportunities. For this, we first look into the performance of AB communication associated with each traffic pattern.

##### A. AB Communication

We first analyze the incident RF signal density at ST from PUs which deliver the  $k$ th traffic application, which is given by

$$P_I^k = p_b^k \sum_{x_i^k \in \Phi_k} \frac{P_{PU} G_{PU}}{4\pi d_0^2} \left( \frac{d_0}{\|x_i^k\|} \right)^\mu h_i \quad (18)$$

where  $P_{PU}$ ,  $G_{PU}$ ,  $d_0$ ,  $\mu$ , and  $h_i$  are the transmit power of PUs, antenna gain of PUs, reference distance, path-loss exponent, and channel gain between the  $i$ th PU and ST, respectively. Here we assume Rayleigh fading channel, and hence  $h_i$  is an exponentially distributed random variable.

The RF power harvested by ST, from the PU signals carrying the  $k$ th traffic pattern, is determined by considering the

effective aperture of antenna  $A_{e,ST}^{(j)}$  and RF-to-DC conversion efficiency  $\eta$ , which is given by

$$P_E^k = \mathbb{E}_j \left[ \eta P_I^k A_{e,ST}^{(j)} \right], \quad j \in \{1, 2\} \quad (19)$$

where  $j$  is the binary symbol index. The effective aperture of antenna at ST is defined as  $A_{e,ST}^{(j)} = \frac{\lambda^2 G_{ST}}{4\pi} [1 - |\Gamma_j|^2]$  where  $\lambda$  is the wavelength of RF signal and  $\Gamma_j \in \{0, -1\}$  is the reflection coefficient for symbol  $j$  [19]. Assume that ST transmits equally-likely binary symbols, then the RF power harvested by ST, from the PU signals carrying the  $k$ th traffic pattern, can be rewritten as

$$P_E^k = 0.5 \eta P_I^k \frac{\lambda^2 G_{ST}}{4\pi}. \quad (20)$$

For binary backscatter communication, the differential radio cross section (RCS) of the antenna determines the power of the backscattered signal [20]. Thus, the transmit power of the reflected signal at ST which utilizes the  $k$ th traffic pattern can be expressed as

$$P_T^k = P_I^k \Delta\sigma \quad (21)$$

where the differential RCS of the antenna is defined as  $\Delta\sigma = \frac{\lambda^2 G_{ST}^2}{4\pi} |\Gamma_1 - \Gamma_2|^2$  [20]. To transmit information, ST should satisfy the following two conditions as:

- Energy causality ( $\mathcal{C}_E$ ): ST should harvest enough energy to operate backscatter communication. If we denote  $\rho_B$  as the threshold for enabling this, then

$$P_E^k \geq \rho_B. \quad (22)$$

- Interference constraint ( $\mathcal{C}_I$ ): ST is allowed to transmit information when the transmit power does not cause critical interference to PUs. For the maximum transmit power denoted by  $P_{max}$ , we set

$$P_T^k \leq P_{max}. \quad (23)$$

If the above two conditions are satisfied, then ST can send information via ambient RF signals, for which the received signal-to-noise ratio (SNR) at SR can be expressed as

$$\nu_B^k = \frac{P_T^k h_{TR} A_{e,SR}}{N_0} \left( \frac{d_0}{d_{TR}} \right)^\mu \quad (24)$$

for the effective aperture at SR  $A_{e,SR}$ , which is defined by  $A_{e,SR} = \frac{\lambda^2 G_{SR}}{4\pi}$ .<sup>3</sup> Here,  $h_{TR}$ ,  $N_0$ ,  $G_{SR}$ , and  $d_{TR}$  are the channel gain between ST and SR, power spectral density (psd) of channel noise, antenna gain of SR, and distance between ST and SR, respectively. We assume that the channel between ST and SR also follows Rayleigh fading. To decode information at SR, the following constraint needs to be satisfied:

- SNR constraint ( $\mathcal{C}_S$ ): If the received SNR  $\nu_B^k$  is greater than a threshold  $\tau_B$ , SR can decode information from ST, which can be expressed as

$$\nu_B^k \geq \tau_B. \quad (25)$$

<sup>3</sup>As for SR, the incident signal is absorbed only for information decoding, so that the reflection coefficient is zero for  $A_{e,SR}$ .

TABLE I. COMPARISON OF STOCHASTIC GEOMETRY MODELS.

Model	Mathematical Tractability	Available Parameter	Repulsiveness
PPP	Closed form	Density	No
MHCP	No closed form	Density, $r_{min}$	Yes
GPP	Closed form	Density, $\alpha$	Yes

### B. Geometric Modeling of PU Network

As a tool for analysis of wireless networks, the Poisson Point Process (PPP) has been widely adopted because of its tractability. However, because ambient RF sources such as mobile sensor network [21] and cellular base stations [22] exhibit repulsion behaviors, the PPP cannot cover practical scenarios. In order to analyze such repulsive behavior, the point process models which reflect repulsive nature such as Matérn hard-core process (MHCP) and  $\alpha$ -GPP should be adopted. Table I shows the distinct features among PPP, MHCP, and  $\alpha$ -GPP. MHCP is the point process where points do not allow other points to be closer than a certain minimum distance  $r_{min}$  [16]. However, it does not yield analytical expressions for the performance of the networks since the Laplace transforms are unknown. The Laplace transform of the received signal at a node is important when we analyze the performance of wireless networks since the distribution of the received signal strength can be obtained by taking the inverse Laplace transform. Therefore, for exact performance analysis, recently  $\alpha$ -GPP has been adopted to model the distribution of cellular base stations.

$\alpha$ -GPP is the point process which can reflect repulsiveness of wireless networks [23]. The degree of repulsion is characterized by the coefficient  $\alpha$ , which is the strongest with  $\alpha = -1$  and disappeared as  $\alpha \rightarrow 0$ . If there is no correlation among nodes (i.e.,  $\alpha \rightarrow 0$ ), the distribution of nodes follows the PPP, which implies that  $\alpha$ -GPP covers the PPP as a special case. Therefore, by adjusting the repulsion factor  $\alpha$  carefully, we can approximate the real networks, for instance, authors in [23] found the deployments of base stations in urban area following -1-GPP closely, while in rural area become more irregular (i.e.,  $\alpha \in [-0.4, -0.2]$ ).

