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/TWC.2018.2796576, IEEE Transactions on Wireless Communications

Heterogeneous Spectrum Aggregation: Coexistence from a Queue Stability Perspective

Yitu Wang, *Student Member*, Wei Wang, *Senior Member, IEEE*, Vincent K. N. Lau, *Fellow, IEEE*, Lin Chen, *Member, IEEE*, Zhaoyang Zhang, *Member, IEEE*

***Abstract*—Spectrum aggregation (SA) across heterogeneous channels, including both dedicated and shared channels, provides the potential for improving spectrum utilization and fulfilling the requirement of broadband services. Heterogeneous SA brings new technical challenges on multi-system coexistence on shared channels and the resource allocation over heterogeneous channels. In this paper, we develop an analytical framework for heterogeneous SA from a queue stability perspective. To make all systems on the shared channels stable, we design a resource allocation algorithm for the coexistence of multiple systems. Specifically, we derive the closed-form modified water-filling power control for the single-pair case by Lyapunov optimization and prove that it achieves the queue stability for all systems. Based on the results, we propose a low-complexity suboptimal resource allocation algorithm for multi-pair SA, which is a NP-hard problem. We partition user pairs into groups by using graph coloring and allocate the shared channels to pair groups according to the maximal weight bipartite matching model. The simulation results verify the queue stability and show that the proposed schemes outperform the conventional schemes.**

I. Introduction

The rapid expansion of wireless services coupled with the increasing scarcity of available spectrum, has intensified the urgency to devise new and more flexible spectrum-access strategies to improve the current spectrum utilization. Spectrum/carrier aggregation (SA) enables the device to bond multiple fragmentary spectrum resources for broadband transmission [1], and the licensed-exempt spectrum (e.g. ISM bands) can be utilized for overcoming the spectrum scarcity. SA across both dedicated and shared channels, referred to as *heterogeneous spectrum aggregation*, embraces the advantages of the above two techniques. Heterogeneous SA also attracts a lot of attentions from the industry as the promising technique for 5G systems. Licensed assisted access (LAA) and LTE-Wi-Fi link aggregation (LWA) [2], [3], proposed for 3GPP Release 13, provide different methods to access the shared channels for heterogeneous SA.

This work is supported in part by National Natural Science Foundation of China (Nos. 61571396, 61725104), Zhejiang Provincial Natural Science Foundation of China (No. LR17F010001), Young Elite Scientist Sponsorship Program by CAST (No. 2016QNRC001), and Talent Project of ZJAST (No. 2017YCGC011). The corresponding author is Wei Wang (Email: wangw@zju.edu.cn).

Yitu Wang, Wei Wang and Zhaoyang Zhang are with College of Information Science and Electronic Engineering, Zhejiang Provincial Key Laboratory of Information Processing, Communication and Networking, Zhejiang University, Hangzhou 310027, China.

Vincent K. N. Lau is with the Department of Electrical and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong.

Lin Chen is with Laboratoire de Recherche en Informatique (LRI), University of Paris-Sud and CNRS, Orsay 91405, France.

A. Motivation

For heterogeneous SA, one of the critical issues is the coexistence of multiple systems on shared channels. It is evaluated in [4] that the Wi-Fi terminal throughput could drop by 70%, if the interference is not properly mitigated. To provide protection to the Wi-Fi users, one of the efficient approaches is opportunistic spectrum access, e.g., listen-before-talk (LBT) and duty-cycle muting (DCM) [5]. Those strategies rely on channel sensing to avoid simultaneous transmissions. However, such strategies with the protocol interference model suffer from one major drawback, i.e., they do not exploit any spare capacity available in lightly-loaded systems, which can actually support the simultaneous transmission in presence of the interference. To exploit more dimensions of network resources for meeting the diverse needs of users, we allow simultaneous transmission under physical interference model with the link adaptation over dedicated and shared channels. We take one step further and try to study *What is the fundamental benefit of aggregating heterogeneous channels with system coexistence through interference management?* Our objective is to construct a theoretical framework to take full advantage of this benefit.

In existing works, the algorithms improving fairness/throughput for the LTE/Wi-Fi coexistence are proposed. In [6], the sum throughput of both systems in the shared channel is maximized. In [7], a novel spectrum sharing paradigm based on spectral proximity is proposed to promote spectral utilization. In [8], some protective fairness measures are proposed for Wi-Fi transmissions. However, those works have all focused on the physical layer performance without the consideration of the bursty data arrivals at the devices. When we have bursty data arrivals for elastic services, the maximization on fairness or throughput performance cannot even guarantee the stability of queues in the systems. It is important to study the coexistence of multiple systems with heterogeneous SA from a queue stability perspective.

B. Main Contributions and Results

In this paper, we develop an analytical framework for heterogeneous SA which aggregates both dedicated and shared channels. To make all systems on shared channels stable, we design a resource allocation algorithm for the coexistence of multiple systems from a queue stability perspective. Our key contributions can be summarized as follows:

1) We propose a closed-form *modified water-filling power allocation* solution for heterogeneous SA in a basic

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/TWC.2018.2796576, IEEE Transactions on Wireless Communications

system consisting of one dedicated channel and one shared channel, where an SA user pair transmits on both channels and the other user pair transmits on the shared channel only. We theoretically quantify how the interference and the channel quality affect the power allocation. The proposed power allocation algorithm can fully exploit the available spare capacity of the shared channel, and is proved to have the capability of stabilizing both systems by controlling the SA user pair only.

2) Through the implementation of queue estimation, we are able to apply the proposed algorithm to general coexisting systems, where the information of the queue length of the sharing system is not precisely known. We further analytically evaluate the performance degradation brought by the imperfect queue estimation.
3) For the general multi-pair case, it is proved that the Lyapunov optimization problem is NP-hard. We propose a low-complexity suboptimal resource allocation solution. Specifically, we partition user pairs into groups by using *graph coloring* and allocate the shared channels to pair groups according to the *maximal weight bipartite matching* model. To provide the weights in the bipartite graph, the modified water-filling power allocation for the single-pair case is extended to the multi-pair case in an iterative way.

The rest of this paper is organized as follows. Section II discusses the related works. Section III presents the system model. In Section IV, we propose the modified water-filling power allocation algorithm for the single-pair case. Based on the results, we study the resource allocation for the general multi-pair case in Section V. The practical implementation issue is discussed in Section VI. The performance of the proposed schemes are evaluated by simulation in Section VII. Finally, this paper is concluded in Section VIII.

II. Related Works

This paper develops an analytical framework for the coexistence of heterogeneous systems from a queue stability perspective. In this section, we briefly review the existing works on the coexistence of heterogeneous systems and stability-based considerations.

A. Coexistence of Heterogeneous Systems

In the literature, the LTE-U/Wi-Fi coexistence has recently drawn extensive attention. Numerous publications investigate on the throughput/fairness issue for the LTE-U/Wi-Fi coexistence based on LBT/DCM. In [9], a fairness based licensed-assisted access and resource scheduling scheme is proposed by optimizing the contention window size. In [10], a cooperative soft combination-based spectrum sensing scheme is designed and the corresponding throughput performance is analyzed. However, such strategies with the protocol interference model do not exploit any spare capacity available in lightly-loaded systems. We take one step further and construct a theoretical framework to take full advantage of the benefit of aggregating heterogeneous channels with system coexistence through interference management.

Interference coordination between multiple systems is also an important issue for system coexistence. In [11], an unlicensed spectrum inter-cell interference coordination mechanism is developed for the coexistence of multiple LTE-U systems. In [12], the hyper access point is introduced to enable better coordination of spectrum allocation and interference management. In [13], efficient resource allocation and spectrum sharing techniques in cognitive radio networks are introduced in a comprehensive manner. In [14], a joint relay selection and power allocation algorithm is proposed for an underlay cooperative cognitive radio system with carrier aggregation to improve the throughput. One limitation of these schemes lies in the need of a centralized controller. We propose a distributed power control algorithm without the need of coordinating with users in multiple systems, where the power is adjusted according to the observed channel quality and the information obtained through monitoring the control signals inherently present in most of the deployed wireless systems.

B. Stability-based Considerations

There are several common approaches to handle delay-aware resource allocation [15]. Large deviation [16] is an approach to convert the delay constraint into an equivalent rate constraint. However, this method achieves good delay performance only for a large delay regime. Stochastic majorization [17] provides a way to minimize the delay for the cases with symmetric arrivals. Markov decision process (MDP) [18] can minimize the delay for general cases but leads to the curse of dimensionality by brute-force value iteration or policy iteration.

Lyapunov optimization [19] is an effective approach on queue stability, which ensures that the queue system is stable as long as the average arrival rates are within the system stability region. In addition, the Lyapunov optimization has two benefits for solving the heterogeneous SA problem in our work. 1) There is no need to coordinate the systems on the shared channel. Such online algorithms are highly suitable for coexisting systems, where we only have limited information about the other sharing systems. 2) The Lyapunov optimization approach has a lower computational complexity, which makes it possible to apply the proposed algorithm to very general scenarios with different traffic models and service rate distributions.

Most existing algorithms using Lyapunov optimization are designed for the dedicated channels which are not applicable to our case with heterogeneous channels. Because it is not possible to control the actions of the devices in the sharing system, and the global information of the queue length of the users in the sharing system cannot be precisely known. There are also some recent publications on spectrum sharing using Lyapunov optimization. In [20], a distributed algorithm is proposed for opportunistic spectrum sharing based on LBT to avoid interference. Different to [20], we allow simultaneous transmission and address the interference issue head-on to fully exploit the available spare capacity of the shared channel.

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/TWC.2018.2796576, IEEE Transactions on Wireless Communications

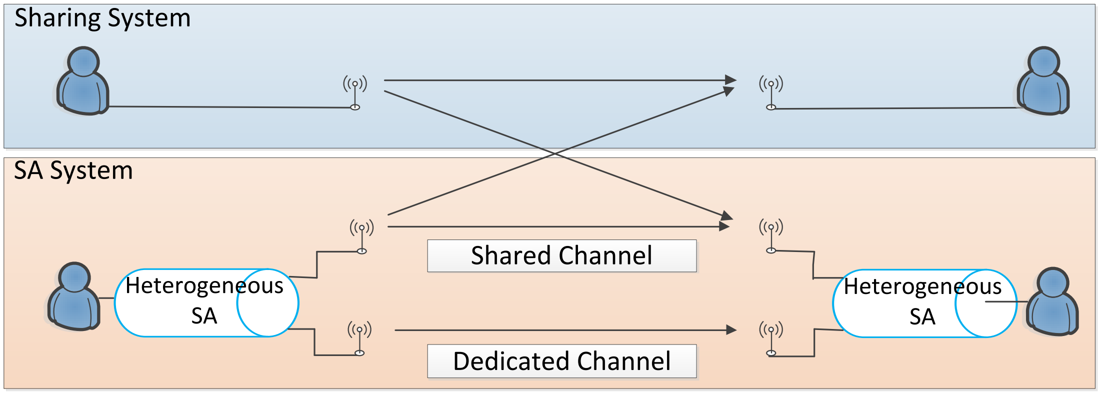


Fig. 1. System model of heterogeneous SA

III. System Model

In this section, we introduce the heterogeneous SA model with the coexistence between the SA system and the sharing system. To address the stability issue, we outline the key resource allocation variables and introduce the queue dynamics. Finally, we formulate the resource allocation problem for the coexistence under both the average power constraints and the stability constraints.

