Energy Efficient Scheduling for Delay-Constrained Spectrum Aggregation

Yitu Wang* Wei Wang* Lin Chen† Zhaoyang Zhang*

*Department of Information Science and Electronic Engineering
Zhejiang Key Lab. of Information Network Technology
Zhejiang University, Hangzhou 310027, P.R. China
Email: wangw@zju.edu.cn

†Lab. de Recherche en Informatique (LRI)
University of Paris-Sud 11
Orsay, France, 91405 Cedex
Email: Lin.Chen@lri.fr

***Abstract*—In this paper, we construct an analytical design framework for energy efficient scheduling for delay-constrained spectrum aggregation (ESSA), where the practical hardware limitations on SA capability bring various technical challenges. Specifically, the conventional water-filling power control cannot be adopted over all the channels, and the delay-aware scheduling solution should interact with the channel allocation. To overcome these challenges, we design the ESSA scheduling scheme in two steps. First, with given rate vector and channel allocation, we minimize the total power consumption for SA, including both the transmit power as well as the circuit power. Due to the properties of delay-constrained SA, we divide the scheduled users into conforming and nonconforming user sets, and design their water-filling power allocation strategies differentially. Second, based on the differentiated water-filling power control, we optimize the channel allocation and rate control iteratively via Lyapunov optimization to minimize the power consumption with the average delay constraint. The proposed ESSA scheme is finally evaluated by simulation results.**

I. Introduction

Spectrum aggregation (SA) [1] has its distinctive value in wireless communication systems, which enables the device to provide homogeneous broadband service by bonding heterogeneous fragmentary spectrum resources. Recently, SA becomes one of the key features during LTE-Advanced standardization.

There have been a few research works on SA. In [2], a heuristic suboptimal algorithm considering both the efficiency and the fairness is proposed for SA by optimizing two metrics separately to lower the complexity. An optimal one is proposed later in [3] in a two-carrier case by using Lagrangian methods. In [4], an energy efficient SA scheme is developed without taking the delay performance into consideration. In [5], a scheduling algorithm is proposed to balance the throughput and the delay for SA, in which the delay is reflected by blocking probability. The issue of hardware limitation for spectrum aggregation is discussed in [6], where the aggregation range is restricted.

Future wireless systems have exigent requirements to support higher data rate and more real-time services [7]. Wireless transmission with high data rate costs enormous amount of energy, and real-time services have a strict delay requirement. The power consumption and the delay performance become a crucial part towards reliability and stability of wireless systems [8]. However, few publication studies the energy efficient schemes for SA with delay consideration, which is an emerging technique to support broadband real-time services.

The tradeoff between power and delay is discussed in [9][10], where the former solves the problem by Markov decision process in a single user case while the latter one addresses the problem using Lyapunov method which can be applied to multi-user but single-channel cases in wireless systems. With the SA capability, a device can adjust the number of channels adaptively according to the service demands, which provides an extra degree of freedom to achieve energy efficiency. To investigate the energy efficient scheduling for delay-constrained SA, there are several technical challenges involved as follows:

- **SA Hardware Limitation**: Due to the practical hardware limitation [6], the aggregation range of SA is restricted. The device can only aggregate a limited number of channels, which brings additional constraint during designing the scheduling scheme. Thus, the conventional water-filling power control cannot be adopted over all the channels.
- **SA Circuit Structure**: The process of SA requires the support of specific circuit structure [11]. The total power consumption varies according to the channel allocation scheme and needs to be taken into consideration for designing the energy efficient scheduling.
- **SA Energy-Delay Tradeoff**: Unlike the conventional wireless systems [10]. the channel bonding in SA leads to complicated rate, power and channel allocation, which are coupled with each other and their effects to the delay is not straightforward.

In spite of the above challenges, SA makes it possible that one user can support the simultaneous transmission over multiple channels. It is obvious that the power consumption is reduced by balancing the water-filling levels across multiple channels used by the same user. If the water-filling levels of more channels are balanced together, the power consumption will be reduced further. In delay-constrained SA systems, it is possible to balance the water-filling levels of the channels across multiple users. Due to the limitation of the aggregation range of SA, we handle the water-filling scheme for the users differentially, i.e., the water-filling levels are balanced across a part of users and are set individually for the other users. The number of users with the same water-filling level depends on the aggregation range of SA.

In this paper, we develop an analytical framework for energy efficient scheduling for delay-constrained SA (ESSA). Taking the SA capability and the circuit power consumption into consideration, ESSA determines the data rate, transmit power and channel allocation to minimize the energy consumption with average delay constraint. The scheduling decisions are made according to both current channel status and current queue backlogs. Specifically, we design the ESSA scheduling algorithm in two steps. First, with given rate vector and channel allocation, we minimize the total power consumption for SA, including both the transmit power as well as the circuit power. The water-filling levels are balanced differentially by partitioning the users into conforming and nonconforming user sets. Second, based on the results of power control, we propose a suboptimal scheduling algorithm based on the Lyapunov optimization to determine the data rate and channel allocation iteratively. The simulation results show that ESSA reduces the power consumption significantly for delay-constrained SA.

The rest of this paper is organized as follows. Section II presents the system model. In Section III, we proposes the design of ESSA algorithm. Following this, the performance of ESSA is evaluated by simulation results in Sections IV. Finally, this paper is concluded in Section V.