In this paper, we adopt  $\alpha$ -GPP to model the PU network for tractable analysis.  $\alpha$ -GPP is kind of determinantal point process (DPP) [24] whose correlation function with respect to the Lebesgue measure on  $\mathbb{R}^2$  is expressed by determinant of a kernel of the point process. If the kernel is Ginibre kernel, the point process is said to be GPP. Let  $\Omega$  follow GPP, then the kernel of  $\Omega$  (i.e.,  $\mathbb{G}_\Omega(\mathbf{x}, \mathbf{y})$ ) is given by

$$\mathbb{G}_\Omega(\mathbf{x}, \mathbf{y}) = \zeta_\Omega \exp \left( \pi \zeta_\Omega \mathbf{x} \bar{\mathbf{y}} - 0.5 \pi \zeta_\Omega (|\mathbf{x}|^2 + |\mathbf{y}|^2) \right), \quad \mathbf{x}, \mathbf{y} \in \mathcal{K} \quad (26)$$

where  $\zeta_\Omega$  is the spatial density of  $\Omega$  and  $\mathcal{K}$  represents an almost surely finite collection of  $\Omega$  located inside an observation window  $\mathbb{O}$ .<sup>4</sup>

Next, we consider the thinning process of GPP. Because GPP is kind of DPP, the resulting GPP after independent thinning is also GPP with transformed kernel [24]. Let  $\hat{\Omega}$  be

<sup>4</sup>Here the observation window  $\mathbb{O}$  is defined as a circular Euclidean space with radius  $R_O$ . In this paper, we restrict our analysis to a point located within the observation window  $\mathbb{O}$ .

obtained as independent thinning of  $\Omega$  with retention probability  $p_{rt}(x)$ , then  $\hat{\Omega}$  follows GPP with the kernel  $\hat{\mathbb{G}}_{\hat{\Omega}}(\mathbf{x}, \mathbf{y})$  which is expressed as [24]

$$\hat{\mathbb{G}}_{\hat{\Omega}}(\mathbf{x}, \mathbf{y}) = \sqrt{p_{rt}(\mathbf{x})} \mathbb{G}_{\Omega}(\mathbf{x}, \mathbf{y}) \sqrt{p_{rt}(\mathbf{y})}. \quad (27)$$

Now we look into the Laplace transform of  $\alpha$ -GPP which will be adopted for analysis. The Laplace transform of  $\alpha$ -GPP can be expressed in terms of Fredholm determinants, which is tractable analytical expressions [1]. The Fredholm determinant is a generalization of determinant of a matrix defined by bounded operators on a Hilbert space. For arbitrary function  $F$  with  $|\alpha| \leq 1$ , the Fredholm determinant is defined by  $\text{Det}(\text{I}_d + \alpha F)$  where  $\text{I}_d$  is the identity matrix. The following proposition [25] represents the Laplace transform of  $\alpha$ -GPP.

**Proposition 1:** Consider independent and identically distributed (i.i.d) random variables  $\{h_n\}$  which are independent of  $\alpha$ -GPP  $\Omega$ . For an arbitrary real-valued function  $\psi$ , such as path-loss function, the Laplace transform of  $\sum_{\mathbf{x}_n \in \mathcal{K}} h_n \psi(\mathbf{x}_n)$  is then expressed by

$$\begin{aligned} \mathcal{L}(s) &= \mathbb{E} \left[ \exp \left( -s \sum_{\mathbf{x}_n \in \mathcal{K}} h_n \psi(\mathbf{x}_n) \right) \right] \\ &= \text{Det} [\text{I}_d + \alpha \mathbb{A}_{h, \psi}(s)]^{-1/\alpha}, \end{aligned} \quad (28)$$

where  $\mathbb{A}_{h, \psi}(s)$  is given by

$$\begin{aligned} \mathbb{A}_{h, \psi}(s) &= \sqrt{1 - M_h(-s\psi(\mathbf{x}))} \mathbb{G}_{\Omega}(\mathbf{x}, \mathbf{y}) \\ &\quad \times \sqrt{1 - M_h(-s\psi(\mathbf{y}))}, \quad \mathbf{x}, \mathbf{y} \in \mathcal{K}. \end{aligned} \quad (29)$$

Here,  $M_X(t) \triangleq \mathbb{E}[\exp(tX)]$  represents the moment generating function (MGF) of a random variable  $X$ .

For the Laplace transform of  $\alpha$ -GPP, we invoke Lemma 3 in [25] to compute the Fredholm determinant, which requires lower computational complexity than conventional one [26].

### C. Performance Analysis

**1) Performance Metrics:** We evaluate the following two metrics: energy outage probability and coverage probability which are defined as:

- **Energy Outage :** The energy outage happens when ST does not harvest enough energy to activate backscatter communication from ambient RF signals. The energy outage probability of traffic application  $k$  is defined as

$$\mathbb{O}_B^k = \Pr[P_E^k < \rho_B]. \quad (30)$$

- **Coverage Probability :** The coverage probability or equivalently the decoding success probability at SR is that the three constraints  $\mathcal{C}_E$ ,  $\mathcal{C}_I$ , and  $\mathcal{C}_S$  above are all satisfied, which is defined as

$$\mathbb{C}_B^k = \Pr[\nu_B \geq \tau_B, P_E^k \geq \rho_B, P_T^k \leq P_{max}]. \quad (31)$$

**2) Analysis:** Since the above constraints (i.e., energy causality, interference and SNR constraints) can be expressed with the signal strength of the incident RF signal, we first evaluate its probability distribution, namely the pdf and cumulative density function (cdf) of  $P_I^k$  through the Laplace transform as  $\mathcal{L}_{P_I^k}(s) = \mathbb{E}[\exp(-sP_I^k)]$ , which is derived in Theorem 1.

**Theorem 1:** Since  $h_i \sim \exp(1)$  and all PUs are distributed in  $\mathbb{R}^2$  plane, the characteristic function of  $P_I^k$  is evaluated as

$$\mathcal{L}_{P_I^k}(s) = \text{Det}(\text{I}_d + \alpha \mathbb{K}_k(s))^{-1/\alpha} \quad (32)$$

where  $\mathbb{K}_k(s)$  is given by

$$\mathbb{K}_k(s) = \sqrt{\frac{sp_k}{|\mathbf{x}|^\mu + sp_k}} \mathbb{G}_k(\mathbf{x}, \mathbf{y}) \sqrt{\frac{sp_k}{|\mathbf{y}|^\mu + sp_k}}. \quad (33)$$

In the above,  $p_k = p_b^k \frac{P_{PU} G_{PU}}{4\pi d_0^{2-\mu}}$ , and  $\mathbb{G}_k(\mathbf{x}, \mathbf{y})$  as the kernel of  $\Phi_k$  can be expressed in terms of transformed kernel as

$$\begin{aligned} \mathbb{G}_k(\mathbf{x}, \mathbf{y}) &= \sqrt{l_k} \mathbb{G}_{\Phi}(\mathbf{x}, \mathbf{y}) \sqrt{l_k} \\ &= \zeta_k \exp(\pi \zeta_k \mathbf{x} \bar{\mathbf{y}} - 0.5\pi \zeta_k (|\mathbf{x}|^2 + |\mathbf{y}|^2)) \end{aligned} \quad (34)$$

where  $\mathbb{G}_{\Phi}(\mathbf{x}, \mathbf{y})$  is the kernel of  $\Phi$ .

