# Cross Layer Scheduling in Wireless Networks

Nitin Salodkar and Abhay Karandikar

**Abstract** *Wireless channel poses significant challenges to resource management due to multipath fading. In this chapter, we review the impact of wireless channel on the design of scheduling algorithms at the link and network layers. It has been well recognized now that significant performance gains can be obtained by designing channel aware scheduling algorithms. Various scheduling schemes catering to different Quality of Service (QoS) objectives such as maximizing throughput, minimizing delay, or minimizing energy have been proposed in the literature. Most of these algorithms can be formulated as control problems within the framework of Constrained Markov Decision Processes. We examine some of the representative work in this area with a specific focus on centralized scheduling in a single cell scenario with single antenna system.*

Over the past decade, there has been a large scale proliferation of wireless communications technologies. While the first and second generation (1G and 2G) cellular systems were driven primarily by voice based cellular telephony, the recent years have witnessed a phenomenal increase in packet data applications like email, web browsing, peer to peer and multimedia applications. These applications require Quality of Service (QoS) guarantees in terms of rate, delay, packet loss etc. The time varying nature of the wireless channel due to multipath fading poses significant challenges to radio resource management for providing QoS. In this chapter, we focus on packet scheduling aspect of resource management.

Recent results suggest that significant performance gains can be obtained by exploiting wireless channel characteristics for the design of scheduling algorithms at the link and network layers. These algorithms are referred to as cross layer scheduling. Cross layer scheduling algorithms can be formulated as optimization problems

---

Nitin Salodkar
Department of Computer Science and Engineering, Indian Institute of Technology Bombay, Mumbai, 400076, India

Abhay Karandikar
Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai, 400076, India. This work was supported by TTSL-IIT Bombay Center for Excellence in Telecom.

where the objective is to optimize a given utility function such as energy subject to QoS constraints such as delay. The scheduler can be viewed as a controller where the control action comprises of determining the user to be scheduled, its transmission rate and power depending upon the system state. These problems can often be cast within the framework of Constrained Markov Decision Process (CMDP). The computational approaches to determine the optimal policy comprises of dynamic programming, linear programming, function approximation, reinforcement learning, stochastic approximation etc. In this chapter, we examine some of the representative formulations to elucidate the nature of the problems being considered in the literature. The focus of this chapter is centralized scheduling in a cellular system with single antenna.

We first begin by reviewing the characteristics of wireless channel and its capacity. We then present a framework for cross layer scheduling algorithms, Finally, we review some techniques in fair opportunistic, power optimal, throughput optimal and delay optimal scheduling algorithms.

## 1 Wireless Channel Characteristics

Wireless channel is characterized by the decay of signal strength due to distance (path loss), obstructions due to objects such as buildings and hills (shadowing) and constructive and destructive interference caused by copies of the same signal received over multiple paths (multipath fading). In this chapter we primarily focus on multipath fading and its impact on the design of scheduling algorithms.

In multipath fading, multiple copies of the signal reach the receiver at various instants of time depending on the length of the path over which the signal traverses. The relative motion between the transmitter and receiver and/or movement of the reflecting objects results in random path length changes; consequently different multipath components have random amplitudes and phases. This results in time varying amplitude and phase of the received signal.

The time varying characteristics of the wireless channel can be modeled as a tapped delay line filter with each tap corresponding to a delay window during which different multipath components arrive at the receiver. In this chapter, we limit our discussion to *flat* fading channels where all the multipath components corresponding to a symbol arrive within the symbol duration. In such a case, the channel can be modeled using a filter with a single tap. Moreover, we also assume that this tap (or channel) gain remains constant for a block of symbols and changes only over block (termed as *slot* in this chapter) boundaries. This model is termed as *Block Fading Model*.

Under this model, if a user transmits a signal $\chi_n$ in slot $n$, then the received signal $Y_n$ is given by:

$$Y_n = H_n \chi_n + Z_n, \tag{1}$$

where $H_n$ corresponds to the time-varying channel (tap) gain due to fading and $Z_n$ is the complex Additive White Gaussian Noise (AWGN) (with zero mean and variance $N_0$). Usually $H_n$ is modeled as a zero mean complex Gaussian random variable. Let $\sigma^2$ denote the variance of $H_n$. Then $|H_n|$ is a Rayleigh random variable and $X_n = |H_n|^2$ is an exponentially distributed random variable with probability density function expressed as:

$$f_X(x) = \frac{1}{\sigma^2} \exp(\frac{-x^2}{2\sigma^2}), \; x \geq 0. \tag{2}$$

This model is called Rayleigh fading model.

We refer to $X_n$ as channel state in slot $n$. Note that the channel state $X_n$ may change from slot to slot either in an independent and identically distributed (i.i.d.) fashion or in a correlated fashion (e.g. may follow a Markov model). Moreover, the channel state $X_n$ is a continuous (exponential) random variable. However, for the scheduling problems considered later in this chapter, we assume that the channel state $X_n$ takes values from a finite discrete set $\mathbb{X}$. This discretization can be achieved by partitioning the channel state into equal probability bins with preselected thresholds. For example, let $x(1) < \ldots < x(L)$ be these thresholds. Then the channel is said to be in state $x_k$ if $x \in [x(k), x(k+1))$, $k = 1, \ldots, L$. The channel state space $\mathbb{X}$ can be represented as $\mathbb{X} = \{x_1, \ldots, x_L\}$.

## *1.1 Capacity of Fading Channel*

For a wireless channel, capacity analysis can be performed both in the presence as well as absence of Channel State Information (CSI) at the transmitter. Throughout this chapter, we assume that the transmitter has the knowledge of perfect CSI. In a Time Division Duplex (TDD) system, due to channel reciprocity, it may be possible for the transmitter to obtain the CSI through channel estimation based on the signal received on the opposite link. In a Frequency Division Duplex (FDD) system, the receiver has to estimate the CSI and feed this information back to the transmitter. In practice, e.g., in IEEE 802.16 [26], the channel related information can be conveyed using ranging request (RNG-REQ) messages. In this chapter, we do not take into account the specific feedback mechanisms; rather, we assume that the transmitter has the knowledge of perfect CSI.

Different notions of capacity of fading channels have been defined in the literature. The classical notion of Shannon capacity defines the maximum information rate that can be achieved over the channel with zero probability of error. This notion involves a coding theorem and its converse, i.e., that there exists a code that achieves the capacity (information can be reliably transmitted using this code at a rate less than or equal to the capacity) and that reliable communication is not possible if information is transmitted at a rate higher than the capacity.

In this chapter, we consider *ergodic* capacity (also termed as *throughput* capacity or *expected* capacity in the literature). This notion of capacity measures the rates achievable in the long run averaged over the channel variations. Here, the channel is assumed to vary sufficiently fast, yet the variations are slow enough such that 'reasonably' long codes can be transmitted. We first derive an expression for the ergodic capacity of a single user point-to-point link under an average power constraint and then extend the notion to multiuser scenario.

### 1.1.1 Point-to-Point Capacity with Perfect Transmitter CSI

Consider a single user wireless channel depicted in Figure 1. We assume transmitter has perfect knowledge of CSI. Let $P_n$ denote the transmission power in slot $n$. Let $X_1 = x_1, \ldots, X_M = x_M$ be a given realization of channel states. We assume that the transmitter has an average power constraint of $\bar{P}$.

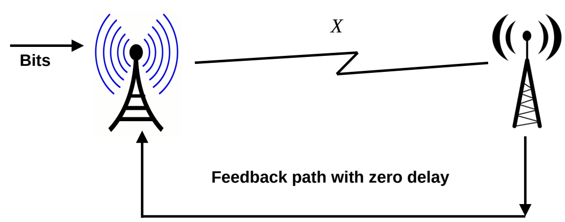


**Fig. 1** Point-to-point transmission model

The restriction on the average power expenditure makes the problem of achieving capacity to be a power allocation problem. Specifically, the problem is to determine the transmission power in each slot that maximizes the information rate (and hence achieves capacity), while keeping the average power expenditure below the prescribed limit. The problem can be stated as:

$$\max_{P_1,\ldots,P_M} \frac{1}{M} \sum_{n=1}^{M} \log\left(1 + \frac{P_n X_n}{N_0}\right), \tag{3}$$

subject to,

$$\frac{1}{M} \sum_{n=1}^{M} P_n = \bar{P}. \tag{4}$$

This is a constrained optimization problem and can be solved using standard Lagrangian relaxation [8]. Let $x^+$ denote $\max(0,x)$. A solution to the optimization problem stated in (3) and (4) is a policy that determines the optimal power in $n^{th}$ slot to be:

$$P_n^* = \left(\frac{1}{\lambda} - \frac{N_0}{X_n}\right)^+, \tag{5}$$

where $\lambda$ is a Lagrange Multiplier which satisfies:

$$\frac{1}{M}\sum_{n=1}^{M}\Big(\frac{1}{\lambda}-\frac{N_0}{X_n}\Big)^+ = \bar{P}. \tag{6}$$

As $M \to \infty$, by ergodicity,

$$\lim_{M\to\infty}\frac{1}{M}\sum_{n=1}^{M}\Big(\frac{1}{\lambda}-\frac{N_0}{X_n}\Big)^+ = \mathbf{E}\Big[\Big(\frac{1}{\lambda}-\frac{N_0}{X}\Big)^+\Big] = \bar{P}, \tag{7}$$

where the expectation is taken with respect to the stationary distribution of the channel states. For a given realization of the channel state $X = x$, the optimal power allocation can be expressed as:

$$P^*(x) = \Big(\frac{1}{\lambda}-\frac{N_0}{x}\Big)^+. \tag{8}$$

Figure 2 provides a pictorial description of the power allocation (termed as *waterfilling*) policy. It can be observed that the transmitter allocates *more* power when the channel is *good* and *less* power when the channel is *poor*. This insight has been used later while designing scheduling schemes at network layer in Section 4. Note that the waterfilling power allocation is in contrast with the traditional power control policy which attempts to invert the channel.

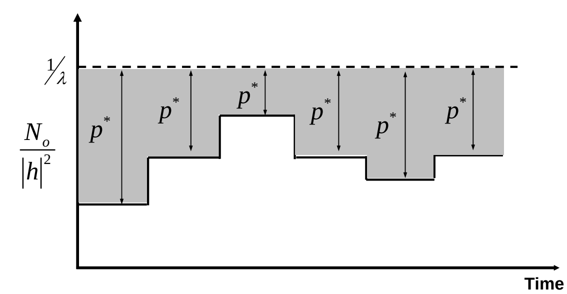


**Fig. 2** Waterfilling power allocation

Once the optimal power is known, the channel capacity with perfect CSI at the transmitter can be expressed as:

$$C = \mathbf{E}\Big[\log\Big(1+\frac{P^*(X)X}{N_0}\Big)\Big]. \tag{9}$$

Though a point-to-point or single user transmission system offers significant insight for transmission over a fading channel, it represents a somewhat restricted scenario. In the next section, we consider a more realistic multiuser scenario where we review generalization of the single user waterfilling power allocation.