A. Heterogeneous SA Model

Consider the coexistence of two systems, including an SA system consisting of N user pairs and a sharing system consisting of K user pairs. Denote the sets of user pairs as $\mathcal{N}$ and $\mathcal{K}$, respectively. There are two types of channels, including N dedicated channels[1] for the SA system only and M shared channels for both systems. Denote the sets of channels as $\mathcal{N}^c$ and $\mathcal{M}$, respectively. The time is slotted and the duration of each time slot is one unit of time. The channel conditions are assumed to be constant within the duration of a slot and may change over slots. The two systems are illustrated in Fig. 1, where the lower part represents an SA system and the upper part represents a sharing system. In the sharing system, each user pair communicates using a channel which can be shared with others, thus the transmissions interfere with each other when multiple pairs transmit. In the SA system, each user pair communicates using a dedicated channel, potentially in addition to an aggregated shared channel.

In the SA system, MAC layer SA is adopted for smooth compatibility with the conventional cellular systems [1], i.e., two physical layer transceivers are used to transmit on a dedicated channel and a shared channel, respectively. Each SA user pair has a fixed dedicated channel, i.e., user pair $i \in \mathcal{N}$ communicates on channel $i \in \mathcal{N}^c$. With heterogeneous SA capability, an SA user pair can aggregate a shared channel $j \in \mathcal{M}$ in addition to its dedicated channel.

In the sharing system, user pairs do not have SA capability and each transmitter transmits on a shared channel. The user pairs do not adjust their transmission actions according to the interference caused by the transmission of the SA system on the shared channels, which is widely assumed in spectrum sharing scenarios [12], [22]. Without awareness of the SA user pairs, the sharing user pairs simply transmit at the maximum power p_m whenever their queues are not empty and stay idle if their queues are empty, to maximize their throughput and ensure the stability in a best-effort way.

Let $\mathbf{G}(t)$ be the global channel gain[2] in slot t, i.e., $\mathbf{G}(t) = \left(g_{ik}^j(t), i,k \in \mathcal{N} \bigcup \mathcal{K}, j \in \mathcal{N}^c \bigcup \mathcal{M}\right)$, where $g_{ik}^j(t)$ is the gain of the channel from transmitter i to receiver k on channel j.

B. Resource Allocation Model

The controller (e.g., base stations) of the SA system can fully control the actions of the user pairs in the SA system but cannot directly control the actions of the user pairs in the sharing system. At the beginning of each slot, the controller determines the channel and power allocation for transmitters in $\mathcal{N}$. The associated control variables are defined as follows:

- *Channel allocation* $\mathbf{b}(t)$: Define $\mathbf{b}(t) = \left\{b_i^j(t), \forall i \in \mathcal{N}, \forall j \in \mathcal{M}\right\}$, where $b_i^j(t) \in \{0,1\}$ and $b_i^j(t) = 1$ represents that transmitter i transmits on channel j in slot t.
- *Power allocation* $\mathbf{P}(t)$: Define $\mathbf{P}(t) = \left\{p_i^j(t), \forall i \in \mathcal{N}, \forall j \in \mathcal{N}^c \bigcup \mathcal{M}\right\}$, where $p_i^j(t)$ is the transmit power of transmitter i on channel j in slot t.

By adjusting the resource allocation variables for transmitters in $\mathcal{N}$, we can indirectly affect the data rates of user pairs in $\mathcal{K}$ by changing the interference caused by the SA user pairs on the shared channels. Denote the data rates as $\mathbf{r}(\mathbf{P}(t), \mathbf{b}(t), \mathbf{G}(t)) = \left\{r_i(\mathbf{P}(t), \mathbf{b}(t), \mathbf{G}(t)), \forall i \in \mathcal{N} \bigcup \mathcal{K}\right\}$, where $r_i(\mathbf{P}(t), \mathbf{b}(t), \mathbf{G}(t))$ is the data rate of user pair i in slot t, which is presented as

$$r_i(\mathbf{P}(t), \mathbf{b}(t), \mathbf{G}(t)) = \begin{cases} r_i^i(t) + \sum\limits_{j \in \mathcal{M}} b_i^j(t) r_i^j(t), \forall i \in \mathcal{N}, \\ r_i^i(t), \forall i \in \mathcal{K}, \end{cases} \tag{1}$$

[1] We assume that each user pair in the SA system has a dedicated channel and mainly focus on the heterogeneous SA issue in this paper. There have been a number of works on aggregating multiple dedicated channels [10][21], with which the proposed algorithm can be easily extended to support flexible spectrum access, i.e., OFDMA.

[2] The SA system can obtain $\mathbf{G}(t)$ through training and feedback from the receivers according to [22].

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/TWC.2018.2796576, IEEE Transactions on Wireless Communications

where $r_i^j(t)$ is the data rate of transmitter i on channel j in t, which can be calculated as

$$r_i^j(t) = \begin{cases} \log_2\left(1 + g_{ii}^i(t)p_i^i(t)\right), \forall i \in \mathcal{N}, i = j, \\ \log_2\left(1 + \dfrac{g_{ii}^j(t)p_i^j(t)}{1 + I_i^j(t)}\right), \forall i \in \mathcal{N}, j \in \mathcal{M}, \\ 1_{\{Q_i(t)>0\}} \log_2\left(1 + \dfrac{g_{ii}^j(t)p_m}{1 + I_i^j(t)}\right), \forall i \in \mathcal{K}, i = j, \\ 0, \text{elsewise}, \end{cases} \tag{2}$$

where $1_{\{Q_i(t)>0\}}$ is an indicator function which is 1 if $Q_i(t) > 0$ and 0 if $Q_i(t) \leq 0$, and $1 + I_i^j(t)$ is the normalized interference [3] on channel j at receiver i, which is

$$I_i^j(t) = \begin{cases} 1_{\{Q_j(t)>0\}} p_m g_{ji}^j(t) + \sum\limits_{l \in \mathcal{N}, l \neq i} b_l^j(t) p_l^j(t) g_{li}^j(t), \forall i \in \mathcal{N}, j \in \mathcal{M}, \\ \sum\limits_{l \in \mathcal{N}} b_l^j(t) p_l^j(t) g_{li}^j(t), \forall i \in \mathcal{K}, i = j. \end{cases} \tag{3}$$

C. Queue Dynamics and Stability

To analyze the stability issue, we first discuss the packet queue backlog. Each transmitter maintains a packet queue, whose length is denoted as $Q_i(t)$ for transmitter i in slot t. Let $\mathbf{A}(t) = \{A_i(t), \forall i \in \mathcal{N} \bigcup \mathcal{K}\}$ be the random packet arrivals, where $A_i(t)$ is the number of arrived bits for transmitter i in slot t. Assume that $\mathbf{A}(t)$ is i.i.d. over time slots, with $\mathbb{E}[A_i(t)] = \lambda_i$, where λ_i is the average arrival rate for transmitter i. The queue dynamics of $Q_i(t)$ is

$$Q_i(t+1) = \max\{Q_i(t) - r_i(t), 0\} + A_i(t). \tag{4}$$

To study the coexistence of heterogeneous SA from a queue stability perspective, according to [19], we define the queue stability as follows:

Definition 1 (Queue Stability). *A queue $Q_i(t)$ is strongly stable [4] if*

$$\lim_{T\to\infty} \frac{1}{T} \left(\sum_{t=0}^{T} \mathbb{E}[Q_i(t)] \right) < \infty. \tag{5}$$

The coexistence of heterogeneous SA is said to be stable if all the queues in both the SA system and the sharing system are strongly stable. ■

D. Problem Formulation

To make sure the coexistence of heterogeneous SA is stabilizable, we first define the capacity region following [19] as follows:

Definition 2 (Capacity Region). *The capacity region Λ is defined as the closure of the set of all input rate vectors λ stabilizable under some power allocation algorithm that conforms to the power constraint $\mathbf{P}(t) \in \Pi$.*

Throughout this paper, we assume that the input rate vector is strictly interior to the capacity region, such that the system is stabilizable.

Our goal for heterogeneous SA is determining the resource allocation to stabilize both the SA system and the sharing system with the transmit power constraints. Define $\pi_{\mathbf{G}(t)}$ as the occurrence probability for each channel state $\mathbf{G}(t)$[5]. For any $\boldsymbol{\lambda} \in \Lambda$, the resource allocation including channel allocation $\mathbf{b}(t)$ and power allocation $\mathbf{P}(t)$ should satisfy

$$\sum_{\mathbf{G}(t)} \pi_{\mathbf{G}(t)} r_i(\mathbf{P}(t), \mathbf{b}(t), \mathbf{G}(t)) \geq \lambda_i, \forall i \in \mathcal{N} \bigcup \mathcal{K}, \tag{6}$$

$$\mathbf{P}(t) \in \Pi, \tag{7}$$

$$\sum_{j \in \mathcal{M}} b_i^j(t) \leq 1, \forall i \in \mathcal{N}, \tag{8}$$

where (6) is the average rate constraints for queue stability, (7) is the power constraint and (8) implies that an SA user pair can aggregate only one shared channel due to the SA capability limitation.

Without a priori knowledge of traffic arrival rates, it is difficult to directly judge whether the stability constraints (6) are satisfied. By rewriting (6) as the queue stability condition in Definition 1 according to Little's law [23], we derive a dynamic resource allocation algorithm, for not only fulfilling stability and power constraints but also optimally balancing the average power consumption and the average queue backlog by adopting Lyapunov optimization.

IV. Heterogeneous SA: A Single-Pair Case

In this section, we first consider a basic system consisting of only one user pair in each system to extract key insights. By adopting the first-order approximation, we provide a closed-form modified water-filling power control solution, which characterizes the influence of the sharing system on the SA system. Furthermore, we prove the stability of both systems based on Lyapunov optimization and discuss the implementation with queue estimation.

A. Modified Water-Filling Power Allocation for Stability

In the basic system model, consisting of two user pairs and two channels, user pair 0 is in the sharing system, who can access channel 0 only, and user pair 1 belongs to the SA system, who can access both channel 0 and channel 1. With the above fixed channel allocation, only the power allocation for transmitter 1 over the shared channel 0 and the dedicated channel 1 needs to be determined.

To satisfy the stability constraints, we adopt the Lyapunov optimization with a commonly used quadratic Lyapunov function [19], which increases quadratically with the increase of

[3]For simplicity of expression, all the channel gains are defined to be normalized over thermal noise level N_0 without loss of generality.

[4]Under mild boundedness assumptions, strong stability implies all of the other forms of stability, such as steady state stable, rate stable and mean rate stable [19]. Therefore, we choose this particular definition to make sure all queues are stable under all forms of stability.

[5]These probabilities of channel states determine the capacity region of the network, but are not necessarily known to the controller.

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/TWC.2018.2796576, IEEE Transactions on Wireless Communications

the queue length and can provide large enough penalty to stabilize the system,

$$L\big(Q_0(t), Q_1(t)\big) = \big(Q_0(t)\big)^2 + \big(Q_1(t)\big)^2. \tag{9}$$

According to Definition 1, both $Q_0(t)$ and $Q_1(t)$ should be stabilized for the stability of the coexistence of heterogeneous SA.