**Proof:** The detailed proof is presented in Appendix A. ■

The pdf of  $P_I^k$  is then derived by taking the inverse Laplace transform of (32) as

$$\begin{aligned} f_{P_I^k}(\rho) &= \mathcal{L}^{-1}\{\mathcal{L}_{P_I^k}(s)\}(\rho) \\ &= \mathcal{L}^{-1}\{\text{Det}(\text{I}_d + \alpha \mathbb{K}_k(s))^{-1/\alpha}\}(\rho) \end{aligned} \quad (35)$$

where  $\mathcal{L}^{-1}(\cdot)$  denotes the inverse Laplace transform. By definition, the cdf of  $P_I^k$  is then evaluated as

$$\begin{aligned} F_{P_I^k}(\rho) &= \int_0^\rho f_{P_I^k}(t) dt \\ &= \int_0^\rho \mathcal{L}^{-1}\{\mathcal{L}_{P_I^k}(s)\}(t) dt \\ &= \mathcal{L}^{-1}\{\mathcal{L}_{P_I^k}(s)/s\}(\rho) \\ &= \mathcal{L}^{-1}\{\text{Det}(\text{I}_d + \alpha \mathbb{K}_k(s))^{-1/\alpha}/s\}(\rho). \end{aligned} \quad (36)$$

Now we can derive the two performance metrics as follows:

**Proposition 2:** The energy outage and coverage probabilities can be evaluated as

$$\mathbb{O}_B^k = F_{P_I^k}(P_{low}), \quad (37)$$

$$\mathbb{C}_B^k = \int_{P_{low}}^{P_{up}} \exp\left(-\frac{\tau_B}{c_0 \rho}\right) f_{P_I^k}(\rho) d\rho \quad (38)$$

where  $P_{low} = \frac{8\pi\rho_B}{\eta\lambda^2 G_{ST}}$  and  $c_0 = \frac{\Delta\sigma A_{e,SR}}{d_{TR}^2 N_0}$ .

**Proof:** The detailed proof is presented in Appendix B. ■



3) *Special Case (PPP) Analysis [13]*: For insightful analysis, we may derive a closed-form solution. Then, we first derive the Laplace transform of  $P_I^k$  when  $\Phi_k$  follows the thinned-PPP (i.e.,  $\alpha \rightarrow 0$ ) as a special case, which gives rise to the following theorem.

*Theorem 2*: If the distribution of PUs follows PPP, the characteristic function is evaluated as

$$\mathcal{L}_{P_I^k}(s) = \exp \left[ -\frac{\pi \zeta_k}{\text{sinc}(2/\mu)} (p_k s)^{2/\mu} \right]. \quad (39)$$

*Proof*: The detailed proof is presented in Appendix C. ■ To gain useful insights here, we define a new parameter  $a_\mu^k$  as

$$a_\mu^k = \left[ \frac{\pi \zeta_k}{\text{sinc}(2/\mu)} \right]^{\frac{\mu}{2}} p_k, \quad (40)$$

which simplifies the characteristic function (39) to

$$\mathcal{L}_{P_I^k}(s) = \exp \left[ - (a_\mu^k s)^{2/\mu} \right]. \quad (41)$$

The parameter  $a_\mu^k$  will play an important role to determine an optimal/suboptimal traffic pattern. Detailed traffic pattern selection criteria are described in Section IV-D.

To gain further useful insights, we proceed to derive closed-form solutions for some special case in Corollary 1 below.

*Corollary 1*: If  $\mu = 4$ ,  $P_I^k$  follows Lévy distributed, whose pdf and cdf can be evaluated as

$$f_{P_I^k}(\rho) = 0.5 \sqrt{\frac{a_4^k}{\pi}} \rho^{-3/2} \exp \left( -\frac{a_4^k}{4\rho} \right), \quad (42)$$

$$F_{P_I^k}(\rho) = \text{erfc} \left( \sqrt{\frac{a_4^k}{4\rho}} \right) \quad (43)$$

where  $\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-t^2) dt$ . The respective energy outage and coverage probabilities can be evaluated as

$$\mathbb{O}_B^k = F_{P_I^k}(P_{low}) = \text{erfc} \left( \sqrt{\frac{a_4^k}{4P_{low}}} \right), \quad (44)$$

$$\mathbb{C}_B^k = \sqrt{\frac{a_4^k}{\pi a^\dagger}} \int_{\sqrt{a^\dagger/P_{up}}}^{\sqrt{a^\dagger/P_{low}}} \exp(-t^2) dt \quad (45)$$

where  $a^\dagger = \frac{\tau_B}{c_0} + \frac{a_4^k}{4}$ .

*Proof*: The detailed proof of Corollary 1 is given in Appendix D. ■

#### D. Traffic Pattern Selection

Based on Corollary 1, we can draw an interesting observation such that the performance metrics of energy outage and coverage probabilities are largely influenced by the key parameter  $a_4^k$  when  $\mu = 4$ . In fact, we observe that the energy outage probability decreases while the coverage probability increasing as  $a_4^k$  increases. This suggests that the PU signal carrying the specific traffic pattern with maximum  $a_4^k$  be selected for optimal AB communication. The detailed procedure

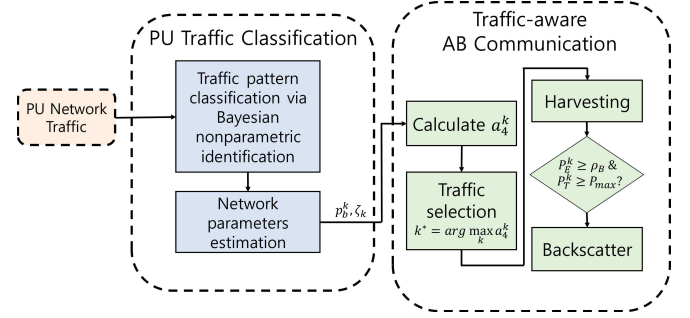


Fig. 4. A flow diagram for traffic-aware backscatter communication.

TABLE II. MEAN VALUES OF MEASURED TRAFFIC DATASET.

Features	VoIP	Game	UDP
Packet length (bytes)	210	69.27	1,512
Packet inter-arrival time ( $\mu\text{sec}$ )	72.4	67,381	3,034
Variance (Packet length)	0	352.46	0

for traffic classification at the first stage and traffic-aware AB communication at the second stage including traffic pattern selection through estimating  $a_4^k$  is illustrated in Fig. 4. If  $\mu = 4$ , we can formulate the criterion for optimal traffic pattern selection as

$$k^* = \arg \max_k a_4^k. \quad (46)$$

In general, it is not allowed to derive the closed-form solutions for the performance metrics analyzed in Corollary 1 when  $\mu = 4$ . Therefore, we develop a general procedure for suboptimal traffic pattern selection below.

*Claim 1*: For a general path-loss exponent  $\mu$ , the suboptimal traffic pattern  $k^*$  can be selected by the following criterion

$$k_C^* = \arg \max_k a_\mu^k. \quad (47)$$

We will prove Claim 1 in terms of the accuracy and usefulness of such suboptimal traffic pattern selection criterion through simulations in the sequel. Furthermore, we will show the claim can be applied to the various distributions of PUs with varying repulsion factor  $\alpha$  as well.

