Machine Learning Meets Point Process: Spatial Spectrum Sensing in User-Centric Networks

Bodong Shang and Lingjia Liu

***Abstract*—This letter introduces a novel machine learning (ML)-based approach to approximate the distributions of the aggregated interference power in wireless networks. We focus on the application of spatial spectrum sensing (SSS) in user-centric networks where Poisson cluster process (PCP) is used to model the primary users. A nonlinear regression method, i.e., kernel regression, is introduced to learn the distributions of the aggregated interference power of the PCP modeled primary user network. Simulation results demonstrate the accuracy of our approach.**

***Index Terms*—Stochastic geometry, Poisson cluster process, cognitive radio networks, machine learning.**

I. Introduction

Spatial spectrum sensing (SSS) motivates the mobile devices to sense the spatial spectrum opportunities, which can improve the wireless networks overall spectrum utilization efficiency by selecting an appropriate sensing radius. The spatial false alarm probability and the spatial miss detection probability of mobile sensing devices (MSDs) are essential to characterize the channel access probability and thus the mutual interference. In order to obtain the spatial false alarm probability and the spatial miss detection probability, one needs to characterize the distributions of the aggregated interference power generated from primary users in false alarm case and miss detection case, respectively. In the previous works [1], [2], Poisson point process (PPP) is widely used to model the primary users due to its tractability in the analysis and the convenience of the well approximate conditional distributions of the aggregated interference power.

However, the assumption of the uniformly distributed primary users is not quite accurate nowadays. In 5G and beyond networks, the cellular small cell base stations (BSs) are suggested to be deployed in the hotspots of high user density such as downtown areas, superstores, and coffee shops. Such user-centric deployment of small cells reuses the wireless spectral resources more aggressively and is proposed by the 3rd Generation Partnership Project (3GPP) [3]. Therefore, Poisson cluster process (PCP) is more suitable to capture the clustering behaviours of the primary users than PPP in wireless networks. However, in order to analyze and optimize the SSS for MSDs in user-centric networks, the analytical conditional distributions of the aggregated interference power remain unknown.

Manuscript received June 25, 2019; revised August 27, 2019; accepted September 1, 2019. Date of publication September 10, 2019; date of current version January 9, 2020. This work was supported in part by the National Science Foundation under Grant ECCS-1802710, Grant ECCS-1811497, and Grant CNS-1811720. The associate editor coordinating the review of this article and approving it for publication was C. Shen. *(Corresponding author: Lingjia Liu.)*

The authors are with the Wireless@VT, Bradley Department of ECE, Virginia Tech, Blacksburg, VA 24061 USA (e-mail: bdshang@vt.edu; ljliu@vt.edu).

Digital Object Identifier 10.1109/LWC.2019.2940442

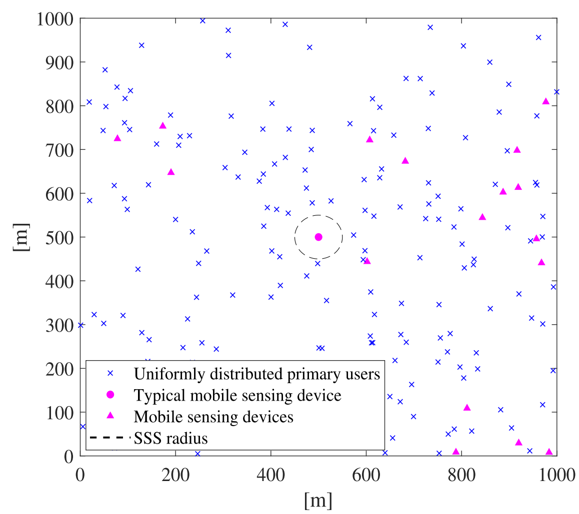

Fig. 1. *Scenario A:* SSS in the PPP modeled primary users networks.

This is because the deficiency of accurate mathematical models and the algorithm deficit of calculating the approximate models.