To incorporate power constraint, we adopt V as the price of power consumption. Thus, we formulate the associated Lyapunov optimization problem as

$$\max_{\mathbf{p}(t)} \; Q_0(t)r_0(t) + Q_1(t)r_1(t) - V\big(p_1^0(t) + p_1^1(t)\big), \tag{10}$$

where

$$r_0(t) = 1_{\{Q_0(t)>0\}} \log_2\left(1 + \frac{p_m g_{00}^0(t)}{1 + p_1^0(t) g_{10}^0(t)}\right), \tag{11}$$

$$\begin{aligned} r_1(t) =& \log_2\left(1 + p_1^1(t) g_{11}^1(t)\right) \\ &+ \log_2\left(1 + \frac{p_1^0(t) g_{11}^0(t)}{1 + 1_{\{Q_0(t)>0\}} p_m g_{01}^0(t)}\right). \end{aligned} \tag{12}$$

To achieve the optimality of the Lyapunov optimization problem, we take partial derivative of (10) with respect to $p_1^0(t)$ and $p_1^1(t)$ respectively and let them equal to zero. For the optimality, the power allocated to the dedicated channel is

$$p_1^1(t) = \frac{Q_1(t)}{V/\ln 2} - \frac{1}{g_{11}^1(t)}, \tag{13}$$

which is a standard water-filling form. The power allocated to the shared channel should satisfy

$$\begin{aligned} &Q_0(t)\frac{g_{10}^0(t)}{1 + p_1^0(t) g_{10}^0(t) + p_m g_{00}^0(t)} - Q_0(t)\frac{g_{10}^0(t)}{1 + p_1^0(t) g_{10}^0(t)} \\ &+ Q_1(t)\frac{g_{11}^0(t)}{1 + 1_{\{Q_0(t)>0\}} p_m g_{01}^0(t) + p_1^0(t) g_{11}^0(t)} - \frac{V}{\ln 2} = 0. \end{aligned} \tag{14}$$

Note that the above equation (14) is a cubic equation, which is too complicated to obtain a closed-form solution and extract key insights on how heterogeneous SA affects these two systems. This is because that the queue dynamics of both transmitters are coupled with each other due to the mutual interference.

To investigate the interactions between the SA system and the sharing system, we exploit the property that the mutual interference of two transmitters sharing the same channel is small[6], i.e., the coupling between the queue dynamics of both transmitters is weak [15], [24]. Denote δ as their largest cross-link channel gain, i.e., $\delta = \max_t g_{10}^0(t)$. We leverage this weak coupling property to derive a reasonable approximate closed-form solution in the following theorem:

[6]Too large mutual interference is actually not suitable for spectrum sharing by simultaneous transmission. In practical systems, the operating SINR is at least 5dB or above for reasonable link-level performance. As such, various interference coordination schemes, such as the eICIC in LTE, are adopted to carefully select the users with small mutual interference for simultaneous transmission. The scenario we consider in this paper focus on the practical application regime in most practical systems.

Theorem 1 (Modified Water-Filling for Heterogeneous SA)**.** *For the optimality of the objective in (10), the power allocated to the shared channel for the transmitter in the SA system is*

$$\begin{aligned} p_1^0(t) =& \frac{Q_1(t)}{\frac{V}{\ln 2} + Q_0(t) g_{10}^0(t) \frac{p_m g_{00}^0(t)}{1 + p_m g_{00}^0(t)}} \\ &- \frac{1 + 1_{\{Q_0(t)>0\}} p_m g_{01}^0(t)}{g_{11}^0(t)} + o(\delta). \end{aligned} \tag{15}$$

Proof: Please refer to Appendix A. ■

Remark 1. *The approximate power allocation solution (15) has a water-filling style with modified water-filling level and sea bed, which is illustrated in Fig. 2. If the queue length of the transmitter in the sharing system is 0, this modified water-filling solution (15) is reduced to*

$$p_1^0(t) = \frac{Q_1(t)}{V/\ln 2} - \frac{1}{g_{11}^0(t)}. \tag{16}$$

Comparing the modified water-filling solution (15) *with the standard water-filling solution* (16)*, we analyze the influences to the water-filling process of SA transmitter* 1 *induced by the parameters of sharing transmitter* 0*:*

- ***Influence due to the queue length*** $Q_0(t)$*:* $Q_0(t)$ *reflects the urgency of user pair* 0 *for transmitting on the shared channel, which lowers the water-filling level by enlarging the parameter* V *with a discount factor* $\frac{p_m g_{00}^0(t)}{1+p_m g_{00}^0(t)}$.
- ***Influence due to the channel gains****: The channel gain* $g_{00}^0(t)$ *implies how good the transmission opportunity of user pair* 0 *is. If user pair* 0 *is suffering poor channel quality, i.e.,* $g_{00}^0(t)$ *is small, transmitter* 1 *can access the shared channel more aggressively. Moreover, large cross-link channel gains, including both* $g_{01}^0(t)$ *and* $g_{10}^0(t)$*, pose negative influence due to the mutual interference and thus lead to a small allocated power on the shared channel.*
- ***Influence due to the transmit power*** p_m*: The influence of the transmit power is two folds, including the interference of the SA user pair and the transmission opportunity of the sharing user pair, which decreases the allocated power by not only rising the sea bed but also lowering the water-filling level.*

As illustrated in Fig. 2, the red bar represents the sea bed risen due to the interference from the sharing system, and the purple bar represents the water-filling level dropped due to the urgency of the user pair in the sharing system. ■

In the following theorem, we prove that the first-order approximation does not affect the stability of both systems.

Theorem 2 (Stability Property)**.** *Suppose the input rate vector* $\boldsymbol{\lambda}$ *is interior to the capacity region* Λ *given the power constraint* $\mathbf{P} \in \Pi$*, then the power allocation scheme in* (13) *and* (15) *stabilizes all queues of the transmitters in both the SA system and the sharing system. Specifically, the average power consumption and the average sum queue length satisfy*

$$\bar{p} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} \mathbb{E}[\mathbf{p}(t)] \le \Phi(\lambda_0 + \epsilon, \lambda_1 + \epsilon) + \frac{B'}{V}, \tag{17}$$

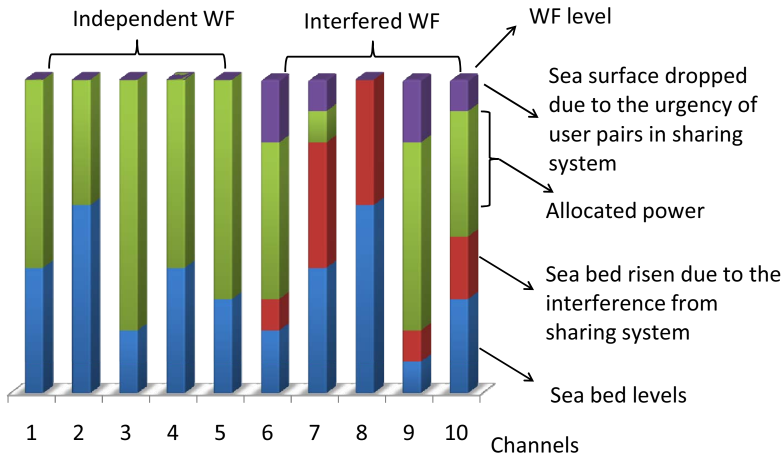


Fig. 2. Illustration of the modified water-filling power allocation

$$\lim_{T\to\infty}\frac{1}{T}\left(\sum_{t=0}^{T}Q_0(t)+\sum_{t=0}^{T}Q_1(t)\right) \leq \frac{B'}{2\epsilon}+\frac{V\left(\Phi(\lambda_0+\epsilon,\lambda_1+\epsilon)-\bar{p}\right)}{2\epsilon}, \tag{18}$$

where $B' = \mathbb{E}[\big(A_0(t)\big)^2] + \mathbb{E}[\big(A_1(t)\big)^2] + \mathbb{E}[\big(r_0(t)\big)^2] + \mathbb{E}[\big(r_1(t)\big)^2] + 2\mathbb{E}[r_0(t)]\lambda_0 + 2\mathbb{E}[r_1(t)]\lambda_1 - o(\delta)$, ϵ *satisfies* $\mathbb{E}[r_0(t)] \geq \lambda_0+\epsilon$ *and* $\mathbb{E}[r_1(t)] \geq \lambda_1+\epsilon$, $\Phi(\lambda_0,\lambda_1)$ *denotes the minimal power consumption stabilizing a system with the average arrival rate vector* $\boldsymbol{\lambda}$ *and* $o(\delta)$ *denotes the approximation error of* $\log_2\left(1+p_1^0(t)g_{10}^0(t)\right)-\ln 2 p_1^0(t)g_{10}^0(t)$, *which is negative.*

Proof: Please refer to Appendix B. ■

Remark 2. *From Theorem 2, the proposed approximate power allocation scheme does not affect the stability but have influence on performance of the average power consumption and the average sum queue length. Specifically, since* $o(\delta)$ *is negative,* B' *is larger than* B *by the order of* $o(\delta)$, *where* $B=\mathbb{E}[\big(A_0(t)\big)^2]+\mathbb{E}[\big(A_1(t)\big)^2]+\mathbb{E}[\big(r_0(t)\big)^2]+\mathbb{E}[\big(r_1(t)\big)^2]+2\mathbb{E}[r_0(t)]\lambda_0+2\mathbb{E}[r_1(t)]\lambda_1$, *which is the parameter in the original optimization problem* (10). *The more accurate the approximation is, the less performance degradation we suffer.*

Remark 3 (Scaling Property)**.** *The proposed scheme can be easily extended to multi-channel cases. Considering the SA user pair which is able to aggregate* N *dedicated channels and* M *shared user pairs/shared channels. Adopting the similar technique, we can easily obtain the power allocated to dedicated channel* j *and shared channel* i *respectively as*

$$\begin{aligned} p_1^j(t) =& \frac{Q_1(t)}{V/\ln 2}-\frac{1}{g_{11}^j(t)}, \\ p_1^i(t) =& \frac{Q_1(t)}{\frac{V}{\ln 2}+Q_i(t)g_{1i}^i(t)\frac{p_m g_{ii}^i(t)}{1+p_m g_{ii}^i(t)}} \\ &-\frac{1+1_{\{Q_i(t)>0\}}p_m g_{i1}^i(t)}{g_{11}^i(t)}+o(\delta). \end{aligned} \tag{19}$$

B. Implementation of Queue Estimation

The proposed power allocation algorithm requires the information of the queue length of the sharing system, which cannot be precisely known in practice. It is necessary for the SA user pair to estimate the queue length of the sharing transmitter. We adopt the following queue estimation algorithm as [22], and further analyze its influence on our proposed power allocation algorithm.

- If transmitter i in the sharing system is busy, we estimate its queue length as

$$\hat{Q}_i(t+1)=\max\left\{\hat{Q}_i(t)-r_i(t),0\right\}+\lambda_i+\omega. \tag{20}$$

  Here, the arrival rate is estimated as $\lambda_i+\omega$ for transmitter i, where ω is an over-estimated slack variable. The queue length is shortened by $r_i(t)$, i.e., the number of successfully transmitted packets, which can be learned by listening to the link layer ACK.
- If transmitter i in the sharing system is idle, we have a perfect alignment such that

$$\hat{Q}_i(t)=Q_i(t)=0, \tag{21}$$

  which practically bounds the error of the estimated queue.

It is not possible for an SA transmitter to obtain the accurate arrival rate $A_i(t)$ of the transmitter in the sharing system. In (20), we estimate the arrival rate $A_i(t)$ using its average arrival rate λ_i plus a positive over-estimated slack variable ω to create a comfortable margin of safety, i.e., providing protection for the user pair in the sharing system.

With the proposed queue estimation scheme, the power allocation algorithm can be modified as

$$\begin{aligned} p_1^0(t) =& \frac{Q_1(t)}{\frac{V}{\ln 2}+\hat{Q}_0(t)g_{10}^0(t)\frac{p_m g_{00}^0(t)}{1+p_m g_{00}^0(t)}} \\ &-\frac{1+1_{\{\hat{Q}_0(t)>0\}}p_m g_{01}^0(t)}{g_{11}^0(t)}+o(\delta), \\ p_1^1(t) =& \frac{Q_1(t)}{V/\ln 2}-\frac{1}{g_{11}^1(t)}. \end{aligned} \tag{22}$$

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/TWC.2018.2796576, IEEE Transactions on Wireless Communications

We evaluate the influence brought by the estimation. Specifically, the influence brought by ω is given in the following theorem.