## V. RESULTS

In this section, we present numerical and simulation results about traffic-aware AB communication with the proposed traffic pattern selection criterion in Claim 1. The frequency band for the PU network is assumed to 1.8GHz. The transmit power and antenna gain for PUs are set to 0.2W and 6dBi, respectively, while the transmit and receive antenna gains for SUs is set to 1.8dBi.  $P_{max}$ ,  $\rho_B$ , and  $\tau_B$  are set to 0.2W, -36dBm [7], and 3dB, respectively. Finally,  $d_{TR}$  and the noise power spectral density are set to 3m and -130dBm/Hz, respectively.

For traffic applications, we use real wireless traces such as VoIP, Game, and UDP whose dataset are available in [27] and [28]. TCPdump is used to capture data, which are captured at campus, subway and bus. Specifically, game traffic reveals large variance in packet length, while VoIP and UDP traffic



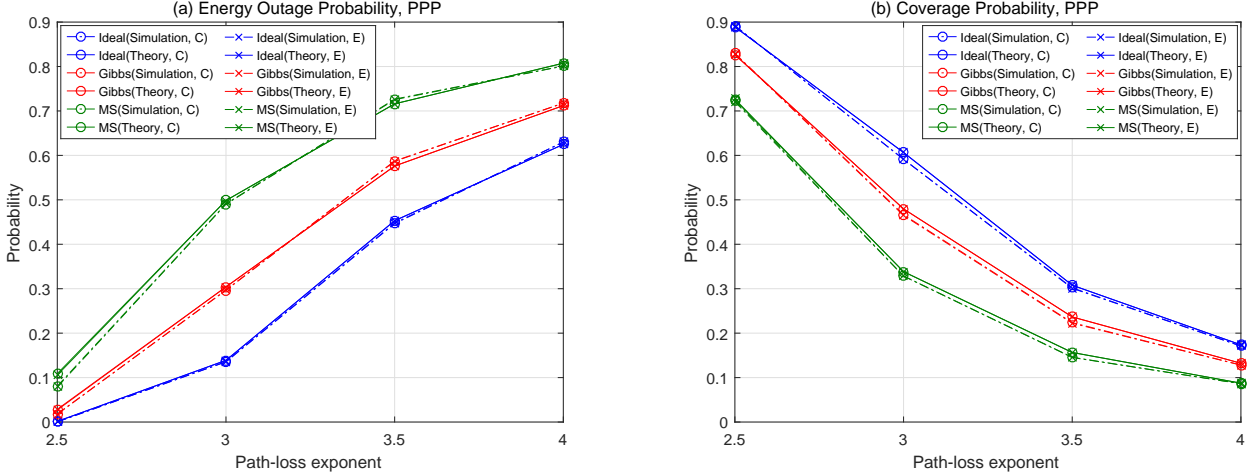


Fig. 5. Energy outage probability (a) and coverage probability (b) with varying path-loss exponent when the distribution of PUs follows PPP. ‘C’ and ‘E’ refer to the traffic pattern selection through Claim 1 and an exhaustive search, respectively.

have stable packet length. Table II shows the feature of the traffic dataset. In average sense, VoIP traffic shows the busiest traffic feature, which is expected to be the most favorable traffic source for AB communication. On the other hand, we can expect the game traffic is not suitable traffic for AB communication.

We assume the density of VoIP, Game, and UDP traffic in the network as 0.005, 0.01, and 0.015, respectively. For the BNP learning algorithm, the hyperparameters are set to  $\mathcal{H} = \{\Lambda_0^{-1}, \nu_0, \bar{\mu}_0, \kappa_0\} = \{\text{Identity}(3), 4, \text{Zeros}(3, 1), 0.5\}$ , and  $\alpha_o = 1$ .

Fig. 5 (a) and (b) show the energy outage probability of ST and the coverage probability of AB communication with increasing path-loss exponent, respectively, when the distribution of PUs follows PPP (i.e.,  $\alpha$ -GPP with  $\alpha \rightarrow 0$ ). We evaluate the performance with Gibbs sampling method applied, and then compare those with nonparametric mean-shift (MS) clustering [29] and an ideal one used, where the ideal one means all traffic information of the PU network *a priori* known to a pair of ST and SR. We also validate the proposed traffic pattern selection criterion (i.e.,  $k_C^*$ ) by comparing with an exhaustive search method which selects traffic patterns with maximum theoretical coverage probability, namely  $k_E^* = \arg \max_k \mathbb{C}_B^k$ . In Fig. 5, ‘C’ and ‘E’ refer to the respective traffic selection criterion.

Here, we see that the performance improves (e.g., increasing the coverage probability and decreasing the energy outage probability) for the low path-loss exponent. We also see that the Gibbs sampling method yields better performance than the MS clustering, as the former classifies traffic applications more accurately than the MS clustering.

Moreover, we see that the performances of the exhaustive search method and proposed traffic pattern selection criterion are almost the same. This confirms that the key parameter  $a_\mu^k$  plays a crucial role in selecting optimal traffic pattern with far reduced complexity in the PPP-distributed PU network. Especially, in view of  $a_\mu^k$ , we see that the performance of

AB (secondary) communication depends largely on the user density and busy period statistics of the selected PU traffic application. This is because AB communication uses the incident RF signal for modulation in the air, which can be harmful interference to the legacy (primary) communication.

Next, we investigate the impact of popularity on the performance, for which the distribution of PUs is changed by adjusting  $\alpha$ . Other simulation parameters are the same as those in Fig. 5. Figs. 6 and 7 show the performance of traffic-aware AB communication with varying path-loss exponent with various distribution of PUs. In Fig. 6, the distribution of PUs follows GPP with  $\alpha = -0.5$ , while  $\alpha = -1$  in Fig. 7. We see that the performance behavior of the MS, Gibbs sampler, and ideal one in GPP environment is similar to that of the PPP. We also see that the performance of the proposed traffic pattern selection criterion and exhaustive search method is almost the same for various  $\alpha$ , which implies that the proposed one can be generalized to other PU distribution. Therefore, we can conclude that  $a_\mu^k$  is also an important parameter in repulsive environment to decide optimal traffic pattern for AB communication.

Fig. 8 shows the energy outage and coverage probabilities with varying  $\alpha$  when  $\mu = 4$ . So far we have shown the validity of the proposed traffic pattern selection criterion relative to the exhaustive search method, and hence omitted the latter. Here we observe that a strong repulsion among nodes (i.e., large  $|\alpha|$ ) yields some performance gain to a pair of SUs. In this case, PUs tend to be scattered, which results in more PUs surrounded near SR. This renders SR receive a strong incident signal which provides more chance to harvest energy and transmit information. In contrast, if the repulsion among nodes disappears, they appear more clustered to have less PUs surrounded near SR, causing less chance for AB communication.