II. System Model

In this section, we first introduce the physical layer model of SA systems and outline the key control variables in scheduling. Considering the specified SA circuit structure, we present the circuit power consumption model for SA. Finally, we formulate the scheduling problem to minimize the power consumption of the SA system with the delay constraint.

A. SA System

Consider a wireless SA system, which includes N users sharing K time-varying channels with the same bandwidth. Denote $\mathcal{N}$ and $\mathcal{K}$ as the set of users and channels, respectively, i.e., $\mathcal{N} = \{1, 2, \cdots, N\}$ and $\mathcal{K} = \{1, 2, \cdots, K\}$. Because of the SA capability restricted by the hardware limitation, a user can transmit over at most M channels simultaneously. The time is slotted and the duration of each time slot is τ.

Let $\mathbf{S}(t)$ be the global channel state in slot t, i.e., $\mathbf{S}(t) = \left(S_j^i(t), j \in \mathcal{K}, i \in \mathcal{N}\right)$, where $S_j^i(t)$ represents the state of channel j when user i is using in slot t, which remains constant within a slot and is i.i.d. over time slots.

At the beginning of each slot, a centralized controller schedules the users to transmit and determines the associated scheduling control variables as follows:

- *Data Rate* $\mathbf{r}(t)$: Define $\mathbf{r}(t) = \left\{r_i(t), \forall i \in \mathcal{N}\right\}$, where $r_i(t)$ is the data rate of user i in slot t.
- *Transmit Power* $\mathbf{P}(t)$: Define $\mathbf{P}(t) = \left\{P_j(t), \forall j \in \mathcal{K}\right\}$, where $P_j(t)$ is the transmit power of channel j in slot t.

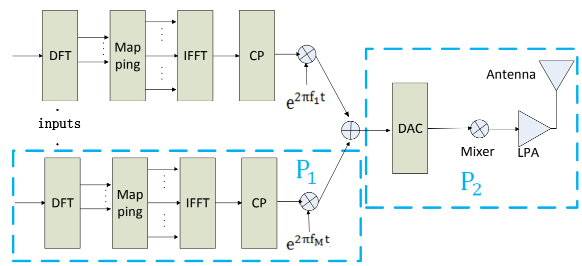


Fig. 1. Illustration of SA circuit structure

- *Channel Allocation* $\mathbf{b}(t)$: Define $\mathbf{b}(t) = \left\{b_j^i(t), \forall i \in \mathcal{N}, j \in \mathcal{K}\right\}$, where $b_j^i(t) \in \{0, 1\}$ and $b_j^i(t) = 1$ represents user i transmits over channel j in slot t.

The data rates $\mathbf{r}_i(t)$ are determined by

$$r_i(t) = \sum_{j \in \mathcal{K}} b_j^i(t) R_j\big(t\big) \tag{1}$$

where R_j is the transmission rate of channel j, which can be calculated as

$$R_j(t) = \log_2\left(1 + S_j^i(t) P_j(t)\right) \tag{2}$$

Note that $\mathbf{b}(t)$ should satisfy $\sum_{j=1}^{K} b_j^i(t) \le M, \forall i \in \mathcal{N}$ with the consideration of SA capability.

B. SA Circuit Structure

Although there are different circuit implementations [11][12] for SA, one of the common characteristics is that the aggregated channels can share a part of circuit modules. Considering the specified structure for SA circuit as illustrated in Fig. 1, we divide the SA circuit into two parts as follows:

- *Individual Modules*: Each set of individual modules provides the processing for a single channel.
- *Shared Modules*: The set of shared modules provides the functions shared by the channels aggregated by a device.

Denote P_1 as the power consumption of a set of individual modules and P_2 as that of a set of shared modules[1]. Based on the specified SA circuit structure, the total circuit power consumption P_c can be modeled as

$$P_c(t) = \sum_{i=1}^{N} \sum_{j=1}^{K} b_j^i(t) P_1 + \sum_{i=1}^{N} \left(1 - \prod_{j=1}^{K} (1 - b_j^i(t))\right) P_2 \tag{3}$$

Note that the shared modules consume the power if at least one of the channels is used.

C. Problem Formulation

Since both power and delay are critical performance metrics in wireless systems, there is an inherent tradeoff between the power consumption and the delay performance.

To analyze the average delay, we first discuss the packet queue backlog, since the average delay can be measured by average queue length according to Little's theorem [14]. Each

[1] SA circuit power consumption can be estimated based on hardware datasheets and time spent by the operations [13].

user possesses a packet queue, whose length is denoted as $U_i(t)$ for user i in slot t. Let $\mathbf{A}(t) = \{A_i(t), \forall i \in \mathcal{N}\}$ be the random packet arrivals, where $A_i(t)$ is the number of arrived bits for user i in slot t. Assume that $\mathbf{A}(t)$ is i.i.d. over time, with $E[A_i(t)] = \lambda_i$, where λ_i is the average arrival rate for user i. The queue dynamics of $U_i(t)$ is

$$U_i(t+1) = \max\left\{U_i(t) - r_i(t), 0\right\} + A_i(t) \tag{4}$$

Our goal is to minimize the energy consumption for delay-constrained SA by scheduling, which can be formulated formally as below:

$$\min_{\mathbf{b}(t),\mathbf{P}(t)} \quad \sum_{j=1}^{K} P_j(t) + P_c(t) \tag{5}$$

$$s.t. \quad \frac{\mathbb{E}\left[\sum_{i=1}^{N} U_i(t)\right]}{N} \leq Q \tag{6}$$

$$\mathbb{E}\left[\sum_{j=1}^{K} b_j^i(t) R_j(t)\right] \geq \lambda_i \tag{7}$$

$$\sum_{j=1}^{K} b_j^i(t) \leq M \tag{8}$$

where the constraint (6) implies the average delay constraint, in which Q denotes the target average queue length corresponding to the average delay, (7) guarantees the system stability, and (8) implies the aggregation range due to SA capability.