Theorem 3 (Stability Property with Queue Estimation)**.** *Suppose the input rate vector $\boldsymbol{\lambda}$ satisfies that $\boldsymbol{\lambda}+\omega\boldsymbol{I}$ is interior to the capacity region Λ given the same power constraint $\mathbf{P}\in\Pi$, where $\boldsymbol{I}$ is an identity matrix with the same rank as $\boldsymbol{\lambda}$. For any positive over-estimated slack variable ω satisfying $0<\omega<\epsilon$, the proposed algorithm using queue estimation with this parameter stabilizes all queues of the system. Specifically, the average power consumption and the average sum queue length satisfy*

$$\bar{p}=\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T}\mathbb{E}[\mathbf{p}(t)]\leq\Phi(\lambda_0+\epsilon,\lambda_1+\epsilon)+\frac{B'}{V},\tag{23}$$

$$\begin{aligned}\lim_{T\to\infty}\frac{1}{T}\left(\sum_{t=0}^{T}Q_0(t)+\sum_{t=0}^{T}Q_1(t)\right)\leq\frac{B'}{2\epsilon-2\omega}\\+\frac{V\big(\Phi(\lambda_0+\epsilon,\lambda_1+\epsilon)-\bar{p}\big)}{2\epsilon-2\omega}.\end{aligned}\tag{24}$$

Proof: Please refer to Appendix C. ■

Remark 4. *With carefully chosen ω satisfying $0<\omega<\epsilon$, the over-estimation does not affect the stability of the system. However, due to the imperfect queue estimation, the proposed power allocation algorithm stabilizes a new system with a larger input rate vector than the actual one, which leads to the shrink of the capacity region by $\omega\boldsymbol{I}$.*

C. Sensitivity to the Channel Estimation

In order to calculate the optimal power allocation, we need to estimate the channel quality of the shared user pair, which may not always be estimated accurately. We determine the amount of power analytically in regard to the variation of the estimation[7]. $\hat{g}_{00}^0(t)$ denotes the estimated channel quality of the shared user pair on the shared channel, and β denotes the percentage change of the variation, i.e., $\hat{g}_{00}^0(t)=\beta g_{00}^0(t)$. According to (33), β will be dominated by $o(\delta)$, hence we derive the power allocation as

$$\begin{aligned}p_1^0(t)=&\frac{Q_1(t)}{\frac{V}{\ln 2}+Q_0(t)g_{10}^0(t)\frac{p_m\beta g_{00}^0(t)}{1+p_m\beta g_{00}^0(t)}}\\&-\frac{1+1_{\{Q_0(t)>0\}}p_m g_{01}^0(t)}{g_{11}^0(t)}+o(\delta).\end{aligned}\tag{25}$$

- A large V reflects expensive unit power price, which results in the drop of the water-filling level. In this context, the power allocation is less sensitive to β.
- A large $Q_0(t)g_{10}^0(t)$ results in the sensitivity to β of the power allocation.
- A small $Q_1(t)$ reflects light transmission load of the SA user pair, which reduces the necessity to occupy the shared channel. In this context, the power allocation is less sensitive to β.

[7]The amount of average queue length in regard to the variation of the estimation is given through simulation, which is demonstrated to be insensitive to β.

V. Resource Allocation for Multi-Pair Heterogeneous SA

In this section, we consider a general system model with multiple user pairs in the SA system and the sharing system. Due to the NP-hard property, we propose a low-complexity suboptimal resource allocation algorithm for multi-pair heterogeneous SA.

For multi-pair heterogeneous SA, the Lyapunov optimization problem can be extended to

$$\begin{aligned}\max_{\mathbf{b}(t),\mathbf{p}(t)}\;&\sum_{i\in\mathcal{N}}Q_i(t)\left(r_i^i(t)+\sum_{j\in\mathcal{M}}b_i^j(t)r_i^j(t)\right)\\&+\sum_{i\in\mathcal{K}}\sum_{j\in\mathcal{M}}Q_i(t)b_i^j(t)r_i^j(t)-V\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}}(p_i^i(t)+p_i^j(t)).\end{aligned}\tag{26}$$

We first discuss the computational complexity of solving (26) in the following theorem:

Theorem 4 (NP-Hard Property)**.** *The optimization problem* (26) *is strongly NP-hard.*

Proof: Please refer to Appendix D. ■

According to Theorem 4, the optimization problem (26) needs an exponential computational complexity for achieving the optimality, which is unacceptable in practice. In the following, we propose a low-complexity suboptimal solution for the problem (26). Specifically, we solve the problem by two steps, as illustrated in Fig. 3.

1) *User Pair Grouping*: The aim is to partition the user pairs in the SA system into groups according to their mutual interference. User pairs in the same group have relatively small interference to each other. The user pair grouping problem is modeled as a graph coloring problem and a suboptimal algorithm is adopted for utility maximization.
2) *Channel and Power Allocation*: The aim is to allocate the shared channels to pair groups. The channel allocation problem is modeled as a maximal weight bipartite matching (MWBM) problem, where the weights are obtained by iterative modified water-filling power allocation.

1) User Pair Grouping

In this step, we partition user pairs in $\mathcal{N}$ into M groups, i.e., $\mathcal{D}=\{\mathcal{F}_1,\mathcal{F}_2,\cdots,\mathcal{F}_M\}$, each of which accesses a shared channel according to the mutual interference between the user pairs in the SA system.

For user grouping, we introduce a conflict graph $\mathcal{H}=(\mathcal{N},\mathcal{E})$, where the vertices are the user pairs in the SA system $\mathcal{N}$ and the edges $\mathcal{E}$ represents that the cross interference between the two user pairs is above a certain threshold, such that the user pairs in the same group have smaller cross interference than the threshold.

Based on the conflict graph $\mathcal{H}$, we adopt the utility-based graph coloring method in [25] for user pair grouping. The graph coloring algorithm aims to maximize the total utility using a given number of channels. Define the utility of transmitter i as the estimated increase of the objective value

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/TWC.2018.2796576, IEEE Transactions on Wireless Communications

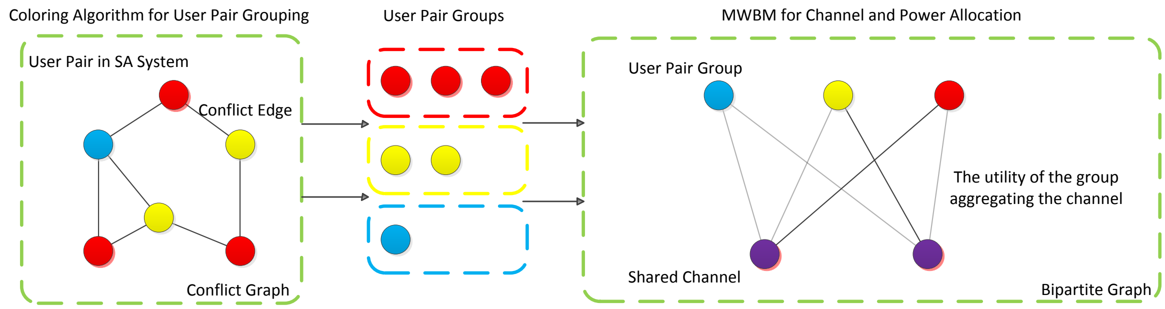


Fig. 3. Example of resource allocation for multi-pair heterogeneous SA

in (26) by assigning an additional shared channel, which can be calculated as

$$z_i(t) = \max_{p_i^i(t),\hat{p}_i(t)} \left(Q_i(t)\big(\hat{r}_i(t) + r_i^i(t)\big)\right) - V\big(\hat{p}_i(t) + p_i^i(t)\big)\big) - \max_{p_i^i(t)} \left(Q_i(t) r_i^i(t) - V p_i^i(t)\right), \tag{27}$$

where $\hat{p}_i(t)$ is the power allocated to the additional shared channel for transmitter i, and $\hat{r}_i(t)$ is the estimated data rate[8] of transmitter i on the additional shared channel, i.e.,

$$\hat{r}_i(t) = \log_2\left(1 + \frac{\hat{p}_i(t)\hat{g}_{ii}}{1 + \hat{I}_i(t)}\right), \tag{28}$$

where $\hat{g}_{ii}$ is the average channel gain over all shared channels and $\hat{I}_i(t)$ is the average interference to transmitter i of the user pairs in the sharing system[9], i.e.,

$$\hat{I}_i(t) = \sum_{j\in\mathcal{M}} \sum_{k\in\mathcal{K}} p_m g_{ki}^j(t)/M. \tag{29}$$

The optimization problem in (27) can be solved by simply letting the derivative equal to zero, and the optimal value is used as the utility of transmitter i.

With the utility values of the user pairs, we can partition the user pairs in $\mathcal{N}$ into M groups using the algorithm in [25]. Note that it is possible that some user pairs in $\mathcal{N}$ are not in these M groups, which means that they can only use their dedicated channel in time slot t due to the limitation of the number of channels and the mutual interference conditions.

In the example in Fig. 3, the user pairs are partitioned into two groups, which are colored red and yellow, respectively. The user pair colored blue is not in any group because there are only 2 shared channels and only 2 user pair groups are formed.

2) Channel and Power Allocation

After user pair grouping, we allocate the shared channels to the formed pair groups. Recall that the user pairs in the SA system are partitioned into M groups, which is equal to the number of the shared channels. Thus, we model the channel allocation problem as an MWBM problem.

We introduce a weighted bipartite graph $\mathcal{G} = (\mathcal{D}, \mathcal{M}, w)$, where the vertices of the two sides are user pair groups $\mathcal{D}$ and shared channel set $\mathcal{M}$, respectively. Any two nodes y, j in different sides have an edge, of which the weight w_{yj} is determined by the Lyapunov optimization problem

$$\begin{aligned} w_{yj} = & \\ & \max_{\mathbf{p}(t)} \sum_{i\in\mathcal{F}_y} Q_i(t)\log_2\left(1 + \frac{p_i^j(t) g_{ii}^j(t)}{1 + p_m g_{ji}^j(t) + \sum_{d\in\mathcal{F}_y, d\neq i} p_d^j(t) g_{di}^j(t)}\right) \\ & - V \sum_{i\in\mathcal{F}_y} (p_i^j(t) + p_i^i(t)) + Q_j(t)\log_2\left(1 + \frac{p_m g_{jj}^j(t)}{1 + \sum_{d\in\mathcal{F}_y} p_d^j(t) g_{dj}^j(t)}\right) \\ & + \sum_{i\in\mathcal{F}_y} Q_i(t)\log_2\left(1 + p_i^i(t) g_{ii}^i(t)\right), \end{aligned} \tag{30}$$

where the first item represents the queue weighted rate of SA user pairs in $\mathcal{F}_y$ on shared channel j, the second item represents the queue weighted rate of sharing user pair j on shared channel j, the third item represents the queue weighted rate of SA user pairs on dedicated channel and the last item represents the cost of power consumption.

In this way, the channel allocation vector $\mathbf{b}(t)$ and the power allocation vector $\mathbf{p}(t)$ are decoupled, i.e., the weights of the edges are computed based on the optimal power allocation $\mathbf{p}(t)$ with each possible $\mathbf{b}(t)$, and the optimal $\mathbf{b}(t)$ is obtained by MWBM after all the weights are determined.