Fig. 9 shows the energy outage and coverage probabilities with varying VoIP traffic density. In the network, VoIP and UDP traffics are mixed with their sum density equal to 0.03

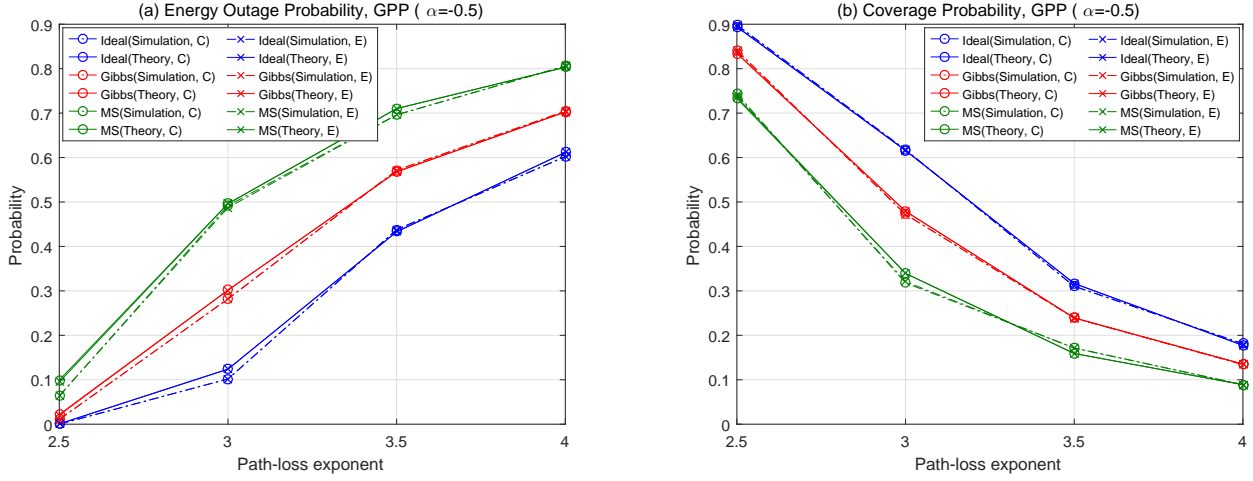


Fig. 6. Energy outage probability (a) and coverage probability (b) with varying path-loss exponent when the distribution of PUs follows GPP with  $\alpha = -0.5$ .

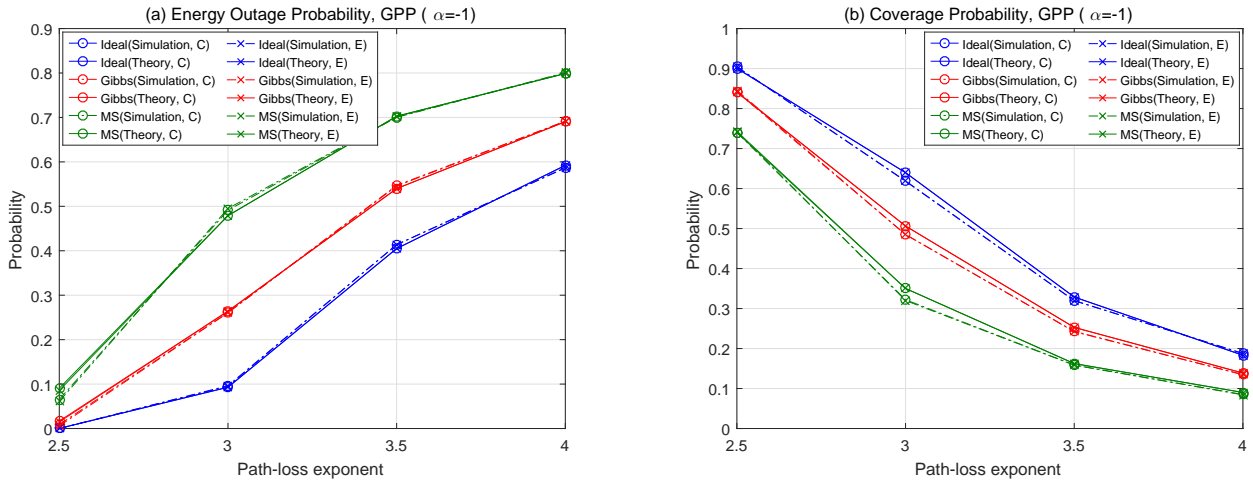


Fig. 7. Energy outage probability (a) and coverage probability (b) with varying path-loss exponent when the distribution of PUs follows GPP with  $\alpha = -1$ .

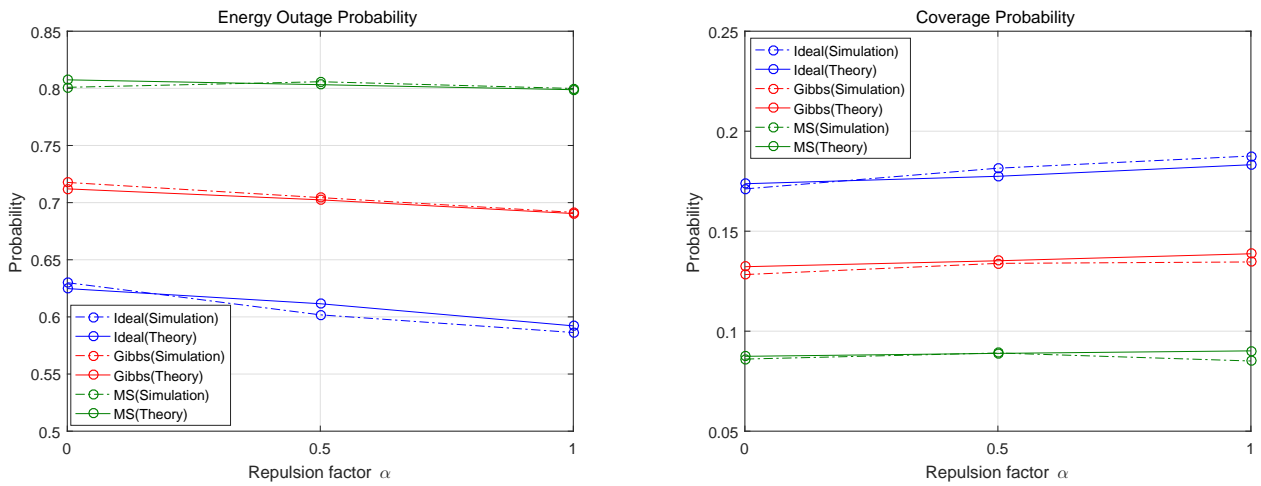


Fig. 8. Energy outage probability (a) and coverage probability (b) with varying repulsion factor  $\alpha$ .