### 1.1.2 Multiuser Capacity with Perfect Transmitter CSI on the Uplink

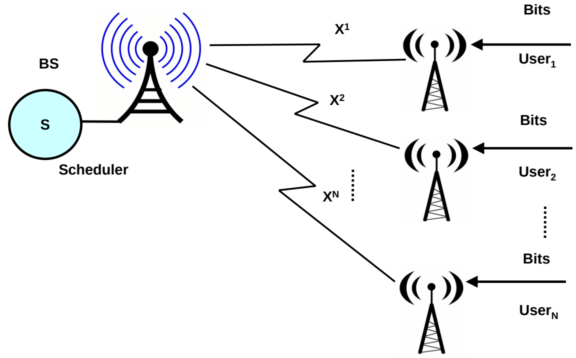


**Fig. 3** Uplink transmission model, infinite backlog of bits at transmitters

In this section, our objective is to determine the ergodic capacity for a multiuser (uplink) fading channel as depicted in Figure 3 where $N$ users communicate with a base station. With the block fading model, the signal $Y_n$ received by the base station in slot $n$ can be described in terms of the transmitted signals $\chi_n^i,\ i = 1,\ldots,N$ as:

$$Y_n = \sum_{n=1}^{N} H_n^i \chi_n^i + Z_n, \tag{10}$$

where $H_n^i$ is the channel gain for user $i$ in slot $n$. Let $X_1^i = x_1^i, \ldots, X_M^i = x_M^i,\ i = 1,\ldots N$, be a given realization of the channel states. User $i$ has an average power constraint of $\bar{P}^i$ ($\bar{\mathbf{P}} = [\bar{P}^1,\ldots,\bar{P}^N]^T$ being the average power constraint vector), the problem is to determine optimal power allocation for each user that maximizes the sum capacity subject to maintaining each user's average power expenditure below its prescribed limit. This problem can be stated as:

$$\max_{P_n^i, i=1,\ldots,N, n=1,\ldots,M} \frac{1}{M} \sum_{n=1}^{M} W \log\left(1 + \frac{\sum_{i=1}^{N} P_n^i X_n^i}{W N_0}\right), \tag{11}$$

subject to the per user power constraint:

$$\frac{1}{M} \sum_{n=1}^{M} P_n^i = \bar{P}^i,\ i = 1,\ldots,N. \tag{12}$$

We first consider *symmetric* scenario where all users have identical channel statistics and power constraints ($\bar{P}^i = \bar{P}, \forall i$). For simplicity, instead of individual power constraints as in (12), we consider the total power constraint:

$$\frac{1}{M}\sum_{n=1}^{M}\sum_{i=1}^{N}P_n^i = N\bar{P}. \tag{13}$$

It turns out that subject to the constraints expressed in (13), the sum capacity in (11) is maximized by allowing only one user with the best channel state to transmit in a slot. The power allocation is expressed as:

$$P^{i_n^*} = \begin{cases} \left(\frac{1}{\lambda} - \frac{WN_0}{x_n^i}\right)^+ & \text{if } x_n^i = \max_j x_n^j, \\ 0 & \text{otherwise,} \end{cases} \tag{14}$$

where $\lambda$ again is a Lagrange Multiplier and is chosen to satisfy the sum power constraint (13) and $i_n^*$ is the index of the best user in slot $n$. Taking $M \to \infty$ and by ergodicity of the fading process, we obtain the capacity-achieving power allocation policy that allocates power $P^{i^*}(\mathbf{x})$ to user $i$ as a function of the joint channel state vector $\mathbf{x} = (x^1, \ldots, x^N)$ where:

$$P^{i^*}(\mathbf{x}) = \begin{cases} \left(\frac{1}{\lambda} - \frac{WN_0}{x^i}\right)^+ & \text{if } x^i = \max_j x^j, \\ 0 & \text{otherwise,} \end{cases} \tag{15}$$

where $\lambda$ is chosen to satisfy the power constraint:

$$\sum_{i=1}^{N}\mathbf{E}\left[P^{i^*}(\mathbf{X})\right] = N\bar{P}. \tag{16}$$

The resulting sum capacity is:

$$C_{sum} = \mathbf{E}\left[W\log\left(1 + \frac{P^{i^*}X^{i^*}}{WN_0}\right)\right], \tag{17}$$

where $X^{i^*}$ is the channel state of the best user indexed by $i^*$. Note that this result is derived by imposing a total power constraint (13). However, because of symmetry and independence between the user channel state processes, the power consumption of all users is same under the optimal power allocation policy. Hence, the per user power constraints in (12) are also satisfied.

The above scheduling policy where the user with the best channel state is scheduled in a slot is called *opportunistic* scheduling. It takes advantage of *multiuser diversity* in order to improve the sum rate (throughput), i.e., in a system with large number of users having independent and diverse channel states, there exists a user having good channel state with a high probability. Moreover, this probability increases with the number of users. The implications of opportunistic scheduling have been investigated in further detail in Section 3.

Let $\mathbb{C}_g(\mathbf{x}, \mathscr{P}(\mathbf{x}))$ denote the set of achievable rates under a policy allocation policy $\mathscr{P}(\mathbf{x})$. It can be expressed as:

$$\mathbb{C}_g(\mathbf{x},\mathbf{P}) = \left\{ \mathbf{R} : \sum_{i \in S} R^i \leq W \log\left(1 + \frac{\sum_{i \in S} x^i P^i}{W N_0}\right) \forall S \in \{1,\ldots,N\} \right\}. \quad (18)$$

A power allocation policy $\mathscr{P}$ is *feasible* if it satisfies the power constraints of all users, i.e., $\mathbf{E}[\mathscr{P}(\mathbf{X})] = \bar{\mathbf{P}}$. Let $\mathbb{F}$ be the set of all feasible power allocation policies. The *throughput capacity region* is defined as the union of the set of rates achievable under all power control policies $\mathscr{P} \in \mathbb{F}$, i.e.,

$$\mathbb{C}(\bar{\mathbf{P}}) = \bigcup_{\mathscr{P} \in \mathbb{F}} \mathbf{E}[\mathbb{C}_g(\mathbf{X}, \mathscr{P}(\mathbf{X}))]. \quad (19)$$

In a general case of asymmetric channels and power constraints, *weighted* rate maximization is a more appropriate metric. Let $\gamma = [\gamma^1, \ldots, \gamma^N]^T$ be a vector of weights assigned to the users. The weighted rate maximization problem can be expressed as:

$$\max \ \gamma \cdot \mathbf{R}, \quad (20)$$

subject to the constraint that the rate vector lies in the capacity region:

$$\mathbf{R} \in \mathbb{C}(\bar{\mathbf{P}}). \quad (21)$$

Using a Lagrangian formulation [8], it can be shown that the optimal power allocation policy can be computed by solving, for each channel state vector $\mathbf{X} = \mathbf{x}$, the following optimization problem:

$$\max_{\mathbf{R},\mathbf{P}} \gamma \cdot \mathbf{R} - \lambda \cdot \mathbf{P}, \quad (22)$$

subject to:

$$\mathbf{R} \in \mathbb{C}_g(\mathbf{X}, \mathbf{P}). \quad (23)$$

The optimal solution to (22) thus provides a power allocation $\mathscr{P}(\mathbf{x})$ and a rate allocation $\mathscr{R}(\mathbf{x})$ at channel state vector $\mathbf{X} = \mathbf{x}$. If the choice of $\lambda = [\lambda^1, \ldots, \lambda^N]^T$ ensures that the power constraint is met then $\mathbf{R}^* = \mathbf{E}[\mathscr{R}(\mathbf{X})]$ is an optimal solution to (22). It can be shown that the optimal solution to (22) is a greedy *successive decoding* scheme where the users are decoded in an order that is dependent on the interference experienced by them.

## 2 A Framework for Cross Layer Scheduling

In the above sections, we have discussed information theoretic capacity notions for multiuser wireless system. These results indicate that significant performance gains can be obtained at the link and network layers by exploiting physical layer information. In this section, we discuss how information from physical layer can be exploited and opportunities created for making scheduling decisions in order to satisfy certain QoS measures.

## 2.1 Opportunistic Scheduling

As we have already studied, for the symmetric case, sum capacity is maximized by scheduling the user with the best channel state in a time slot. This suggests that users should transmit at opportunistic time. This leads to the foundation of *opportunistic scheduling*. Since the radio channel conditions vary independently for each user, in a given slot, there is a high probability of having a user whose channel state is near its peak. Scheduling such a user leads to high sum throughput. The gains are larger if the channel variations are larger which in turn are indeed larger if the number of users is large. Thus, the traditional view that rapid variations in the wireless channel pose a significant challenge for efficient communication has been converted into an opportunity for exploiting multiuser diversity.

Thus, we have a scheduling scheme where the scheduler picks up the user $i_n$ in slot $n$ such that

$$i_n = \arg\max_j x_n^j. \tag{24}$$

This 'pure' opportunistic scheduling, though, maximizes overall sum throughput, is not necessarily fair. It may starve the users who have poor average channel states. For example in Figure 4, user 2 is starved. This problem can be addressed by imposing fairness constraints in (24). We will review such algorithms later in the chapter.

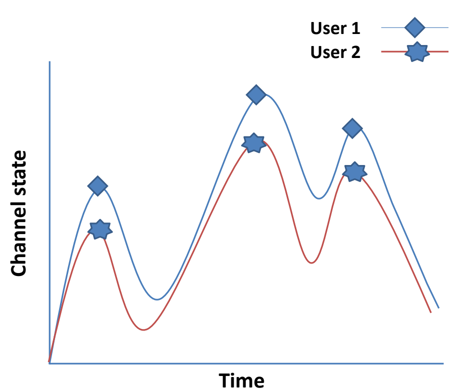


**Fig. 4** User with perennially poor channel condition may be starved by a pure opportunistic scheduler

### 2.2 Energy Efficient Scheduling

Apart from maximizing throughput, energy efficiency is also an important concern in wireless communications. As is evident from the foregoing discussion on capacity, the power required for error free communication in Shannon's sense at a rate $u$ when the channel state is $x$ is given by

$$P(x,u) = \frac{N_0}{x}(e^u - 1), \tag{25}$$

where $N_0$ is the thermal noise power spectral density. Note that for a giver $x$, the transmission power is an increasing and strictly convex function of $u$.

Even for practical digital communication systems, the power required to transmit at a rate $u$ for a specified bit error rate (BER) is given by

$$P(x,u) = \frac{N_0}{\Gamma(BER)\,x}(e^u - 1), \tag{26}$$

where $\Gamma(BER)$ denotes the signal to noise ratio (SNR) gap corresponding to a practical modulation and coding setting. Even in this case, the power is convex and strictly increasing function of $u$.