To address the above issue, this letter introduces a novel approach which is based on machine learning (ML) to obtain the approximate conditional distributions of the aggregated interference power in user-centric networks. We verify that the conditional probability density functions (PDF) of the aggregated interference power in PCP modeled primary users networks can be well approximated by inverse Gaussian (IG) distributions by correcting the mean and the shape parameters. Before conducting the training process, we generate the training dataset based on Monte Carlo method. Furthermore, a kernel-based nonlinear regression is introduced to learn the parameters of IG distributions. The proposed approach can be applied to many complex wireless networks where the distributions of the interference power are unknown.

The remainder of this letter is organized as follows. In Section II, system model is presented. Section III discusses the analytical results for SSS in PCP modeled networks. Section IV introduces the ML-based approach to approximate the conditional distributions of aggregated interference power. Simulation results are shown in Section V. Finally, conclusions are drawn in Section VI.

II. System Model

A. Network Layout

Scenario A: In the networks where the primary users are modeled by PPP shown in Fig. 1, the primary users are uniformly distributed in a 2-dimension plane $\mathbb{R}^2$ denoted by a set

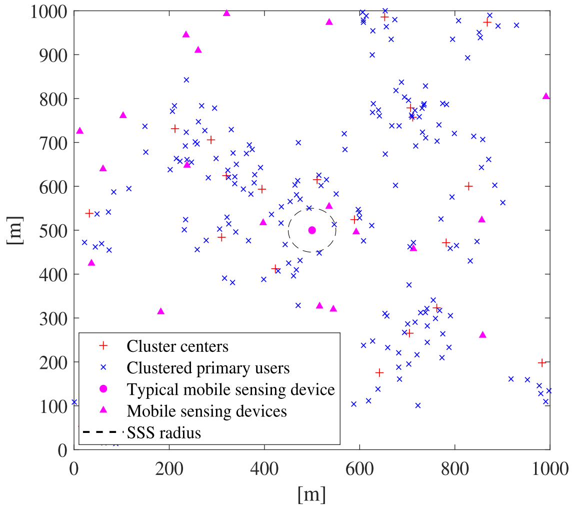

Fig. 2. *Scenario B:* SSS in the PCP modeled primary users networks.

of $\Phi_u = \{u_j\}$ with the density of λ_u, where u_j indicates the j^{th} primary user. The MSDs are modeled by a PPP with the density of λ_d and the set of MSDs is denoted by $\Phi_d = \{d_k\}$ where d_k represents the k^{th} MSD.

Scenario B: In the networks where the primary users are modeled by PCP shown in Fig. 2, the locations of cluster centers, being the parent point process of clusters, are modeled as a PPP with the density of λ_p denoted by $\Phi_p = \{x_j\}$. The clustered primary users are modeled as a PCP with the cluster centers Φ_p. The set of clustered primary users is $\Phi_u^C = \{\Phi_u^{x_j}\}$, where $\Phi_u^{x_j}$ represents the set of primary users attached to the cluster center x_j. The number of primary users per cluster can be fixed or random following Poisson distribution averaged on n_u. The PDF of each primary user conditioned on its cluster center can be modeled by a symmetric normal distribution or a symmetric uniform spatial distribution named Thomas cluster process or Matern cluster process, respectively [4].

B. Spatial Spectrum Sensing

Assume that each MSD has a sensing radius $\mathcal{R}_s$. The sensing region $\mathcal{A}_{d_k}$ of a typical MSD d_k is given by $\mathcal{A}_{d_k} = \mathcal{B}(d_k, \mathcal{R}_s)$, where $\mathcal{B}(d_k, \mathcal{R}_s) = \{x \in \mathbb{R}^2 | \|d_k - x\| \le \mathcal{R}_s\}$ and $\|\mathrm{x} - \mathrm{y}\|$ denotes the distance between x and y. Let $\mathcal{H}^0$ be the event that there is no primary user in the sensing region $\mathcal{A}_{d_k}$, and let $\mathcal{H}^1$ be the event that there is at least one primary user in $\mathcal{A}_{d_k}$.