To calculate the weights, we solve the Lyapunov optimization problem in (30). Since the cross interference within a group is small, the first-order approximation can be adopted as Theorem 1. Substituting those approximations into (30), the transmit power can be obtained as where the power allocation $p_i^j(t)$ are coupled and cannot be solved directly. Similar to the iterative water-filling power allocation [26], we propose the iterative modified water-filling power allocation by updating the power vector according to (31) iteratively. Substituting the converged power allocation vector into (30), we obtain the weights of the edges. After calculating the edge weights, we solve the MWBM problem for channel allocation adopting a modified shortest path search in augmenting path algorithm using Dijkstra algorithm with Fibonacci heap [27].

[8] Note that the shared channels have not allocated yet during the user pair grouping process, so we can only consider the average case over all shared channels.

[9] Here, we ignore the cross interference brought by the other user pairs in the SA system, because the cross interference within a group is sufficiently small, i.e., smaller than a threshold.

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/TWC.2018.2796576, IEEE Transactions on Wireless Communications

$$\begin{aligned} p_i^j(t) = & \frac{b_i^j(t)Q_i(t)}{\frac{V}{\ln 2} + Q_j(t)g_{ij}^j(t)\left(\frac{p_m g_{jj}^j(t)}{\left(1+p_m g_{jj}^j(t)+\sum_{d\in\mathcal{N},d\neq i} b_d^j(t)p_d^j(t)g_{di}^j(t)\right)\left(1+\sum_{d\in\mathcal{N},d\neq i} b_d^j(t)p_d^j(t)g_{dj}^j(t)\right)}\right)} \\ & - \frac{b_i^j(t)\left(1+1_{\{Q_j(t)>0\}}p_m g_{ji}^j(t)+\sum_{d\in\mathcal{N},d\neq i} b_d^j(t)p_d^j(t)g_{di}^j(t)\right)}{g_{ii}^j(t)}, i\in\mathcal{N}, j\in\mathcal{M}, \\ p_i^i(t) = & \frac{Q_i(t)}{V/\ln 2} - \frac{1}{g_{ii}^i(t)}, i\in\mathcal{N}, \end{aligned} \tag{31}$$

In the example in Fig. 3, the user pairs in the SA system are partitioned into two groups, then a two-by-two weighted bipartite graph can be obtained accordingly. The MWBM algorithm finds an optimal match achieving the maximal utility, where the red group aggregates the channel on the left, and the yellow group aggregates the other one.

Algorithm 1 Resource Allocation Algorithm for Multi-Pair Heterogeneous SA

```
1:  loop
2:      Observe parameters g_ik^j(t).
3:      Choose interference threshold s and parameter V.
4:      Generate conflict graph H.
5:      Compute the utility according to (27).
6:      Use maximal utility graph coloring algorithm in [25] to
        obtain M groups.
7:      Construct the weighted bipartite graph G consisting by
        vertices corresponding to M user pair groups and M
        respectively, and the edges are ψ.
8:      for i = 1 to |ψ| do
9:          Obtain the approximated water-filling power vector
            using iterative water-filling according to (31).
10:         Calculate the weight w(u_i, v_j) according to (30).
11:     end for
12:     Apply an MWBM algorithm to find a maximum
        weighted matching in graph G using a modified shortest
        path search in augmenting path algorithm [27].
13:     Queues are updated according to (4).
14: end loop
```

We provide the details of the proposed resource allocation algorithm using pseudo codes in Algorithm 1, which is launched at the beginning of each time slot. In the pseudo-codes, Lines 2–3 provide the initialization, Lines 4–6 establish the conflict graph to partition user pairs into M groups, Lines 7–12 obtain the channel and power allocation solution using the MWBM algorithm.

On the computational complexity of Algorithm 1, the most time-consuming part is the MWBM algorithm in Line 12, in which the running time is $O\big(\min\{N,K\}(N+K)\log(N+K+\psi)\big)$ [28], where $|\psi| = NK + O(NK)$ in Algorithm 1. Thus, the computational complexity of Algorithm 1 is bounded by $O(e^2\log(e))$, where $e = NK$.

VI. Practical Implementation

In this section, we discuss that the proposed framework can be applied to the coexistence of LTE-U/Wi-Fi systems with considering a few implementation issues.

The LTE-U system uses a centralized MAC protocol that allocates a non-overlapping set of physical resource blocks in the time-frequency domain to users, where there is no contention among the devices with dedicated channels. The Wi-Fi system implements an exclusive channel occupancy policy among the devices by adopting the LBT-based protocol for avoiding contentions. To be compatible with the LBT-based protocol, we can deploy our proposed framework with considering the following implementation issues:

1) The sensing time for the LTE-U network is set slightly longer than that of the Wi-Fi network in each time slot, such that the Wi-Fi network is not aware of the existence of the LTE-U network. The LTE-U users access the network with the proposed resource allocation to guarantee the stability of both systems.

2) Since the Wi-Fi users with the LBT-based protocol do not stay in the same channel, the users sharing the same channel vary over time slots. Fortunately, the Lyapunov drift approach that we adopt divides the queue stability into the effort of minimizing the drift in each time slot. Thus, a LTE-U user can just make the resource allocation decisions by taking the current sharing user into account even if the sharing users are not the same over different time slots.

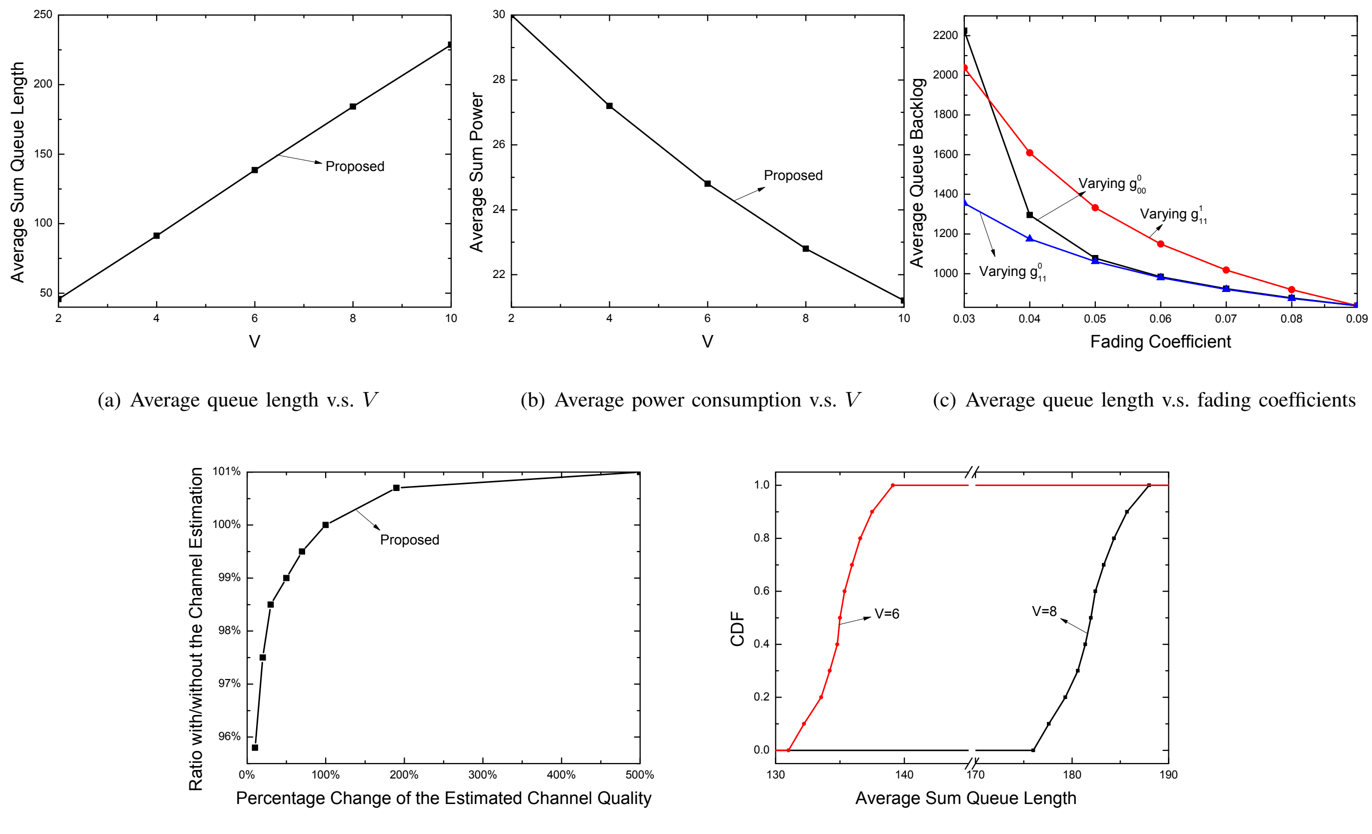
Furthermore, we analyze the performance of the proposed framework for a sharing system with the LBT-based protocols. In order to avoid collisions with the sharing user pairs, we adopt a slightly longer sensing period $\tau' > \tau$ and utilize the following MAC rule,

- If in the sensing period, no transmission is sensed. Then we update the estimated queue $\hat{Q}(t) = 0$, and allocate the transmit power as a standard water-filling form.
- Otherwise, we allocate the transmit power as a modified water-filling form.

However, due to the added sensing time, we will suffer a performance degradation. We evaluate the influence brought by queue estimation and sensing. Specifically, the influence brought by ω and τ' is given in the following theorem.

Theorem 5 (Stability Property with Queue Estimation and Sensing)**.** *Suppose the input rate vector* $\boldsymbol{\lambda}$ *satisfies that* $\boldsymbol{\lambda} + \omega\boldsymbol{I} + (\epsilon\boldsymbol{I} + \boldsymbol{\lambda})\frac{\tau'}{T}$ *is interior to the capacity region* Λ *given the same power constraint* $\mathbf{P}\in\Pi$*, where* $\boldsymbol{I}$ *is an identity matrix with the same rank as* $\boldsymbol{\lambda}$*, and* T *is the length of the frame. For any positive over-estimated slack variable* ω *and sensing time* τ' *satisfying* $0 < \omega + (\epsilon + \lambda_{\max})\frac{\tau'}{T} < \epsilon$*, where* $\lambda_{\max} =$

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/TWC.2018.2796576, IEEE Transactions on Wireless Communications



(a) Average queue length v.s. V

(b) Average power consumption v.s. V

(c) Average queue length v.s. fading coefficients

(d) Sensitivity to the estimation of the modeling parameter

(e) CDFs of different V

Fig. 4. Performance of the proposed scheme with different parameters

$\max_{i\in\mathcal{N}}\{\lambda_i\}$, *the proposed algorithm using queue estimation and sensing with these parameters stabilizes all queues of the system. Specifically, the average power consumption and the average sum queue length satisfy*

$$\bar{p} = \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T}\mathbb{E}[\mathbf{p}(t)] \leq \Phi(\lambda_0+\epsilon,\lambda_1+\epsilon)+\frac{B'}{V},$$

$$\lim_{T\to\infty}\frac{1}{T}\left(\sum_{t=0}^{T}Q_0(t)+\sum_{t=0}^{T}Q_1(t)\right) \leq \frac{B'}{2\epsilon-2\omega-2(\epsilon+\lambda_{\max})\frac{\tau'}{T}} + \frac{V\big(\Phi(\lambda_0+\epsilon,\lambda_1+\epsilon)-\bar{p}\big)}{2\epsilon-2\omega-2(\epsilon+\lambda_{\max})\frac{\tau'}{T}}.$$

■

Note that the performance loss due to the sensing time is suffered by almost all LBT-based algorithms.