The convex power-rate relationship implies that if we want to, say, double the rate, we may have to transmit at more than double the power. Thus to transmit $u$ packets, the scheduler can transmit $\frac{u}{2}$ packets in one slot and $\frac{u}{2}$ packets in the next slot instead of $u$ packets in one slot (if the rate permits). This strategy saves the power, albeit, at a cost of delay of one slot. This suggest that the scheduler should transmit the data in opportunistic chunks for energy efficiency.

Note that packets (hence bits) arrive randomly and may be subjected to buffering. Hence to maximize throughput or to minimize energy, opportunistic or energy efficient scheduling has to contend with network layer issues like fairness, packet delay and queue stability. This leads to formulating the scheduling problem as a control problem which exploits fading state information to maximize (or minimize) a given utility function subject to some constraints such as fairness, delay, stability. In the rest of the chapter, we focus on such formulations. We begin by discussing the system model in more detail in the next section.

### 2.3 System Model for Centralized Scheduling

We consider a multiuser wireless system where $N$ users communicate with a base station. On the uplink, as depicted in Figure 5, users communicate with the base station using TDMA, i.e., time is divided into slots of equal duration and only one user can transmit in a slot. We assume that the slot duration is normalized to unity. The base station is the centralized entity that makes the scheduling decision and the user scheduled by the base station transmits in a slot.

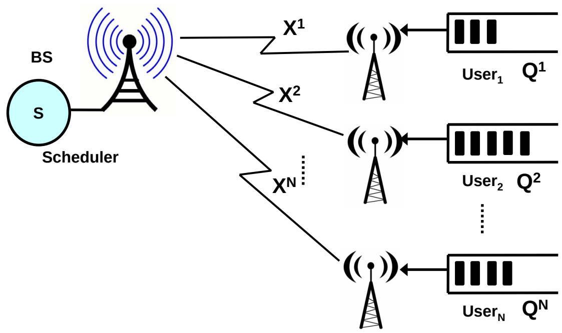


**Fig. 5** Uplink transmission model, finite buffer at each user

The channel state as discussed in Section 1.1.2 is assumed to remain constant for the duration of a slot and to change in an i.i.d. manner across slots. We assume that packets arrive randomly into the user buffer and are queued in the buffer until they are transmitted.

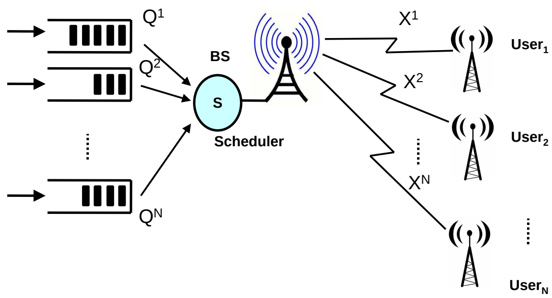


**Fig. 6** Downlink transmission model, finite buffer for each user at base station

On the downlink, as depicted in Figure 6, we assume that the base station multiplexes the transmissions corresponding to $N$ users using Time Division Multiplexing (TDM). The base station maintains a queue for each user.

Let $A_n^i \in \mathbb{A} = \{0, \ldots, A\}$ denote the number of arrivals into user $i$ buffer at the *beginning* of slot $n$[1]. We also assume that the user's packets are of equal size, say, $\ell$ bits. The packet arrival process is assumed to be stationary and is independent of the channel fading and noise processes.

Let $Q_n^i \in \mathbb{Q}$ denote the queue length corresponding to user $i$ at the beginning of slot $n$. Let $U_n^i$ be the number of packets transmitted from user $i$ buffer in slot $n$. Since the slot duration is normalized to 1, $U_n^i \in \mathbb{U}$ also denotes the transmission rate of user

[1] Random variables are denoted with capital letters while their values are denoted with small letters.

$i$ in slot $n$. Let $R_n^i \in \mathbb{U}$ denote the number of packets that user $i$ should transmit in slot $n$ if it is scheduled. Then $U_n^i = I_n^i R_n^i$, where $I_n^i$ is an indicator variable that is set to 1 if the user $i$ is scheduled in slot $n$, otherwise it is set to 0.

The queue dynamics for user $i$ can be expressed as (for $U_n^i$):

$$Q_{n+1}^i = \max(0, Q_n^i + A_{n+1}^i - I_n^i R_n^i). \tag{27}$$

Since the scheduler can at most schedule all the bits in a buffer in any slot, $U_n^i \leq Q_n^i$. Moreover, $Q_{n+1}^i \geq A_{n+1}^i$, $\forall n$. We assume that the scheduler can choose $U_n^i$ based on the queue length $Q_n^i$, the channel state $X_n^i$ and the number of arrivals (source arrival state) $A_n^i$. More generally, the scheduler can determine $U_n^i$ based on entire history of queue lengths, channel states and source arrival states.

Different formulations make various assumptions on the arrival process $\{A_n^i\}$ and control action process $\{U_n^i\}$. We state these assumptions later while formulating different scheduling problems.

The average queue length of a user $i$ can be expressed as:

$$\bar{Q}^i = \limsup_{M \to \infty} \frac{1}{M} \sum_{n=1}^{M} Q_n^i. \tag{28}$$

Average delay $\bar{D}^i$ can be treated to be equivalent to the average queue length $\bar{Q}^i$ because of the Little's law (Chapter 3, [9]) as follows:

$$\bar{Q}^i = \bar{a}^i \bar{D}^i, \tag{29}$$

where $\bar{a}^i$ is the average packet arrival rate of user $i$. Due to this relationship, queue length measure is usually considered to be synonymous with delay measure.

Similarly, the sum throughput over a long period of time can be expressed as:

$$\bar{T} = \liminf_{M \to \infty} \frac{1}{M} \sum_{n=1}^{M} \sum_{i=1}^{N} I_n^i R_n^i. \tag{30}$$

The average power consumed by a user $i$ over a long period of time can be expressed as:

$$\bar{P}^i = \limsup_{M \to \infty} \frac{1}{M} \sum_{n=1}^{M} P(X_n^i, I_n^i R_n^i), \tag{31}$$

where $P(X_n^i, I_n^i R_n^i)$ is the power required by $i$ when the channel state is $X_n^i$ and user transmits $I_n^i R_n^i$ packets.

A number of scheduling algorithms have been proposed in the literature that focus on the above measures or variations of these measures. Broadly, these algorithms can be classified into three types:

1. *Maximize sum throughput subject to fairness constraint.*
2. *Maximize sum throughput subject to delay and queue stability constraint.*

3. *Minimize average power subject to delay constraint.*

We begin by first discussing fair scheduling algorithms.

## 3 Fair Scheduling

Exploiting multiuser diversity in an opportunistic manner by scheduling the user with the best channel state might introduce unfairness. Users who are closer to the base station might experience perennially better channel conditions and thereby obtain a higher share of the system resources at the expense of users who are farther away from the base station. On the other hand, scheduling users with poor channel states results in a reduction in the overall throughput. Thus, there exists a fairness-sum throughput tradeoff. One of the earliest systems to exploit this tradeoff in order to improve the sum throughput is the Code Division Multiple Access/High Data Rate (CDMA/HDR) system.

Different scheduling algorithms provide fairness over different time intervals. A scheduling algorithm is *long term* fair if it provides a *fair share* of a certain quantity such as fraction of time slots or throughput to all users over a long period of time. As outlined earlier, the average throughput achieved by a user $i$ over a long period of time can be expressed as:

$$\bar{T}^i = \liminf_{M\to\infty} \frac{1}{M} \sum_{n=1}^{M} I_n^i R_n^i. \tag{32}$$

The fraction of slots allocated to a user in the long run can be expressed as:

$$\bar{I}^i = \liminf_{M\to\infty} \frac{1}{M} \sum_{n=1}^{M} I_n^i. \tag{33}$$

Long term fair algorithms allocate the quantities such as $\bar{T}^i$ and $\bar{I}^i$ $\forall i$ in a fair manner over a long period of time.

On the other hand, a scheduling algorithm is *short term* fair if it provides a fair allocation of a certain quantity such as fraction of time slots or throughput to all users in an interval of $M$ slots. The average throughput by a user $i$ over a window of $M$ slots can be expressed as:

$$T^i(M) = \frac{1}{M} \sum_{n=1}^{M} I_n^i R_n^i. \tag{34}$$

The fraction of slots allocated to a user $i$ in a window of $M$ slots can be expressed as:

$$I^i(M) = \frac{1}{M} \sum_{n=1}^{M} \sum_{i=1}^{N} I_n^i. \tag{35}$$

Short term fair algorithms allocate the quantities such as $T^i(M)$ and $I^i(M)\ \forall i$ in a fair manner over a window of $M$ slots.

### 3.1 Notions of Fairness

There are various fairness measures that have been considered in the literature. Let $\phi = [\phi^1, \ldots, \phi^N]^T$ be a weight vector associated with the users indicating their relative priorities.

- Minimum Allocation: Under this notion of fairness, the scheduling scheme attempts to provide a certain minimum throughput or fraction of time slots to each user. Let $\bar{\Psi} = [\bar{\Psi}^1, \ldots, \bar{\Psi}^N]^T$ be a vector indicating certain minimum throughput that must be achieved by the users. Let $\bar{\varepsilon} = [\bar{\varepsilon}^1, \ldots, \bar{\varepsilon}^N]^T$ be a vector indicating minimum fraction of time slots that must be allocated to a user. Then the scheme is said to be fair if $\bar{\mathbf{T}} \geq \bar{\Psi}$ (minimum throughput allocation) or $\bar{\mathbf{I}} \geq \bar{\varepsilon}$ (minimum time slot allocation).
- Fair Relative Throughput/Time Slot Allocation: The system attempts to provide equal weighted throughput/fraction of time slots to all users under this notion of fairness. The scheme is said to be fair if $\frac{\bar{T}^i}{\phi^i} = \frac{\bar{T}^j}{\phi^j},\ \forall i, j$ (fair relative throughput allocation) or $\frac{\bar{I}^i}{\phi^i} = \frac{\bar{I}^j}{\phi^j},\ \forall i, j$ (fair relative time slot allocation).
- Proportional Fair Allocation: The fraction of slots allocated to a user is proportional to the average channel state of that user. Better the channel state perceived by a user on an average, higher is the fraction of slots allocated to such a user. The proportional fair scheduling algorithm is discussed in the next section.

Note that each notion of fairness defined above can have a probabilistic extension, where the system is allowed to be unfair with a certain probability.

