Statistical QoS Provisioning over Uncertain Shared Spectrums in Cognitive IoT Networks: A Distributionally Robust Data-Driven Approach

Xuanheng Li, *Member, IEEE*, Haichuan Ding, Miao Pan, *Senior Member, IEEE*, Jie Wang, *Senior Member, IEEE*, Haixia Zhang, *Senior Member, IEEE* and Yuguang Fang, *Fellow, IEEE*

***Abstract*—With the soaring wireless traffic for Internet of Things (IoT), spectrum shortage becomes an extremely serious problem, leading to the paradigm shift in spectrum usage from an exclusive mode to a sharing mode. However, how to guarantee the quality of service (QoS) when using the shared spectrum is not straightforward due to its uncertain availability. In this paper, from a session-based view, we propose a metric to evaluate how much data can be delivered via a shared band during a session period, named probabilistic link capacity (PLC), which offers us an effective way to guarantee the QoS statistically. Different from most existing works where the distributional information is assumed exactly known, we develop a distributionally robust (DR) data-driven approach to estimate the value of the PLC based on the first and second order statistics. Two cases are considered that the statistics are exact or uncertain with estimation errors. For each case, to calculate the DR-PLC, we formulate it into a semidefinite programming problem based on the worst-case of conditional-value-at-risk. With the proposed metric, we further design a service-based spectrum-aware data transmission scheme, which allows us to efficiently use different kinds of spectrums to satisfy the diverse IoT service requirements.**

***Index Terms*—Spectrum sharing, spectrum uncertainty, distributionally robust optimization, data-driven approach, IoT.**

Manuscript received December 21, 2018; revised April 29, 2019 and July 23, 2019; accepted September 16, 2019. The work of X. Li was supported in part by the National Natural Science Foundation of China under Grant 61801080, the Fundamental Research Funds of Dalian University of Technology under Grant DUT18RC(3)012, and the Doctoral Research Initiation Foundation of Liaoning Province under Grant 2019-BS-049. The work of M. Pan was supported in part by the U.S. National Science Foundation under Grant CNS-1350230 (CAREER). The work of J. Wang was supported in part by the National Natural Science Foundation of China under Grant 61671102. The work of H. Zhang and Y. Fang was supported in part by the National Natural Science Foundation of China under Grant 61860206005. The work of H. Ding and Y. Fang was supported in part by the Natural Science Foundation of China under 61672106 and the U.S. National Science Foundation under Grant CNS-1717736.

X. Li and J. Wang is with the School of Information and Communication Engineering and the Faculty of Electronic Information and Electrical Engineering, respectively, Dalian University of Technology, Dalian 116024, China. Email: xhli@dlut.edu.cn, wangjie@dlut.edu.cn. H. Ding is with the Department of Electrical Engineering and Computer Science, the University of Michigan, Ann Arbor, MI 48109-2121, USA. E-mail: dhcbit@gmail.com. M. Pan is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204, USA. Email: mpan2@uh.edu. H. Zhang is with the School of Control Science and Engineering, Shandong University, Jinan 250002, China. Email: haixia.zhang@sdu.edu.cn. Y. Fang is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, USA. Email: fang@ece.ufl.edu.

The preliminary version has been published in IEEE WCNC 2019 [1].

I. Introduction

A. Backgrounds

RECENTLY, with the rapid development of Internet of Things (IoT) services, a tremendous number of devices are being connected to the Internet [2]. As the connections increase dramatically, wireless data traffic has shown an explosive growth. According to the forecast of Cisco Visual Networking Index, by 2021, the number of mobile-connected devices will reach 11.6 billion globally, and the data traffic will grow sevenfold between 2016 and 2021 [3]. Such an unprecedent growth on traffic load will soon surpass the capacity of our telecommunications networks, calling for more spectrums as support [4].

High frequency spectrums can provide substantial bandwidth and mitigate the spectrum shortage problem. However, the notorious propagation characteristics over high frequency bands make them only applicable to some specific cases with limited coverage. Therefore, more spectrum in medium and low frequency ranges is also needed to help support IoT services [4]. Although most of the sub-6GHz spectrum has been occupied, many measurement campaigns have shown that under the current static spectrum assignment policy, lots of licensed spectrums are in fact significantly under-utilized in both spatial and temporal domains, leading to a low spectrum efficiency [5]–[7].

As a promising solution, spectrum sharing has received increasing attention recently [8]–[10]. During the sharing, users can be divided into two groups. One is primary users (PUs) who own the spectrum license and have the highest accessing right. The other is secondary users (SUs) who are allowed to opportunistically use the spectrum when PUs are inactive. Such a hierarchical sharing can well comply with the current spectrum assignment situation and improve the spectrum utilization effectively.

To enable such a hierarchical spectrum sharing, cognitive radio (CR) has been regarded as an effective technology, and many CR-based IoT (CR-IoT) frameworks have been proposed recently to support the huge number of IoT devices considering the insufficient spectrum resource [11], [12]. By employing the CR technology, some IoT services could be offloaded to the shared spectrums, i.e., the under-utilized spectrums shared by other parties, and the spectrum shortage problem in IoT networks would be mitigated accordingly [13].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2946834, IEEE Transactions on Vehicular Technology

B. Motivations

Although the CR-IoT framework could create much more transmission opportunities, providing services over shared spectrums is very challenging for operators. Since the shared spectrum has to be vacated when the corresponding PUs reclaim it, the availability of each shared band is actually uncertain, which is different from the traditional licensed one. Such an unique feature causes many new problems for service provisioning, for example, what kind of services can be provided by using the uncertain bands, which bands should be employed to fulfill the transmission task, how to guarantee the quality of service (QoS) considering the uncertainty, etc. Although the spectrum access opportunities might be identified accurately through some advanced sensing or database-based approaches, if the aforementioned problems cannot be well addressed, we might not be able to efficiently utilize these spectrum resources for service delivery. Hence, it is extremely important to study the impact of uncertain availability of shared spectrums on service provisioning.

C. Contributions

In this paper, to facilitate the shared spectrum based service provisioning, from a session-based view, we propose a metric called probabilistic link capacity (PLC), which can be used to evaluate how much data can be delivered via a shared band during a session period and thus provide us with an effective way to guarantee the session accomplished statistically. Although how the PLC is mathematically defined is similar to the outage capacity (OC) and the effective capacity (EC) in [14], where a probability is introduced to quantify the uncertain achievable rate, they are actually conceptually different. The PLC proposed in this paper corresponds to an equivalent achievable rate from the view on time average for a session period, not represents the actual instantaneous achievable rate. It is developed based on a different uncertainty model from that considered in OC and EC, and is especially suitable for the CR based spectrum sharing system, where the uncertainty comes from the hierarchical sharing rule, reflected in the available duration of a shared band within a session period which is modeled as a random variable.

To describe PLC, we formulate it as a stochastic optimization problem. Instead of assuming an exact distribution function for the random variable as what has been widely adopted in existing works, which, however, might be hardly obtained in practice, we use the first and second order statistical information to construct an ambiguity set, representing all possible distributions, and develop a distributionally robust solution. Specifically, we model the random variable with an ambiguous distribution subject to the statistics [15], [16]. We focus on the worst-case regarding to all possible distributions with the same statistics and make the PLC a distributionally robust one (DR-PLC). Such a lower bound corresponds to a conservative measurement and can make the QoS guaranteed in fact with a higher probability than the pre-defined confidence level. Since the formulated DR-PLC optimization problem is intractable, we first make a conservative approximation by employing the concept of conditional-value-at-risk (CVaR) [17] and take its worst-case to get an approximated DR-PLC. Then, we reformulate the problem into a tractable semidefinite programming (SDP) problem, so that the DR-PLC can be achieved. With the proposed metric, by considering different types of IoT services (delay-sensitive and delay-tolerant) and different features of spectrums (licensed and shared), we design a service-based spectrum-aware data transmission scheme to utilize the spectrums efficiently for service provisioning. Compared with our preliminary work [18], in this paper, an additional case that the statistical information is uncertain with estimation errors has been further studied, which is more common in practice, and a concrete theoretical proof has been included to show that the developed worst-case CVaR based approximation is in fact exact. Our major contributions are listed as follows.

- By modeling the available duration of a shared band during a service period as a random variable, we propose a metric, PLC, to evaluate the equivalent achievable rate of a link within the period under certain confidence level. Such a metric brings us a way to offer a statistical QoS guarantee when using a shared band for service provisioning, and frequency switching can be avoided during the service provisioning if a high confidence level is adopted for the performance measurement.
- Unlike most existing works where certain specific probability distribution is assumed for the random variable to make the stochastical constraint tractable, which, however, might be difficult to obtain in practice, we adopt its first and second order statistical information, and design a distributionally robust data-driven approach to make the obtained PLC achievable under all possible distributions subject to the statistics. We studied it under both accurate and inaccurate statistical information. For each case, to solve the DR-PLC, we leverage the worst-case of CVaR with the analysis for the accuracy on approximation, and develop a SDP-based reformulation.
- Based on the newly proposed metric, we design a service-based spectrum-aware data transmission scheme by considering both delay-sensitive and delay-tolerant IoT services and both licensed and shared spectrums, where different kinds of spectrums can be used efficiently to meet as many service requirements as possible. For such a schematic design, we formulate it as a two-step optimization problem, including an admission control step and a spectrum re-allocation step.

The rest of paper is organized as follows. Related works are reviewed in Section II. In Section III, the network model is presented, including the new metric PLC. Then, we develop a distributionally robust data-driven approach to calculate the PLC based on the statistical information in Section IV. In Section V, by leveraging the new metric, we design a service-based spectrum-aware transmission scheme. Finally, numerical results are discussed in Section VI and conclusions are drawn in Section VII.

II. Related Works

Spectrum allocation is an important issue for wireless networks, especially for the IoT network considering the massive

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2946834, IEEE Transactions on Vehicular Technology

and heterogeneous data traffic. Taking the software defined networking as the solution for IoT, in [19] and [20], Tang *et. al.* investigated the channel assignment issue and proposed different novel deep-learning-based intelligent assignment approaches, where future traffic loads of switches are predicted based on the historical data and the channel resources are allocated intelligently. Facing the spectrum shortage problem in IoT networks, many research efforts have been dedicated to investigating how to embrace the CR technology for IoT networks, where the CR enabled devices could dynamically access the vacant spectrums by perceiving the radio environment [12]. In [21], Ejaz *et. al.* designed a multiband cooperative sensing scheme for CR-IoT networks with reduced energy consumption, and a cross-layer reconfiguration scheme was further developed for dynamic resource allocation. In [22], Zhang *et. al.* proposed a blind joint sub-Nyquist sensing scheme by employing the surround IoT devices to jointly sample the spectrum. Based on these advanced spectrum sensing methods, a new kind of spectrum would be brought to the CR-IoT network, that is the shared spectrum, i.e., the captured under-utilized spectrum shared by other parties. However, how to effectively use these spectrums to provide services is not straightforward because of their uncertain availability, making the QoS difficult to guarantee.