In *Scenario B*, the received signals $y[n]$ from the primary users during SSS at d_k under $\mathcal{H}^0$ and $\mathcal{H}^1$ are given as follows

$$\mathcal{H}^0 : y[n] = \sum_{x_j \in \Phi_p} \sum_{u_{i,j} \in \Phi_u^{x_j}, u_{i,j} \notin \mathcal{A}_{d_k}} \Upsilon_{u_{i,j}}^B [n] + n_0[n], \quad (1)$$

$$\mathcal{H}^1 : y[n] = \sum_{x_j \in \Phi_p} \sum_{u_{i,j} \in \Phi_u^{x_j}, \Phi_u^C \cap \mathcal{A}_{d_k} \neq \emptyset} \Upsilon_{u_{i,j}}^B [n] + n_0[n], \quad (2)$$

where $\Upsilon_{u_{i,j}}^B [n] = \sqrt{P_u h_{u_{i,j} d_k} \|u_{i,j} - d_k\|^{-\alpha}} s_{u_{i,j}}[n]$ is the received signal from $u_{i,j}$ in n^{th} sample in *Scenario B*, P_u is the transmit power of primary users, $h_{u_{i,j} d_k}$ is the channel power gain, α is the path-loss exponent, $s_{u_{i,j}}[n]$ is the n^{th} sample from $u_{i,j}$, and $n_0[n]$ is i.i.d. circularly symmetric complex Gaussian with mean 0 and variance ${\sigma_n}^2$. The distribution of the test statistics $\Gamma | I^\chi = \frac{1}{N} \sum_{n=0}^{N-1} |y[n]|^2$ approaches to Gaussian distribution according to central limit theorem, i.e., $\Gamma | I^\chi \sim \mathcal{N}(I^\chi + {\sigma_n}^2, \frac{(I^\chi + {\sigma_n}^2)^2}{N}), \chi = 0, 1$, where

$$I^0 = \sum_{x_j \in \Phi_p} \sum_{u_{i,j} \in \Phi_u^{x_j}, u_{i,j} \notin \mathcal{A}_{d_k}} \frac{P_u h_{u_{i,j} d_k}}{\|u_{i,j} - d_k\|^\alpha}, \quad (3)$$

$$I^1 = \sum_{x_j \in \Phi_p} \sum_{u_{i,j} \in \Phi_u^{x_j}, \Phi_u^C \cap \mathcal{A}_{d_k} \neq \emptyset} \frac{P_u h_{u_{i,j} d_k}}{\|u_{i,j} - d_k\|^\alpha}. \quad (4)$$

Note that I^0 and I^1 rely on various network parameters. In SSS, the probability of spatial false alarm and the probability of spatial miss detection are given by

$$P_{fa} = \mathbb{E}_{I^0} \Big\{ \mathbb{P}\Big(\Gamma > \varepsilon | \mathcal{H}^0, I^0 \Big) \Big\} = \int_0^\infty \Theta(x) f_{I^0}(x) dx, \quad (5)$$

$$P_{md} = \mathbb{E}_{I^1} \Big\{ \mathbb{P}\Big(\Gamma < \varepsilon | \mathcal{H}^1, I^1 \Big) \Big\} = \int_0^\infty \Xi(x) f_{I^1}(x) dx, \quad (6)$$

where ε is the energy detection threshold, $\Theta(x) = Q(\frac{\varepsilon - x - {\sigma_n}^2}{x + {\sigma_n}^2} \sqrt{N})$, $Q(\cdot)$ is the Q-function, $\Xi(x) = 1 - \Theta(x)$.

If the test statistic Γ at a MSD is greater than ε, the MSD will transmit with probability β_1, otherwise, it will transmit with probability β_0, where $\beta_0 > \beta_1$. Under event $\mathcal{H}^0$, a MSD will access the channel with probability $P_d^0 = P_{fa}\beta_1 + (1 - P_{fa})\beta_0$. Under event $\mathcal{H}^1$, a MSD will access the channel with probability $P_d^1 = (1 - P_{md})\beta_1 + P_{md}\beta_0$. To evaluate the SSS network performance, the distributions of the aggregated interference power conditioned on the false alarm and miss detection cases ($f_{I^0}(x)$ and $f_{I^1}(x)$) are required.

III. Stochastic Geometry-Based Results

In this section, we discuss the analytical results for the distributions of the aggregated interference power, i.e., I^0 and I^1, in *Scenario B* in SSS.