### 3.2 Fair Scheduling Algorithms

Let $T_n^i$ be the average throughput of a user $i$ in an exponentially averaged window of length $t_c$. The proportional fair scheduling algorithm schedules the user $i$ in a slot $n$ where:

$$i = \arg\max_j \frac{U_n^j}{T_n^j}. \tag{36}$$

The average throughput $T_n^j$ is updated using exponential averaging:

$$T_{n+1}^j = \begin{cases} (1 - \frac{1}{t_c})T_n^j + (\frac{1}{t_c})U_n^j, & j = i, \\ T_n^j & j \neq i. \end{cases} \tag{37}$$

Users having the same channel statistics tend to have the same average throughput and consequently the scheduling policy reduces to the opportunistic policy, i.e., in each slot, the user with the highest rate is scheduled. On the other hand, if the channel statistics of the users are not identical, then the users compete for resources based on their rates normalized by their respective throughputs. Note that the algorithm schedules a user when its channel state is high relative to its own average channel state over the time scale $t_c$. The proportional fair scheduler has the following property: For large $t_c$, i.e., for $t_c \to \infty$, the algorithm maximizes $\sum_{i=1}^{N} \log \bar{T}^i$.

Another fair scheduling policy is to maximize the sum throughput while providing minimum fraction of time slots to the users. This optimization problem can be expressed as:

$$\max \bar{T} \tag{38}$$

subject to:

$$\bar{I}^i \geq \bar{\varepsilon}^i, \;\; i = 1, \ldots, N. \tag{39}$$

One possible approach to designing such a scheduling policy can be based on stochastic approximation where one determines the throughput maximizing time slot allocation in an iterative fashion.

The opportunistic scheduling problem with short term fairness constraints (under the minimum time slot allocation criterion) can be expressed as the following optimization problem: in any window of $M$ slots,

$$\max \sum_{i=1}^{N} \bar{T}^i(M), \tag{40}$$

subject to:

$$\bar{I}^i(M) \geq M\bar{\varepsilon}^i. \tag{41}$$

## 4 Power Optimal Scheduling

In this section, we discuss the problem of scheduling such that the average transmit power is minimized under constraints on the delay. As already discussed, the power required to transmit packets is a convex and increasing function of the rate (i.e., number of packets being transmitted in a slot). From an energy efficiency point of view, the scheduler should transmit packets opportunistically, i.e. when the channel condition is favorable. This leads to buffering the packets and consequent delay. Average delay may be a QoS metric for some applications and the packets need to be scheduled intelligently. We thus have the average cost scheduling problem where the objective is to minimize average power subject to constraint on average delay.

The system model has been already discussed in Section 2.3. In this section, the packet arrival process for each user $\{A_n^i\}$ is assumed to be i.i.d. across users. $A_n^i$ takes values from a finite and discrete set $\mathbb{A} \stackrel{\Delta}{=} \{0, \ldots, A\}$. Since we consider TDMA, only one user can be scheduled in a slot. The scheduling problem is to determine the user to be scheduled in each time slot and also the number of packets that the user should transmit (i.e., rate) such that the average power $\bar{P}^i$ for each user $i$ is minimized subject to a constraint on individual queue length $\bar{Q}^i$. Thus the scheduling problem can be stated as:

$$\text{Minimize } \bar{P}^i \text{ subject to } \bar{Q}^i \leq \delta^i \text{ for } i = 1, \ldots, N. \quad (42)$$

Since the scheduling decision in a slot affects the buffer occupancy of all users, these $N$ problems are not independent.

Before we proceed with the multiuser case, we first consider the case of a point-to-point link, i.e., there is only one transmitter.

## 4.1 Single User Scheduling

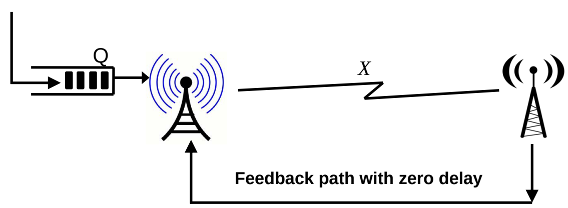


**Fig. 7** Point-to-point transmission model with finite buffer

In this section, we assume that $\mathbb{Q} = \{0, \ldots, B\}$. For the point to point link, the queue evolution can be expressed as:

$$Q_{n+1} = Q_n + A_{n+1} - U_n. \quad (43)$$

Since we consider a single user, we drop the superscript $i$ in the notation in this section. Let us define the state of the system in slot $n$ by $S_n \stackrel{\Delta}{=} (Q_n, X_n, A_n)$, i.e., the state comprises of queue length, channel state and source arrival state. The system state space $\mathbb{S} = \mathbb{Q} \times \mathbb{X} \times \mathbb{A}$ is finite and discrete. In each slot, the control or scheduling action corresponds to the number of packets transmitted $U_n \leq Q_n$. $U_n$ takes values from a finite action space $\mathbb{U} = \{0, \ldots, B\}$. Note that the channel state $X_n$ is independent of the action, queue and arrival state. Similarly, the source arrival state is also independent of channel state and action. The control policy is a sequence of functions $\{\mu_1, \mu_2, \ldots\}$, where $\mu_n$ specifies $U_n$ (or the probability of taking action $U_n$) given the past history of the system state and control actions.

We consider the arrivals to be i.i.d. across slots, in such a case, the state of the system simplifies to $S_n = (Q_n, X_n)$.

Let $c_p(S_n,U_n) = P(X_n,U_n)$ be the 'immediate' cost in terms of power required in transmitting $U_n$ packets when the state is $S_n$. Let $c_q(S_n,U_n) \stackrel{\Delta}{=} Q_n$ denote the 'immediate' cost due to buffering. Let $\mu = \{\mu_1, \mu_2, \ldots\}$ be the control policy. We would like to determine the policy $\mu$ that minimizes

$$\bar{P} = \limsup_{N\to\infty} \frac{1}{N} \mathbf{E} \sum_{n=1}^{N} c_p(S_n,U_n), \tag{44}$$

subject to

$$\bar{Q} = \limsup_{N\to\infty} \frac{1}{N} \mathbf{E} \sum_{n=1}^{N} c_q(S_n,U_n) \le \bar{\delta}. \tag{45}$$

It can be easily argued that this is a CMDP with average cost and finite state and action spaces. For average cost CMDP with finite state and action space, it is well known that a optimal stationary randomized policy exists. Let $\bar{P}^*$ denote the optimal cost i.e.,

$$\bar{P}^* = \min_{\mu} \bar{P}^{\mu}, \tag{46}$$

where $\bar{P}^{\mu}$ is the cost (44) under policy $\mu$.

Let $\mu(\cdot|s) : s \in \mathbb{F}$ be the probability measure on $\mathbb{U}$. For each state $s$, $\mu(\cdot|s)$ specifies the distribution with which the control in that state is applied. We assume that $\{S_n\}$ is an ergodic Markov chain under such policies and thus has a unique stationary distribution $\rho^{\mu}$.

Let $\mathbf{E}^{\mu}$ denote the expectation with respect to (w.r.t.) $\rho^{\mu}$. Under a randomized policy $\mu$, the costs in (44) can be expressed as:

$$\bar{P}^{\mu} \stackrel{\Delta}{=} \mathbf{E}^{\mu}\Big[c_p(S_n,\mu(S_n))\Big] = \sum_{u,s} \rho^{\mu}(s)\mu(u|s)c_p(s,\mu(s)), \tag{47}$$

and,

$$\bar{Q}^{\mu} \stackrel{\Delta}{=} \mathbf{E}^{\mu}\Big[c_q(S_n,\mu(S_n))\Big] = \sum_{u,s} \rho^{\mu}(s)\mu(u|s)c_q(s,\mu(s)), \tag{48}$$

respectively. Then the scheduler objective can be stated as:

$$\text{Minimize } \ \bar{P}^{\mu} \ \text{ subject to } \ \bar{Q}^{\mu} \le \delta. \tag{49}$$

We now demonstrate that the optimal average cost and policy can be determined using an unconstrained Markov Decision Process (MDP) problem and Lagrangian approach.

### 4.1.1 The Lagrangian Approach

Let $\lambda \ge 0$ be a real number. Define $c : \mathbb{R}^+ \times \mathbb{S} \times \mathbb{U} \to \mathbb{R}$ as follows,

$$c(\lambda,s,u) = c_p(s,u) + \lambda(c_q(s,u) - \delta). \tag{50}$$

Note that the function $c(\cdot,\cdot,u)$ is a strictly convex function of $u$ (as the power required to transmit $u$ packets is a strictly convex function of $u$). The unconstrained problem is to determine an optimal stationary policy $\mu^*(\cdot)$ that minimizes

$$L(\mu,\lambda) = \mathbf{E}^{\mu}\Big[c(\lambda, S_n, \mu(S_n))\Big], \tag{51}$$

for a particular value of $\lambda$ called the Lagrange Multiplier (LM). $L(\cdot,\cdot)$ is called the Lagrangian.

Let $p(s,u,s')$ be the probability of reaching state $s'$ upon taking action $u$ in state $s$. Let $V(s)$ denote the optimal value function (i.e. expected cost) for a state $s$. The following dynamic programming equation provides the necessary condition for optimality of the policy.

$$V(s) = \min_{u}\Big[c(\lambda,s,u) - \beta + \sum_{s'} p(s,u,s')V(s')\Big], \;\; s' \in \mathbb{S}, \tag{52}$$

where $\beta \in \mathbb{R}$ is uniquely characterized as the corresponding optimal cost (power) per stage. If we impose $V(s^0) = 0$ for any pre-designated state $s^0 \in \mathbb{S}$, then $V$ is unique. Furthermore, an optimal policy $\mu^*$ must satisfy,

$$\text{support}(\mu^*(\cdot|s)) \subseteq \arg\min\Big[c(\lambda,s,u) - \beta + \sum_{s'} p(s,u,s')V(s')\Big] \;\; \forall s \in \mathbb{S}. \tag{53}$$

It follows that the constrained problem has a stationary optimal policy which is also optimal for the unconstrained problem considered in (51) for a particular choice of $\lambda = \lambda^*$ (say). In general, this optimal policy may be a randomized policy. In fact, it can be shown that the optimal stationary policy is *deterministic* for all states but at most one $s$, i.e., there exists a unique $u^*(s)$ such that $\mu^*(u^*(s)|s) = 1$ and $u^*$ is the solution to the following equation,

$$u^*(s) = \arg\min\Big[c(\lambda^*,s,u) - \beta + \sum_{s'} p(s,u,s')V(s')\Big] \;\; \forall s \in \mathbb{S}. \tag{54}$$

Furthermore, for the single (if any) state $s$ for which this fails, $\mu(\cdot|s)$ is supported on exactly two points. The optimal average cost $\beta$ gives the minimum power consumed $\bar{P}^*$ subject to the specified queue length constraint $\delta$. Moreover, the following *saddle point condition* holds:

$$L(\mu^*,\lambda) \leq L(\mu^*,\lambda^*) \leq L(\mu,\lambda^*). \tag{55}$$

### 4.1.2 Structural Properties of the Optimal Policy

Before we discuss the computational issues in determining the optimal policy through dynamic programming equation (52), we discuss some structural properties of the optimal policy. The result for i.i.d. arrival and channel state processes can be stated as:

**Theorem 1.** *The optimal policy $\mu^*(s) = \mu^*(q,x)$ is non decreasing in channel state $x$ and non-decreasing in queue length $q$.*

The proof of this theorem is based on supermodularity and increasing differences properties of the value function. These properties establish the monotonicity of the optimal policy in channel state and buffer occupancy.