For the transmission phase, to further protect the QoS of the sharing user pairs, the SA user pairs can transmit on the shared channel with a power constraint, i.e.,

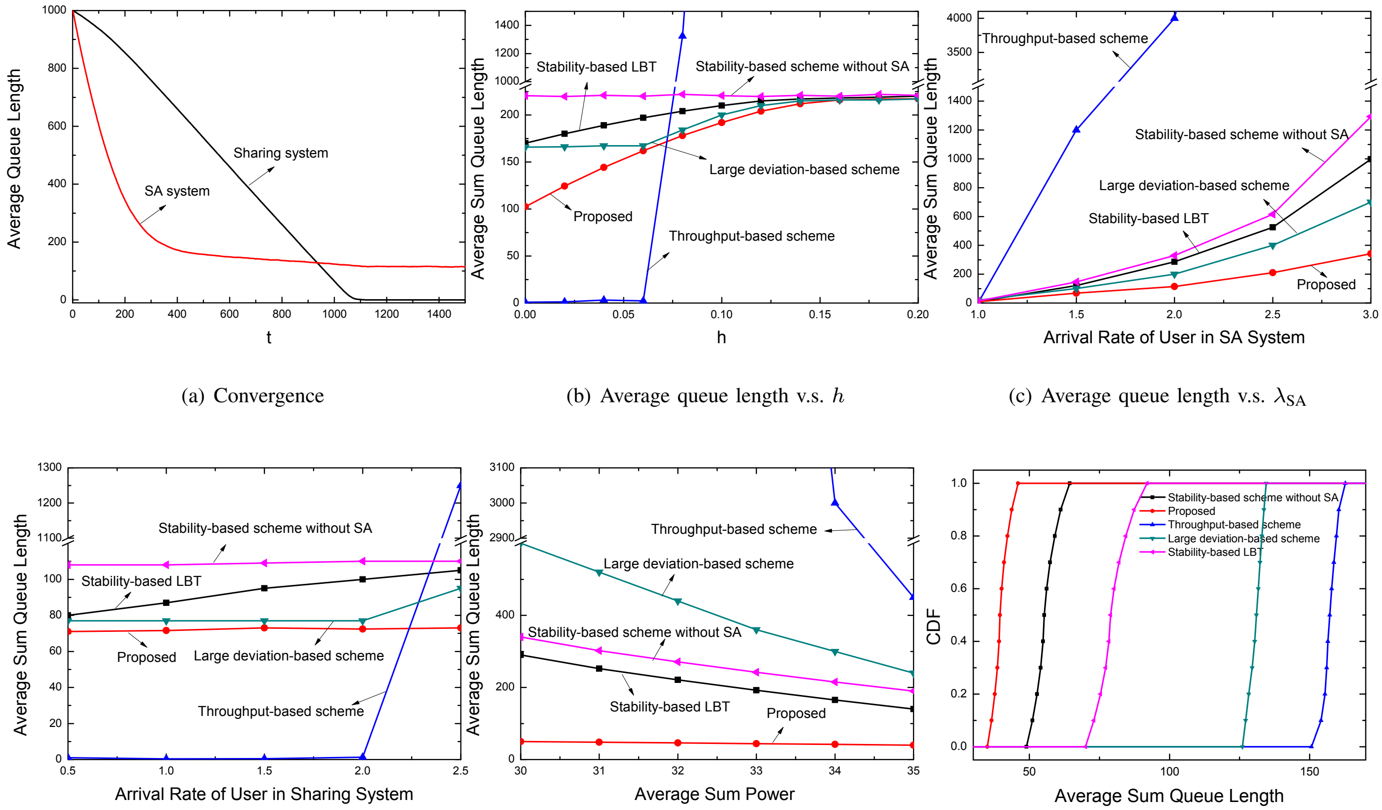
$$p_1^0(t)1_{Q_0(t)>0} \leq d,$$

where d represents the upper bound of the power allocated to the shared channel for the SA user pairs when the shared channel is engaged. When d is set to zero, we do not allow simultaneous transmission on the shared channel. By incorporating such constraint, the capacity region of the SA user pairs will shrink, while that of the sharing user pairs will be larger.

VII. Simulation

In this section, we evaluate the performance of the proposed schemes by simulation. First, the characteristics of the proposed schemes are analyzed, including the queue stability property and the influence of key parameters. Second, the performance of the proposed schemes are compared with those of conventional schemes. For performance comparison, we adopt four baseline schemes:

- *Baseline 1 (Stability-based LBT)*: The SA user pairs access the shared channels according to LBT protocol [8], and allocate the power for queue stability using Lyapunov optimization.
- *Baseline 2 (Throughput-based scheme)*: The SA user pairs access the shared channels to maximize the total throughput without considering the queue information [6].
- *Baseline 3 (Stability-based scheme without SA)*: The SA user pairs stabilize the queues by only using the dedicated channels with the conventional Laypunov optimization [19].
- *Baseline 4 (Large deviation-based scheme)*: The SA user pairs access the shared channels to maximize the total



(a) Convergence

(b) Average queue length v.s. h

(c) Average queue length v.s. λ_{SA}

(d) Average queue length v.s. λ_{sharing}

(e) Average queue length v.s. average power

(f) CDFs of different schemes

Fig. 5. Performance comparison for the single-pair case

throughput, while ensuring the QoS of the sharing user pairs using large deviation [16].

To better connect the proposed analytical framework to the real-world system, we further perform a system-level Monte Carlo simulation to demonstrate that the proposed framework can be applied to the coexistence of LTE-U/Wi-Fi systems.

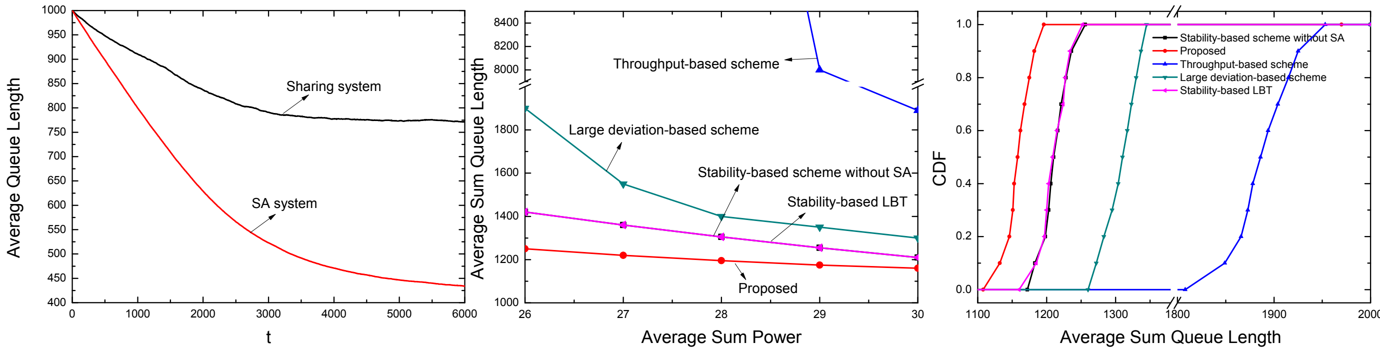
A. Single-Pair Case

In this simulation, we consider the single-pair case where Poisson packet arrival is assumed with the average arrival rates are 2.5 packets per time slot for the SA user pair, and 1.5 packets per time slot for the sharing user pair. The channels of the transmission links obey the Rayleigh distribution with the fading coefficient 0.09 and are i.i.d. over time slots. The cross link interference is set to 0.01 [29]. The transmit power of the sharing user pair is fixed to be 40dBm [30]. To obtain the average and the cumulative distribution functions (CDFs) of the performance, we run the simulation for 100 times, each of which includes 6000 time slots.

Fig. 4 analyzes the effects of some key parameters. It is shown that the average queue length grows almost linearly with V, which is quite consistent with the result in (18). We can adjust the power consumption by adjusting V to fulfill certain constraints on average power consumption. Since the transmission of the SA user pair mainly relies on the dedicated channel, the dedicated channel quality has a huge impact on performance. The average queue length changes slightly when varying the percentage change of the estimated channel quality, which demonstrates the insensitivity to β of the proposed algorithm. The average sum queue lengths of the instances are concentrated to a small region by adopting Lyapunov optimization.

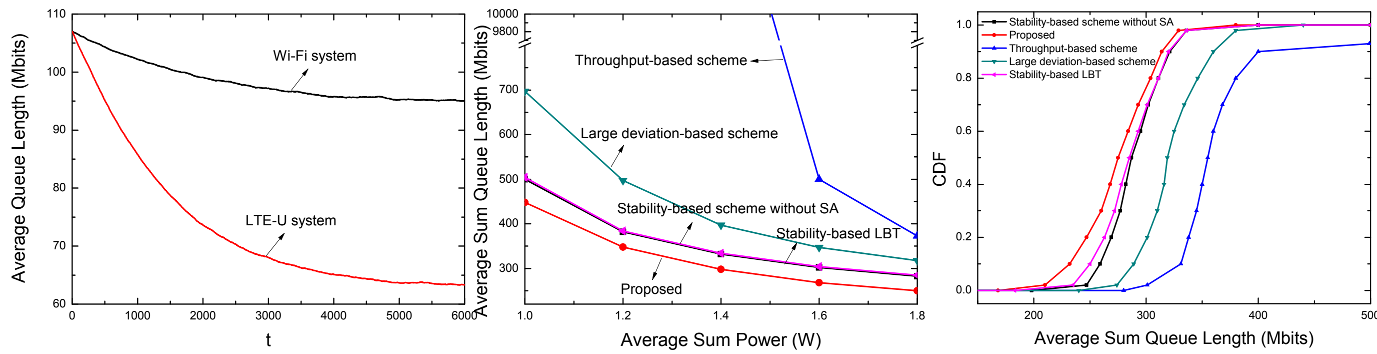
Fig. 5 compares the performance of the proposed scheme and the four baselines. The stability of both systems is verified. The performance of the proposed scheme outperforms the four conventional schemes, because it dynamically allocates the power to each channel. Even though both the proposed scheme and stability-based LBT allow dynamic power control and incorporate SA technique, we can see a larger performance gap with a smaller cross link interference or a larger average arrival rate of the sharing system. The performance of stability-based scheme without SA is worse than the proposed scheme and stability-based LBT, which implies that the performance can be improved by incorporating SA technique. For the cases with high cross interference, low transmit power or high arrival rate for the user pair in the SA system, the throughput-based scheme is not stable, which signifies the importance of devising a queue-based stable scheme. The large deviation-based scheme does not consider the queue length of the SA user pair when allocating power. As a result, it consumes more power than schemes based on Lyapunov optimization.

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/TWC.2018.2796576, IEEE Transactions on Wireless Communications



(a) Convergence (b) Average queue length v.s. average power consumption (c) CDFs of different schemes

Fig. 6. Stability and performance for the multi-pair case



(a) Convergence (b) Average queue length v.s. average power consumption (c) CDFs of different schemes

Fig. 7. Stability and performance for the coexistence of LTE-U/Wi-Fi systems

B. Multi-Pair Case

In this simulation, there are 6 user pairs in the sharing system and 4 user pairs in the SA system. 2 shared channels are used for the coexistence of both systems. Poisson packet arrival is assumed with the average arrival rates are 1, 1.5, 1.5, 2 packets per time slot respectively for the SA user pairs, and 1, 1, 1.5, 1.5, 2, 2 packets per time slot respectively for the sharing user pairs, which are within the capacity region. The rest parameters are the same as those in the single-pair case.

Fig. 6 demonstrates the stability of the proposed scheme and provides the performance comparison of the four schemes for the multi-pair case. Since the average queue length of the sharing system will not go to zero, the stability-based LBT scheme is almost reduced to the stability-based scheme without SA. The slope of the CDF of the throughput-based scheme is less than that of the Lyapunov-based schemes, because the queue lengths are not taking into consideration under the throughput-based scheme. The proposed scheme also outperforms the baseline schemes. Compared to the single-pair case, the performance gap is smaller, due to the higher cross interference brought by multiple user pairs.