Actually, regarding the channel fading or the imperfect channel estimation as the spectrum uncertainty, many works have been proposed to guarantee the transmission performance against such an uncertainty issue, which can trace back to the classic concept of outage capacity. In [14], Wu *et. al.* considered the impact of small-scale channel fading and proposed a link-layer channel model called effective capacity, which could be utilized for QoS guarantee, such as delay bounds. Similarly, in [23], Zhang *et. al.* studied the QoS provisioning issue for IoT in LTE-A heterogeneous networks, where the concept of effective bandwidth was leveraged to provide users with probabilistic QoS guarantee under the fast fading channel gain. In [24], Li *et. al.* designed a probabilistic robust beamforming scheme for MISO-SWIPT systems, where the obtained channel vectors are described based on the Gaussian uncertainty model considering the imperfect channel estimation. Also considering the imperfect CSI problem, in [25], Sun *et. al.* developed a novel robust beamforming scheme for the NOMA-based CR system from the perspective on energy efficiency, where both bounded and Gaussian error models are employed. In [26], Zhou *et. al.* proposed an artificial-noise-aided cooperative scheme for the CR system, where the uncertainty on the jamming signal is well utilized to improve the security of the primary network. In [27], Zhang *et. al.* investigated the resource allocation problem in a two-tier OFDM-based cognitive heterogeneous cellular network, where the channel gain is considered uncertain and the interference is limited in the sense of probability. Although these works have well addressed the spectrum uncertainty issue, it is noteworthy that the uncertainty is actually different from the one considered in this paper. Theirs are from the channel fading or the imperfect channel estimation, whereas, ours is caused by the hierarchical sharing rule, i.e., the uncertain availability, which is a particular issue for the CR-IoT system. However, these excellent research works really inspire us a lot for the design on PLC, where a similar approach is adopted that introduce a probability to quantify the uncertain achievable rate.

As for the uncertain availability of shared spectrums, many research works have considered it when designing spectrum sharing schemes. In [28], Zhang *et. al.* jointly investigated spectrum access, power allocation and user scheduling schemes when both licensed and harvested channels exist, where the harvested ones are assumed to be available with certain specific probabilities. Similarly, the uncertain availability is characterized by a probability in [29], where PU activities on the shared channels are assumed to follow an i.i.d. Bernoulli distribution in each time slot, and a spectrum allocation framework was proposed accordingly. Based on the probability assumption, Markov chain has been used to model the spectrum state transition, such as the spectrum sensing and access protocol developed in [30], and the channel allocation scheme designed in [31]. In [32], Cheng *et. al.* studied an opportunistic spectrum access scheme for CR vehicular ad hoc networks, where the length of an idle period of a shared channel is modeled as an exponential random variable. Similarly, considering the idle time distribution, Sharma *et. al.* designed a stochastic model-based opportunistic transmission scheme in [33]. In [34], Yu *et. al.* developed an optimal spectrum investment strategy, including spectrum sensing and leasing decisions, where the amount of available spectrum is modeled as a random variable with a given distribution. Also by modeling the available bandwidth as a random variable, Pan *et. al.* investigated the joint routing and frequency scheduling problem in multi-hop CR networks under uncertain spectrum supply [35]. All above works characterize the uncertain availability of the shared spectrum through specific probabilities or distribution functions. Unfortunately, in practice, such information is usually difficult to be precisely obtained. Even if it might be derived from the historical data, the distribution is more likely to follow a complex expression, rather than the simplified theoretical modeling assumptions adopted in most existing works. Moreover, estimation errors of the distribution might lead to an over-optimistic solution. Hence, we develop a distributionally robust approach to deal with the ambiguity on distribution.

Besides the statistics based approach, as one of the most popular data-driven approaches, machine learning (ML) plays an important role for spectrum prediction and has been employed for CR systems. As described in [36], the ML approach can learn the PUs' behavioral patterns by invoking a set of spectrum data and rely on it for making forecasts. In [37], Xing *et. al.* employed a popular feedforward artificial neural network model, called multilayer perceptron (MLP), for spectrum prediction, which was also adopted in [38], where each CR user predicts the future channel states by using an MLP based predictor and senses only those channels that are predicted to be idle. In [39], Zhang *et. al.* utilized an online support vector regression (SVR) based method to predict the channels' statuses of PUs to help CR users make the channel selection decision. Although many existing works have embraced the ML approach to help CR users make a better strategic decision, like spectrum sensing and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2946834, IEEE Transactions on Vehicular Technology

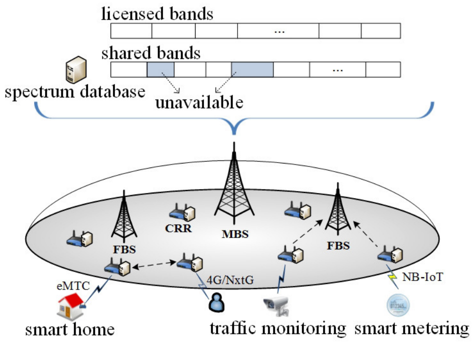


Fig. 1. Cognitive capacity harvesting network architecture for IoT service delivery by using both licensed and shared spectrums.

spectrum accessing, most of them focus on how to predict the spectrum state for the next time slot, instead of solving the uncertain availability problem from the perspective of service provisioning. With regard to service provisioning, we are interested in the spectrum availability within the whole service period, rather than a time slot. Thus, it is necessary to find a way to evaluate how much data can be carried over a shared band within a period, which depends on not only the current available bandwidth but also the future available duration during the period. In fact, for the uncertain available duration, the ML approach might be used to predict how long it is, which, however, needs to feed in massive spectrum data and causes high computation complexity. According to the existing spectrum measurement works [5]–[7], we find that for some bands, the available duration within a period can be well modeled by a random variable, which inspires us to adopt the statistics based modelling to develop the PLC concept and the study on ML based approach for QoS provisioning in CR-IoT systems will be one of our future works.

III. Network Model

A. Architectural Enhancement for IoT Data Transmissions based on the Shared Spectrum

In general, most light-weighted IoT devices have no capability to implement spectrum sensing to determine the availability of the shared spectrum, even hardly directly work on those non-contiguous under-utilized bands belonging to other parties. Thus, before introducing the proposed PLC, we first present a cognitive capacity harvesting network (CCHN) architecture as shown in Fig. 1 [11], which can facilitate the IoT data transmissions over shared spectrums and is taken as the network scenario in this paper.

To be specific, the CCHN can be regarded as an architectural enhancement on the existing cellular network, making it an ultra-dense network as advocated in the 5G vision. Three kinds of entities are included, namely, macro-cell base station (MBS), femto-cell base station (FBS), and cognitive radio router (CRR). *a*) MBS. The MBS provides a basic coverage and is mainly in charge of the control signalling. *b*) FBS. The FBSs are connected to the Internet, acting as data aggregation points for CRRs to deliver data to IoT service providers. *c)* CRR. The CRRs are equipped with multiple radio interfaces and deployed at the edge of the network. On the one hand, they serve as gateways for IoT devices with the support for various accessing technologies, such as 3G/4G/NxtG, NB-IoT, eMTC, etc., so that any IoT device type can access the network. On the other hand, CRRs can connect IoT devices to FBSs, or directly connect two IoT devices to enable an extended device-to-device (D2D) communication, using the non-contiguous shared bands. Thus, the IoT data can be delivered through the shared bands even though the IoT devices cannot directly use them. Two modes of data transmission exist in the network. One is Device $\leftrightarrow$ CRR $\leftrightarrow$ FBS, such as the pollution monitoring data transmissions from sensors to a data center for analysis, and the other is Device $\leftrightarrow$ CRR $\leftrightarrow$ CRR $\leftrightarrow$ Device, such as the surveillance video data transmissions from a house to a user's device for home security systems.

We consider two types of services in the network. One is the delay-sensitive (DS) service, e.g., online video conference, and the other is the delay-tolerant (DT) service, e.g., movie downloading, which are denoted by $\mathcal{L}_{\mathrm{DS}}$ and $\mathcal{L}_{\mathrm{DT}}$, respectively. Note that the shared spectrum can only support the DT services because its availability is uncertain and thus cannot guarantee the maximum delay on each packet for DS services. For a DS service $l \in \mathcal{L}_{\mathrm{DS}}$, assume that its service request includes a rate requirement $r_{\mathrm{DS}}(l)$, a source node $s_{\mathrm{DS}}(l)$, and a destination node $d_{\mathrm{DS}}(l)$. Such a DS service can only be served by the licensed bands with stable availability, denoted as $\mathcal{M}_{\mathrm{l}}$. For a DT service $l \in \mathcal{L}_{\mathrm{DT}}$, assume that its service request is transmitting certain amount of data, denoted as $z_{\mathrm{DT}}(l)$, from a source node $s_{\mathrm{DT}}(l)$ to a destination node $d_{\mathrm{DT}}(l)$ within a period $T(l)$. Then, each DT service l actually corresponds to have an equivalent rate requirement as $r_{\mathrm{DT}}(l) = \frac{z_{\mathrm{DT}}(l)}{T(l)}$. For example, if the request is transmitting 6Gbits monitoring data from a collecting CRR to an FBS within 10 minuets, then it is equivalent to having a requirement on average rate as 10Mbps. Such a DT request can be achieved by either licensed bands or shared bands $\mathcal{M}_{\mathrm{s}}$. When using the shared bands, how to satisfy the rate requirement is a problem because of the uncertain availability of the bands during the session period, which is what we intend to address in this paper.

Suppose that there are $\mathcal{N}$ nodes (including CRRs and FBSs) in the network. For any node $i \in \mathcal{N}$, all licensed bands are available, whereas, only part of shared bands can be used depending on the sensing result, denoted as $\mathcal{M}_{\mathrm{s}}^{i} \subseteq \mathcal{M}_{\mathrm{s}}$. Then, all available bands at this node can be represented by $\mathcal{M}^{i} = \mathcal{M}_{\mathrm{s}}^{i} \cup \mathcal{M}_{\mathrm{l}}$, and the common band set for two different nodes i and j is expressed as $\mathcal{M}^{ij} = \mathcal{M}^{i} \cap \mathcal{M}^{j}$. Upon receiving service requests, MBS will check which shared bands can be used, which can be sensed and submitted by CRRs along with the service request information. Then, according to the aggregated information, by evaluating the link capacity over different bands, MBS will make the data transmission scheduling to satisfy the diverse service requirements by using the different kinds of spectrums[1].

[1]In our scheme, the links will stay on the allocated bands instead of implementing spectrum switching, i.e., keep silence when PUs reclaim them and wait for them available again, because the data volume that can be transmitted during the session period has been well evaluated based on the proposed PLC model, which will be introduced later.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2946834, IEEE Transactions on Vehicular Technology

B. Related Models for Data Transmissions

Transmission Range and Interference Range: For the data transmission from node i to node j, it is considered to be successful only if the received power at node j exceeds certain threshold P_{th}^{T}. Therefore, each node in the network has a transmission range. To be specific, we adopt a widely used model to represent the power propagation gain from node i to node j, described as $g_{ij} = \tau \cdot d_{ij}^{-\alpha}$, in which τ is an antenna related parameter, α is the path loss factor, and d_{ij} is the distance between these two nodes. Suppose the transmission power at node i on band $m \in \mathcal{M}^i$ is p_i^m. Then the received power at node j can be calculated as $p_i^m \cdot g_{ij}$, and the transmission range of node i on band m can be obtained by $R_{i,m}^{\text{T}} = \left(\tau \cdot p_i^m / P_{\text{th}}^{\text{T}}\right)^{1/\alpha}$. Accordingly, for any node j which can use band m, i.e., $m \in \mathcal{M}^j$, if it is located within the transmission range of node i, i.e., $d_{ij} \leq R_{i,m}^{\text{T}}$, we define it as the transmission neighbor of node i on band m, and the corresponding neighbor set is described as

$$\mathcal{T}_{i,m} = \left\{ j \in \mathcal{N} | d_{ij} \leq R_{i,m}^{\text{T}}, j \neq i, m \in \mathcal{M}^{ij} \right\}. \tag{1}$$

Similarly, for the received interference power at each node from an unintended transmitter, it can be ignored only if it is below certain threshold P_{th}^{I} ($P_{\text{th}}^{\text{I}} < P_{\text{th}}^{\text{T}}$). Hence, for any node $i \in \mathcal{N}$ transmitting data on band m with transmission power p_i^m, there exists an interference range as well denoted as $R_{i,m}^{\text{I}} = \left(\tau \cdot p_i^m / P_{\text{th}}^{\text{I}}\right)^{1/\alpha}$. Then, all nodes receiving data on band m located within this area will be interfered by the node i, and similar to (1), the interfering neighbor set of node i on band m can be defined as

$$\mathcal{I}_{i,m} = \left\{ j \in \mathcal{N} | d_{ij} \leq R_{i,m}^{\text{I}}, j \neq i, m \in \mathcal{M}^{ij} \right\}. \tag{2}$$

Link Capacity: Considering a link from node i to node j on band m, $j \in \mathcal{T}_{i,m}$, the link capacity can be expressed as

$$c_{m,ij} = W_m \cdot \log_2 \left(1 + \frac{\hat{p}_i \cdot g_{ij}}{\gamma} \right), \tag{3}$$

where γ is the ambient Gaussian noise density at node j, W_m is the bandwidth of band m, and $\hat{p}_i$ is the transmission power density at node i. Interference is not considered here because it can be eliminated by the transmission scheduling as introduced in Section IV.