The structural results imply that the optimal decision is to transmit a certain number of packets in a given slot where this number is an increasing function of the current queue length and channel state. Thus for a fixed channel state, the greater the queue length, the more will be the number of packets that will be transmitted. Similarly, for a fixed queue length, the better the channel, the more will be the number of packet transmissions. Thus, the optimal policy always transmits at the highest rate when channel condition is the most favorable and queue length is the largest.

Similar structural results also hold for Markovian arrivals and Markovian channel state and continuous state processes. However, for these general state spaces, the dynamic programming equation (52) for the long run average cost problem requires to be rigorously justified. See bibliographic notes and [1] for a discussion of this.

### 4.1.3 Learning Algorithm for Scheduling

In this section, we discuss the computation of optimal policy discussed in the preceding section. For a fixed $\lambda$, the Relative Value Iteration Algorithm (RVIA) can be used for solving the dynamic programming equation in an iterative fashion. The average cost RVIA for determining the value function such that (52) is satisfied can be written as:

$$V_{n+1}(s) = \min_{u \in U(s)} [c(\lambda, s, u) + \sum_{s'} p(s, u, s') V_n(s')] - V_n(s^0), \tag{56}$$

where $s, s', s^0 \in \mathbb{S}$ and $s^0$ is any fixed state. $V_n(\cdot)$ is an estimate of the value function after $n$ iterations for a fixed LM $\lambda$.

Due to the large size of the state space $\mathbb{S}$, solution of this algorithm is computationally expensive. One approach for addressing this issue is to utilize the structural properties of the optimal policy outlined above to develop efficient heuristics that are computationally less expensive and hence can be implemented in a practical system. In [54], one such approach has been discussed. Techniques based on function approximation can also be employed for dimensionality reduction. The structural results of the policy may be used to choose the basis functions in function approximation.

Another technique is to use ‘primal-dual’ type approach with conventional iteration schemes for the value function (primal) and Lagrange multiplier (dual) [13]. However, even with this approach (and for that matter, with all other approaches discussed above for computing (56)), one major issue is that it requires the knowledge of transition probabilities $p(s,u,s')$. Note that $p(s,u,s')$ depends upon the channel state and packet arrival distributions. Both these distributions are unknown. In practice, it is usual to assume packet arrivals to be Poisson and channel state to be exponential (assuming Rayleigh fading as discussed in Section 1) but these are only modeling assumptions and it is difficult to obtain accurate information even about the parameters of these distributions. Reinforcement Learning (RL) algorithms are useful in such scenarios. We present here one such approach.

Let $\zeta$ be the law for the arrivals and $\kappa(\cdot|\cdot)$ the transition probability function for the channel state process[2]. To address the issue of these unknown probability laws, we can employ stochastic approximation to perform averaging in real time w.r.t unknown laws. Unfortunately, however, RVIA (56) is not amenable to real time implementation because of the occurrence of *min* operator outside the *averaging* operation. This problem is addressed by rewriting RVIA in a novel manner as discussed below.

We define *post-decision* state to be the virtual state of the system immediately *after* taking a decision but *before* the action of the noise (arrivals). If the transmitter transmits $U=u$ packets in a slot, then the post-decision state denoted by $\tilde{s}$, $\tilde{s}\in\mathbb{S}$ is $(q-u,x)$. We rewrite RVIA (56) in the form of post-decision state. Thus $\tilde{V}$ (i.e, the value function based on the post decision state) satisfies the following dynamic programming equation: for $\tilde{s}=(q,x)$,

$$\tilde{V}(\tilde{s})=\sum_{a,x'}\zeta(a)\kappa(x'|x)(\min_{u\leq q+a}[c(\lambda,(q+a,x'),u)+\tilde{V}(q+a-u,x')])-\beta. \quad (57)$$

From (56) and (57), we get the following RVIA:

$$\tilde{V}_{n+1}(\tilde{s})=\sum_{a,x'}\zeta(a)\kappa(x'|x)(\min_{u\leq q+a}[c(\lambda,(q+a,x'),u)+\tilde{V}_n(q+a-u,x')])-\tilde{V}_n(\tilde{s}^0);$$
$$\tilde{V}_{n+1}(\tilde{s}'')=\tilde{V}_n(\tilde{s}'')\ \ \forall\,\tilde{s}''\neq\tilde{s}. \quad (58)$$

The important thing to note here is that we update only the $\tilde{s}$-th component, not the rest. Note (58) has a useful structure in the sense that the averaging operation has been moved outside of the min operator. We can now employ stochastic approximation to perform averaging in real time.

We set up 'primal-dual' iterations for the value function (of RVI (58)) and LM. To solve both value function and LM iteratively, let us first choose the sequences $\{f_n\}$ and $\{e_n\}$ that have the following properties,

$$\sum_n (f_n)^2, \quad \sum_n (e_n)^2<\infty, \quad (59)$$

$$\sum_n f_n=\infty, \quad \sum_n e_n=\infty, \quad (60)$$

$$\lim_{n\to\infty}\frac{e_n}{f_n}\to 0. \quad (61)$$

[2] Here we allow the channel state process to be Markovian also.

The complete primal-dual RVI algorithm can be expressed as: for $S_n = \tilde{s} = (q,x)$,

$$\begin{aligned}
\tilde{V}_{n+1}(\tilde{s}) &= \tilde{V}_n(\tilde{s}) + f_n \Big\{ \min_u [c(\lambda_n, (q+A_{n+1}, X_{n+1}), u) + \tilde{V}_n((q+A_{n+1}-u, X_{n+1}))] \\
&\quad - \tilde{V}_n(\tilde{s}) - \tilde{V}_n(\tilde{s}^0) \Big\}, \qquad (62)\\
\tilde{V}_{n+1}(\tilde{s}'') &= \tilde{V}_n(\tilde{s}) \ \forall \ \tilde{s}'' \neq \tilde{s}, \qquad (63)\\
\lambda_{n+1} &= \Lambda[\lambda_n + e_n (Q_n - \delta)], \qquad (64)
\end{aligned}$$

where we use the projection operator $\Lambda$ to project the LM onto interval $[0,L]$ for large enough $L > 0$, to ensure boundedness of the LM. These iterations can be performed in real time at every slot.

The sequences $f_n$ and $e_n$ are chosen appropriately to ensure that the sequences converge to 0 neither too fast nor too slow. Since we have assumed $\lim_{n\to\infty} \frac{e_n}{f_n} \to 0$, it induces two time scales, a fast one for (62) and a slow one for (64). Using the theory of two time scale stochastic approximation, it can be proved that these iterates indeed converge to optimal values.

The convergence of value function and LM is analyzed by first freezing $\lambda_n \approx$ a constant $\lambda^i$ and then proving that the value function converges to its optimal value $\tilde{V}$, i.e., $\tilde{V}_n \to \tilde{V}$.

Let $h(\tilde{V}) = [h_{q,x}(\tilde{V})]$ be given by:

$$h_{q,x}(\tilde{V}) = \sum_{a,x'} \zeta(a)\kappa(x'|x) \times \min_u [c(\lambda, q+a, x, u) + \tilde{V}(q+a-u, x') - \tilde{V}(q^0, x^0)],$$

where $(q^0, x^0)$ is any pre-designated state. The limiting o.d.e. for (62) is given by

$$\dot{\tilde{V}}(t) = \Lambda(t)(h(\tilde{V}(t)) - \tilde{V}(t)), \qquad (65)$$

where $\Lambda(t)$ is a diagonal matrix with nonnegative elements summing to 1 on the diagonal. Then the following Lemma can be proved.

**Lemma 1.** *If the diagonal elements of $\Lambda(t)$ remain uniformly bounded away from zero, $\tilde{V}_n^i \to \tilde{V}$.*

Note that the above analysis treats $\lambda_n \approx$ a constant, so what this Lemma states is that $\{\tilde{V}_n\}$ closely tracks $\{\tilde{V}^{\lambda_n}\}$, where $\tilde{V}^{\lambda}$ is $\tilde{V}$ with its $\lambda$-dependence made explicit. Note that the $\tilde{V}_n$ and $\lambda_n$ iterations are primal-dual iterations. The primal iterations perform relative value iteration and determine a minimum of the Lagrangian (51) with respect to the policy for an almost constant LM.

To prove the convergence of $\lambda_n$, we consider the limiting o.d.e of (64). It can be shown that the limiting o.d.e. for the $\lambda_n$'s is a steepest ascent for the Lagrangian minimized over the primal variables. By standard results for stochastic gradient ascent for concave functions, it can be proved that this o.d.e converges to the optimal LM $\lambda^*$. We thus have the following lemma.

**Lemma 2.** *The LM iterates $\lambda_n$ converge to optimal value $\lambda^*$.*

Lemmas 1 and 2 imply that $(\tilde{V}_n, \lambda_n) \rightarrow (\tilde{V}, \lambda^*)$ as required.

In practice, the power optimal scheduling algorithm can be implemented using the above primal-dual online algorithm. In each time slot, the scheduler observes the channel state $x$ and determines $u$, the number of packets to be transmitted, that minimizes the right hand side of (62). The value function and LM are updated as in (62), (64) and the algorithm proceeds. Though, the convergence results are asymptotic, for most practical systems, it has been observed that the algorithm converges in reasonable number of iterations (time slots). The readers are referred to references in bibliographic notes for details of the implementation.

In the next section, we consider multiuser scheduling where a centralized scheduler has the responsibility of determining the user to be scheduled in a slot in addition to determining the number of bits (or packets) to be transmitted and the corresponding transmission power.

## *4.2 Multiuser Uplink Scheduling*

As outlined earlier in (42), the multiuser power optimal scheduling problem is to design an algorithm that minimizes the average power expenditure of each user subject to a constraint on the individual queuing delay. This problem is an optimization problem with $N$ objectives and $N$ constraints and can also be formulated within the framework of CMDP. Let $\mathbf{Q}_n$ and $\mathbf{X}_n$ denote the vectors $[Q_n^1, \ldots, Q_n^N]$ and $[X_n^1, \ldots, X_n^N]$ respectively. The system state is then defined by $\mathbf{S_n} = (\mathbf{Q}_n, \mathbf{X}_n)$. Since we consider TDMA where only one user is scheduled in a slot, control action comprises of the index of the user and its transmission rate.

However, since the system state consists of vector of queue lengths and channel states of all users, even for moderate number of users, the size of the state space is very large. Thus numerical approaches for determining the optimal policy are computationally infeasible. Moreover, to compute the optimal policy, the base station needs to know the queue length information of each user. This requires significant communication overhead on the uplink where each user needs to communicate its queue length. Finally, as discussed in point to point link case, even the packet arrival and channel state distributions are unknown. One, therefore, has to consider suboptimal, albeit, efficient approaches.