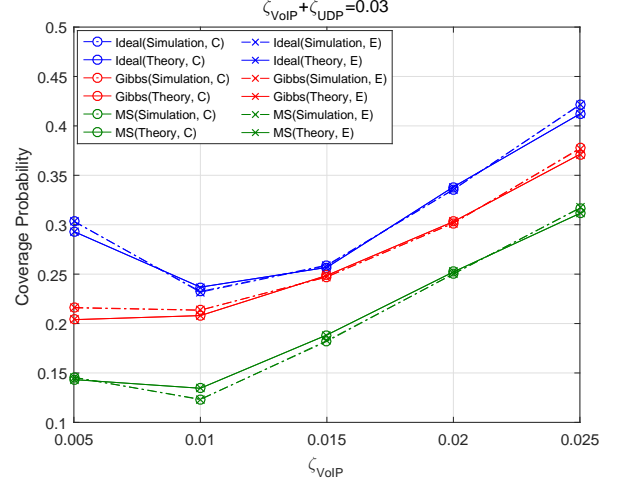
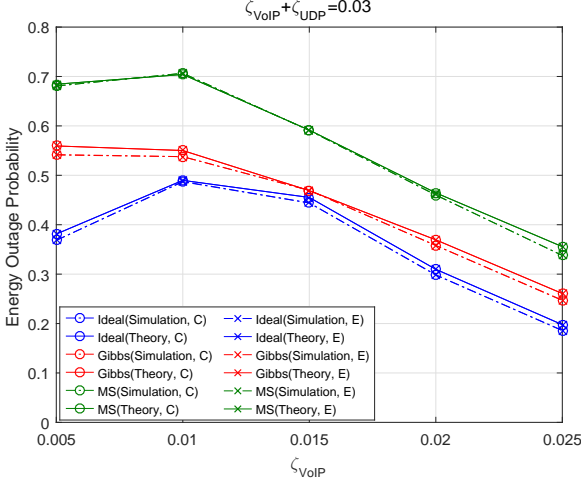


Fig. 9. Energy outage probability (a) and coverage probability (b) with varying VoIP traffic density.

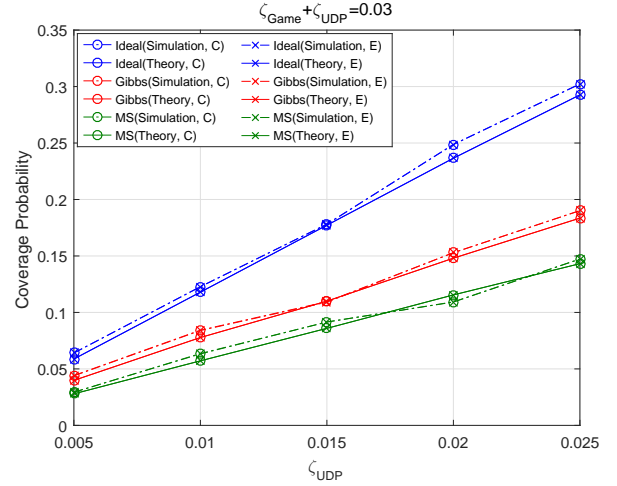
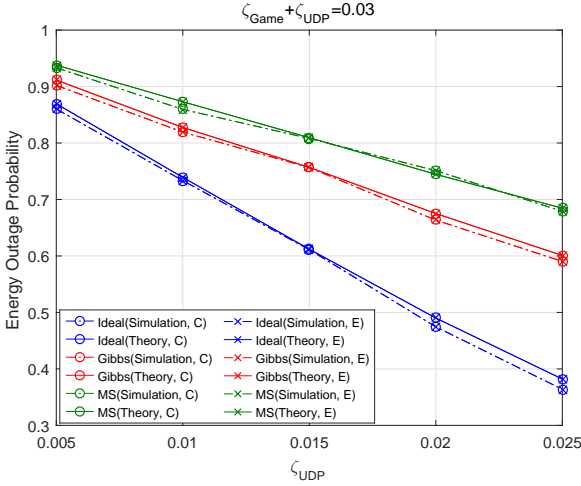


Fig. 10. Energy outage probability (a) and coverage probability (b) with varying UDP traffic density.

(i.e.,  $\zeta_{VoIP} + \zeta_{UDP} = 0.03$ ). PUs are distributed according to -0.5-GPP with  $\mu = 4$ . First, we confirm the validity of the proposed traffic pattern selection criterion even with varying traffic density, as the performance of the proposed one well matches with that of the exhaustive search method. We notice that the performance improves as the portion of one traffic becomes dominant. This is because SU utilizes more likely the traffic source for AB communication as the density of one traffic increases. Moreover, the performance of AB communication with VoIP traffic, which has high busy period channel statistics, is superior to that of AB communication with UDP traffic, which implies that VoIP traffic influences AB communication more than UDP traffic.

Fig. 10 shows the energy outage and coverage probabilities with varying UDP traffic density. The network environment in Fig. 10 is the same as that in Fig. 9, except the traffic applications: one is Game, the other is UDP. Unlike Fig. 9, we see the performance gain as the density of UDP traffic

increases. This is because Game traffic exhibits low busy period statistics, and SU utilizes UDP traffic more for AB communication, even if its density is rather lower than that of Game traffic. This also confirms that AB communication with UDP traffic is superior to that with Game traffic. Through Figs. 9 and 10, we can conclude that the frequent traffic source (e.g., VoIP traffic) provides better chance for traffic-aware AB communication, compared with the less occupied traffic source (e.g., Game traffic).

## VI. CONCLUSION

In this paper, we have proposed traffic-aware AB communication for WPHetNets. In order to operate AB communication, A pair of SUs utilized the traffic sources of PU network whose deployment follows the repulsive point process. Toward this, the BNP learning algorithm was employed to classify traffic applications, and then the optimal traffic pattern selection criterion was obtained by the stochastic geometrical analysis.

The validity of the BNP learning algorithm was shown through numerical analysis by comparing with the well-known MS clustering method. Because of intractable analysis, we further developed a general procedure for suboptimal traffic pattern selection criterion whose validity was demonstrated by comparing with the exhaustive search method through simulations. We confirmed the validity of the proposed selection criterion with various popularity and the impact of traffic sources. It was shown that traffic application with the busiest channel statistics such as VoIP traffic is appropriate for AB communication. Future work will perform the analysis with multiple SU pairs for practical scenario.

#### APPENDIX A PROOF OF THEOREM 1

By applying Proposition 1, the Laplace transform of  $P_I^k$  can be expressed as

$$\begin{aligned}\mathcal{L}_{P_I^k}(s) &= \mathbb{E} \left[ \exp \left( -s \sum_{x_i^k \in \Phi_k} h_i \left( \frac{p_k}{|x_i^k|^\mu} \right) \right) \right] \\ &= \text{Det} [\mathbf{I}_d + \alpha \mathbb{K}_k(s)]^{-1/\alpha}\end{aligned}\quad (48)$$

where the kernel  $\mathbb{K}_k(s)$  is written as

$$\mathbb{K}_k(s) = \sqrt{1 - M_h \left( -s \frac{p_k}{|\mathbf{x}|^\mu} \right)} \mathbb{G}_k(\mathbf{x}, \mathbf{y}) \sqrt{1 - M_h \left( -s \frac{p_k}{|\mathbf{y}|^\mu} \right)}.\quad (49)$$

Since  $h_i \sim \exp(1)$ , the MGF  $h_i$  is given by  $M_h(t) = 1/(1-t)$ . Therefore, the kernel  $\mathbb{K}_k(s)$  has the expression as given in (33).