III. Energy Efficient Scheduling for Delay Constrained Spectrum Aggregation

In this section, we propose a scheduling framework to minimize power consumption with an average delay constraint. We design the scheduling scheme in two steps. First, under data rate vector $\mathbf{r}(t)$ and channel allocation matirx $\mathbf{b}(t)$, we minimize the total power consumption of the SA system. Due to the limitations in delay-constrained SA systems, we partition the scheduled users into conform and nonconform user sets, and determine their water-filling power allocation strategies separately. Second, based on the results of the minimum power consumption, we propose a sub-optimal scheduling algorithm ESSA by Lyapunov optimization, to minimize the power consumption with an average delay constraint.

A. Differentiated Water-Filling Power Allocation

We consider the minimization of power consumption under rate vector $\mathbf{r}(t)$ and channel allocation matrix $\mathbf{b}(t)$, and obtain the minimum power consumption, which provides a basis for designing ESSA.

In SA systems, one user can support the simultaneous transmission over multiple channels, which makes it possible that the power consumption is reduced by balancing the water-filling levels across multiple channels used by the same user[2]. Moving one step ahead, we balance the water-filling levels across users to reduce the power consumption further. In the proposed power control scheme, we consider the total power consumption including both the transmit power and the SA circuit power.

Under rate vector $\mathbf{r}(t)$ and channel allocation matrix $\mathbf{b}(t)$, we adopt the rate constraint[3] instead of the queue length constraint in (6), and rewrite the power minimization problem as

$$\min_{\mathbf{p}(t)} \sum_{j=1}^{K} \sum_{i=1}^{N} b_j^i P_j(t) + P_c(t) \tag{9}$$

$$s.t. \sum_{j=1}^{K} \sum_{i=1}^{N} b_j^i \log_2\left(1 + S_j^i(t) P_j(t)\right) = \sum_{i=1}^{N} r_i(t) \tag{10}$$

$$P_c(t) = \sum_{i=1}^{N} \sum_{j=1}^{K} b_j^i(t) P_1 + \sum_{i=1}^{N} \left(1 - \prod_{j=1}^{K} (1 - b_j^i(t))\right) P_2 \tag{11}$$

$$\mathbb{E}\left[\sum_{j=1}^{K} b_j^i(t) R_j(t)\right] \geq \mathbb{E}[r_i(t)] = \lambda_i \tag{12}$$

$$\sum_{j=1}^{K} b_j^i(t) \leq M \tag{13}$$

where the transmit power $P_j(t)$ can be calculated according to (2), (10) guarantees that K channels provide enough capacity to support transmission and (11) captures the circuit power.

To solve the above optimization problem, we first treat an easier case in which only the constraint (10) is considered, to obtain some insight of the problem. We establish the Lagrangian function according to (9) and (10) as

$$\begin{aligned} Z\big(P_j(t), \gamma\big) = &\sum_{j=1}^{K} \sum_{i=1}^{N} b_j^i(t) P_j(t) \\ &- \gamma \left(\sum_{j=1}^{K} \sum_{i=1}^{N} b_j^i(t) \log_2\left(1 + S_j^i(t) P_j(t)\right) - \sum_{i=1}^{N} r_i(t) \right) \end{aligned} \tag{14}$$

where γ is the Lagrangian multiplier.

By Lagrangian method, i.e., letting the partial derivative of $P_j(t)$ with respect to j equal to 0, we obtain the transmit power over channel j as

$$P_j(t) = \frac{2^{\frac{\sum_{i=1}^{N} r_i(t)}{K}}}{\prod_{k=1}^{K} \sum_{i=1}^{N} b_k^i(t) S_k^i(t)^{\frac{1}{K}}} - \frac{1}{\sum_{i=1}^{N} b_j^i(t) S_j^i(t)} \tag{15}$$

Essentially, (15) provides a water-filling power allocation over all channels.

The corresponding transmission rate of channel j is

$$\begin{aligned} R_j(t) = &\frac{\sum_{i=1}^{N} r_i(t)}{K} \\ &- \sum_{l=1}^{K} \frac{1}{K} \log_2 \sum_{i=1}^{N} b_l^i(t) S_l^i(t) + \log_2 \sum_{i=1}^{N} b_j^i(t) S_j^i(t) \end{aligned} \tag{16}$$

[2] According to Jensen's inequation [15], the energy consumption of satisfying the sum rate over multiple channels is not more than that of satisfying the rate requirement over each corresponding channel, because the data rate is an increasing concave function of the transmit power [16].

[3] The rate vector will be given according to the current queue length and the target Q in the next subsection. Thus, the rate constraint can be adopted to guarantee the queue length constraint.