One such strategy to solve this problem consists of decomposing the problems into $N$ dependent single user problems and a base station problem. In a user problem, each user $i$ determines a rate at which it should transmit in a slot as if it were being controlled by a single user policy. Since the channel and arrival statistics are not known, each user employs learning algorithm described in Section 4.1 (rate determination algorithm). The users' rates are then conveyed to the base station. The base station schedules the user with the highest rate (user selection algorithm). The rationale behind choosing the highest rate is that this strategy will favor the user either with good channel condition or large queue length. The queue length of a user, who is not scheduled for a while, will keep increasing thereby increasing its rate requirement as well. This will ensure that the user will eventually be scheduled. It can be shown that the algorithm indeed converges to equilibrium and the delay constraints are satisfied. The complete strategy is explained below.

### 4.2.1 Rate Determination Algorithm

Each user implements the single user algorithm ((62)–(64)) discussed in Section 4.1 Thus user $i$ determines its rate $R_n^i$ (or the number of packets it should transmit) that maximizes the r.h.s of primal value function iteration (62). Note that the value function iteration (62) considers the following power cost:

$$\bar{P}_e^i = \limsup_{M\to\infty} \frac{1}{M} \sum_{n=1}^{M} P(X_n^i, R_n^i). \tag{66}$$

This is the same cost as discussed for the point to point link (44) (in this case $U_n^i$ is replaced by $R_n^i$ since $U_n^i$ in multiuser case now refers to the packets actually transmitted by user $i$). User $i$ transmits at this rate if the channel is allocated to it by the base station otherwise it is unable to proceed with the transmission. Thus the queue evolution equation for user $i$ is given by,

$$Q_{n+1}^i = Q_n^i + A_{n+1}^i - I_n^i R_n^i, \tag{67}$$

where $I_n^i = 1$ if user $i$ is scheduled in a slot $n$ by the base station, else, $I_n^i = 0$. The control variables are $\mathbf{I}_n$, $\mathbf{R}_n = [R_n^1, \ldots, R_n^N]$. Note that users independently choose the corresponding components of $\mathbf{R}_n$ and the base station chooses $\mathbf{I}_n$ subject to the constraints that $I_n^i \in \{0,1\}$ and $\sum_i I_n^i = 1$.

The actual average power consumed by user $i$ is given by,

$$\bar{P}^i = \limsup_{M\to\infty} \frac{1}{M} \sum_{n=1}^{M} P(X_n^i, I_n^i R_n^i). \tag{68}$$

The relative value iteration is performed not for the actual power cost (68) but exactly as discussed in single user case. However, the difference is reflected in the queue length update (67) (and thereby the LM update (64)) that depends on whether a user is scheduled in a slot or not. It can be proved that value function iteration will converge to the single user optimum rate required for the current quasi-static value of the LM. The difference with the single user case comes from the fact that the relative value iteration for each user is coupled through LM iterations also and the LM updates are indeed affected by the actual transmission through queue length. The convergence of LM would imply that the delay constraints are satisfied.

### 4.2.2 User Selection Algorithm

The base station schedules the user with the highest rate $R_n^i$ in a slot, i.e, $I_n^i = 1$ if $i = \arg\max_j R_n^j$ and all other $I_n^j = 0$. If more than one user has the highest rate, one of them is selected at random with uniform probability.

This strategy can be thought of as an auction where users bid their rates to the base station which then schedules the user bidding the highest rate. Note that it is not in the interest of the users to bid unnecessarily high rates as this might result in higher power consumption.

### 4.2.3 Algorithm Analysis

*Convergence Analysis:* The convergence of value functions and LM follow as in the single user case. Since the value function of each user is updated in each slot regardless of whether the user is scheduled in that slot or not; the value functions are decoupled across users. This is in the spirit of the decoupling of static formulations of network flow problems via the Lagrange multiplier as in [22], except that here it is *mandated* by our algorithm. The decoupling is facilitated by the fact that the users compute their value function as though the cost is (66) and not (68). Convergence of the LMs implies that the delay constraints are satisfied and vice versa. This implies that if there is sufficient capacity, the multiuser scheduling satisfies the delay constraints of all the users.

*Queue Stability:* Assuming that the users and base station have already learned their policies, one can establish the queue stability. A sketch of the proof is as follows.

$\bar{a}^i = \mathbf{E}[A_n^i]$. Let $\pi$ denote the (unique) stationary distribution of the Markov chain $\{\mathbf{X}_n\}$. Let $R \stackrel{\Delta}{=} \min_i \sum_{\mathbf{x}} \pi(\mathbf{x}) \hat{R}^i(x^i)$. Recall that $R_n^i$ depends on $(Q_n^i, X_n^i)$. Suppose that $R_n^i = \ell^i(Q_n^i, X_n^i)$ for some $\ell^i(\cdot,\cdot)$. Write $\ell(\mathbf{Q}_n, \mathbf{X}_n)$ for $[\ell^1(Q_n^1, X_n^1), \ldots, \ell^N(Q_n^N, X_n^N)]$.

For proof of stability, we assume a more general scheme where the channel is allocated to user $i$ with a probability,

$$F^i(\mathbf{R}_n) := g(R_n^i)^m / (\sum_j g(R_n^j)^m)$$

,

where $g$ is a monotone increasing and smooth function and $m >> 1$. Note that with this user selection algorithm, the base station uses a smooth approximation $\mathbf{F} := [F^1, \ldots, F^N]$ of the maximum function for channel allocation, i.e., it allocates the channel to the user with the highest rate (say, $i$th) with a probability $F^i(\mathbf{R}_n)$ close to one, but does allocate it to others also with a small but nonzero probability $F^j(\mathbf{R}_n), j \neq i$. We also assume that the $\ell^i(\cdot,\cdot)$ above is continuously differentiable.

The queue evolution (67) for user $i$ can be rewritten as:

$$Q_{n+1}^i = Q_n^i + (\bar{a}^i - F^i(\mathbf{R}_n)R_n^i) + \Big((A_{n+1}^i - \bar{a}^i) + (F^i(\mathbf{R}_n)R_n^i - U_n^i)\Big).$$

Let $\widetilde{\mathbf{F}} = [\widetilde{F}^1, \ldots, \widetilde{F}^N]$ be defined by: $\widetilde{F}^i(r) = F^i(r) r^i \ \forall i$. Note that this is continuously differentiable. Then the queue evolution is of the form:

$$Q^i_{n+1} = Q^i_n + (\bar{a}^i - \widetilde{F}^i(\mathbf{R}_n)) + M_{n+1}, \ n \geq 0, \tag{69}$$

where $\{M_n\}$ is a martingale difference sequence.

Let $\mathbf{q} = [q^1, \ldots, q^N]$. Define $\widehat{\mathbf{F}} = [\widehat{F}^1, \ldots, \widehat{F}^N]$ by:

$$\widehat{F}^i(\mathbf{q}) = \sum_{\mathbf{x}} F^i(\ell(\mathbf{q}, \mathbf{x})) \ell^i(q^i, x^i) \pi(\mathbf{x}).$$

Consider the scaled version of (69), given by

$$Q^i_{n+1} = Q^i_n + \eta[(\bar{a}^i - \widetilde{F}^i(\mathbf{R}_n)) + M_{n+1}], \ n \geq 0, \tag{70}$$

for a small $\eta > 0$. If we consider smaller and smaller time slots of width $\eta$ with $\bar{a}^i$, and $\widetilde{F}^i$ being ‘rates per unit time’ rather than ‘per slot’ quantities, then we obtain the ‘fluid’ approximation of (69) , which is given by the o.d.e.

$$\dot{q}^i(t) = \bar{a}^i - \widehat{F}^i(\mathbf{q}(t)), \tag{71}$$

It can be proved that the trajectories of the o.d.e. in (71) converge to an equilibrium for all almost all initial conditions. This in turn proves that the actual queue lengths will concentrate near the equilibrium set of this o.d.e with very high probability. The readers are referred to [41] for more details.

The above strategy thus ensures a stable energy efficient scheduling algorithm that also satisfies the delay constraints of the users. In the uplink scenario, the base station does not have information about queue lengths of users. In the above approach, each user needs to communicate only the desired rate (instead of the queue length information). In a practical system, we may have few possible rates say 16. In such a case, we may need only 4 bits of communications overhead. This strategy does not require the knowledge of packet arrival and channel state distributions. It thus provides a powerful framework for implementing multiuser packet scheduling algorithms.

## 5 Scheduling Schemes that Maximize Throughput

In the preceding section, we have considered power optimal scheduling algorithms that minimize power cost under *delay constraints*. In this section, we consider scheduling algorithms under the *power constraints*. We first study scheduling algorithms that consider queue stability as a notion of QoS. While some of these algorithms are throughput optimal, they do not necessarily ensure small average queue lengths and hence small delays. Subsequently, we consider scheduling algorithms that address this issue while achieving high sum throughput.

## 5.1 Throughput Optimal Scheduling

We first review feasible rate and power allocation with an objective of stabilizing the queues of the users. We define the overflow function as follows:

$$f^i(\xi) = \limsup_{M\to\infty} \frac{1}{M} \sum_{n=1}^{M} I_{Q_n^i > \xi}, \tag{72}$$

where $I_{Q_n^i > \xi}$ is an indicator variable that is set to 1 if $Q_n^i > \xi$, else it is set to 0. We say that the system is *stable* if $f^i(\xi) \to 0$ as $\xi \to \infty$ for all $i = 1, \ldots, N$. Let $\bar{\mathbf{a}} = [\bar{a}^1, \ldots, \bar{a}^N]^T$ denote the arrival vector, $\bar{a}^i$ being the average arrival rate for user $i$. In this section, in addition to the average power constraints, we also consider the peak power constraints, i.e., a user $i$ can transmit at a maximum power $\hat{P}^i$ in any slot. Let $\hat{\mathbf{P}} = [\hat{P}^1, \ldots, \hat{P}^N]^T$ denote the peak power constraint vector.