C. Probabilistic Link Capacity

Considering a DT service l which intends to transmit $z(l)$ data within a period $T(l)$, if we want to use shared bands to fulfill this transmission task, we need to evaluate how much data that can be supported by a shared band within the period and then determine which bands should be selected to guarantee the session accomplished. Nevertheless, as aforementioned, due to the uncertain PUs' activities, the available duration of a shared band m within the period, denoted as $\tilde{T}_m(l)$, is uncertain and might be less than $T(l)$, so that the actual data volume that can be transmitted over this band during this period is also uncertain calculated as

$$\tilde{z}(l) = \tilde{T}_m(l) \cdot W_m \cdot \log_2 \left(1 + \frac{\hat{p}_i \cdot g_{ij}}{\gamma} \right). \tag{4}$$

To capture the uncertain availability of the shared band, inspired by the findings in some existing spectrum measurement works [5]–[7], we model the available duration $\tilde{T}_m(l)$ as a random variable, making $\tilde{z}(l)$ also a random variable associated with it. Divided by $T(l)$, we can get an equivalent link capacity with regard to this session. By introducing a probability, we can quantify this capacity under a confidence level, which is defined as the probabilistic link capacity of the link from i to j on this band for this session as

$$\tilde{c}_{m,ij} = \max \left\{ c : \Pr\left\{ \tilde{W}_m^l \cdot \log_2 \left(1 + \frac{\hat{p}_i \cdot g_{ij}}{\gamma} \right) \geq c \right\} \geq \alpha \right\}, \tag{5}$$

i.e., the maximal value of c satisfying the stochastical constraint in (5), where $\tilde{W}_m^l = W_m \cdot \frac{\tilde{T}_m(l)}{T(l)}$ represents the average available bandwidth in terms of the session, and $0 < \alpha < 1$ is the confidence level.

Remark 1: By employing such a metric PLC, the operator can measure the impact of uncertain availability on the shared band effectively when scheduling transmissions, and offer a statistical QoS guarantee for the requested DT services with certain confidence level. Note that when a high confidence level is adopted, it would not be necessary to implement spectrum switching when the current band turns to be unavailable during the transmission, where waiting until it is available again might be a better decision, because it is with a high probability that the transmission task could be completed via this band within the period. Thus, the considerable overhead brought by the spectrum handoff could be avoided.

Remark 2: Note that the PLC proposed here is different from the classic concept of effective capacity (EC) developed in [14]. First, they are targeted to different scenarios. EC offers a characterization of the link where source continuously sending delay constrained packets, and it is especially suitable for delay-sensitive applications. In contrast, the PLC is designed for a scenario where certain amount of delay-tolerant data is to be delivered in a specific period, and it is proposed mainly for the delay-tolerant services. Second, the EC is developed from a per packet based point of view, which can be used to facilitate the determination of the source rate at which the violation probability of packet delay can be constrained below certain threshold. Whereas, the PLC is developed from a session-based point of view, which is used to evaluate whether or not a session could be accomplished via certain bands at the end of the session period considering the uncertain transmission duration. Third, the uncertainty issues addressed in both works are different, which are the spectrum reliability under the small-scale fading and the spectrum availability under the hierarchical sharing rule, respectively. Actually, the PLC corresponds to an equivalent achievable rate from the view on time average for a session period, not reflects the actual instantaneous achievable rate, and is tailored to CR based spectrum sharing systems.

IV. A Distributionally Robust Data-Driven Approach for Probabilistic Link Capacity

Considering a link on a shared band $m \in \mathcal{M}_{\text{s}}$, to obtain its PLC defined as (5), obviously, the key is to evaluate the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2946834, IEEE Transactions on Vehicular Technology

stochastical constraint as

$$\Pr\left\{\tilde{W}_m^l \cdot \log_2\left(1+\frac{\hat{p}_i \cdot g_{ij}}{\gamma}\right) \geq c\right\} \geq \alpha. \tag{6}$$

Generally speaking, the widely adopted approach in the current literature is to utilize the probability distribution of the random variable $\tilde{W}_m^l$ based on certain structural assumption [34], [35]. However, whether such an assumed distribution really fits for the variable is questionable. In practice, the full and accurate information about the distribution is hardly obtainable. Even if it could be approximated based on the historical data, replacing the unknown distribution with an estimated one to evaluate the stochastical constraint (6) may lead to an over-optimistic solution to (5). Thus, the obtained PLC may fail to satisfy the constraint under the true distribution. Furthermore, the expression of the estimated distribution might be complicated, making the stochastically constrained optimization problem (5) intractable. To address these challenges, we will develop a data-driven approach by using the first and second order statistical information and achieve a distributionally robust solution.

A. Worst-Case for Distributional Robustness

Recalling the stochastical constraint (6), we rewrite it as

$$\Pr\left(\varphi\left(c, \delta_m\right) \leq 0\right) \geq \alpha, \tag{7}$$

where

$$\varphi\left(c, \delta_m\right) = c - \delta_m = c - \tilde{W}_m^l \cdot \log_2\left(1+\frac{\hat{p}_i \cdot g_{ij}}{\gamma}\right). \tag{8}$$

As aforementioned, the accurate probability distribution function of the random variable δ_m associated with $\tilde{W}_m^l$, denoted as f_{δ_m}, is hardly obtainable in practice, which makes it difficult to deal with such a stochastical constraint. Thus, we consider the worst-case, which can be regarded as its distributionally robust counterpart, and describe it as

$$\min_{f_{\delta_m} \in \mathcal{U}_f} \left\{\Pr\left(\varphi\left(c, \delta_m\right) \leq 0\right)\right\} \geq \alpha. \tag{9}$$

$\mathcal{U}_f$ represents all possible probability distributions that are consistent with certain known statistical characteristics of the random variable δ_m. By replacing the original constraint (6) with its worst-case (9), we can obtain a conservative estimation and achieve a distributionally robust solution.

We employ the expectation and the variance of δ_m as the known statistics to characterize the probability distribution function set $\mathcal{U}_f$, which can be easily extracted from historical observations and are denoted as μ and σ^2, respectively. We consider two cases. One is that the estimates of statistics are derived from sufficient data and close to the real ones. Then the distribution function set can be expressed as

$$\mathcal{U}_f^1 = \left\{ f_{\delta_m} \geq 0 \;\middle|\; \begin{array}{c} \mathbb{E}_f\left(\delta_m\right) = \mu, \mathbb{E}_f\left(1\right) = 1 \\ \mathbb{E}_f\left(\left(\delta_m - \mu\right)^2\right) = \sigma^2 \end{array} \right\}, \tag{10}$$

where the constraints are used to make the functions in the set $\mathcal{U}_f^1$ subject to the known statistical information, i.e., μ and σ^2, with the integral sum as 1. In other words, $\mathcal{U}_f^1$ represents a set of all possible distribution functions exhibiting the same statistics, and the actual distribution f_{δ_m} might be in any shape belonging to this family.

The other case is to treat the estimates as uncertain ones due to the limited data or measurement noise, and the set can be re-presented as

$$\mathcal{U}_f^2 = \left\{ f_{\delta_m} \geq 0 \;\middle|\; \begin{array}{c} \left(\mathbb{E}_f\left(\delta_m\right) - \mu\right)^2 \leq \theta_\mu \sigma^2, \mathbb{E}_f\left(1\right) = 1 \\ \left|\mathbb{E}_f\left(\left(\delta_m - \mu\right)^2\right) - \sigma^2\right| \leq \theta_\sigma \sigma^2 \end{array} \right\}, \tag{11}$$

where θ_μ and θ_σ are the parameters to quantify the confidence in μ and σ, respectively.

Hence, based on (5) and (9), we can obtain the distributionally robust PLC (DR-PLC) by solving the following problem

$$\hat{c}_{m,ij} = \max\left\{c : \min_{f_{\delta_m}} \left\{\Pr\left(\varphi\left(c, \delta_m\right) \leq 0\right)\right\} \geq \alpha\right\}, \tag{12}$$

subject to the constraint (10) or (11). In this problem, the determination on f_{δ_m} actually corresponds to a density allocation on each possible point δ_m, and the distribution uncertainty constraint (10) and (11) can be regarded as a set of linear constraints with infinite number of decision variables.

B. Approximation Based on Conditional Value at Risk

From the definition of PLC, we observe that it is similar to the risk measurement for a portfolio to keep the loss below certain level with certain probability. This motivates us to adopt a metric for risk measurement in portfolio optimizing problems, namely, conditional-value-at-risk (CVaR), and focus on its worst-case to solve the distributionally robust stochastically constrained problem as in (12).

Before introducing CVaR, we first present a fundamental definition called value-at-risk (VaR).

Definition 1: with respect to certain probability level α, the α-value-at-risk (VaR$_\alpha$) of a portfolio is defined as the lowest amount ξ such that the loss is no more than ξ with at least α probability. Mathematically, let $h\left(\mathbf{x}, \mathbf{y}\right)$ be the loss function, where $\mathbf{x}$ is a decision vector representing a portfolio and $\mathbf{y}$ is a random vector standing for the uncertainties in the market influencing the loss. Then, the VaR$_\alpha$ can be described as follows

$$\mathrm{VaR}_\alpha\left(h\left(\mathbf{x}, \mathbf{y}\right)\right) = \min\left\{\xi : \int\limits_{h(\mathbf{x},\mathbf{y}) \leq \xi} f\left(\mathbf{y}\right) d\mathbf{y} \geq \alpha\right\}, \tag{13}$$

where $f\left(\cdot\right)$ represents the probability distribution function.

Although the VaR defined as (13) is a very popular measurement method for the portfolio risk, it has been proved that it has undesirable mathematical characteristics and is ill-behaved as a function for portfolio optimizations [17]. Thus, CVaR is introduced based on the VaR concept as follows.

Definition 2: CVaR, also known as mean excess loss or tail VaR, is the conditional expectation of the loss above the amount of VaR ξ. Mathematically, based on (13), the CVaR$_\alpha$ can be expressed as

$$\mathrm{CVaR}_\alpha\left(h\left(\mathbf{x}, \mathbf{y}\right)\right) = \frac{1}{1-\alpha} \int\limits_{h(\mathbf{x},\mathbf{y}) \geq \mathrm{VaR}_\alpha(h(\mathbf{x},\mathbf{y}))} h\left(\mathbf{x}, \mathbf{y}\right) f\left(\mathbf{y}\right) d\mathbf{y}. \tag{14}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2946834, IEEE Transactions on Vehicular Technology

To better understand these two definitions, we present a toy example as in Fig. 2, where $\alpha = 0.9$. Obviously, for a loss function $h(\mathbf{x}, \mathbf{y})$, $\text{CVaR}_{\alpha}\left(h(\mathbf{x}, \mathbf{y})\right) \geq \text{VaR}_{\alpha}\left(h(\mathbf{x}, \mathbf{y})\right)$, and minimizing CVaR_{α} is closely related to minimizing VaR_{α} [40].