Now, we take the constraints on the system stability and SA limitation in (12) and (13) into consideration. In this case, if we balance the water-filling levels for all users, it is possible that the total transmit rates of M channels cannot support the arrival rate λ_i, which leads to the instability of the system. To address this issue, we classify the users into two categories in the following definition:

Definition 1 (Conform/Nonconform User Set): A set of users $\mathcal{U}$ is said to be a conform user set if it is satisfied that

$$0 < \lambda_i \leq M \frac{\sum_{l \in \mathcal{U}} r_l(t)}{|\mathcal{C}|}, \forall i \in \mathcal{U} \tag{17}$$

where $\mathcal{C}$ is the set of channels allocated to the users in $\mathcal{U}$. On the other hand, the scheduled users with a positive rate but not in $\mathcal{U}$ are called nonconform users. The set of nonconform users is denoted as $\mathcal{U}^-$ and the set of channels allocated to user $i \notin \mathcal{U}$ is denoted as $\mathcal{C}_i$. ■

Remark 1 (Partition of Conform/Nonconform Users): To determine the conform user set $\mathcal{U}$, we can initialize $\mathcal{U} = \mathcal{N}$ and remove the users who satisfy $\lambda_i > M \frac{\sum_{l \in \mathcal{U}} r_l(t)}{|\mathcal{C}|}$ iteratively until all of the users in $\mathcal{U}$ satisfy the condition of conform user set in Definition 1. ■

Considering the power consumption of SA circuit (11), it is not always to transmit over all the channels. Define $\mathcal{C}'$ as the set of transmitting channels in the set $\mathcal{C}$ and $\mathcal{C}'_i$ as the set of transmitting channels in the set $\mathcal{C}_i$.

Because of the system stability requirement (12) and the limitation of SA capability (13), the water-filling levels cannot be balanced across all channels but just the channels in $\mathcal{C}'$ for the conform users in $\mathcal{U}$. By handling the conform users and the nonconform users respectively, we obtain the minimum total power consumption in the following theorem.

Theorem 1 (Minimum Power Consumption): Under the rate vector $\mathbf{r}(t)$ and the channel allocation matrix $\mathbf{b}(t)$, the minimum power consumption $\phi\big(\mathbf{r}(t), \mathbf{b}(t)\big)$ of SA system is

$$\begin{aligned} \phi\big(\mathbf{r}(t), \mathbf{b}(t)\big) =& |\mathcal{C}'| \frac{2^{\frac{\sum_{i \in \mathcal{U}} r_i(t)}{|\mathcal{C}'|}}}{\prod_{l \in \mathcal{C}'} \sum_{i \in \mathcal{U}} b_l^i(t) S_l^i(t)^{\frac{1}{|\mathcal{C}'|}}} \\ &+ \sum_{i \in \mathcal{U}^-} |\mathcal{C}'_i| \frac{2^{\frac{r_i(t)}{|\mathcal{C}'_i|}}}{\prod_{l \in \mathcal{C}'_i} \sum_{i \in \mathcal{U}^-} b_l^i(t) S_l^i(t)^{\frac{1}{|\mathcal{C}'_i|}}} \\ &- \sum_{j \in \mathcal{C}'} \frac{1}{\sum_{i \in \mathcal{U}} b_j^i(t) S_j^i(t)} \\ &- \sum_{i \in \mathcal{U}^-} \sum_{j \in \mathcal{C}'_i} \frac{1}{\sum_{i \in \mathcal{U}^-} b_j^i(t) S_j^i(t)} \\ &+ \big(|\mathcal{C}'| + \sum_{i \in \mathcal{U}^-} |\mathcal{C}'_i|\big) P_1 + \big(|\mathcal{U}| + |\mathcal{U}^-|\big) P_2 \end{aligned} \tag{18}$$

Remark 2 (Balancing Water-Filling Levels Across Users): In SA systems, the water-filling power allocation achieves an extra power reduction by balancing water-filling levels across users. Specifically, the water levels of the users in $\mathcal{U}$ are balanced together while those of the users in $\mathcal{U}^-$ are balanced for each individual user, as illustrated in Fig. 2. Obviously, the more channels allocated to the users in $\mathcal{U}$, the less power is consumed because of the benefit achieved by balancing the water-filling levels across more users and channels. Note that a large SA capability M can significantly increase the number of channels with a balanced water-filling level, which contributes to reducing the power consumption. ■

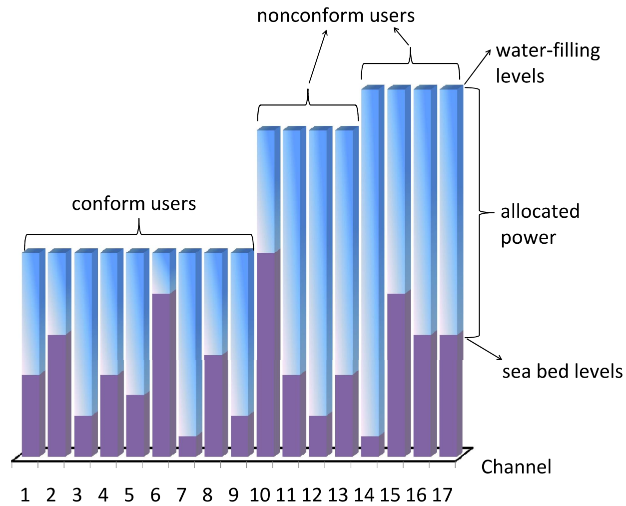


Fig. 2. Differentiated water-filling across users

B. ESSA Algorithm Design

To satisfy the average delay constraint, we adopt the Lyapunov optimization method. The Lyapunov function includes two parts as follows [19]:

$$\psi\big(\mathbf{U}(t), \mathbf{X}(t)\big) = L\big(\mathbf{U}(t)\big) + J\big(\mathbf{X}(t)\big) \tag{19}$$

The function $L\big(\mathbf{U}(t)\big)$ is designed to be exponential, which reaches its minimum when $U_i(t) = Q, \forall i \in \mathcal{N}$, and increases exponentially with the increase of the difference between $U_i(t)$ and Q.

$$L\big(\mathbf{U}(t)\big) = \sum_{i \in \mathcal{N}} \big(e^{\omega(U_i(t) - Q)} + e^{\omega(Q - U_i(t))} - 2\big) \tag{20}$$

where ω is a positive value affecting the rate of exponential increase. This Lyapunov function provides a large enough penalty to push the queue length $U_i(t)$ to the threshold Q.