Note that, since the objective is to keep the queues stable, the power and rate allocation policies have to be cognizant of the queue lengths of the users in each slot. A power allocation policy $\mathscr{P}$ is a mapping from the joint channel state and queue length vector $(\mathbf{x}, \mathbf{q})$ to a power allocation vector $\mathbf{P}$. A rate allocation policy $\mathscr{R}$ is a mapping from the joint channel state and queue length vector $(\mathbf{x}, \mathbf{q})$ to a rate allocation vector $\mathbf{R}$. As noted previously in Section 1.1.2, a feasible rate allocation policy allocates rates within the multi-access capacity region $\mathbb{C}_g(\mathbf{x}, \mathbf{P})$. The *stability region* of the multi-access system is the set of all arrival vectors $\bar{\mathbf{a}}$ for which there exists some feasible power allocation policy and rate allocation policy under which the system is stable. The stability region of a multi-access system can be shown to be given by:

$$\mathbb{C}_s(\bar{\mathbf{P}}, \hat{\mathbf{P}}) = \bigcup_{\mathscr{P} \in \mathbb{F}} \mathbf{E}\left[\mathbb{C}_g(\mathbf{X}, \mathscr{P}(\mathbf{X}))\right]. \tag{73}$$

Note that the power control policy $\mathscr{P}(\mathbf{X})$ depends only on the channel state vector $\mathbf{X}$. More importantly, this stability region of the multi-access system is same as the throughput capacity region under power control defined in Section 1.1.2.

If the joint arrival process $\{\mathbf{A}_n\}$ and joint channel state process $\{\mathbf{X}_n\}$ are ergodic Markov chains, then the system can be stabilized by a power and rate allocation policy if $\bar{\mathbf{a}} \in \mathbb{C}_s(\bar{\mathbf{P}}, \hat{\mathbf{P}})$. In practice, one does not have a knowledge of the arrival vector $\bar{\mathbf{a}}$ and this can only be estimated over time. Power and rate allocation policies that do not assume knowledge of the arrival vector $\bar{\mathbf{a}}$ and stabilize the system as long as $\bar{\mathbf{a}} \in \mathbb{C}_s(\bar{\mathbf{P}}, \hat{\mathbf{P}})$ are referred to as *throughput optimal* policies. Throughput optimal scheduling policies have been explored in several papers in literature, see bibliographic notes for further details. Longest Connected Queue (LCQ), Exponential (EXP), Longest Weighted Queue Highest Possible Rate (LWQHPR) and Modified Longest Weighted Delay First (M-LWDF) are some well known examples of throughput optimal scheduling policies. We now review some of these scheduling rules that are throughput-optimal under a power allocation policy $\mathscr{P}$.

- LWQHPR: Let $\alpha = [\alpha^1, \ldots, \alpha^N]^T$ be a vector of weights. The throughput optimal rate allocation policy is obtained by maximizing $\sum_{i=1}^N \alpha^i Q_n^i U_n^i$ over $\mathbb{C}_g(\mathbf{x}, \mathscr{P}(\mathbf{x}))$.

The solution $\mathbf{r}^*$ is obtained by successively decoding the users in an increasing order of their weights $\alpha^i Q_n^i$, i.e., shorter queues are decoded before longer queues. This implies that longer queues are given preference over shorter queues.

- M-LWDF: Let $\bar{\Upsilon}^i$ and $\bar{D}^i$ be the delay requirement and achieved delay for user $i$ respectively. The M-LWDF scheduler attempts to satisfy the delay constraints of the form,
$$Pr\big(\bar{D}^i > \Upsilon^i\big) \leq \bar{\rho}^i, \tag{74}$$
where $\bar{\rho}^i$ is an upper bound on the probability with which $\bar{D}^i$ is allowed to exceed $\Upsilon^i$. The M-LWDF scheme achieves this by scheduling a user $i$ in a slot $n$ where:
$$\frac{-\log(\bar{\rho}^i) \times Q_n^i \times U_n^i}{\bar{\Upsilon}^i} = \max_j \frac{-\log(\bar{\rho}^j) \times Q_n^j \times U_n^j}{\bar{\Upsilon}^j}. \tag{75}$$
Note that higher the queue length and better the channel state (and hence higher the rate) of a user in a slot, higher is the probability of scheduling the user in the slot.
- EXP: Let $\gamma = [\gamma^1, \ldots, \gamma^N]^T$, $\mathbf{b} = [b^1, \ldots, b^N]^T$ be an arbitrary set of positive constants. Let $\alpha$ and $\eta \in (0,1)$ be fixed. The Exponential (EXP) rule schedules a user $i$ in a slot $n$ where:
$$i = \arg\max_j \gamma^j U_n^j \exp\left(\frac{b^j Q_n^j}{\alpha + [\hat{Q}_n]^\eta}\right), \tag{76}$$
where $\hat{Q}_n \triangleq \frac{1}{N}\sum_{i=1}^N b^i Q_n^i$. Thus, a user with better channel state and hence higher rate and higher queue length has a higher probability of being scheduled.

## 5.2 Delay Optimal Scheduling

While throughput optimal scheduling policies maintain the stability of the queuing system, they do not necessarily guarantee small queue lengths and consequently lower delays. Delay optimal scheduling deals with optimal rate and power control such that the average queue length and hence average delay are minimized for arrival rates within the stability region under average and peak power constraints. Due to the nature of the constraints, there is no loss of optimality in choosing the rate and power control policies separately. Hence, to simplify the problem, one can choose any stationary power control policy that satisfies the peak and average power constraints. The delay optimal policy, therefore, deals with optimal rate allocation for minimizing delays under a given power allocation policy. The objective is to maximize a weighted combination of the rates expressed in (20), while at the same time minimizing the achieved delay $\bar{Q}^i$, $i = 1, \ldots, N$. Note that this problem is a multi-objective optimization problem.

We now study a scheme that is throughput optimal and delay optimal under certain assumptions on the arrival and channel state processes for both multi-access (uplink) and broadcast (downlink) channels. Before outlining these assumptions, we define a symmetric channel state process. The channel state process is called *symmetric* or *exchangeable* if for all $n$ and $\mathbf{x} = [x^1, \ldots, x^N]^T$ in the channel state space $\mathbb{X}^N$,

$$Pr\left(X_n^1 = x^1, \ldots, X_n^N = x^N\right) = Pr\left(X_n^1 = x^{\pi(1)}, \ldots, X_n^N = x^{\pi(N)}\right), \tag{77}$$

for any permutation $\pi \in \Pi$, where $\Pi$ is the set of all permutations on the set $\{1, \ldots, N\}$. A power control policy $\mathscr{P}$ that is a function of the channel state vector only is symmetric if for all $\mathbf{x} \in \mathbb{X}$,

$$\mathscr{P}^i\left(x^1, \ldots, x^N\right) = \mathscr{P}^{\pi^{-1}(i)}\left(x^{\pi(1)}, \ldots, x^{\pi(N)}\right). \tag{78}$$

Intuitively, under a symmetric power control policy, the power allocated to a given user is determined by the channel state perceived by that user relative to the channel states perceived by the other users and not on the identity of that user. In [56], the authors consider symmetric channel state and power control. Moreover, they assume Poisson arrivals and exponentially distributed packet lengths. Under these assumptions, they prove that the Longest Queue Highest Possible Rate (LQHPR) policy, besides being throughput optimal, also minimizes delay.

The problem of maximizing the sum throughput subject to constraints on the individual user delays has the structure of a Constrained Markov Decision Process (CMDP). However, the primary difficulty in computing optimal policy lies in large state space size that increases exponentially with the number of users. A simple heuristic would be to compute indices for all users in each slot based on their channel state and queue length. The base station schedules the user with the largest index. The indices are carefully updated so that the delay constraints of the users are satisfied and the system achieves a high sum throughput.

In this chapter, we have considered scheduling algorithms that take advantage of the opportunities provided by the fading wireless channel. These channel aware algorithms are termed as cross layer scheduling algorithms. Most of these algorithms can be considered as control problem where the objective is to optimize a given utility such as throughput, energy, queue stability subject to some constraints. These policies can be computed using dynamic programming tools. However, the curse of dimensionality is a major impediment in determining optimal solutions within this framework. MDP framework also needs knowledge of transition probability mechanism of the underlying Markov chain. In practice, this depends on the fading process and arrival process. In most cases, this knowledge is difficult to possess thereby requiring learning approaches. While significant progress has been made in the scheduling literature, multiuser scheduling still continues to be a challenging problem. In this chapter, we have given few formulations that outlined the nature of problems being considered in the literature.

## Bibliographic Notes

An overview of fading in wireless channel has been provided in [44, 45]. The block fading channel model has been suggested in [33]. The Markov model for modeling Rayleigh channel has been suggested in [55]. An excellent review of information theoretic limits over fading channels can be found in [11]. The waterfilling power allocation policy has been suggested in the work by Goldsmith and Varaiya [19]. Further insights on transmission in multiuser wireless channel has been provided in the works by Knopp and Humblet [23] and Tse and Hanly [50, 51]. The treatment in Section 1.1.2 borrows from [50].

For the multiuser scheduling with symmetric channel fading and symmetric constraints, it has been shown in [23] that the optimal policy schedules the user with the best channel condition in each slot. A generalized version of the problem has been considered in [50] where the authors show that the optimal policy is a greedy *successive decoding* scheme where the users are decoded in an order that is dependent on the interference experienced by them. 'Opportunistic scheduling' has been considered in [30]. Multiuser diversity has been discussed in a lucid way in [52]. Expressions for multiuser capacity with perfect transmitter CSI on the downlink have been derived in [49, 28].

The single user power optimal scheduling with delay constraints discussed in Section 4.1 has been formulated in a seminal paper by Berry and Gallager [6]. The paper has also quantified power-delay tradeoff and demonstrates the convexity of the power delay curve. Structural properties of the optimal policy under various assumptions on the channel and arrival processes have been proved in [5, 20, 18, 1, 37]. These problems have the structure of CMDP. [2] is an excellent monograph on CMDP. While there have been a number of studies to explore structural properties of the power optimal policy, numerical computation using dynamic programming is hard. Heuristic algorithms have been developed in [54]. Interestingly, Delay constrained power optimal scheduling has also been considered under non-fading gaussian noise channel [38] since the power-rate convexity relation also holds for additive white gaussian channel.

The learning scheme of Section 4.1.3 is from [39]. General references on reinforcement learning algorithms are [10, 47] while [7, 36] can be referred for dynamic programming. The result bounding the number of states at which the optimal stationary policy for a constrained problem takes randomized decisions can be found in [12]. The post-decision state concept used in this chapter was introduced in [53]. Similar ideas have been around for quite some time, see, e.g., [32]. For further details on stochastic approximation, refer to [46, 14, 25]. Techniques used to prove convergence of stochastic approximation schemes as well as multi-time scale and asynchronous stochastic approximation (which can be employed to prove the convergence of learning scheme of this chapter) have been discussed in [14].

Multiuser power optimal scheduling under delay constraint is a somewhat difficult problem. This problem has been considered for sum power minimization subject to delay constraints in [34]. However, this scenario corresponds to the downlink scheduling. For an uplink case, weighted power minimization with a special case of two users has been studied in [5]. Power minimization of each user on the uplink subject to individual delay constraints has been explored in [41]. The discussion in Section 4.2 borrows from [41]. For a general discussion on theory of multiobjective optimization see [17, 42].