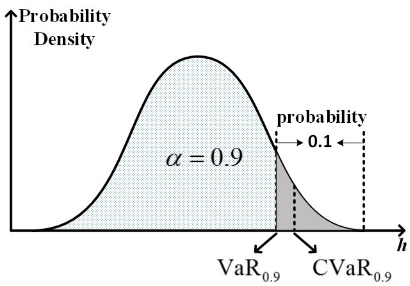


Fig. 2. A toy example for VaR and CVaR.

As a result, according to *Definition 2*, we can obtain

$$\Pr\left(\varphi\left(c, \delta_m\right) \leq \text{CVaR}_{\alpha}\left(\varphi_m\right)\right) \geq \alpha, \tag{15}$$

where φ_m is the simplified notation of $\varphi\left(c, \delta_m\right)$. Therefore, $\text{CVaR}_{\alpha}\left(\varphi_m\right) \leq 0$ is sufficient to imply the original stochastical constraint (6), also described as (7). Then, considering the worst-case, we have

$$\max_{f_{\delta_m} \in \mathcal{U}_f} \left\{\text{CVaR}_{\alpha}\left(\varphi_m\right)\right\} \leq 0 \Rightarrow \min_{f_{\delta_m} \in \mathcal{U}_f} \left\{\Pr\left(\varphi\left(c, \delta_m\right) \leq 0\right)\right\} \geq \alpha. \tag{16}$$

Thus, the DR-PLC as (12) can be approximated by solving the following problem

$$\bar{c}_{m,ij} = \max\left\{c : \max_{f_{\delta_m}} \left\{\text{CVaR}_{\alpha}\left(\varphi_m\right)\right\} \leq 0\right\}, \tag{17}$$

subject to the distribution uncertainty constraint (9) or (10).

Based on the Theorem 1 in [17], $\text{CVaR}_{\alpha}\left(\varphi_m\right)$ can be achieved from the following formula as

$$\begin{aligned}
\text{CVaR}_{\alpha}\left(\varphi_m\right) &= \min_{\beta} F_{\alpha}\left(\beta, c\right) \\
&= \min_{\beta} \left\{\beta + \frac{1}{1-\alpha} \int\limits_{\delta_m \in \mathcal{R}} \left[\varphi\left(c, \delta_m\right) - \beta\right]^{+} f_{\delta_m} d\delta_m\right\} \\
&= \min_{\beta} \left\{\beta + \frac{1}{1-\alpha} \mathbb{E}_f\left(\left[\varphi\left(c, \delta_m\right) - \beta\right]^{+}\right)\right\},
\end{aligned} \tag{18}$$

in which $[x]^{+} = \max\{x, 0\}$. Then, according to (17), we can obtain the worst-case CVaR based approximation for the distributionally robust stochastical constraint (9) as

$$\min_{\beta} \left\{(1-\alpha)\beta + \max_{f_{\delta_m} \in \mathcal{U}_f} \mathbb{E}_f\left(\left[c - \delta_m - \beta\right]^{+}\right)\right\} \leq 0. \tag{19}$$

So, determining the approximated DR-PLC (ADR-PLC) described as (17) is equivalent to finding the maximal value of c satisfying the constraint (19), subject to one of the two distribution uncertainty constraints, re-presented as

$$\begin{aligned}
&\max \;\; c \\
&\text{s.t.} \;\; \{(18), (9)\} \;\text{ or }\; \{(18), (10)\}
\end{aligned} \tag{20}$$

C. Tractable Data-Driven Reformulation Based on Semidefinite Programming

To obtain the ADR-PLC under the two cases with exact and uncertain statistics, constrained by (10) and (11), respectively, in this subsection, we will reformulate the two optimization problems as in (20) into two tractable SDP problems. Since the adopted statistical characteristics, μ and σ^2, are derived from the historical data, we call them tractable data-driven reformulations.

1) Distribution Uncertainty with Exact Statistics: Considering the ADR-PLC solving problem with exact statistical information, i.e.,

$$\bar{c}^{1}_{m,ij} = \max\left\{c : (19) \text{ and } (10)\right\}, \tag{21}$$

we first give the following proposition, which shows the problem (21) can be reformulated as a tractable SDP problem.

Proposition 1: For the following optimization problem as

$$\begin{aligned}
\pi_1 = &\min_{f_{\delta_m} \geq 0} \mathbb{E}_f\left(-\left[c - \delta_m - \beta\right]^{+}\right) \\
\text{s.t. } &\mathbb{E}_f\left(\delta_m\right) = \mu, \mathbb{E}_f\left(\left(\delta_m - \mu\right)^2\right) = \sigma^2, \mathbb{E}_f\left(1\right) = 1.
\end{aligned} \tag{22}$$

it is equivalent to solving the SDP problem below

$$\begin{aligned}
\pi_1 = &\max_{\lambda, \eta, k} -\lambda\mu - k - \left(\sigma^2 + \mu^2\right)\eta \\
\text{s.t. } &\begin{bmatrix} \eta & \frac{\lambda}{2} \\ \frac{\lambda}{2} & k \end{bmatrix} \succeq 0, \begin{bmatrix} \eta & \frac{\lambda+1}{2} \\ \frac{\lambda+1}{2} & k - c + \beta \end{bmatrix} \succeq 0,
\end{aligned} \tag{23}$$

in which λ, η, and k are Lagrangian multipliers associated with the three statistic constraints in (22).

Proof: By introducing three Lagrangian multipliers, λ, η and k, we can get the Lagrangian function as in (24) on the next page, and the dual problem of (22) can be expressed as

$$\max_{\lambda, \eta, k} \left(\min_{f_{\delta_m} \geq 0} \Gamma\left(f_{\delta_m}, \lambda, \eta, k\right)\right). \tag{25}$$

The feasibility and convexity of the problem (22) can ensure the strong duality [16], [41]. Considering the expectation component in (24), if it is negative, since there is no constraint on the shape of the distribution function $f_{\delta_m} \geq 0$, it is possible that f_{δ_m} allocates very large mass on a single point so that $\min\limits_{f_{\delta_m} \geq 0} \Gamma\left(f_{\delta_m}, \lambda, \eta, k\right)$ is unsolvable. Thus, the following condition needs to be satisfied for any possible δ_m

$$\left[c - \delta_m - \beta\right]^{+} \leq \eta {\delta_m}^2 + \lambda\delta_m + k. \tag{26}$$

Then, the dual problem as (25) turns to be the following constrained optimization problem

$$\begin{aligned}
\pi_1 = &\max_{\lambda, \eta, k} -\lambda\mu - k - \left(\sigma^2 + \mu^2\right)\eta \\
\text{s.t. } &\eta {\delta_m}^2 + \lambda\delta_m + k \geq 0, \quad \forall \delta_m \in \mathcal{R}, \\
&\eta {\delta_m}^2 + \lambda\delta_m + k \geq c - \delta_m - \beta, \quad \forall \delta_m \in \mathcal{R}.
\end{aligned} \tag{27}$$

For the first constraint in (27), it can be rewritten as

$$\begin{bmatrix} \delta_m & 1 \end{bmatrix} \begin{bmatrix} \eta & \frac{\lambda}{2} \\ \frac{\lambda}{2} & k \end{bmatrix} \begin{bmatrix} \delta_m & 1 \end{bmatrix}^{\mathrm{T}} \geq 0, \quad \forall \delta_m \in \mathcal{R}, \tag{28}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2946834, IEEE Transactions on Vehicular Technology

$$\Gamma\left(f_{\delta_m}, \lambda, \eta, k\right) = \mathbb{E}_f\left(-\left[c-\delta_m-\beta\right]^+ + \eta{\delta_m}^2 + \lambda\delta_m + k\right) - \lambda\mu - k - \left(\sigma^2+\mu^2\right)\eta. \tag{24}$$

$$\begin{aligned}\tilde{\Gamma}\left(f_{\delta_m}, \nu, \mathbf{Z}, \varsigma\right) &= \mathbb{E}_f\left(-[c-\delta_m-\beta]^+\right) + \nu\left(\mathbb{E}_f\left({\delta_m}^2\right) - 2\mu\mathbb{E}_f\left(\delta_m\right) + \mu^2 - (1+\theta_\sigma)\sigma^2\right) - \mathbb{E}_f\begin{bmatrix}\sigma^2 & \delta_m-\mu\\ \delta_m-\mu & \theta_\mu\end{bmatrix}\otimes\mathbf{Z} + \mathbb{E}_f\left(\varsigma\right) - \varsigma\\ &= \mathbb{E}_f\left(-[c-\delta_m-\beta]^+ + \nu{\delta_m}^2 - 2\left(\mu\nu+\omega\right)\delta_m + \varsigma\right) + \left(\mu^2 - (1+\theta_\sigma)\sigma^2\right)\nu - \xi\sigma^2 + 2\mu\omega - s\theta_\mu - \varsigma, \end{aligned}\tag{35}$$

which is equivalent to the first constraint in (23). Similarly, for the second constraint in (27), it can be reformulated as

$$\begin{bmatrix}\delta_m & 1\end{bmatrix}\begin{bmatrix}\eta & \frac{\lambda+1}{2}\\ \frac{\lambda+1}{2} & k-c+\beta\end{bmatrix}\begin{bmatrix}\delta_m & 1\end{bmatrix}^{\mathrm{T}} \geq 0, \forall\delta_m \in \mathcal{R}, \tag{29}$$

corresponding to the second constraint in (23). Then, we can achieve the SDP problem as (23) which is equivalent to the optimization problem as (22). ■

Since $\max\limits_{f_{\delta_m}\in\mathcal{U}_f^1} \mathbb{E}_f\left(\left[c-\delta_m-\beta\right]^+\right) = -\pi_1$, based on *Proposition 1*, we find that the constraint (19) with the exact statistical information is equivalent to a feasibility check:

$$\begin{aligned}&\exists\left(\beta, \lambda, \eta, k\right)\\ &\text{s.t. } \left(1-\alpha\right)\beta + \lambda\mu + k + \left(\sigma^2+\mu^2\right)\eta \leq 0,\\ &\begin{bmatrix}\eta & \frac{\lambda}{2}\\ \frac{\lambda}{2} & k\end{bmatrix}\succeq 0, \begin{bmatrix}\eta & \frac{\lambda+1}{2}\\ \frac{\lambda+1}{2} & k-c+\beta\end{bmatrix}\succeq 0.\end{aligned}\tag{30}$$

Consequently, the ADR-PLC solving problem (21) can be reformulated into a tractable SDP problem as

$$\bar{c}_{m,ij}^1 = \max_{\beta,\lambda,\eta,k} c \ , \tag{31}$$

subject to the constraints as in (30).

2) Distribution Uncertainty with Uncertain Statistics: As a more general case, the statistical information extracted from the historical data is usually with errors due to the limited observations, and the ADR-PLC solving problem in this case can be formulated as

$$\bar{c}_{m,ij}^2 = \max\left\{c : (19) \text{ and } (11)\right\}. \tag{32}$$

Similarly, we first present a proposition, and then reformulate the problem (32) into an SDP problem accordingly.

Proposition 2: For the following optimization problem as

$$\begin{aligned}\pi_2 = &\min_{f_{\delta_m}\geq 0} \mathbb{E}_f\left(-[c-\delta_m-\beta]^+\right)\\ &\text{s.t. } \left(\mathbb{E}_f(\delta_m)-\mu\right)^2 \leq \theta_\mu\sigma^2,\\ &\mathbb{E}_f\left(\left(\delta_m-\mu\right)^2\right) \leq (1+\theta_\sigma)\sigma^2, \mathbb{E}_f\left(1\right) = 1,\end{aligned}\tag{33}$$

it is equivalent to solving the SDP problem below

$$\begin{aligned}\pi_2 = &\max_{\nu,\xi,\omega,s,\varsigma} \left(\mu^2 - (1+\theta_\sigma)\sigma^2\right)\nu - \xi\sigma^2 + 2\mu\omega - s\theta_\mu - \varsigma\\ \text{s.t. } &\begin{bmatrix}\nu & -\left(\mu\nu+\omega\right)\\ -\left(\mu\nu+\omega\right) & \varsigma\end{bmatrix}\succeq 0,\\ &\begin{bmatrix}\nu & \frac{-2(\mu\nu+\omega)+1}{2}\\ \frac{-2(\mu\nu+\omega)+1}{2} & \varsigma-c+\beta\end{bmatrix}\succeq 0,\\ &\nu\geq 0, \mathbf{Z}\succeq 0,\end{aligned}\tag{34}$$

where $\mathbf{Z} = \begin{bmatrix}\xi & \omega\\ \omega & s\end{bmatrix}$, ν, and ς are the Lagrangian multipliers associated with the three statistical constraints in (33), respectively.