#### APPENDIX B PROOF OF PROPOSITION 2

By the definition (30), the outage probability can be written as

$$\mathbb{O}_B^k = \Pr[P_E^k < \rho_B] = \Pr[P_I^k < P_{low}] = F_{P_I^k}(P_{low}).\quad (50)$$

Similarly, beginning with the definition (31), and using Bayes' rule, the coverage probability can be derived as

$$\begin{aligned}\mathbb{C}_B^k &= \Pr[P_{low} \leq P_I^k \leq P_{up}, \nu_B^k \geq \tau_B] \\ &= \Pr \left[ P_{low} \leq P_I^k \leq P_{up}, h_{TR} \geq \frac{\tau_B}{c_0 P_I^k} \right] \\ &= \Pr \left[ h_{TR} \geq \frac{\tau_B}{c_0 P_I^k} \middle| P_{low} \leq P_I^k \leq P_{up} \right] \\ &\quad \times \Pr[P_{low} \leq P_I^k \leq P_{up}] \\ &= \int_{P_{low}}^{P_{up}} \Pr \left[ h_{TR} \geq \frac{\tau_B}{c_0 P_I^k} \middle| P_I^k \right] f_{P_I^k}(\rho) d\rho \\ &= \int_{P_{low}}^{P_{up}} \exp \left( -\frac{\tau_B}{c_0 \rho} \right) f_{P_I^k}(\rho) d\rho.\end{aligned}\quad (51)$$

#### APPENDIX C PROOF OF THEOREM 2

The GPP becomes PPP if there is no repulsion among PUs, namely  $\alpha \rightarrow 0$ . Then we can apply the following expansion [30]

$$\text{Det} [\mathbf{I}_d + \alpha \mathbb{K}_k(s)]^{-1/\alpha} \xrightarrow{\alpha \rightarrow 0} \exp \left[ -\int_{\mathbb{O}} \mathbb{K}_k(\mathbf{x}, \mathbf{x}) d\mathbf{x} \right].\quad (52)$$

Hence, (32) can be written as

$$\mathcal{L}_{P_I^k}(s) = \exp \left[ -2\pi\zeta_k \int_0^{R \rightarrow \infty} \frac{r}{1 + r^\mu/(sp_k)} dr \right].\quad (53)$$

The integral term in (53) can be rewritten as

$$\begin{aligned}& \int_0^{R \rightarrow \infty} \frac{r}{1 + r^\mu/(sp_k)} dr \\ &= \int_0^{R \rightarrow \infty} \left( 1 - \frac{1}{1 + sp_k r^{-\mu}} \right) r dr \\ &= \int_0^{R \rightarrow \infty} \int_0^\infty \left[ \exp(-h) - \exp[-(sp_k r^{-\mu} + 1)h] \right] r dh dr \\ &= \int_0^\infty \int_0^{R \rightarrow \infty} \left[ 1 - \exp(-sp_k h r^{-\mu}) \right] r dr \exp(-h) dh \\ &= \mathbb{E}_h \left[ \int_0^{R \rightarrow \infty} \left[ 1 - \exp(-sp_k h r^{-\mu}) \right] r dr \right] \\ &= \mathbb{E}_h \left[ 0.5 sp_k h \int_0^\infty t^{2/\mu} \exp(-sp_k h t) dt \right] \\ &= \mathbb{E}_h \left[ 0.5 (sp_k h)^{2/\mu} \Gamma(1 - 2/\mu) \right] = \frac{(p_k s)^{2/\mu}}{2 \text{sinc}(2/\mu)}.\end{aligned}\quad (54)$$

Therefore, combining (53) with (54) leads to (39).

#### APPENDIX D PROOF OF COROLLARY 1

When  $\mu = 4$ , we can evaluate the pdf of the strength of the incident RF signal through the inverse Laplace transform. By applying Mellin's inverse formula, the pdf can be written as

$$\begin{aligned}f_{P_I^k}(\rho) &= \mathcal{L}^{-1} \left\{ \exp \left( -\sqrt{a_4^k s} \right) \right\}(\rho) \\ &= \frac{1}{2\pi i} \lim_{T \rightarrow \infty} \int_{z-iT}^{z+iT} \exp \left( \rho s - \sqrt{a_4^k s} \right) ds \\ &\stackrel{(a)}{=} \frac{1}{2\pi i} \int_0^\infty \exp(-\rho u) \left[ \exp \left( i\sqrt{a_4^k u} \right) - \exp \left( -i\sqrt{a_4^k u} \right) \right] du \\ &= \frac{2}{a_4^k \pi} \int_0^\infty \exp \left( -\frac{\rho}{a_4^k} v^2 \right) v \sin(v) dv \\ &\stackrel{(b)}{=} 0.5 \sqrt{\frac{a_4^k}{\pi}} \rho^{-3/2} \exp \left( -\frac{a_4^k}{4\rho} \right)\end{aligned}\quad (55)$$

where  $i$  in (55) represents imaginary unit (i.e.,  $i = \sqrt{-1}$ ), (a) follows the Bromwich inversion theorem with the modified contour [31] and (b) is given in [32].

By definition, the cdf of the strength of the incident RF signal when  $\mu = 4$  can be written as

$$\begin{aligned} F_{P_I^k}(\rho) &= \int_0^\rho f_{P_I^k}(t) dt \\ &= \int_0^\rho 0.5 \sqrt{\frac{a_4^k}{\pi}} t^{-3/2} \exp\left(-\frac{a_4^k}{4t}\right) dt \\ &\stackrel{(c)}{=} \int_{\sqrt{\frac{a_4^k}{4\rho}}}^\infty \frac{2}{\sqrt{\pi}} \exp(-x^2) dx \\ &= \operatorname{erfc}\left(\sqrt{\frac{a_4^k}{4\rho}}\right) \end{aligned} \quad (56)$$

where (c) applies integration by substitution (i.e.,  $x^2 = \frac{a_4^k}{4t}$ ).

Now we proceed to evaluate the performance metrics. First, we can easily prove the outage probability  $\mathbb{O}_B^k$  by applying the cdf of  $P_I^k$  in (56) to (37). For the coverage probability  $\mathbb{C}_B^k$ , starting from (38) in Proposition 2, it can be rewritten as

$$\begin{aligned} \mathbb{C}_B^k &= \int_{P_{low}}^{P_{up}} \exp\left(-\frac{\tau_B}{c_0\rho}\right) f_{P_I^k}(\rho) d\rho \\ &= 0.5 \sqrt{\frac{a_4^k}{\pi}} \int_{P_{low}}^{P_{up}} \rho^{-3/2} \exp\left[-\left(\frac{\tau_B}{c_0} + \frac{a_4^k}{4}\right) \rho^{-1}\right] d\rho \\ &\stackrel{(d)}{=} \sqrt{\frac{a_4^k}{\pi a^\dagger}} \int_{\sqrt{a^\dagger/P_{up}}}^{\sqrt{a^\dagger/P_{low}}} \exp(-t^2) dt \end{aligned} \quad (57)$$

where  $a^\dagger = \frac{\tau_B}{c_0} + \frac{a_4^k}{4}$  and (d) applies integration by substitution (i.e.,  $t^2 = \frac{a^\dagger}{\rho}$ ).

#### ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government under Grant 2014R1A5A1011478.

#### REFERENCES

- [1] X. Lu, H. Jiang, D. Niyato, D. I. Kim, and Z. Han, "Wireless-powered device-to-device communications with ambient backscattering: Performance modeling and analysis," *IEEE Trans. Wireless Commun.*, vol. 17, pp. 1528-1544, Mar. 2018.
- [2] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Commun. Surv. Tuts.*, vol. 17, no. 2, pp. 757-789, Second Quart. 2015.
- [3] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless charging technologies: Fundamentals, standards, and network applications," *IEEE Commun. Surveys & Tutorials*, vol. 18, no. 2, pp. 1413-1452, Second Quarter 2016.
- [4] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 13, pp. 418-428, Jan. 2014.
- [5] S. H. Kim and D. I. Kim, "Hybrid backscatter communication for wireless-powered heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 16, pp. 6557-6570, Oct. 2017.
- [6] X. Lu, D. Niyato, H. Jiang, D. I. Kim, Y. Xiao, and Z. Han, "Ambient backscatter assisted wireless powered communications," *IEEE Wireless Commun.*, vol. 25, no. 2, pp. 170-177, Apr. 2018.
- [7] V. Liu, A. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. R. Smith, "Ambient backscatter: Wireless communication out of thin air," *ACM SIGCOMM '13*, pp. 39-50, Hong Kong, Aug. 2013.
- [8] A. N. Parks, A. Liu, S. Gollakota, and J. R. Smith, "Turbocharging ambient backscatter communication," in *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 4, pp. 619-630, Oct. 2014.
- [9] V. Iyer, V. Talla, B. Kellogg, S. Gollakota and J. R. Smith, "Inter-technology backscatter: Towards Internet connectivity for implanted devices," *Proc. of ACM SIGCOMM '16*, Florianopolis, Brazil, Aug. 2016.
- [10] K. Han and K. Huang, "Wirelessly powered backscatter communication networks: Modeling, coverage and capacity," *IEEE Trans. Wireless Commun.*, vol. 16, pp. 2548-2561, Apr. 2017.
- [11] D. T. Hoang, D. Niyato, P. Wang, D. I. Kim, and Z. Han, "Ambient backscatter: A new approach to improve network performance for RF-powered cognitive radio networks," *IEEE Trans. Commun.*, vol. 65, pp. 3659-3674, Sep. 2017.
- [12] N. V. Huynh, D. T. Hoang, X. Lu, D. Niyato, P. Wang, and D. I. Kim, "Ambient backscatter communications: A contemporary survey," to appear in *IEEE Commun. Surv. Tuts.*
- [13] S. H. Kim and D. I. Kim, "Traffic-aware backscatter communications in wireless-powered heterogeneous networks," accepted for publication, *IEEE 88th Vehicular Technology Conference: VTC 2018-Fall*, Chicago, USA, Aug. 2018.
- [14] M. E. Ahmed, D. I. Kim, J. Y. Kim, and Y. Shin, "Energy-arrival-aware detection threshold in wireless-powered cognitive radio networks," *IEEE Trans. Vehic. Technol.*, vol. 66, pp. 9201-9213, Oct. 2017.
- [15] M. E. Ahmed, D. I. Kim, and K. W. Choi, "Traffic-aware optimal spectral access in wireless powered cognitive radio networks," *IEEE Trans. Mobile Computing.*, vol. 17, pp. 733-745, Mar. 2018.
- [16] M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge University Press, 2012.
- [17] T. L. Griffiths and Z. Ghahramani, "Infinite latent feature models and the Indian buffet process," *Gatsby Computational Neuroscience Unit*, Tech. Rep. 2005-001, 2005.
- [18] H. S. Lee, M. E. Ahmed, and D. I. Kim, "Optimal Spectrum Sensing Policy in RF-Powered Cognitive Radio Networks," (invited) 23rd Asia-Pacific Conference on Communications (APCC) 2017, Perth, Australia, Dec. 2017.
- [19] V. Talla and J. R. Smith, "Hybrid analog-digital backscatter: A new approach for battery-free sensing," *2013 IEEE International Conference on RFID (RFID)*, Penang, Malaysia, Apr. 2013.
- [20] P. V. Nikitin, K. V. S. Rao, and R. D. Martinez, "Differential RCS of RFID tag," *Electronic Letters.*, vol. 43, pp. 431-432, Apr. 2007.
- [21] T.-Y. Lin, H. A. Santoso, and K. -R. Wu, "Global sensor deployment and local coverage-aware recovery schemes for smart environments," *IEEE Trans. Mobile Comput.*, vol. 14, pp. 1382-1396, Jul. 2015.
- [22] S.-R. Cho and W. Choi, "Energy-efficient repulsive cell activation for heterogeneous cellular networks," *IEEE J. Select. Areas in Commun.*, vol. 31, pp. 870-882, May 2013.
- [23] N. Deng, W. Zhou, and M. Haenggi, "The Ginibre point process as a model for wireless networks with repulsion," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 107-121, Jan. 2015.
- [24] F. Lavancier, J. Møller, and E. Rubak, "Determinantal point process models and statistical inference," *Journal of the Royal Statistical Society*, vol. 77, pp. 853-877, Sep. 2015.
- [25] H. Kong, I. Flint, P. Wang, D. Niyato, and N. Privault, "Exact performance analysis of ambient RF energy harvesting wireless sensor networks with Ginibre point process," *IEEE J. Select. Areas in Commun.*, vol. 34, pp. 3769-3784, Dec. 2016.

- [26] F. Boremann, "On the numerical evaluation of Fredholm determinants," *Mathematics of Computation.*, vol. 79, pp. 871-915, Apr. 2010.
- [27] [Online]. Available FTP: [http://crawdad.cs.dartmouth.edu/meta.php?name=snu/wow\\_via\\_wimax](http://crawdad.cs.dartmouth.edu/meta.php?name=snu/wow_via_wimax)
- [28] [Online]. Available FTP: <http://crawdad.cs.dartmouth.edu/meta.php?name=kaist/wibro>
- [29] Y. Cheng, "Mean shift, mode seeking, and clustering," *IEEE Trans. Pattern Analysis and Machine Intelligence.*, vol. 17, pp. 790-799, Aug. 1995.
- [30] T. Shirai, and Y. Takahashi, "Random point fields associated with certain Fredholm determinants I: Fermion, Poisson and Boson point process," *Journal of Functional Analysis.*, vol. 205, pp. 414-463, Dec. 2003.
- [31] A. M. Cohen, *Numerical methods for Laplace transform inversion*. New York, NY, USA: Springer-Verlag, 2007.
- [32] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 7th ed. Orlando, FL: Academic, 2007.