The function $J\big(\mathbf{X}(t)\big)$ is designed for the stability of the virtual queue $\mathbf{X}(t)$ as

$$J\big(\mathbf{X}(t)\big) = \sum_{i \in \mathcal{N}} X_i^2(t) \tag{21}$$

Both the actual and virtual queues should be stabilized for the whole system stability.

The dynamics of the actual queue $U_i(t)$ and the virtual queue $X_i(t)$ are presented respectively as

$$U_i(t+1) = U_i(t) - r_i(t) + A_i(t) \tag{22}$$

$$\begin{aligned} X_i(t) = \max\big\{&X_i(t) - \big(r_i(t) + \epsilon 1_{U_i(t) < Q}(t)\big), 0\big\} \\ &+ A_i(t) + \epsilon 1_{U_i \geq Q}(t) \end{aligned} \tag{23}$$

To minimize the total power consumption with the average delay constraint, we adopt V as the weight of power consumption using the Lyapunov optimization, which is formulated as

$$\begin{aligned}\min_{\mathbf{r}(t),\mathbf{b}(t)} Y(t) =&V\Bigg(|\mathcal{C}'| \frac{2^{\frac{\sum_{i\in\mathcal{U}} r_i(t)}{|\mathcal{C}'|}}}{\prod_{j\in\mathcal{C}'}\sum_{i\in\mathcal{U}} b_j^i(t)S_j^i(t)^{\frac{1}{|\mathcal{C}'|}}} \\ &+ \sum_{i\in\mathcal{U}^-}|\mathcal{C}'_i| \frac{2^{\frac{r_i(t)}{|\mathcal{C}'_i|}}}{\prod_{j\in\mathcal{C}'_i} b_j^i(t)S_j^i(t)^{\frac{1}{|\mathcal{C}'_i|}}} \\ &- \sum_{j\in\mathcal{C}'}\frac{1}{\sum_{i\in\mathcal{U}} b_j^i(t)S_j^i(t)} - \sum_{j\in\mathcal{C}'_i}\frac{1}{\sum_{i\in\mathcal{U}^-} b_j^i(t)S_j^i(t)} \\ &+ \big(|\mathcal{C}'|+\sum_{i\in\mathcal{U}^-}|\mathcal{C}'_i|\big)P_1 + \big(|\mathcal{U}|+|\mathcal{U}^-|\big)P_2\Bigg) \\ &- \sum_{i=1}^{N} 1_{U_i(t)\geq Q}(t)\big(\omega e^{\omega(U_i(t)-Q)}+2X_i(t)\big)r_i(t) \\ &- \sum_{i=1}^{N} 1_{U_i(t)<Q}(t)\big(-\omega e^{\omega(Q-U_i(t))}+2X_i(t)\big)r_i(t)\end{aligned} \tag{24}$$

Since the rate vector $\mathbf{r}(t)$ and the channel allocation matrix $\mathbf{b}(t)$ are coupled with each other, we propose an algorithm to optimize $\mathbf{r}(t)$ and $\mathbf{b}(t)$ iteratively.

1) Rate Vector Optimization:

For a given channel allocation matrix $\mathbf{b}(t)$, we optimize the rate vector $\mathbf{r}(t)$ to minimize $Y(t)$ in (24).

Considering the transmit powers of the users in $\mathcal{U}$ are allocated according to the same water-filling level, the scheduled rate of each user $i \in \mathcal{U}$ is a function of $\mathbf{b}(t)$ as follows,

$$\begin{aligned} r_i(t) = &\sum_{j\in\mathcal{C}'} b_j^i(t)\frac{\sum_{k\in\mathcal{U}} r_k(t)}{|\mathcal{C}'|} + \sum_{k\in\mathcal{C}'} b_k^i(t) \\ &\times \Bigg(\log\big(\sum_{i\in\mathcal{U}} b_k^i(t)S_k^i(t)\big) - \sum_{j\in\mathcal{C}'}\frac{1}{N}\log\big(\sum_{i\in\mathcal{U}} b_j^i(t)S_j^i(t)\big)\Bigg)\end{aligned} \tag{25}$$

Therefore, the value of $r_i(t)$ only depends on $\sum_{k\in\mathcal{U}} r_k(t)$. Substituting (25) into (24), and using Lagrangian method, we obtain

$$\begin{aligned}\sum_{k\in\mathcal{U}} r_k(t) =&|\mathcal{C}'|\Bigg(\log\Big(\sum_{i\in\mathcal{U}}\big(1_{U_i(t)\geq Q}(t)\big(\omega e^{\omega(U_i(t)-Q)}+2X_i(t)\big) \\ &+ 1_{U_i(t)<Q}(t)\big(-\omega e^{\omega(Q-U_i(t))}+2X_i(t)\big)\big)\frac{\sum_{j\in\mathcal{C}'} b_j^i(t)}{|\mathcal{C}'|} \\ &+ \sum_{j\in\mathcal{C}'}\log\Big(\sum_{i\in\mathcal{U}}\big(b_j^i(t)S_j^i(t)\big)^{\frac{1}{|\mathcal{C}'|}}\Big)\Big) - \log(V)\Bigg)\end{aligned} \tag{26}$$