Proof: Considering the first constraint in (33), we can rewrite it as

$$\mathbb{E}_f\begin{bmatrix}\sigma^2 & \delta_m-\mu\\ \delta_m-\mu & \theta_\mu\end{bmatrix}\succeq 0. \tag{35}$$

Similar to (24), by introducing three Lagrangian multipliers, $\mathbf{Z}\succeq 0$, $\nu\geq 0$, and ς, we can obtain the Lagrangian function shown in (35) where $\otimes$ denotes the Frobenius product. Then, the dual problem of (33) can be expressed as

$$\max_{\nu,\mathbf{Z},\varsigma}\left(\min_{f_{\delta_m}\geq 0}\tilde{\Gamma}\left(f_{\delta_m},\nu,\mathbf{Z},\varsigma\right)\right). \tag{36}$$

Similar to (25), to make $\min\limits_{f_{\delta_m}\geq 0}\tilde{\Gamma}\left(f_{\delta_m},\nu,\mathbf{Z},\varsigma\right)$ solvable, we have

$$[c-\delta_m-\beta]^+ \leq \nu{\delta_m}^2 - 2\left(\mu\nu+\omega\right)\delta_m + \varsigma, \forall\delta_m\in\mathcal{R}, \tag{37}$$

and the dual problem as (36) can be rewritten as

$$\begin{aligned}\pi_2 = &\max_{\nu,\xi,\omega,s,\varsigma} \left(\mu^2 - (1+\theta_\sigma)\sigma^2\right)\nu - \xi\sigma^2 + 2\mu\omega - s\theta_\mu - \varsigma\\ &\text{s.t. } \nu{\delta_m}^2 - 2\left(\mu\nu+\omega\right)\delta_m + \varsigma \geq 0 \quad \forall\delta_m\in\mathcal{R},\\ &\nu{\delta_m}^2 - 2\left(\mu\nu+\omega\right)\delta_m + \varsigma \geq c-\delta_m-\beta \quad \forall\delta_m\in\mathcal{R},\\ &\nu\geq 0, \begin{bmatrix}\xi & \omega\\ \omega & s\end{bmatrix}\succeq 0.\end{aligned}\tag{38}$$

Similar to (28) and (29), the first two constraints can be transformed into two linear matrix inequalities as in (34) and thus the conclusion can be achieved. ■

For the ADR-PLC solving problem in (32), the worst-case distribution is expected to have a larger variance, so that the distribution uncertainty constraint set (11) turns to be

$$\mathcal{U}_f^2 = \left\{f_{\delta_m}\geq 0 \left| \begin{array}{c}\left(\mathbb{E}_f\left(\delta_m\right)-\mu\right)^2 \leq \theta_\mu\sigma^2, \mathbb{E}_f\left(1\right) = 1\\ \mathbb{E}_f\left(\left(\delta_m-\mu\right)^2\right) \leq (1+\theta_\sigma)\sigma^2\end{array}\right.\right\}. \tag{39}$$

Similar to (30), based on *Proposition 2*, the constraint (19) in this case is equivalent to the following feasibility check:

$$\begin{aligned}&\exists\left(\beta,\nu,\xi,\omega,s,\varsigma\right)\\ &\text{s.t. } (1-\alpha)\beta + \left((1+\theta_\sigma)\sigma^2 - \mu^2\right)\nu + \xi\sigma^2 - 2\mu\omega + s\theta_\mu + \varsigma \leq 0\\ &\begin{bmatrix}\nu & -\left(\mu\nu+\omega\right)\\ -\left(\mu\nu+\omega\right) & \varsigma\end{bmatrix}\succeq 0,\\ &\begin{bmatrix}\nu & \frac{-2(\mu\nu+\omega)+1}{2}\\ \frac{-2(\mu\nu+\omega)+1}{2} & \varsigma-c+\beta\end{bmatrix}\succeq 0,\\ &\nu\geq 0, \begin{bmatrix}\xi & \omega\\ \omega & s\end{bmatrix}\succeq 0.\end{aligned}\tag{40}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2946834, IEEE Transactions on Vehicular Technology

Therefore, the ADR-PLC solving problem in this case written as (32) can be reformulated as a tractable SDP problem with the objective

$$\bar{c}_{m,ij}^{2} = \max_{\beta,\nu,\xi,\omega,s,\varsigma} c \ , \tag{41}$$

subject to the constraints as in (40).

D. Accuracy of the Worst-Case CVaR Based Approximation

To calculate the DR-PLC formulated as (12), we have made a conservative approximation for the stochastical constraint based on the worst-case CVaR and turned it to be an ADR-PLC formulated as (17). In this subsection, we will demonstrate that such an approximation is actually accurate, i.e., DR-PLC and ADR-PLC are equivalent. That is, we will prove that (16) is in fact an equivalence, i.e.,

$$\max_{f_{\delta_m}\in\mathcal{U}_f}\{\mathrm{CVaR}_\alpha(\varphi_m)\}\le 0 \Leftrightarrow \min_{f_{\delta_m}\in\mathcal{U}_f}\{\Pr(\varphi(c,\delta_m)\le 0)\}\ge\alpha. \tag{42}$$

Due to the limited space, we only consider the first case where the statistical information is treated as an exact one, i.e., $\mathcal{U}_f = \mathcal{U}_f^1$, which can be easily extended to the other case.

Based on the definition of VaR in (13), we can express the worst-case VaR of φ_m as

$$\text{WC-VaR}_\alpha(\varphi_m) = \min\left\{\rho : \min_{f_{\delta_m}\in\mathcal{U}_f^1}\{\Pr(\varphi_m \le \rho)\} \ge \alpha\right\}. \tag{43}$$

To prove the equivalence (42), we will first show that

$$\min_{f_{\delta_m}\in\mathcal{U}_f^1}\{\Pr(\varphi_m \le 0)\} \ge \alpha \Leftrightarrow \text{WC-VaR}_\alpha(\varphi_m) \le 0, \tag{44}$$

and then demonstrate that

$$\text{WC-VaR}_\alpha(\varphi_m) = \max_{f_{\delta_m}\in\mathcal{U}_f^1}\{\mathrm{CVaR}_\alpha(\varphi_m)\}. \tag{45}$$

For the equivalence (44), if the left side is satisfied, obviously, $\rho = 0$ is feasible in (43), which means that the right side of (44) can be satisfied as well. Conversely, if the right side is met, given that $\text{WC-VaR}_\alpha(\varphi_m) = \hat{\rho} \le 0$, then we can derive that for any fixed $f_\delta \in \mathcal{U}_f^1$,

$$\Pr(\varphi_m \le 0) \ge \Pr(\varphi_m \le \hat{\rho}) \ge \alpha, \tag{46}$$

which means the left side is also met. Therefore, the equivalence (44) follows.

Next, we will show that the equivalence (45) holds so that (42) can be proved to be true. Firstly, We present two Lemmas as follows.

Lemma 1 (Farkas Lemma) Let $f, g_1, \ldots, g_m$ be convex functions. Assume that there exists a feasible point $\bar{x}$ satisfying $g_i(\bar{x}) < 0,\ i = 1, \ldots, m$. Then, $f(x) \ge 0$ for all x with $g_i(x) \le 0,\ i = 1, \ldots, m$, if and only if there are $y_1, \ldots, y_m \ge 0$ such that

$$f(x) + \sum_{i=1}^{m} y_i g_i(x) \ge 0, \forall x \in \mathcal{R}. \tag{47}$$

Proof: Refer to Theorem 2.1 in [42]. ■

Lemma 2 Denote $\pi = \max_{f_\delta\in\mathcal{U}_f^1}\{\Pr(\varphi_m \ge \rho)\}$. Then,

$$\begin{aligned} \pi = \min_{\lambda,\eta,k}\ & \lambda\mu + k + (\sigma^2 + \mu^2)\eta \\ \text{s.t.}\ & \eta{\delta_m}^2 + \lambda\delta_m + k \ge 0, \quad \forall \delta_m \in \mathcal{R}, \\ & \eta{\delta_m}^2 + \lambda\delta_m + k \ge 1, \quad \forall \delta_m : \varphi_m \ge \rho. \end{aligned} \tag{48}$$

Proof: Define an indicator function as

$$I(\delta_m) = \begin{cases} 1 & \varphi_m \ge \rho \\ 0 & \varphi_m < \rho \end{cases}. \tag{49}$$

We can find that $\pi = \max_{f_\delta\in\mathcal{U}_f^1}\mathbb{E}(I(\delta_m))$. Then, by applying the similar manipulations as in the proof of *Proposition 1*, we can achieve the conclusion. ■

Rewrite (43) as

$$\text{WC-VaR}_\alpha(\varphi_m) = \min\left\{\rho : \max_{f_\delta\in\mathcal{U}_f^1}\{\Pr(\varphi_m \ge \rho)\} \le \theta\right\}, \tag{50}$$

where $\theta = 1 - \alpha$. According to *Lemma 1* and *Lemma 2*, we can find that (50) can be reformulated as

$$\begin{aligned} \text{WC-VaR}_\alpha(\varphi_m) = \min\ & \rho \\ \text{s.t.}\ & \lambda\mu + k + (\sigma^2 + \mu^2)\eta \le \theta, \\ & \eta{\delta_m}^2 + \lambda\delta_m + k \ge 0, \quad \forall \delta_m \in \mathcal{R}, \\ & \eta{\delta_m}^2 + \lambda\delta_m + k - 1 + y(\rho - \varphi_m) \ge 0, \ \forall \delta_m \in \mathcal{R}, y \ge 0. \end{aligned} \tag{51}$$

If $y = 0$, then $\eta{\delta_m}^2 + \lambda\delta_m + k \ge 1$. Let $\delta_m = \mu$ and we can find that $\lambda\mu + k + (\sigma^2 + \mu^2)\eta \ge 1$ which is in conflict with the first constraint in (51). Thus, we have $y \ne 0$. Dividing both sides of the third constraint by y and performing variable substitutions, we can rewrite (51) as

$$\begin{aligned} \text{WC-VaR}_\alpha(\varphi_m) = \min\ & \tilde{\rho} + \tilde{y} \\ \text{s.t.}\ & \tfrac{1}{\theta}\left(\tilde{\lambda}\mu + \tilde{k} + (\sigma^2 + \mu^2)\tilde{\eta}\right) \le \tilde{y}, \\ & \tilde{\eta}{\delta_m}^2 + \tilde{\lambda}\delta_m + \tilde{k} \ge 0, \quad \forall \delta_m \in \mathcal{R}, \\ & \tilde{\eta}{\delta_m}^2 + \tilde{\lambda}\delta_m + \tilde{k} + \tilde{\rho} - \varphi_m \ge 0, \ \forall \delta_m \in \mathcal{R}, \ \tilde{y} > 0, \end{aligned} \tag{52}$$

where $\tilde{y} = \frac{1}{y}$, $\tilde{\eta} = \frac{\eta}{y}$, $\tilde{k} = \frac{k}{y}$, $\tilde{\lambda} = \frac{\lambda}{y}$, and $\tilde{\rho} = \rho - \tilde{y}$. Substituting $\delta_m = \mu$ into the second constraint, we can find that $\tilde{\lambda}\mu + \tilde{k} + (\sigma^2 + \mu^2)\tilde{\eta} \ge 0$, which means that the constraint $\tilde{y} > 0$ is redundant and can be removed. Furthermore, for the first constraint, to achieve the optimality, it should satisfy that $\tilde{y} = \frac{1}{\theta}\left(\tilde{\lambda}\mu + \tilde{k} + (\sigma^2 + \mu^2)\tilde{\eta}\right)$. Consequently, (52) can be further rewritten as

$$\begin{aligned} \text{WC-VaR}_\alpha(\varphi_m) = \min\ & \tilde{\rho} + \tfrac{1}{\theta}\left(\tilde{\lambda}\mu + \tilde{k} + (\sigma^2 + \mu^2)\tilde{\eta}\right) \\ \text{s.t.}\ & \tilde{\eta}{\delta_m}^2 + \tilde{\lambda}\delta_m + \tilde{k} \ge 0, \quad \forall \delta_m \in \mathcal{R}, \\ & \tilde{\eta}{\delta_m}^2 + \tilde{\lambda}\delta_m + \tilde{k} + \tilde{\rho} - \varphi_m \ge 0, \ \forall \delta_m \in \mathcal{R}. \end{aligned} \tag{53}$$