In *Scenario A*, $f_{I^0}(x)$ can be approximated as IG distribution or log-normal distribution with the first and the second cumulants of I^0 [5]. In addition, $f_{I^1}(x)$ can be obtained using the inverse Laplace transform by [6, eq. (3.24)].

In *Scenario B*, the Laplace transform of I^0 is given in (7) at the top of the next page, where we consider the Thomas cluster process, $f_U(r|x) = \frac{r}{{\sigma_u}^2} \exp(-\frac{r^2 + x^2}{2{\sigma_u}^2}) I_0(\frac{rx}{{\sigma_u}^2})$, σ_u is the scattering variance of the locations of primary users around each cluster center, $I_0(\cdot)$ is the modified Bessel function of the first kind with order zero.

The Laplace transform of I^1 is given in (8) at the top of the next page, where we assume that the nearest primary user is located in the nearest cluster with the cluster center x_1, and $f_{X_1}(x_1) = 2\pi\lambda_p x_1 e^{-\pi\lambda_p {x_1}^2}$, $U_{i,1} = \|u_{i,1} - d_k\|$, $F_U(r|x_1) = 1 - Q_1(\frac{x_1}{\sigma_u}, \frac{r}{\sigma_u})$, $Q_1(a,b) = \int_b^\infty t e^{-\frac{t^2 + a^2}{2}} I_0(at) dt$ is the Marcum Q-function, $f_{U_{1,1}}(r) = n_u (1 - F_U(r|x_1))^{n_u - 1} f_U(u_{1,1}|x_1)$, $f_{U_{i,1}}(r|x_1) = \frac{(1 - F_U(r|x_1))^{n_u - i}}{B(i, n_u - i + 1)} {F_U}^{i-1}(r|x_1) f_U(r|x_1)$, $B(\cdot,\cdot)$ is the beta function.

Based on the inverse Laplace transform, the PDF of I^0 and I^1 can be expressed as

$$f_{I^\chi}(x) = \mathcal{L}^{-1}\{\mathcal{L}_{I^\chi}(s)\}, \chi = 0, 1. \quad (9)$$

Remark 1: The analysis in PCP modeled networks is more complex than that in PPP modeled networks. It is not tractable to calculate (9) since there are many integrals, which leads

$$\mathcal{L}_{I^0}(s) = \mathbb{E}\left\{e^{-sI^0}\right\} = \exp\left\{-2\pi\lambda_p \int_0^\infty \left[1 - \sum_{n=0}^{\infty}\left(\int_{\mathcal{R}_s}^{\infty} \frac{f_U(r|x)}{1+sP_u r^{-\alpha}}dr\right)^{n_u} \frac{(n_u)^n e^{-n_u}}{n!}\right] x\,dx\right\} \tag{7}$$

$$\mathcal{L}_{I^1}(s) = \mathbb{E}\left\{e^{-sI^1}\right\} \approx \mathbb{E}\left\{\frac{e^{-\sum_{u_{i,1}\in\Phi_u^{x_1}, u_{i,1}\in\mathcal{A}_{d_k}} \frac{sP_u h_{u_{1,j}d_k}}{\|u_{i,1}-d_k\|^\alpha}}}{\mathbb{P}\left\{u_{i,1}\in\mathcal{A}_{d_k}\middle|x_1\right\}}\right\}\mathbb{E}\left\{\exp\left[-\sum_{x_j\in\Phi_p, j\neq 1}\sum_{u_{i,j}\in\Phi_{\mathrm{u}}^{x_j}} \frac{sP_u h_{u_{i,j}d_k}}{\|u_{i,j}-d_k\|^\alpha}\right]\right\}$$

$$= \int_0^\infty \int_0^{\mathcal{R}_s} \frac{f_{X_1}(x_1) f_{U_{1,1}}(t|x_1) \int_0^{\mathcal{R}_s} \frac{f_{U_{1,1}}(r|x_1)}{1+sP_u r^{-\alpha}}dr \prod_{l=2}^{n_u} \int_t^\infty \frac{f_{U_{i,1}}(r|x_1)}{1+sP_u r^{-\alpha}}dr}{\int_0^{\mathcal{R}_s} f_{U_{1,1}}(r|x_1)dr \exp\left\{2\pi\lambda_p \int_{x_1}^\infty x\left[1-\sum_{n=0}^{\infty}\left(\int_{\mathcal{R}_s}^\infty \frac{f_U(r|x)}{1+sP_u r^{-\alpha}}dr\right)^{n_u}\frac{(n_u)^n e^{-n_u}}{n!}\right]dx\right\}} dt\,dx_1 \tag{8}$$

to the algorithm deficit. In addition, the accurate mathematical model for the Laplace transform of I^1 in PCP modeled networks remains unknown, which leads to the model deficit.