According to (17), if $\exists i \in \mathcal{U}, \lambda_i > M\frac{\sum_{k\in\mathcal{U}} r_k(t)}{|\mathcal{C}'|}$, then drop this user into nonconform set $\mathcal{U}^-$. The procedure is iterated until $\forall i \in \mathcal{U}, \lambda_i < M\frac{\sum_{k\in\mathcal{U}} r_k(t)}{|\mathcal{C}'|}$. As for those users in nonconform set $\mathcal{U}^-$, the scheduled rate can be obtained

Algorithm 1 ESSA

Initialize parameters V, M, ω, Q and c
$\mathcal{U} = \{1, 2, \cdots, N\}$ and $\mathcal{C} = \{1, 2, \cdots, K\}$
Initialize $\mathbf{b}(t)$ in a greedy manner
for $l = 1 : c$ **do**
  Allocate $b_j^i = 0, r_i = 0, \forall j$ to users $U_i(t) < Q, \forall i$
  (Rate Vector Optimization)
  Determine $\mathcal{U}$ and $\mathcal{U}^-$ according to (17) and (26)
  Rate for user $i \in \mathcal{U}$ is obtained according to (25) and (26)
  Rate for user $i \in \mathcal{U}^-$ is obtained according to (27)
  (Channel Allocation Matrix Optimization)
  Calculate $f(j)$, $F(j)$ for each channel, and calculate $g(i, j)$, $G(i)$ for each user according to (28) and (29), respectively
  for From j with the highest $f(j)$ to the lowest **do**
    Choose a user with highest $g(i, j)$ satisfying $\sum_{j\in\mathcal{K}} b_j^i(t) < M$
  end for
end for
Update the actual queue and virtual queue according to (22) and (23), respectively

through Lagrangian method as

$$\begin{aligned} r_i =&|\mathcal{C}'_i|\Bigg(\log\Big(\sum_{i\in\mathcal{U}^-}\big(1_{U_i(t)\geq Q}(t)\big(\omega e^{\omega(U_i(t)-Q)}+2X_i(t)\big) \\ &+ 1_{U_i(t)<Q}(t)\big(-\omega e^{\omega(Q-U_i(t))}+2X_i(t)\big)\big)\frac{\sum_{j\in\mathcal{C}'_i} b_j^i(t)}{|\mathcal{C}'_i|} \\ &+ \sum_{j\in\mathcal{C}'_i}\log\Big(\sum_{i\in\mathcal{U}^-}\big(b_j^i(t)S_j^i(t)\big)^{\frac{1}{|\mathcal{C}'_i|}}\Big)\Big) - \log(V)\Bigg)\end{aligned} \tag{27}$$

2) Channel Allocation Matrix Optimization:

For a given rate vector $\mathbf{r}(t)$, we optimize the channel allocation matrix $\mathbf{b}(t)$ to minimize $Y(t)$ in (24).

We propose a heuristic scheme similar to [21], which studies the channel allocation for OFDMA systems, to find a sub-optimal allocation vector $\mathbf{b}(t)$ as follows:

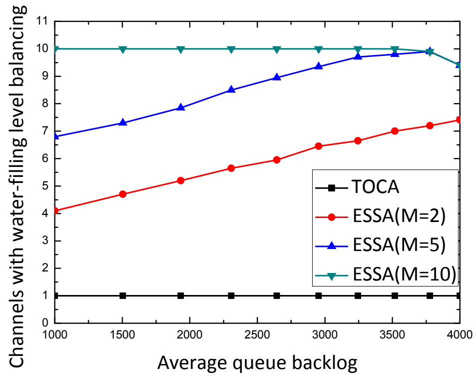
- The channels are allocated one by one according to the descending order of the channel priority values $f(j)$. Each channel j is allocated to the user with the highest priority value $g(i, j)$ according to (28). The priority values for the users and the channels are

$$\begin{aligned} g(i, j) &= -Y(t)|_{b_j^i(t)=1} \\ f(j) &= -\min_{i\in\mathcal{N}} Y(t)|_{b_j^i(t)=1}\end{aligned} \tag{28}$$

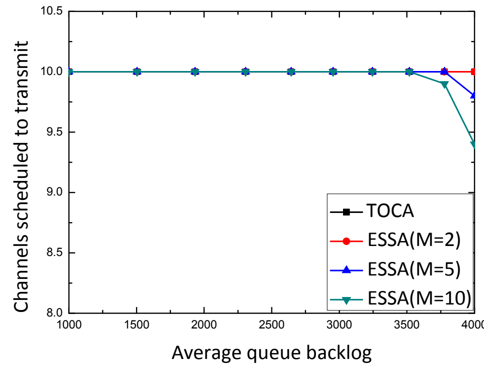
- If two channels have the same priority value or two users have the same priority value, their priorities are determined by the minor priority values, which are defined as

$$\begin{aligned} G(i) &= \prod_{j\in\mathcal{K}} Y(t)|_{b_j^i(t)=1} \\ F(j) &= \prod_{i\in\mathcal{N}} Y(t)|_{b_j^i(t)=1}\end{aligned} \tag{29}$$

- Once channel j is allocated to user i, we set $b_j^i(t) = 1$ and $b_j^k(t) = 0, \forall k \in \mathcal{N}$. $\sum_i b_j^i(t)S_j^i(t)$ is temporarily set to 1 when channel j has not been allocated yet.