Based on (18) and *Proposition 1*, we can achieve that

$$\begin{aligned} \max_{f_{\delta_m}\in\mathcal{U}_f^1}\{\mathrm{CVaR}_\alpha(\varphi_m)\} = \min\ & \beta + \tfrac{1}{\theta}\left(\lambda\mu + k + (\sigma^2 + \mu^2)\eta\right) \\ \text{s.t.}\ & \eta{\delta_m}^2 + \lambda\delta_m + k \ge 0, \quad \forall \delta_m \in \mathcal{R}, \\ & \eta{\delta_m}^2 + \lambda\delta_m + k \ge c - \delta_m - \beta, \quad \forall \delta_m \in \mathcal{R}. \end{aligned} \tag{54}$$

Obviously, (53) and (54) are equivalent. Hence, we can obtain the conclusion (45).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2946834, IEEE Transactions on Vehicular Technology

So far, we have demonstrated that both (44) and (45) are satisfied. As a result, we achieve the equivalence (42), so that the worst-case CVaR based approximation for DR-PLC is actually accurate, i.e., $\bar{c}_{m,ij} = \hat{c}_{m,ij}$.

V. Service-Based Spectrum-Aware Data Transmission Scheme in the CCHN

In this section, by adopting the proposed metric, i.e., DR-PLC, to evaluate the achievable rate of a link on a shared band with uncertain availability, we will develop a service-based spectrum-aware (S^2) transmission scheme in the CCHN to meet as many service requests as possible by using the least licensed bands. To be specific, we consider two kinds of services in the network, i.e., DS services $\mathcal{L}_{\text{DS}}$ and DT services $\mathcal{L}_{\text{DT}}$. Only the DT ones can be carried by the shared bands $\mathcal{M}_{\text{s}}$ with uncertain availability. Their QoS can be statistically guaranteed by using the proposed metric DR-PLC when making resource allocation decisions.

For such an S^2 transmission scheme, it has two steps. Step-1 corresponds to an admission control, aiming at maximizing the number of admitted services with the consideration of their different importance levels and rate requirements, subject to the available resources. Step-2 focuses on the spectrum efficiency with the objective of minimizing the used bandwidth of the licensed band, i.e., offloading the transmission of DT services to the shared bands which are relatively sufficient compared with the licensed ones. For the transmission scheduling, we only consider the links transmitting the aggregated data, i.e., CRR $\leftrightarrow$ FBS and CRR $\leftrightarrow$ CRR, and ignore the accessing link, Device $\leftrightarrow$ CRR, which can work on various accessing technologies and be managed by the CRRs in a distributed way instead of being controlled by the MBS in a centralized manner. The end-to-end hybrid transmission scheduling will be investigated in our future works.

A. Step-1: Admission Control

We exploit a binary variable $k(l)$ to describe whether or not the service request $l \in \{\mathcal{L}_{\text{DT}} \cup \mathcal{L}_{\text{DS}}\}$ can be admitted as

$$k(l) = \begin{cases} 1, & \text{if service request } l \text{ is admitted,} \\ 0, & \text{otherwise.} \end{cases} \tag{55}$$

Denote $w(l)$ as the weight attributed to the service request l corresponding to its importance level. Then, the objective of the step-1 can be expressed as

$$\max \sum_{l \in \{\mathcal{L}_{\text{DT}} \cup \mathcal{L}_{\text{DS}}\}} w(l)\, k(l). \tag{56}$$

Next, we will formulate the constraints for the spectrum allocation, satisfying the admitted services' requirements. We employ a binary variable x_{ij}^m to denote whether or not band $m \in \mathcal{M}^{ij}$ is allocated on the link from i to j, $i \neq j \in \mathcal{N}$, as

$$x_{ij}^m = \begin{cases} 1, & \text{if band } m \text{ is allocated on the link } i \rightarrow j, \\ 0, & \text{otherwise.} \end{cases} \tag{57}$$

First, we need to avoid the mutual interference among different links. To be specific, for any node $i \in \mathcal{N}$, it cannot transmit to or receive from other nodes on the same band, which can be described as

$$\sum_{j \in \mathcal{T}_{i,m}} x_{ij}^m \leq 1, \sum_{\{h | i \in \mathcal{T}_{h,m}\}} x_{hi}^m \leq 1, \forall i \in \mathcal{N}. \tag{58}$$

Furthermore, any node cannot transmit and receive on the same band simultaneously due to the self-interference, which can be expressed as

$$x_{ij}^m + \sum_{q \in \mathcal{T}_{j,m}} x_{jq}^m \leq 1, \forall i \in \mathcal{N}, \forall j \in \mathcal{T}_{i,m}. \tag{59}$$

Moreover, if $x_{ij}^m = 1$, all interfering neighbors of node i cannot receive data on band m, and we can obtain the constraint as

$$x_{ij}^m + \sum_{\{k | g \in \mathcal{T}_{k,m}, k \neq i\}} x_{kg}^m \leq 1, \forall i \in \mathcal{N}, \forall g \in \mathcal{I}_{i,m}. \tag{60}$$

Second, considering the rate requirements and different types of the admitted services, we need to guarantee: 1) the admitted DS services' data should be carried by the licensed bands; 2) the achievable rate of the link should be able to support the data traffic generated by the admitted services. To quantify the achievable rate of a link on a shared band with uncertain availability, we use DR-PLC to make the requirement statistically satisfied with a high probability. Then, we can get the following two constraints as

$$\sum_{\{l \in \mathcal{L}_{\text{DS}} | s_{\text{DS}}(l) = i, d_{\text{DS}}(l) = j\}} r_{\text{DS}}(l)\, k(l) \leq \sum_{m \in \mathcal{M}_l} x_{ij}^m c_{m,ij}, \tag{61}$$

and

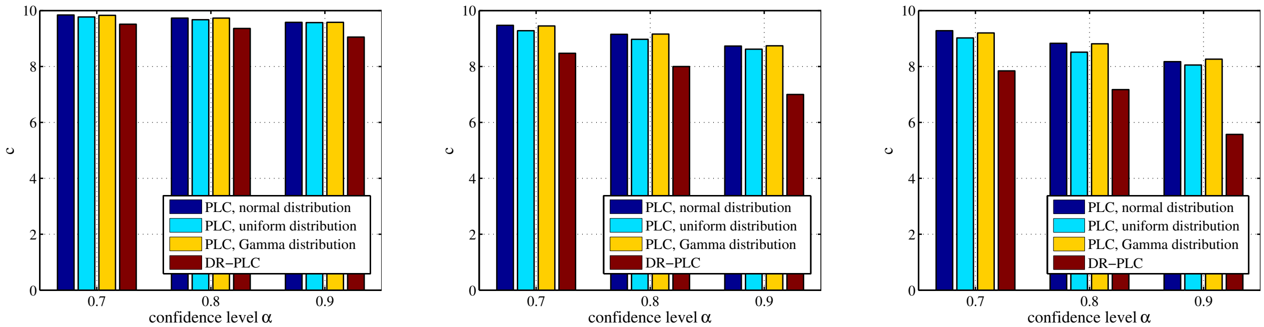
$$\begin{aligned} &\sum_{\{l \in \mathcal{L}_{\text{DS}} | s_{\text{DS}}(l) = i, d_{\text{DS}}(l) = j\}} r_{\text{DS}}(l)\, k(l) + \sum_{\{l \in \mathcal{L}_{\text{DT}} | s_{\text{DT}}(l) = i, d_{\text{DT}}(l) = j\}} r_{\text{DT}}(l)\, k(l) \\ &\qquad \leq \sum_{m \in \mathcal{M}_l} x_{ij}^m c_{m,ij} + \sum_{m \in \mathcal{M}_s^i} x_{ij}^m \hat{c}_{m,ij}, \end{aligned} \tag{62}$$

where $c_{m,ij}$ and $\hat{c}_{m,ij}$ are the capacity and the DR-PLC of the link from i to j on band $m \in \mathcal{M}_l$ and $m \in \mathcal{M}_s^i$, calculated by (3) and (12), respectively.

By solving the integer linear programming (ILP) problem with the objective as (56), subject to the constraints from (58) to (62), we can obtain the optimal admission control result $k^*(l)$ for any service request $l \in \{\mathcal{L}_{\text{DT}} \cup \mathcal{L}_{\text{DS}}\}$, which can support as many service requests as possible based on their importance levels and the available resources in the network.

B. Step-2: Spectrum Re-allocation

Although the step-1 has optimally determined which service requests are admitted by the network, it is noteworthy that the spectrum allocation result, i.e., the x_{ij}^m obtained by solving the aforementioned optimization problem, may not be the best decision, leading to a low spectrum efficiency. On the one hand, some links with low rate requirements may be assigned with a wide bandwidth. On the other hand, some DT services may be carried by the scare licensed bands, leaving the relative sufficient shared bands underutilized. Hence, to improve the spectrum efficiency, we introduce the step-2 for S^2 transmission scheme to re-allocate bands on the links which will undertake the data transmissions for the admitted services.



(a) $\sigma^2 = 0.1$ (b) $\sigma^2 = 1$ (c) $\sigma^2 = 2$

Fig. 3. PLC and DR-PLC under different probability distributions, where exact statistical information is considered.

The objective of the step-2 is minimizing the total occupied bandwidth of the licensed bands. By using the same notations as those in step-1, we can formulate the objective as

$$\min \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{T}_{i,m}} \sum_{m\in\mathcal{M}_l} x_{ij}^m W_m. \tag{63}$$

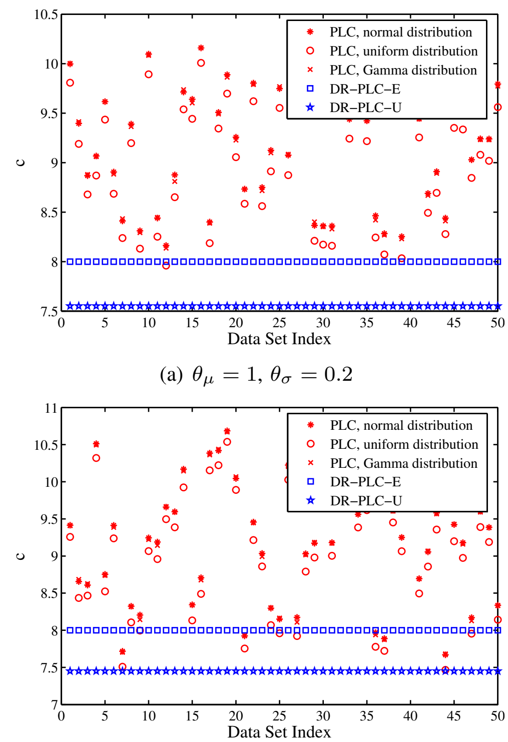
Similar to step-1, the constraints for step-2 also involve two aspects. One is to avoid the co-channel interference, which can be described by (58), (59) and (60) as well. The other is to guarantee the QoS for the admitted services that scheduled in step-1, which can be formulated by substituting $k\left(l\right)$ with $k^*\left(l\right)$ in (61) and (62). Then, by solving such an ILP problem, we can get the optimal spectrum allocation solution.