C. Coexistence of LTE-U/Wi-Fi systems

We perform a system-level Monte Carlo simulation to demonstrate that the proposed framework can be applied to the coexistence of LTE-U/Wi-Fi systems. Following the settings in [12], [31], [32], we consider a system with one LTE-U user pair and 5 Wi-Fi user pairs, which are randomly deployed in a rectangular field of $500m \times 500m$. The transmit power is 30dBm for the Wi-Fi user pairs, and that for the LTE-U user pair is dynamically controlled according to the proposed scheme. The additive Gaussian noise power is $N_0 = -174$ dBm/Hz and the bandwidth is $B = 20$MHz. The channels of the transmission links follow the Rayleigh distribution with the fading coefficient 0.09 and are i.i.d. over time slots. The channel gain $g_{ij}^k(t)$ from transmitter i to receiver j using channel k is calculated according to $g_{ij}^k(t) = q_{ij}^k(t)(l_{ij})^{-4}/(BN_0)$, where $q_{ij}^k(t)$ represents the Rayleigh fading, and l_{ij} is the physical distance between transmitter i and receiver j. The average arrival rates are 4Mbit/s, 4.5Mbit/s, 5Mbit/s, 5.5Mbit/s, 6Mbit/s respectively for the Wi-Fi user pairs, and 30Mbit/s for the LTE-U user pair. The spectrum sensing duration of each Wi-Fi user pair is 20ms, and that of the LTE-U user pair is 20.5ms, which is set slightly longer than that of the Wi-Fi user pairs, such that the Wi-Fi user pairs are not aware of the existence of the LTE-U user pair. The transmission time is 60ms. We run the simulation in 500 randomly generated scenarios to obtain the average and the CDFs [5] of system performance.

Fig. 7 demonstrates the stability of the proposed scheme and provides the performance comparison of LTE-U/Wi-Fi systems. The gap between the CDF of the proposed scheme and those of the stability-based scheme without SA/stability-based LBT is large when the average sum queue length is small, which is consistent with the results in Fig. 5(b). Small average sum queue length reflects good channel condition and small cross interference, which leads to a significant performance improvement of the proposed scheme.

VIII. Conclusions

In this paper, we develop an analytical framework for heterogeneous SA which aggregates both dedicated and shared channels. To make all systems on the shared channels stable, we design a resource allocation algorithm for the coexistence of multiple systems from a stability perspective. For the single-pair case, we propose the modified water-filling power allocation by the first-order approximation, and analyze the influence of different parameters. We prove the queue stability with the proposed scheme even if the approximation is adopted and the queue information is estimated. For the multi-pair case, we propose a two-step suboptimal resource allocation algorithm by combining the graph coloring and bipartite matching model. By simulation, we discuss the influence of the cross interference and the power price to the system performance, and verify the queue stability with the proposed scheme. The simulation results also show the superiority of the proposed schemes.

Appendix A
Proof of Theorem 1

We adopt the first-order approximation according to Taylor expansion [33] that

$$\log_2\left(1+p_1^0(t)g_{10}^0(t)\right)=\ln 2p_1^0(t)g_{10}^0(t)+o(\delta), \tag{32}$$

$$\begin{aligned}&\log_2\left(1+p_1^0(t)g_{10}^0(t)+p_m g_{00}^0(t)\right)=\\&\log_2\left(1+p_m g_{00}^0(t)\right)+\frac{\ln 2}{1+p_m g_{00}^0(t)}p_1^0(t)g_{10}^0(t)+o(\delta).\end{aligned} \tag{33}$$

Substituting (32) and (33) into the optimization problem (10), we have

$$\begin{aligned}\max_{p_1^0(t),p_1^1(t)}\ & Q_0(t)\left(\log_2\left(1+p_m g_{00}^0(t)\right)+\frac{\ln 2}{1+p_m g_{00}^0(t)}p_1^0(t)g_{10}^0(t)\right)\\&-\ln 2Q_0(t)(p_1^0(t)g_{10}^0(t))+o(\delta)\\&+Q_1(t)\log_2\left(1+p_1^1(t)g_{11}^1(t)\right)\\&+Q_1(t)\log_2\left(1+1_{\{Q_0(t)>0\}}p_m g_{01}^0(t)+p_1^0(t)g_{11}^0(t)\right)\\&-Q_1(t)\log_2\left(1+1_{\{Q_0(t)>0\}}p_m g_{01}^0(t)\right)-V\left(p_1^0(t)+p_1^1(t)\right).\end{aligned} \tag{34}$$

Taking partial derivative of (34) with respect to $p_1^0(t)$ and let it equal to zero,

$$\begin{aligned}&Q_0(t)\frac{1}{1+p_m g_{00}^0(t)}g_{10}^0(t)-Q_0(t)g_{10}^0(t)\\&+Q_1(t)\frac{g_{11}^0(t)}{1+1_{\{Q_0(t)>0\}}p_m g_{01}^0(t)+p_1^0(t)g_{11}^0(t)}+o(\delta)-\frac{V}{\ln 2}=0.\end{aligned} \tag{35}$$

Note that (35) is a linear equation of $p_1^0(t)$. By solving this equation, Theorem 1 is proved.

Appendix B
Proof of Theorem 2

According to the approximation (32), since the first derivative of the function $\log_2(1+x)$ is less than 1 for all positive x, we obtain that $o(\delta)<0$. According to the approximation (33), since $\frac{1}{1+p_m g_{00}^0(t)}<1$, we obtain that $|o(\delta)|$ in (32) is greater than $|o(\delta)|$ in (33). After adopting those approximations, $o(\delta)$ in (34) satisfies $o(\delta)<0$. Those approximations actually approximate $r_0(t)$ by $r_0'(t)-o(\delta)$. Instead of optimizing the original optimization problem (10), we optimize an approximated optimization problem (34) as

$$\max_{p_1^0(t),p_1^1(t)} Q_0(t)r_0'(t)+Q_1(t)r_1(t)-V\left(p_1^0(t)+p_1^1(t)\right). \tag{36}$$

To implement Lyapunov optimization, we adopt the min-drift algorithm in [19]. The corresponding Lyapunov drift plus penalty of the original problem is

$$\begin{aligned}&\Delta\left(Q_0(t),Q_1(t)\right)+V\mathbb{E}[\mathbf{p}(t)|Q_0(t),Q_1(t)]\\&\le B+2Q_0(t)\lambda_0-2Q_0(t)\mathbb{E}[r_0(t)]+2Q_1(t)\lambda_1\\&\quad-2Q_1(t)\mathbb{E}[r_1(t)]+V\Phi(\lambda_0+\epsilon,\lambda_1+\epsilon).\end{aligned} \tag{37}$$

By substituting $r_0'(t)=r_0(t)+o(\delta)$ into (37), we obtain the corresponding Lyapunov drift plus penalty of the approximate problem, it follows that

$$\begin{aligned}&\Delta\left(Q_0(t),Q_1(t)\right)+V\mathbb{E}[\mathbf{p}(t)|Q_0(t),Q_1(t)]\\&\le B'-2Q_0(t)\epsilon-2Q_1(t)\epsilon+V\Phi(\lambda_0+\epsilon,\lambda_1+\epsilon).\end{aligned} \tag{38}$$

Summing over t from 0 to T and taking expectation, we have

$$\begin{aligned}&\mathbb{E}[L\left(Q_0(T),Q_1(T)\right)]-\mathbb{E}[L\left(Q_0(0),Q_1(0)\right)]+V\sum_{t=0}^{T}\mathbb{E}[\mathbf{p}(t)]\\&\le TB'-2\epsilon\sum_{t=0}^{T}Q_0(t)-2\epsilon\sum_{t=0}^{T}Q_1(t)+V\Phi(\lambda_0+\epsilon,\lambda_1+\epsilon).\end{aligned} \tag{39}$$

Thus, we have

$$\frac{1}{T}\sum_{t=0}^{T}\mathbb{E}[\mathbf{p}(t)]\le\Phi(\lambda_1+\epsilon,\lambda_2+\epsilon)+\frac{B'}{V}+\frac{\mathbb{E}[L\left(Q_0(0),Q_1(0)\right)]}{2VT}, \tag{40}$$

$$\begin{aligned}&\frac{1}{T}\sum_{t=0}^{T}Q_0(t)+\frac{1}{T}\sum_{t=0}^{T}Q_1(t)\\&\le\frac{B'+V(\Phi(\lambda_0+\epsilon,\lambda_1+\epsilon)-\sum_{t=0}^{T}\mathbb{E}[\mathbf{p}(t)])}{2\epsilon}+\frac{\mathbb{E}[L\left(Q_0(0),Q_1(0)\right)]}{2\epsilon T}.\end{aligned} \tag{41}$$

Letting $T\to\infty$, the theorem is proved.

Appendix C
Proof of Theorem 3

Despite of the relationship between $Q_0(t)$ and $\hat{Q}_0(t)$, we always have $1_{\{\hat{Q}_0(t)>0\}}=1_{\{Q_0(t)>0\}}$, which implies that the estimation error in $1_{\{\hat{Q}_0(t)>0\}}$ does not affect the output rates of both user pairs. Thus, we only need to evaluate the influence brought by ω on the estimation of $\hat{Q}_0(t)$ in the first term of $p_1^0(t)$ in (22).

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/TWC.2018.2796576, IEEE Transactions on Wireless Communications

Using the queue estimation, we are stabilizing a new system with arrival rate $\boldsymbol{\lambda}+\omega\boldsymbol{I}$. As for the new system, the network capacity region is the set Λ of all non-negative rate vectors $\boldsymbol{\lambda}+\omega\boldsymbol{I}+\alpha\boldsymbol{I}$ for which $\alpha \geq 0$, where the value of α represents a measure of the distance between the rate vector $\boldsymbol{\lambda}+\omega\boldsymbol{I}$ and the capacity region boundary [19].

For any $\omega > 0$, if it satisfies that $\boldsymbol{\lambda}+\omega\boldsymbol{I}$ is interior to the capacity region Λ, then the original rate vectors $\boldsymbol{\lambda}$ is also interior to the capacity region Λ. Therefore, we can safely draw the conclusion that if the new system can be stabilized, the original system is also stabilized.

For the new system, we optimize an approximate problem with queue estimation as

$$\max_{p_1^0(t),p_1^1(t)} \hat{Q}_0(t)r_0'(t)+Q_1(t)r_1(t)-V\big(p_1^0(t)+p_1^1(t)\big). \quad (42)$$

The corresponding Lyapunov drift plus penalty of the problem becomes

$$\begin{aligned}&\Delta\big(\hat{Q}_0(t),Q_1(t)\big)+V\mathbb{E}[\mathbf{p}(t)|\hat{Q}_0(t),Q_1(t)]\\ \leq& B+2\hat{Q}_0(t)\lambda_0-2\hat{Q}_0(t)\mathbb{E}[r_0(t)]+2Q_1(t)\lambda_1\\ &-2Q_1(t)\mathbb{E}[r_1(t)]+V\Phi(\lambda_0+\epsilon+\omega,\lambda_1+\epsilon+\omega).\end{aligned} \quad (43)$$

Following a similar proof of Theorem 2, summing over t from 0 to T, taking expectation, and letting $T\rightarrow\infty$, we have the theorem proved.