(a) Channels water-filling level balancing



(b) Channels scheduled to transmit

Fig. 3. Channel utilization

Remark 3 (Complexity Analysis): The complexity is mainly brought by channel allocation vector optimization whose complexity is $O(NK)$ in each iteration. Therefore, the complexity of ESSA is $O(cNK)$, where c represents the iteration round.

We provide the details of ESSA using pseudo codes in Algorithm 1, which is launched at the beginning of each time slot. In the pseudo-codes, Lines 1–3 initialize the system parameters, Lines 7–9 find the optimal rate vector $\mathbf{r}(t)$ for a given channel allocation matrix $\mathbf{b}(t)$, and Lines 11–15 find a sub-optimal channel allocation matrix $\mathbf{b}(t)$ for a given rate vector $\mathbf{r}(t)$ and Line 16 updates the queue information for the next slot.

IV. Simulation

In this section, we evaluate the performance of the proposed ESSA algorithm by simulation. In this simulation, there are 10 users in two categories, including users with heavy traffic[4] and users with light traffic. These users share 10 time-varying channels, which obey the Rayleigh distribution and are i.i.d. over time slots. The circuit power refers to that of a specific wireless base stations in practice, i.e., $P_1 = 20.40$ and $P_2 = 40.63$, according to [20]. For performance comparison, we adopt TOCA [10] as a baseline, which also adopts Lyapunov optimization but does not consider the characteristics of SA.

With different ratios of users with heavy traffic, we consider 3 scenarios [20] that reflect the expected share of mobile broadband subscribers:

- **Scenario 1**: 20 percents of the subscribers are classified as users with heavy traffic.
- **Scenario 2**: 10 percents of the subscribers are classified as users with heavy traffic.
- **Scenario 3**: No user with heavy traffic.

Fig. 3(a) demonstrates the relationship between the number of the channels in $\mathcal{C}'$ which have the water-filling levels balanced across the users in $\mathcal{U}$. From the results, it can be found that with the increase of SA capability M and the average queue backlog, the number of the channels with water-filling level balancing increases.

Fig. 3(b) demonstrates the number of the channels scheduled to transmit. When the average delay is large, it is not necessary to use all channels, because the saved circuit power by using less channels is larger than the increased transmit power.

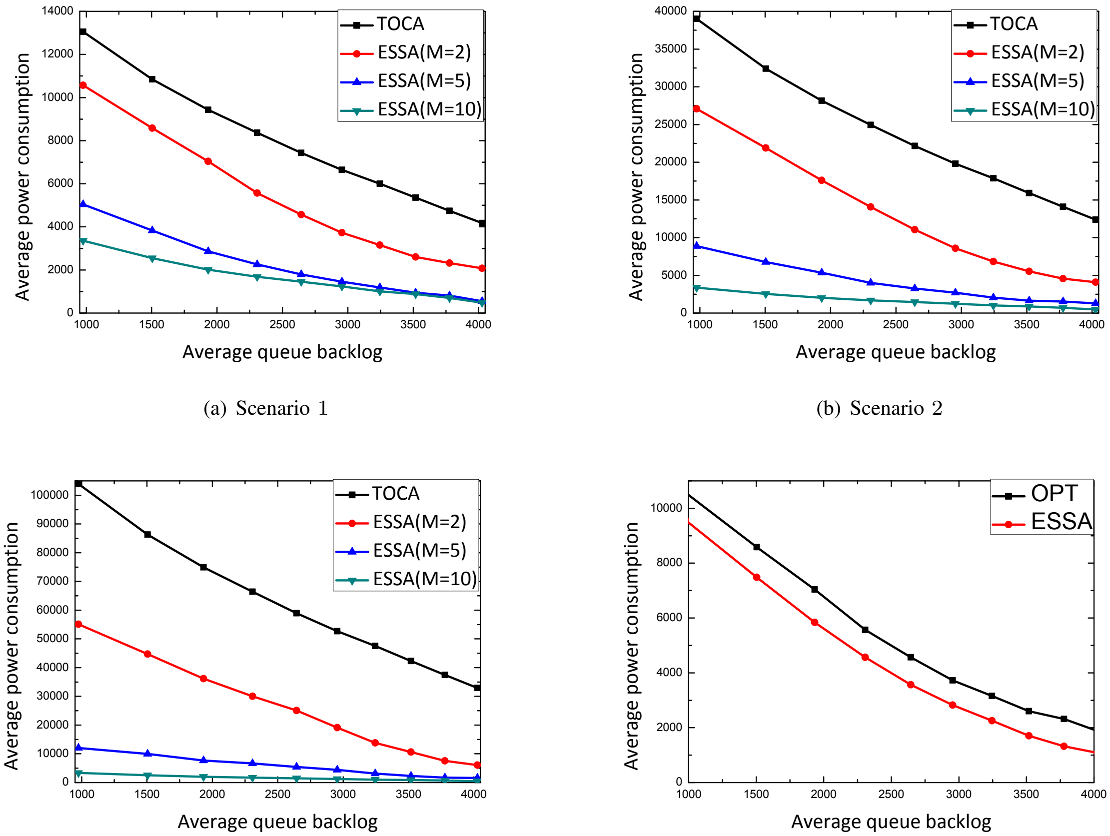
Fig. 4 illustrates the average power consumption versus average queue backlog. The proposed ESSA algorithm outperforms the baseline TOCA significantly in all 3 scenarios. The performance gain is achieved out of two reasons. First, the water-filling levels of channels are balanced across users in ESSA. The large number of the channels with balanced water-filling levels leads to the reduction of power consumption. Second, the circuit power is considered in ESSA to schedule an appropriate number of channels to minimize the total power.