VI. Numerical Results

A. Evaluation for the Distributionally Robust Data-Driven Approach With Exact or Uncertain Statistics

In this subsection, by directly assuming that δ_m follows certain probability distribution, we evaluate the performance of the proposed DR approach through comparing the PLC as (5) and the DR-PLC as (31) or (41) under exact or uncertain data-driven statistics, respectively[2]. Three common distributions are employed: normal distribution, uniform distribution, and Gamma distribution. First, we consider the case that the statistical information is accurate, i.e., the DR-PLC is obtained based on (31), and present the DR-PLC and the PLC under the three distributions with different confidence levels in Fig. 3. To be specific, we set the same statistics for each distribution in each experiment, where the expectation is 10 and the variance is 0.1, 1, and 2 in (a), (b), and (c), respectively. Three confidence levels are considered, i.e., $\alpha = 0.7$, $\alpha = 0.8$, and $\alpha = 0.9$. From Fig. 3, we observe that the DR-PLC is lower than the PLC under any distribution because it is in fact a lower bound of PLC, robust to any possible distributions with the same statistics. When the variance is small, the difference between DR-PLC and PLC is small. With the variance increases, the gap becomes more obvious, as showing in Fig. 3(a) and Fig. 3(c), because it is more difficult to achieve the distributional robustness when the random variable fluctuates more dramatically with a larger variance.

[2]As conducted in [6], [35], Monte Carlo experiment as an effective method is adopted to calculate the PLC under different distributions.



(a) $\theta_\mu = 1,\ \theta_\sigma = 0.2$

(b) $\theta_\mu = 2,\ \theta_\sigma = 0.3$

Fig. 4. PLC and DR-PLC under different probability distributions, where uncertain statistical information is considered.

Furthermore, for the different confidence levels, the higher it is, the lower PLC and DR-PLC can be achieved, because it has to make the result achievable with a higher probability. Moreover, the gap between DR-PLC and PLC increases when the adopted confidence level increases.

Next, we consider the case that the statistical information of δ_m is uncertain, i.e., the real expectation and variance are not equal to the one extracted from the sampled data, and compare PLC and DR-PLC in Fig. 4. In each subfigure, we have conducted 50 experiments. In each experiment, the confidence level is set to 0.8, and the real statistics are generated randomly within an uncertainty interval as shown in (11), where $\mu = 10$ and $\sigma^2 = 1$ with uncertainty parameters as $\{\theta_\mu = 1,\ \theta_\sigma = 0.2\}$ and $\{\theta_\mu = 2,\ \theta_\sigma = 0.3\}$ in Fig. 4(a) and Fig. 4(b), respectively. The DR-PLC-E represents the DR-PLC obtained based on (31) by treating the sampled statistics exact,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2946834, IEEE Transactions on Vehicular Technology

and the DR-PLC-U is achieved based on (41) considering the estimation errors. From Fig. 4, we observe that if we directly use the inaccurate data-driven statistics to calculate DR-PLC, i.e., the DR-PLC-E, it will result in an overestimated solution in some cases due to the errors in the statistics, especially when the estimation errors are high as shown in Fig. 4(b) where many results are larger than the corresponding PLC. Such an overestimation on the achievable performance might lead to the violation of the QoS. By taking the estimation errors into account, DR-PLC-U can be regarded as the lower bound of PLC, robust to any possible distributions subject to the same uncertain statistics[3].

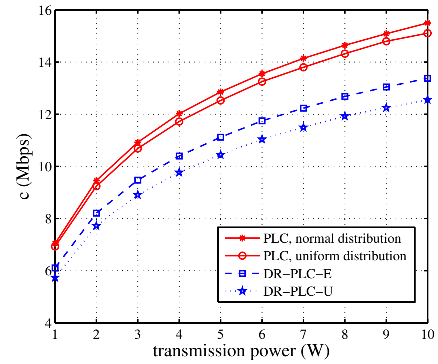
B. Case Study 1: PLC and DR-PLC of a Link on a Shared Band With Uncertain Availability

As a case study, we consider a specific transmission link with the distance as 200m. The transmission related parameters are set as follows: $\tau = 4$ and $\alpha = 4$. The noise density power at the receiver is set to $\gamma = 10^{-16}$ W/Hz. Assume that the link is using a shared band m with bandwidth $W_m = 5$MHz, whose actual available bandwidth $\tilde{W}_m^l$ follows a uniform or truncated normal distribution within $[0, 5]$, two widely adopted theoretical models in the literature [35], [43]. Fig. 5 shows PLC and DR-PLC of this link under 0.8 confidence level with different transmission power, which is assumed to be allocated uniformly on this band. The expectation of $\tilde{W}_m^l$ is set to $\mu_m = 3$, and its variance is set to $\sigma_m^2 = 0.1$ and $\sigma_m^2 = 0.6$ in Fig. 5(a) and Fig. 5(b), respectively, in terms of MHz. The DR-PLC-E and DR-PLC-U have the same meaning as those in Fig. 4, and for the DR-PLC-U, the parameters for the statistics of δ_m are set to $\theta_\mu = 4$, $\theta_\sigma = 0.2$. From Fig. 5, we can get some similar conclusions as what we have observed from Fig. 4. As the lower bound, both DR-PLC-E and DR-PLC-U are smaller than the corresponding PLC under both distributions. Compared with DR-PLC-E, DR-PLC-U corresponds to a more conservative evaluation because the estimation errors of statistics are considered. Furthermore, under the two different variances, we can see that when the available bandwidth has a larger variance, i.e., the historical data fluctuates more seriously, both PLC and DR-PLC will become smaller to achieve the same confidence level, and the gap between PLC and DR-PLC becomes larger.

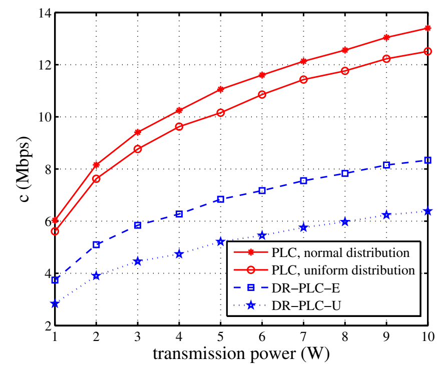
C. Case Study 2: Service-Based Spectrum-Aware Transmission Scheme

We consider a grid network with $N = 9$ nodes as shown in Fig. 6. The transmission power of each node on each band is assumed to be 2W, which is allocated on the band uniformly, and the transmission and interference range is 210m and 260m, respectively. The transmission related parameters are set the same as those in Fig. 5. Suppose that there are three DS requests and five DT requests as exhibited in Fig. 6. The weights of the three DS ones are set the same

[3]Noticing that we do not aim at comparing the PLC under the different distributions, but try to show the distributional robustness of the proposed DR-PLC. Since we cannot present all possible distributions, we only take these three common ones as examples.



(a) $\sigma_m^2 = 0.1$



(b) $\sigma_m^2 = 0.6$

Fig. 5. PLC and DR-PLC of a link on the shared band m, whose available bandwidth follows normal or uniform distribution.

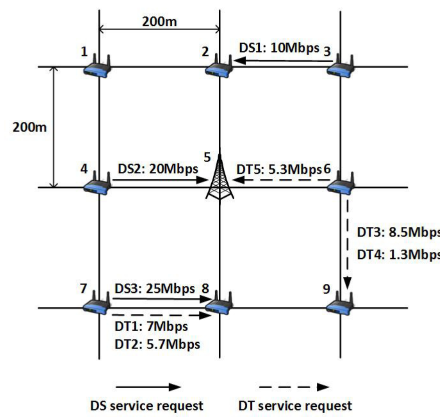


Fig. 6. Case study: a grid network with 9 nodes, 3 DS service requests and 5 DT service requests.

as 6, and those of the five DT ones are set to 5, 4, 3, 1, 2, respectively[4]. Assume that there are two licensed bands, $\mathcal{M}_l = \{1, 2\}$, and three shared bands, $\mathcal{M}_s = \{3, 4, 5\}$, and all nodes can use any of them. The bandwidth of the two licensed ones are $W_1 = 3$MHz and $W_2 = 4$MHz, and that of the three shared ones are $W_3 = 3$MHz, $W_4 = 4$MHz and $W_5 = 5$MHz, respectively. Due to the spectrum uncertainty, the actual available bandwidth of each shared band is a random variable as $\tilde{W}_m \le W_m$, $m = 3, 4, 5$. Suppose that they follow the uniform distribution with the expectation as 1.5, 2.5, 3.5,

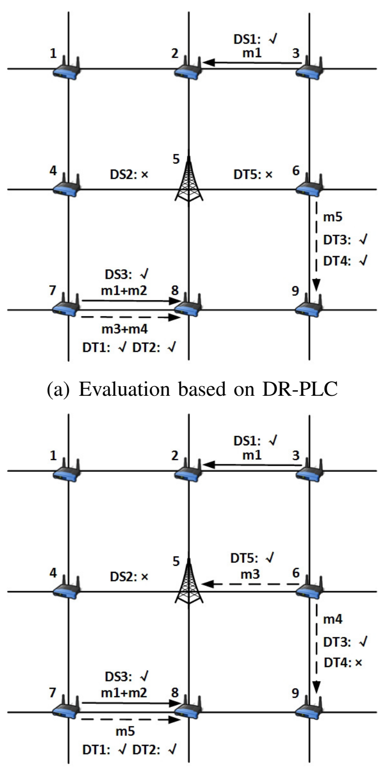
[4]The weight of a service request corresponds to its importance level, which depends on the payment or the application.

respectively, and the same variance as 0.2. Since the distance and the transmission related parameters of each transmission link have been set the same, we can calculate the link capacity (LC) and the DR-PLC of a link on a band as shown in Table I, where the confidence level is set to 0.7 and the statistics used to determine the DR-PLC are considered exact. The

TABLE I
DIFFERENT KINDS OF CAPACITY OF A LINK ON DIFFERENT BANDS

capacity \ band	m_1	m_2	m_3	m_4	m_5
LC (Mbps)	12.4	15.0			
PLC-F (Mbps)			5.3	8.5	11.3
DR-PLC (Mbps)			3.4	6.9	9.8

PLC-F in Table I represents the PLC calculated under a false distribution, e.g., by learning from insufficient observations or employing a wrong theoretical model for analysis. We assume that the false distribution is normal distribution with the same statistics as the real one. By adopting PLC-F and DR-PLC to evaluate the achievable rate of a link on a shared band, the scheduling result of the S^2 scheme is shown in Fig. 7. We can



(a) Evaluation based on DR-PLC

(b) Evaluation based on PLC-F

Fig. 7. Scheduling result of the S^2 scheme.

see that for the three DS requests which can only be carried by the licensed bands, DS1 and DS3 are admitted, working on $\{m_1\}$ and $\{m_1, m_2\}$, respectively. No mutual interference exists between them because the transmitter of DS1/DS3 is far enough away from the receiver of DS3/DS1. DS2 is rejected for the total weight maximization because it has conflicting relationship with both DS1 and DS3. Considering the five DT requests, when DR-PLC is adopted to evaluate the link capacity achieved by the shared band, DT1, DT2, DT3 and DT4 will be admitted as depicted in Fig. 7(a). Notice that since the capacity provided by the licensed bands m_1 and m_2 exceeds the demand of DS3, some data of DT1 is actually carried by the licensed bands. If the operator takes the PLC-F to evaluate the link capacity, DT5, which has a higher weight than DT4, will become the admitted one as shown in Fig. 7(b) because of the overestimation of the achievable rate of each shared band. In Table II, we present the probability that the QoS for each admitted DT service is satisfied under the two cases based on DR-PLC and PLC-F, respectively, by taking 1000 experiments. In each experiment, the available bandwidth of

TABLE II
PROBABILITY THAT THE QOS OF EACH ADMITTED SERVICE IS SATISFIED

	DT1	DT2	DT3	DT4	DT5
DR-PLC	1	99.3%	1	93.7%	
PLC-F	1	80.4%	63.7%		62.7%

each shared band is generated randomly, following the uniform distribution with the aforementioned settings on statistics. We can see that if we use a specific distribution to handle the spectrum uncertainty, due to the possible false estimation as in this case study, the QoS might be hardly guaranteed. When the DR-PLC metric is adopted, the QoS can be satisfied with a high probability. Thus, DR-PLC offers us an effective way to guarantee the QoS when using the shared spectrum, even if the specific probability distribution of the available bandwidth is unknown. Noticing that although the confidence level is set to 0.7, the probability that the QoS is satisfied is higher than 90%. The reason is that such a DR-PLC corresponds to a conservative evaluation, as a lower bound of PLC under any possible distributions, so that the probability that DR-PLC is achievable should be higher than the confidence level, which could achieve a higher QoS guarantee level.