Various notions of fairness have been explored (See Chapter 8 [27] and the references therein). The proportional fair scheduler has been proposed in [16]. For a discussion on ‘proportional fairness’ and associated properties see [21]. Long term sum throughput maximization subject to providing minimum throughput or fraction of slots to users has been variously considered in [29, 15, 4, 31]. Formulation (38) has been considered in [29]. Short term fairness has been investigated in [24]. Formulation (40) in this chapter is from [24].

Throughput optimal policies have been considered in [50, 35]. LCQ has been suggested in [48], EXP in [43], LWQHPR in [56] and M-LWDF is discussed in [3]. Delay optimality of LQHPR policy has been proved in [56]. The indexing heuristic of Section 5.2 has been proposed in [40].

## References

1. Agarwal, M., Borkar, V.S., Karandikar, A.: Structural Properties of Optimal Transmission Policies over a Randomly Varying Channel. IEEE Transactions on Automatic Control **53**(6), 1476–1491 (2008)
2. Altman, E.: Constrained Markov Decision Processes. Chapman and Hall/CRC Press, Boca Raton, FL (1999)
3. Andrews, M., Kumaran, K., Ramanan, K., Stolyar, A., Whiting, P., Vijayakumar, R.: Providing Quality of Service over a Shared Wireless Link. IEEE Communications Magazine **39** (1996)
4. Berggren, F., Jantti, R.: Asymptotically Fair Transmission Scheduling over Fading Channels. IEEE Transactions on Wireless Communication **3**(1), 326–336 (2004)
5. Berry, R.: Power and Delay Trade-offs in Fading Channels (2000). PhD Thesis, Massachusetts Institute of Technology
6. Berry, R.A., Gallager, R.G.: Communication over Fading Channels with Delay Constraints. IEEE Transactions on Information Theory **48**(5), 1135–1149 (2002)
7. Bertsekas, D.P.: Dynamic Programming and Optimal Control, vol. 1. Athena Scientific, Belmont, MA (1995)
8. Bertsekas, D.P.: Nonlinear Programming. Athena Scientific, Belmont, MA (1999)
9. Bertsekas, D.P., Gallager, R.: Data Networks. Prentice Hall (1987)
10. Bertsekas, D.P., Tsitsiklis, J.N.: Neuro-Dynamic Programming. Athena Scientific, Belmont, MA (1996)
11. Biglieri, E., Proakis, J., Shamai, S.: Fading Channels: Information-Theoretic and Communications Aspects. IEEE Transactions on Information Theory **44**(6), 2619–2692 (1998)
12. Borkar, V.S.: Convex Analytic Methods in Markov Decision Processes. In: E. A. Feinberg, A. Schwartz (Eds.) Handbook of Markov Decision Processes, pp. 347–375. Kluwer Academic Publishers, Dordrechet (2001)
13. Borkar, V.S.: An Actor-Critic Algorithm for Constrained Markov Decision Processes. Systems and Control Letters **54**, 207–213 (2005)
14. Borkar, V.S.: Stochastic Approximation - A Dynamical Systems Viewpoint. Hindustan Publishing Agency and Cambridge University Press (2008)
15. Borst, S., Whiting, P.: Dynamic Channel-Sensitive Scheduling Algorithms for Wireless Data Throughput Optimization. IEEE Transactions on Vehicular Technology **52**(3), 569–586 (2002)

16. Chaponniere, E.F., Black, P., Holtzman, J.M., Tse, D.: Transmitter directed Multiple Receiver System using Path Diversity to Equitably Maximize Throughput (2002). U.S. Patent No. 6449490
17. Coello, C.A.C.: A Comprehensive Survey of Evolutionary-Based Multiobjective Optimization Techniques. Knowledge and Information Systems **3**(1), 269–308 (1999)
18. Djonin, D.V., Krishnamurthy, V.: Structural Results on Optimal Transmission Scheduling: A Constrained Markov Decision Process Approach. The IMA Volumes in Mathematics and its Applications **143** (2007)
19. Goldsmith, A., Varaiya, P.: Capacity of Fading Channels with Channel Side Information. IEEE Transactions on Information Theory **43**(6), 1986–1992 (1997)
20. Goyal, M., Kumar, A., Sharma, V.: Power Constrained and Delay Optimal Policies for Scheduling Transmissions over a Fading Channel. In: Proceedings of IEEE INFOCOM, vol. 1, pp. 311–320 (2003)
21. Kelly, F.: Charging and Rate Control for Elastic Traffic. European Transactions on Telecommunications **8**, 33–37 (1997)
22. Kelly, F.P., Maulloo, A.K., Tan, D.K.H.: Rate Control for Communication Networks: Shadow Prices, Proportional Fairness and Stability. The Journal of the Operational Research Society **49**(3), 237–252 (1998)
23. Knopp, R., Humblet, P.A.: Information Capacity and Power Control in Single-Cell Multiuser Communications. In: Proceedings of IEEE ICC, vol. 1, pp. 331–335. Seattle, USA (1995)
24. Kulkarni, S., Rosenberg, C.: Opportunistic Scheduling Policies for Wireless Systems with Short Term Fairness Constraints. In: Proceedings of IEEE GLOBECOM, vol. 1, pp. 533–537 (2003)
25. Kushner, H.J., Yin, G.G.: Stochastic Approximation Algorithms and Applications. Springer-Verlag, New York (1997)
26. LAN/MAN Committee: IEEE 802.16-2004: IEEE standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access Systems. IEEE Computer Society (2004)
27. Lau, V.K.N., Kwok, Y.K.R.: Channel-Adaptive Technologies and Cross-Layer Designs for Wireless Systems with Multiple Antennas. Wiley-Interscience (2006)
28. Li, L., Goldsmith, A.: Capacity and Optimal Resource Allocation for Fading Broadcast Channels: Part I: Ergodic Capacity . IEEE Transactions on Information Theory **47**(3), 1083–1102 (2001)
29. Liu, X., Chong, E., Shroff, N.: Opportunistic Transmission Scheduling with Resource-Sharing Constraints in Wireless Networks. IEEE Journal on Selected Areas in Communication **19**(10), 2053–2064 (2001)
30. Liu, X., Chong, E., Shroff, N.: Transmission Scheduling for Efficient Wireless Resource Utilization with Minimum-Performance Guarantees. In: Proceedings of IEEE VTC'2001, vol. 2, pp. 824–828 (2001)
31. Liu, Y., Knightly, E.: Opportunistic Fair Scheduling over Multiple Wireless Channels. In: Proceedings of IEEE INFOCOM, vol. 2, pp. 1106–1115. San Francisco, USA (2003)
32. Magirou, V.F.: Stockpiling under Price Uncertainty and Storage Capacity Constraints. European Journal of Operational Research pp. 233–246 (1982)
33. McEliece, R., Stark, W.: Channels with Block Interference. IEEE Transactions on Information Theory **30**, 44–53 (1984)
34. Neely, M.J.: Optimal Energy and Delay Tradeoffs for Multi-user Wireless Downlinks. In: Proceedings of IEEE INFOCOM, pp. 1–13 (2006)
35. Neely, M.J., Modiano, E., Rohrs, C.: Power and Server Allocation in Multi-beam Satellite with Time Varying Channels. In: Proceedings of IEEE INFOCOM, pp. 138–152. San Francisco, USA (2002)
36. Puterman, M.: Markov Decision Processes. Wiley, New York (1994)
37. Rajadhyaksha, G., Borkar, V.S.: Transmission Rate Control over Randomly Varying Channels. Probability in the Engineering and Informational Sciences **19**(1), 73–82 (2005)
38. Rajan, D., Sabharwal, A., Aazhang, B.: Delay-Bounded Packet Scheduling of Bursty Traffic over Wireless Channels. IEEE Transactions on Information Theory **50**(1), 125–144 (2004)

39. Salodkar, N., Bhorkar, A., Karandikar, A., Borkar, V.S.: An On-Line Learning Algorithm for Energy Efficient Delay Constrained Scheduling over a Fading Channel. IEEE JSAC, Special Issue on Control and Communications **26**(4), 732–742 (2008)
40. Salodkar, N., Karandikar, A.: An Indexing Scheduler for Delay Constrained Scheduling with Applications to IEEE 802.16. In: Proceedings of IEEE WCNC'2008, pp. 1471–1476 (2008)
41. Salodkar, N., Karandikar, A., Borkar, V.S.: A Stable On-line Algorithm for Energy Efficient Multi-user Scheduling. URL `http://www.ee.iitb.ac.in/~karandi/pubs_dir/submitted/uplink_tmc_revised_july2009.pdf`. To appear in IEEE Transactions on Mobile Computing, 2010
42. Sawargi, Y.: Theory of Multiobjective Optimization. Academic Press, Orlando (1985)
43. Shakkottai, S., Stolyar, A.: Scheduling for Multiple Fows Sharing a Time-varying Channel: the Exponential Rule. AMS Translations Series 2 **207** (2002)
44. Sklar, B.: Rayleigh Fading Channels in Mobile Digital Communication Systems. I. Characterization. IEEE Communications Magazine **35**(7), 90–100 (1997)
45. Sklar, B.: Rayleigh Fading Channels in Mobile Digital Communication Systems. II. Mitigation. IEEE Communications Magazine **35**(7), 102–109 (1997)
46. Spall, J.C.: Introduction to Stochastic Search and Optimization. Wiley (2003)
47. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. MIT Press, Cambridge MA, USA (1998)
48. Tassiulas, L., Ephremides, A.: Dynamic Server Allocation to Parallel Queues with Randomly Varying Connectivity. IEEE Transactions on Information Theory **39**(2), 466–478 (1993)
49. Tse, D.: Optimal Power Allocation over Parallel Gaussian Broadcast Channels. In: Proceedings of IEEE ISIT, p. 27 (1997)
50. Tse, D., Hanly, S.: Multi-access Fading Channels - Part I: Polymatroid Structure, Optimal Resource Allocation and Throughput Capacities. IEEE Transactions on Information Theory **44**(7), 2796–2815 (1998)
51. Tse, D., Hanly, S.: Multi-access Fading Channels - Part II: Delay-Limited Capacities. IEEE Transactions on Information Theory **44**(7), 2816–2831 (1998)
52. Tse, D., Viswanath, P.: Fundamentals of Wireless Communication. Cambridge University Press (2005)
53. W. B. Powell: Approximate Dynamic Programming: Solving the Curses of Dimensionality. Wiley-Interscience (2007)
54. Wang, H., Mandayam, N.: A Simple Packet Transmission Scheme for Wireless Data over Fading Channels. IEEE Transcations on Communications **52**(7), 1055–1059 (2004)
55. Wang, H.S., Moayeri, N.: Finite-State Markov Channel – A Useful Model for Radio Communication Channels. IEEE Transactions on Vehicular Technology **44**(1), 163–171 (1995)
56. Yeh, E., Cohen, A.: Delay Optimal Rate Allocation in Multiaccess Fading Communications. In: Proceedings of IEEE Workshop on Multimedia Signal Processing, pp. 404–407 (2002)