Fig. 4(d) shows the performance gap between ESSA and the optimal performance which is obtained by exhaustive searching. Due to its heuristic nature, ESSA suffers approximately 10% performance loss.

V. Conclusions

In this paper, we propose a scheduling algorithm called ESSA for delay-constrained SA. The ESSA algorithm is designed in two steps. First, we minimize the total power consumption for SA by differentiated water-filling across users, and partition the scheduled users into conform and nonconform user sets due to the limitations brought by SA systems. Second, we propose a sub-optimal algorithm using Lyapunov optimization to achieve the minimal power consumption with the delay constraint. The simulation results exploit the relationship between the differentiated water-filling power allocation and the SA capability and show the performance improvements of ESSA compared to existing TOCA scheme.

[4] A user is with heavy traffic if it satisfies $\lambda_i \geq \sum_{j=1}^{10} \lambda_j/3$ in our simulation.



(a) Scenario 1

(b) Scenario 2

(c) Scenario 3

(d) Performance gap

Fig. 4. Average power consumption versus average delay

REFERENCES

[1] Z. Shen, A. Papasakellariou, J. Montojo, et al, "Overview of 3GPP LTE-advanced carrier aggregation for 4G wireless communications," *IEEE Communications Magazine*, vol. 50, no. 2, pp. 122-130, Feb. 2012

[2] F. Wu, Y. Mao, S. Leng, et al, "A carrier aggregation based resource allocation scheme for pervasive wireless networks," *Proc. of IEEE DASC 2011*, pp. 196-201, Aug. 2011

[3] H. Shajaiah, A. Abdel-Hadi, C. Clancy, "Utility proportional fairness resource allocation with carrier aggregation in 4G-LTE," *Proc. of IEEE Milcom 2013*, pp. 412-417, Jun. 2013.

[4] F. Liu, K. Zheng, W. Xiang, et al, "Design and performance analysis of an energy-efficient uplink carrier aggregation scheme," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 2, pp. 197 - 207, Dec. 2014

[5] J. Lee, S. Jaewoo, "Analysis of cognitive radio networks with channel aggregation," *Proc. of IEEE WCNC 2010*, pp. 1-6, Nov. 2010

[6] L. Wu, W. Wang, Z. Zhang, et al, " A rollout-based joint spectrum sensing and access policy for cognitive radio networks with hardware limitations," *Proc. of IEEE Globecom 2012*, pp. 1277-1282, May 2012

[7] M. Iwamura, K. Etemad, M. Fong, et al, "Carrier aggregation framework in 3GPP LTE-advanced," *IEEE Communications Magazine*, vol. 48, no. 8, pp. 60-67, Jan. 2010

[8] H. Fattah, C. Leung, "An overview of scheduling algorithms in wireless multimedia networks," *IEEE Wireless Communications*, vol. 9, no. 5, pp. 76-83, Mar. 2012

[9] X. He, A. Yener, "On the energy-delay trade-off of a two-way relay network," *Information Sciences and Systems*, pp. 865-870, Apr. 2008

[10] M. J. Neely , "Optimal energy and delay tradeoffs for multiuser wireless downlinks," *IEEE Transactions on Information Theory*, vol. 53, no. 9, pp. 3095-3113, Jun. 2007

[11] R. Ratasuk, D. Tolli, A. Ghosh, "Carrier aggregation in LTE-Advanced," *Proc. of IEEE VTC 2010-Spring*, May 2010

[12] K. I. Pedersen, F. Frederiksen, C. Rosa, et al, "Carrier aggregation for LTE-advanced: functionality and performance aspects," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 89-95, Apr. 2011

[13] M. Calle, J. Kabara, "Measuring energy consumption in wireless sensor networks using GSP," *Proc. of IEEE PIMRC 2006*, pp. 1-5, Aug. 2006

[14] W. A. Rosenkrantz, "Little's theorem, A stochastic integral approach," *Queueing systems*, vol. 12, no. 3, pp. 319-324, Jan. 1992

[15] M. Kuczma, *An introduction to the theory of functional equations and inequalities: Cauchy equation and Jensen's inequality*, Springer, Oct. 2008

[16] L. Lovsz, "On the Shannon capacity of a graph," *IEEE Transactions on Information Theory*, vol. 25, no. 1, pp. 1-7, Dec. 1979

[17] L. M. Graves, "Riemann integration and *Taylor* theorem in general analysis," *Amer. Math. Soc*, vol. 1, no. 29, pp.163-177, Mar. 1927

[18] R. Berry and R. Gallager, "Communication over fading channels with delay constraints," *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1135-1149, May 2002

[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] G. Auer, et al, "Energy efficiency analysis of the reference systems, areas of improvements and target breakdown," *EARTH*, Nov. 2010

[21] S. Hwang, J. Park, Y. S. Jang, et al, "A heuristic method for channel allocation and scheduling in an OFDMA system" *ETRI journal*, vol. 30, no. 5, pp. 741-743, Aug. 2008