VII. CONCLUSIONS

In this paper, we have investigated how to support the increasing IoT services based on shared spectrums. Considering the uncertain availability of the shared spectrum, we have modeled the average available bandwidth of a shared band within a service period as a random variable, and proposed a new metric, PLC, to evaluate the achievable rate of a link with certain confidence level, which offers us an effective way to guarantee the QoS statistically when using the shared spectrum for service delivery. Considering that the accurate distributional information is hardly obtainable in practice, we have designed a distributionally robust data-driven approach by using the first and second order statistics and achieved a conservative evaluation on PLC. With the proposed metric, we have developed a service-based spectrum-aware data transmission scheme, so that licensed and shared spectrums can be efficiently used to fulfill diverse service requirements.

REFERENCES

[1] X. Li, H. Ding, M. Pan, J. Wang, H. Zhang, and Y. Fang, "Data-driven service provisioning over shared spectrums with statistical QoS guarantee," in *IEEE Wireless Communications and Networking Conference (WCNC'2019)*, Marrakech, Morocco, April 15-18 2019.

[2] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1125–1142, Oct. 2017.

[3] Cisco, "Cisco visual networking index: global mobile data traffic forecast update, 2016-2021, white paper," Mar. 2017.

[4] T. Wang, G. Li, B. Huang, Q. Miao, J. Fang, P. Li, H. Tan, W. Li, J. Ding, J. Li, and Y. Wang, "Spectrum analysis and regulations for 5G," 5G Mobile Communications, Springer, Cham, 2017.

[5] Y. Chen and H. S. Oh, "A survey of measurement-based spectrum occupancy modeling for cognitive radios," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 848–859, Firstquarter 2016.

[6] S. Yin, D. Chen, Q. Zhang, M. Liu, and S. Li, "Mining spectrum usage data: a large-scale spectrum measurement study," *IEEE Transactions on Mobile Computing*, vol. 11, no. 6, pp. 1033–1046, Jun. 2012.

[7] M. L. Benitez and F. Casadevall, "Time-dimension models of spectrum usage for the analysis, design, and simulation of cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 5, pp. 2091–2104, Jan. 2013.

[8] S. Bhattarai, J. M. J. Park, B. Gao, K. Bian, and W. Lehr, "An overview of dynamic spectrum sharing: Ongoing initiatives, challenges, and a roadmap for future research," *IEEE Transactions on Cognitive Communications and Networking*, vol. 2, no. 2, pp. 110–128, Jun. 2016.

[9] X. Li, H. Ding, Y. Fang, M. Pan, P. Li, X. Huang, Y. Sun, and S. Glisic, "Collaborative spectrum trading and sharing for cognitive radio networks," Handbook of Cognitive Radio, Springer, 2017.

[10] X. Li, H. Ding, M. Pan, Y. Sun, and Y. Fang, "Users first: Service-oriented spectrum auction with a two-tier framework support," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 11, pp. 2999–3013, Nov. 2016.

[11] H. Ding, Y. Fang, X. Huang, M. Pan, P. Li, and S. Glisic, "Cognitive capacity harvesting networks: Architectural evolution toward future cognitive radio networks," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1902–1923, Thirdquarter 2017.

[12] A. A. Khan, M. H. Rehmani, and A. Rachedi, "Cognitive-radio-based internet of things: Applications, architectures, spectrum related functionalities, and future research directions," *IEEE Wireless Communications*, vol. 24, no. 3, pp. 17–25, Jun. 2017.

[13] M. D. P. Guirao, A. Wilzeck, A. Schmidt, K. Septinus, and C. Thein, "Locally and temporary shared spectrum as opportunity for vertical sectors in 5G," *IEEE Network*, vol. 31, no. 6, pp. 24–31, Dec. 2017.

[14] D. Wu and R. Negi, "Effective capacity: A wireless link model for support of quality of service," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 630–643, Jul. 2003.

[15] S. Zymler, D. Kuhn, and B. Rustem, "Distributionally robust joint chance constraints with second-order moment information," *Mathematical Programming*, vol. 137, no. 1-2, pp. 167–198, Feb. 2013.

[16] S. Gong, L. Duan, and P. Wang, "Robust optimization of cognitive radio networks powered by energy harvesting," in *Proc. IEEE IEEE International Conference on Computer Communications (INFOCOM)*. Hong Kong, China, Apr. 2015, pp. 612–620.

[17] R. T. Rochafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of Risk*, vol. 2, pp. 21–41, 2000.

[18] X. Li, H. Ding, M. Pan, J. Wang, H. Zhang, and Y. Fang, "Data-driven service provisioning over shared spectrums with statistical qos guarantee," *Accepted by IEEE Wireless Communications and Networking Conference (WCNC)*, 2019.

[19] F. Tang, B. Mao, Z. M. Fadlullah, and N. Kato, "On a novel deep-learning-based intelligent partially overlapping channel assignment in SDN-IoT," *IEEE Wireless Communications*, vol. 56, no. 9, pp. 80–86, Sep. 2018.

[20] F. Tang, Z. M. Fadlullah, B. Mao, and N. Kato, "An intelligent traffic load prediction-based adaptive channel assignment algorithm in SDN-IoT: A deep learning approach," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 5141–5154, Dec. 2018.

[21] W. Ejaz and M. Ibnkahla, "Multiband spectrum sensing and resource allocation for IoT in cognitive 5G networks," *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 150–163, Feb. 2018.

[22] X. Zhang, Y. Ma, H. Qi, Y. Gao, Z. Xie, Z. Xie, M. Zhang, X. Wang, G. Wei, and Z. Li, "Distributed compressive sensing augmented wideband spectrum sharing for cognitive IoT," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 3234–3245, Aug. 2018.

[23] R. Zhang, M. Wang, X. Shen, and L.-L. Xie, "Probabilistic analysis on QoS provisioning for Internet of Things in LTE-A heterogeneous networks with partial spectrum usage," *IEEE Internet of Things Journal*, vol. 3, no. 3, pp. 354–365, Jun. 2016.

[24] Q. Li, Q. Zhang, and J. Qin, "Optimal probabilistic robust beamforming for MISO SWIPT systems with gaussian channel uncertainties," *Wireless Personal Communications*, vol. 82, no. 4, pp. 2407–2415, Feb. 2015.

[25] H. Sun, F. Zhou, R. Q. Hu, and L. Hanzo, "Robust beamforming design in a NOMA cognitive radio network relying on SWIPT," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 1, pp. 142–155, Jan. 2019.

[26] F. Zhou, Z. Chu, H. Sun, R. Q. Hu, and L. Hanzo, "Artificial noise aided secure cognitive beamforming for cooperative MISO-NOMA using SWIPT," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 4, pp. 918–931, Apr. 2018.

[27] Y. Zhang and S. Wang, "Resource allocation for cognitive radio-enabled femtocell networks with imperfect spectrum sensing and channel uncertainty," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 7719–7728, Sep. 2016.

[28] N. Zhang, S. Zhang, J. Zheng, X. Fang, J. W. Mark, and X. Shen, "QoE driven decentralized spectrum sharing in 5G networks: potential game approach," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 9, pp. 7797–7808, Mar. 2017.

[29] S. Li, Z. Zheng, E. Ekici, and N. B. Shroff, "Maximizing social welfare in operator-based cognitive radio networks under spectrum uncertainty and sensing inaccuracy," in *Proc. IEEE International Conference on Computer Communications (INFOCOM)*. Turin, Italy, Apr. 2013, pp. 953–961.

[30] J. Lai, E. Dutkiewicz, R. Liu, and R. Vesilo, "Opportunistic spectrum access with two channel sensing in cognitive radio networks," *IEEE Transactions on Mobile Computing*, vol. 14, no. 1, pp. 126–138, Jan. 2015.

[31] M. J. Piran, N. H. Tran, D. Y. Suh, J. Song, C. S. Hong, and Z. Han, "QoE-driven channel allocation and handoff management for seamless multimedia in cognitive 5G cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 7, pp. 6569–6585, Jul. 2017.

[32] N. Cheng, N. Zhang, N. Lu, X. Shen, J. W. Mark, and F. Liu, "Opportunistic spectrum access for CR-VANETs: a game-theoretic approach," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 1, pp. 237–251, Jan. 2014.

[33] M. Sharma and A. Sahoo, "Stochastic model based opportunistic channel access in dynamic spectrum access networks," *IEEE Transactions on on Mobile Computing*, vol. 13, no. 7, pp. 1625–1639, Jul. 2014.

[34] J. Yu, M. H. Cheung, and J. Huang, "Spectrum investment under uncertainty: a behavioral economics perspective," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 10, pp. 2667–2677, Oct. 2016.

[35] M. Pan, H. Yue, Y. Fang, and H. Li, "The X loss: band-mix selection for opportunistic spectrum accessing with uncertain spectrum supply from primary service providers," *IEEE Transactions on Mobile Computing*, vol. 11, no. 12, pp. 2133–2144, Dec. 2012.

[36] G. Ding, Y. Jiao, J. Wang, Y. Zou, Q. Wu, Y. Yao, and L. Hanzo, "Spectrum inference in cognitive radio networks: algorithms and applications," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 150–182, Firstquarter 2018.

[37] X. Xing, T. Jing, W. Cheng, Y. Huo, and X. Cheng, "Spectrum prediction in cognitive radio networks," *IEEE Wireless Communications*, vol. 20, no. 2, pp. 90–96, Apr. 2013.

[38] V. K. Tumuluru, P. Wang, and D. Niyato, "Channel status prediction for cognitive radio networks," *Wireless Communications and Mobile Computing*, vol. 12, no. 10, pp. 862–874, Jul. 2012.

[39] Z. Zhang, K. Zhang, F. Gao, and S. Zhang, "Spectrum prediction and channel selection for sensing-based spectrum sharing scheme using online learning techniques," in *Proc. IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. Hong Kong, China, Aug. 2015, pp. 355–359.

[40] A. Ben-Tal, L. E. Ghaoui, and A. Nemirovski, "Robust optimization," Princeton: Princeton University Press, 2009.

[41] A. Shapiro, "On duality theory of conic linear problems," *Nonconvex Optimization & Its Applications*, vol. 57, pp. 135–165, 2001.

[42] I. Pólik and T. Terlaky, "A survey of the S-lemma," *Siam Review*, vol. 49, no. 3, pp. 371–418, 2007.

[43] L. Duan, J. Huang, and B. Shou, "Investment and pricing with spectrum uncertainty: a cognitive operator's perspective," *IEEE Transactions on Mobile Computing*, vol. 10, no. 11, pp. 1590–1604, Nov. 2011.