IV. Machine Learning-Based Approach

In this section, we introduce an ML-based approach, i.e., kernel regression, to obtain the approximate analytical distributions of I^0 and I^1 in *Scenario B*. The obtained distributions are closed-form which are useful in the network analysis and simulations.

Proposition 1: By observing from the simulation results, we propose that the aggregated received signal strength generated from PCP modeled primary users can be approximated by IG distributions with appropriate mean and shape parameters.

In *Scenario B*, the network parameters which can impact the SSS performance are the density of parent point process of clusters λ_p, the spatial spectrum sensing radius $\mathcal{R}_s$, the average number of primary users per cluster n_u, the scattering variance σ_u and the path-loss exponent α. Due to the algorithm deficit and the model deficit for SSS in the PCP modeled primary users networks, we apply the data-driven ML method to characterize the relationships between the input values of network parameters and the output values of the parameters of distributions (the mean parameter μ^χ and the shape parameter λ^χ, $\chi = 0, 1$ for events $\mathcal{H}^0$ and $\mathcal{H}^1$).

To generate the dataset, we can use Monte Carlo simulations or practical field tests. In this letter, we generate the dataset by Monte Carlo simulations with various network parameters to guarantee the diversity of the data. The parameters μ^χ and λ^χ, $\chi = 0, 1$ can be determined as follows

$$\mu^\chi = \mathbb{E}\{I^\chi\}, \tag{10}$$

$$\lambda^\chi = \frac{1}{\mathbb{E}\left\{\frac{1}{I^\chi}\right\} - \frac{1}{\mathbb{E}\{I^\chi\}}}. \tag{11}$$

For SSS in *Scenario B*, the input variables are $\mathbf{x}^{(i)} = (x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, x_4^{(i)}, x_5^{(i)})$ (corresponding to the network parameters λ_p, $\mathcal{R}_s$, n_u, σ_u, α) and the output variable is $y^{(i)}$ (corresponding to the mean parameter μ^χ or the shape parameter λ^χ of the IG distributions), where (i) indicates the dataset index. Since the values of different inputs in $\mathbf{x}^{(i)}$ vary greatly, we can use the normalization method (scaling inputs values between 0 and 1) to statistically equivalent the impact of each input in the training process as follows

$$\widetilde{x}_j^{(i)} = \frac{x_j^{(i)} - \min(\mathbf{x}_j)}{\max(\mathbf{x}_j) - \min(\mathbf{x}_j)}, j = 1, \ldots, 5, \tag{12}$$

where $\mathbf{x}_j = [x_j^{(1)}, \ldots, x_j^{(L)}]$ and L is the number of input-output pairs in the dataset. Then we have the normalized input variables denoted by $\widetilde{\mathbf{x}}^{(i)} = (\widetilde{x}_1^{(i)}, \widetilde{x}_2^{(i)}, \widetilde{x}_3^{(i)}, \widetilde{x}_4^{(i)}, \widetilde{x}_5^{(i)})$.

Remark 2: If primary users have same transmit power, when the transmit power varies, we can update the trained distributions by the scaling property of IG distribution, i.e., if $I^\chi \sim IG(\mu^\chi, \lambda^\chi)$, we have $tI^\chi \sim IG(t\mu^\chi, t\lambda^\chi)$. If transmit powers of primary users are different, we need to add the new network parameters which can impact the primary user's transmit power to input variables of the training process.