Appendix D
Proof of Theorem 4

According to [34], for the sum rate utility function, the spectrum power management problem is strongly NP-hard, where the problem is defined as

$$\begin{aligned}&\max \ \sum_{k=1}^{K}\log_2\left(1+\frac{s_k}{\sigma_k+\sum_{j\neq k}\alpha_{kj}s_j}\right)\\ &s.t.\ 0\leq s_k\leq P_k, \forall k\in\mathcal{K},\end{aligned} \quad (44)$$

where s_k denotes the power of transmitter k, α_{kj} denotes cross interference from transmitter k to receiver j and σ_k denotes the noise level.

By letting $Q_i(t)=0, \forall i\in\mathcal{K}$ and $Q_i(t)=1, \forall i\in\mathcal{N}$ and give $\mathbf{b}(t)$, the problem (26) can be reduced to a similar problem to (44) and hence the complexity of solving (26) is not less than that of (44). Therefore, the problem (26) is strongly NP-hard.

References

[1] W. Wang, Z. Zhang and A. Huang, “Spectrum aggregation: overview and challenges,” *Network Protocols and Algorithms*, vol. 2, no. 1, pp. 184-196, May 2010.

[2] Qualcomm, ”Making the best use of unlicensed spectrum for 1000x,” Sept. 2015. https://www.qualcomm.com/documents/whitepaper-making-best-use-unlicensed-spectrum

[3] B. Chen, J. Chen, Y. Gao and J. Zhang, “Coexistence of LTE-LAA and Wi-Fi on 5 GHz with corresponding deployment scenarios: A survey,” *IEEE Commun. Surveys Tuts.,* vol. 19, no. 1, pp. 7-32, Mar. 2017.

[4] A. Cavalcante, E. Almeida, R. Vieira, S. Choudhury, E. Tuomaala and K. Doppler, “Performance evaluation of LTE and Wi-Fi coexistence in unlicensed bands,” *Proc. of IEEE VTC 2013-Spring*, Jun. 2013.

[5] O. Holland, A. Aijaz, F. Kaltenberger, F. Foukalas, M. Buczkowski and S. Pietrzyk, “Management architecture for aggregation of heterogeneous systems and spectrum bands,” *IEEE Commun. Mag.*, no. 54, vol. 9, pp. 112-118, Sept. 2016.

[6] S. Singh, S. Yeh, N. Himayat, and S. Talwar, “Optimal traffic aggregation in multi-RAT heterogeneous wireless networks,” *Proc. of. IEEE ICC 2016*, pp. 626-631, May 2016.

[7] M. Srinivasan, V. J. Kotagi and C. S. R. Murthy, “A Q-learning framework for user QoE enhanced self-organizing spectrally efficient network using a novel inter-operator proximal spectrum sharing,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 11, pp. 2887-2901, Nov. 2016.

[8] Y. Wu, W. Guo, H. Yuan, L. Li, S. Wang, X. Chu and J. Zhang, “Device-to-device (D2D) meets LTE-unlicensed,” *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 154-159, May 2016.

[9] Q. Zhang, Q. Wang, Z. Feng and T. Yang, “Design and performance analysis of a fairness-based license-assisted access and resource scheduling scheme,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 11, pp. 2968-2980, Oct. 2016.

[10] Y. Liu, G. Wang, M. Xiao and Z. Zhong, “Spectrum sensing and throughput analysis for cognitive two-way relay networks with multiple transmit powers,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 11, pp. 3038-3047, Oct. 2016.

[11] M. R. Khawer, J. Tang and F. Han, “usICIC - A proactive small cell interference mitigation strategy for improving spectral efficiency of LTE networks in the unlicensed spectrum,” *IEEE Trans. Wireless Commun.*, vol. 15, pp. 2303-2311, Mar. 2016.

[12] Q. Chen, G. Yu, and Z. Ding, “Optimizing unlicensed spectrum sharing for LTE-U and WiFi network coexistence,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 10, pp. 2562-2574, Oct. 2016.

[13] G. I. Tsiropoulos, O. A. Dobre, M. H. Ahmed and K. E. Baddour, “Radio resource allocation techniques for efficient spectrum access in cognitive radio networks,” *IEEE Commun. Surveys Tuts.,* vol. 18, no. 1, pp. 824-847, Mar. 2016.

[14] P. D. Diamantoulakis, K. N. Pappi, S. Muhaidat, G. K. Karagiannidis and T. Khattab, “Carrier aggregation for cooperative cognitive radio networks,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 5904-5918, Jul. 2017.

[15] W. Wang and V. K. N. Lau, “Delay-aware cross-layer design for device-to-device communications in future cellular systems,” *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 133-139, Jun. 2014.

[16] Y. Cui, V. K. N. Lau, R. Wang, H. Huang and S. Zhang, “A survey on delay-aware resource control for wireless systems: Large derivation theory, stochastic Lyapunov drift and distributed stochastic learning,” *IEEE Trans. Info Theory*, vol. 58, no. 3, pp. 1677-1700, Mar. 2012.

[17] E. M. Yeh, “Multiaccess and Fading in Communication Networks,” *Ph.D. dissertation*, MIT, Sept. 2001.

[18] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd Ed. Athena Scientific, 2007.

[19] M. J. Neely, “Stochastic network optimization with application to communication and queueing systems,” *Synthesis Lectures on Communication Networks*, vol. 3, no. 1, pp. 1-211, May 2010.

[20] Z. Jiang and S. Mao, “Inter-operator opportunistic spectrum sharing in LTE-unlicensed,” *IEEE Trans. Veh. Technol.*, pp. 1-11, Sept. 2016.

[21] Y. Wang, W. Wang, L. Chen, P. Zhou and Z. Zhang, “Energy efficient scheduling for delay-constrained spectrum aggregation: A differentiated water-filling approach,” *IEEE Trans. Green Commun. Netw.*, Dec. 2017.

[22] F. E. Lapiccirella, X. Liu and Z. Ding, “Distributed control of multiple cognitive radio overlay for primary queue stability,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 112-122, Jul. 2013.

[23] W. A. Rosenkrantz, “Little’s theorem, a stochastic integral approach,” *Queueing Systems*, vol. 12, pp. 319-324, Jan. 1992.

[24] W. Wang, F. Zhang and V. K. N. Lau, “Dynamic power control for delay-aware device-to-device communications,” *IEEE J. Sel. Areas Commun.*, vol. 33, no. 1, pp. 14-27, Jan. 2015.

[25] C. Peng, H. Zheng and B. Zhao, “Utilization and fairness in spectrum assignment for opportunistic spectrum access,” *Mobile Networks and Applications,* vol. 11, no. 4, pp. 555-576, May 2006.

[26] W. Yu, “Multiuser water-filling in the presence of crosstalk,” *Proc. of IEEE ITA 2007*, Jan. 2007.

[27] D. Yang, X. Fang, and G. Xue, “OPRA: Optimal relay assignment for capacity maximization in cooperative networks,” *Proc. of IEEE ICC 2011*, Jun. 2011.

[28] W. Cook, W. Cunningham, W. Pulleyblank, and A. Schrijver, *Combinatorial Optimization,* John Wiley and Sons, 1998.

[29] S. Movassaghi, M. Abolhasan, D. Smith and A. Jamalipour, “Aim: adaptive internet-work interference mitigation amongst co-existing wireless

body area networks," *Proc. of IEEE GLOBECOM 2014*, pp. 2460-2465, Dec. 2014.

[30] B. S. Park, J. G. Yook and H. K. Park, "The determination of base station placement and transmit power in an inhomogeneous traffic distribution for radio network planning," *Proc. of IEEE VTC-Fall 2002,* pp. 2051-2055, Sept. 2012.

[31] S. Ping, A. Aijaz, O. Holland, and A. H. Aghvami, "Energy and interference aware cooperative routing in cognitive radio ad-hoc networks," *Proc. of IEEE WCNC 2014*, Apr. 2014.

[32] S. Ping, A. Aijaz, O. Holland, and A. H. Aghvami, "SACRP: A spectrum aggregation-based cooperative routing protocol for cognitive radio ad-hoc networks," *IEEE Trans. Commun.*, vol. 63, no. 6 pp. 2015-2030, Jun. 2015.

[33] C. R. Wylie, *Advanced Engineering Mathematics,* 4th ed. McGraw-Hill Book Co., 1975.

[34] Z. Luo and S. Zhang, "Dynamic spectrum management: complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57-73, Feb. 2008.

Yitu Wang (S'16) received the B.S. degree from Zhejiang University, Hangzhou, China, in 2013. From August to November 2014, he was a visiting student with the University of Paris-Sud, Orsay, France. He is currently a Ph.D candidate at Zhejiang University, Hangzhou, China. His research interests mainly focus on stochastic optimization for cross-layer resource allocation in wireless networks and cache-enabled networks.

Wei Wang (S'08-M'10-SM'15) received the B.S. and Ph.D. degrees from the Beijing University of Posts and Telecommunications, China, in 2004 and 2009, respectively. From 2007 to 2008, he was a Visiting Student with the University of Michigan, Ann Arbor, USA. From 2013 to 2015, he was a Hong Kong Scholar with the Hong Kong University of Science and Technology, Hong Kong. He is currently an Associate Professor with the College of Information Science and Electronic Engineering, Zhejiang University, China. His research interests mainly focus on stochastic optimization for cross-layer resource allocation in wireless networks, and caching and computing in wireless networks. He is the Editor of the book entitled Cognitive Radio Systems, and serves as an Editor of the IEEE Access, Transactions on Emerging Telecommunications Technologies, and KSII Transactions on Internet and Information Systems.

Vincent K. N. Lau (F?11) received the B.Eng. degree (Hons.) from the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, in 1992, and the Ph.D. degree from the University of Cambridge, Cambridge, U.K., in 1997. He was a member of technical staff with Bell Labs, Lucent Technologies Inc., NJ, USA, from 1997 to 2004. He is currently a Chair Processor with the Department of ECE, The Hong Kong University of Science and Technology (HKUST), and the Founding Director of the Huawei-HKUST Joint Innovation Laboratory. His research interests include the delay-sensitive cross-layer optimization of MIMO/OFDM wireless systems, cooperative communications, stochastic approximation, and the Markov decision process.

He is a fellow of HKIE, Changjiang Chair Professor, and Croucher Senior Research Fellow. He received four IEEE best paper awards. He served as a Senior Area Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the IEEE SIGNAL PROCESSING LETTERS, a Guest Editor of the IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS, a Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, a Book-Series Editor of the Information and Communications Technologies book series (John Wiley and Sons).

Lin Chen (S'07-M'10) received his B.E. degree in Radio Engineering from Southeast University, China in 2002 and the Engineer Diploma from Telecom ParisTech, Paris in 2005. He also holds a M.S. degree of Networking from the University of Paris 6. He currently works as associate professor in the department of computer science of the University of Paris-Sud. He serves as Chair of IEEE Special Interest Group on Green and Sustainable Networking and Computing with Cognition and Cooperation, IEEE Technical Committee on Green Communications and Computing. His main research interests include modeling and control for wireless networks, distributed algorithm design and game theory.

Zhaoyang Zhang (M'10) received the Ph.D. degree in communication and information systems from Zhejiang University, Hangzhou, China, in 1998. He is currently a Full Professor with the College of Information Science and Electronic Engineering, Zhejiang University. He has coauthored more than 150 refereed international journal and conference papers, as well as two books in his areas of interest. His research interests are mainly focused on information theory and coding theory, signal processing techniques, and their applications in wireless communications and networking.

Dr. Zhang coreceived three conference Best Paper Awards/Best Student Paper Award. He is currently serving as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, IET Communications, and several other international journals. He has served as a Technical Program Committee Cochair or a Symposium Cochair for many international conferences, such as the 2013 International Conference on Wireless Communications and Signal Processing and the 2014 IEEE Global Communications Conference Wireless Communications Symposium.