We use the kernel regression method to fit the data which utilizes the local information of the dataset at a query point. The query point is given by $\mathbf{q} = (q_1, q_2, q_3, q_4, q_5)$. In this letter, we use the Gaussian kernel function $f(x) = \frac{1}{\sqrt{2\pi\sigma_G^2}} \exp\left[-\frac{(x-\mu_G)^2}{2\sigma_G^2}\right]$, where μ_G is the mean and σ_G is the standard variance which controls the degree of the influence of local information at a query point. Note that the out term $\frac{1}{\sqrt{2\pi\sigma_G^2}}$ will not impact the weighted average. Therefore, the modified kernel function is expressed as

$$K\left(\mathbf{q}, \widetilde{\mathbf{x}}^{(i)}\right) = \exp\left(-\frac{D\left(\mathbf{q}, \widetilde{\mathbf{x}}^{(i)}\right)^2}{2\sigma_G^2}\right), \tag{13}$$

where $\widetilde{\mathbf{x}}^{(i)}$ is the normalized input variables given in (12), $D(\mathbf{q}, \widetilde{\mathbf{x}}^{(i)})$ is the distance between the query point $\mathbf{q}$ and the normalized input variables (data point) $\widetilde{\mathbf{x}}^{(i)}$ from dataset which is given by

$$D\left(\mathbf{q}, \widetilde{\mathbf{x}}^{(i)}\right) = \sqrt{\left(\mathbf{q} - \widetilde{\mathbf{x}}^{(i)}\right)\left(\mathbf{q} - \widetilde{\mathbf{x}}^{(i)}\right)^{\mathrm{T}}}. \tag{14}$$

To leverage the local information of the dataset at a query point, we can obtain the approximate output values at the query point $\mathbf{q}$ by the following equation

$$\widehat{y}(\mathbf{q}) = \frac{\sum_{i=1}^{L}\left(K\left(\mathbf{q}, \widetilde{\mathbf{x}}^{(i)}\right) y^{(i)}\right)}{\sum_{i=1}^{L} K\left(\mathbf{q}, \widetilde{\mathbf{x}}^{(i)}\right)}. \tag{15}$$

Note that the kernel function $K(\mathbf{q}, \widetilde{\mathbf{x}}^{(i)})$ reflects the weight value for the input-output pair $(\widetilde{\mathbf{x}}^{(i)}, y^{(i)})$ regarding to the query point $\mathbf{q}$. Due to the nature of the Gaussian kernel function, the data points which are near the query point will contribute significantly to the estimate output value $\widehat{y}(\mathbf{q})$.

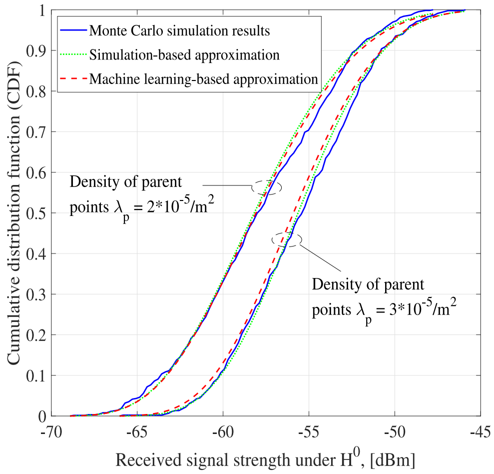

Fig. 3. CDF of I^0 in *Scenario B*.

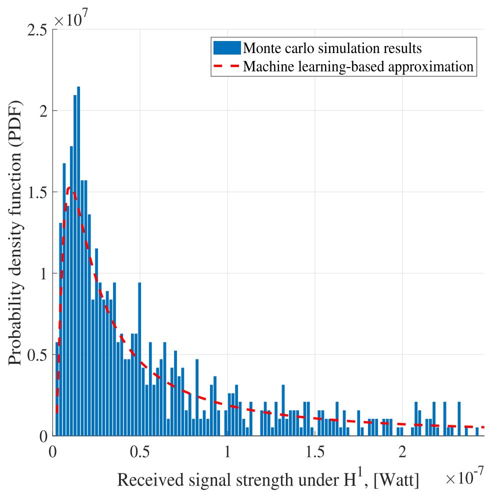

Fig. 4. PDF of I^1 in *Scenario B* where $\mathcal{R}_s = 60$m.

To obtain the approximate values of μ^χ and λ^χ, $\chi = 0, 1$, given the network parameters, we need to evaluate the (15) over a range of query points. When we input a combination of network parameters $(\lambda_p, \mathcal{R}_s, n_u, \sigma_u, \alpha)$, we search its nearest query point and obtain the corresponding output value.

V. Simulation Results and Discussions

In this section, we examine the proposed ML-based approach on the distributions of the aggregated interference power under the events $\mathcal{H}^0$ and $\mathcal{H}^1$ in *Scenario B*, where $n_u = 5$, $\sigma_u = 30$, $\alpha = 4$, $\lambda_p = 3 \times 10^{-5}$/m^2, $P_u = 0.05$W, the number of Monte Carlo simulations is 2000 which refers to the number of simulations for testing ML-based approach in a specific network setting, unless specified otherwise. In the process of dataset collection, the input parameters $\mathbf{x}^{(i)}$ should guarantee the data diversity and also reflect the practical network settings.

In Fig. 3, the CDF of I^0 in *Scenario B* is presented, where $\mathcal{R}_s = 50$m. The results of the simulation-based approximation are obtained by the IG distribution which parameters are generated by (10) and (11) from Monte Carlo results. We observe that the ML-based approximation can approximate the simulation results. Particularly, the ML-based approximation performs well in the top twenty percent, which is helpful in the design of practical engineering system since the false alarm probability is usually set to a small value. It's worth noting that the ML-based approximation needs to collect dataset in the early stage before the training process, but the trained distributions can be used permanently. For simulation-based approximation, we need to do multiple simulations when the network parameters change.

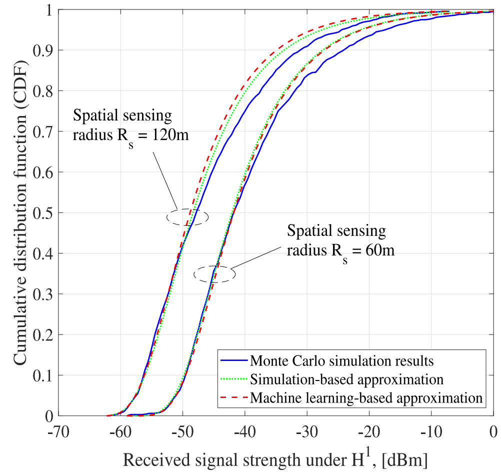

Fig. 5. CDF of I^1 in *Scenario B*.

In Fig. 4 and Fig. 5, the PDF and CDF of I^1 in *Scenario B* are shown, respectively. We observe that the ML-based approximation can well approximate the simulation results. In SSS, if Neyman-Pearson criteria is applied, the values of I^1 in the bottom twenty percent are usually used to characterize the miss detection probability which can be well estimated by the ML-based approximation.

VI. Conclusion

This letter focuses on the SSS in PCP modeled primary users networks and discusses the technical issues on obtaining the distributions of the aggregated interference power analytically. Furthermore, a ML-based approach to approximate the distributions are introduced. Simulation results verify the accuracy of the proposed approach.

In future work, similar approaches are encouraged to be applied to various realistic scenarios that may have more complex models, such as unmanned aerial vehicle (UAV) networks, vehicle-to-everything (V2X) networks, and heterogeneous user-centric deployed cellular networks, etc.

References

[1] H. Chen, L. Liu, T. Novlan, J. D. Matyjas, B. L. Ng, and J. Zhang, "Spatial spectrum sensing-based device-to-device cellular networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7299–7313, Nov. 2016.

[2] H. Chen, L. Liu, H. S. Dhillon, and Y. Yi, "QoS-aware D2D cellular networks with spatial spectrum sensing: A stochastic geometry view," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3651–3664, May 2019.

[3] C. Saha, M. Afshang, and H. S. Dhillon, "3GPP-inspired HetNet model using poisson cluster process: Sum-product functionals and downlink coverage," *IEEE Trans. Commun.*, vol. 66, no. 5, pp. 2219–2234, May 2018.

[4] M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[5] Z. Chen *et al.*, "Aggregate interference modeling in cognitive radio networks with power and contention control," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 456–468, Feb. 2012.

[6] M. Haenggi and R. K. Ganti, "Interference in large wireless networks," *Found. Trends® Netw.*, vol. 3, no. 2, pp. 127–248, 2009.