Combined Flow Control and Interference Cancellation for Packet Data Transmission in Wideband CDMA Systems

Amine Maaref, Sonia Aïssa, Sofiène Affes
INRS-Telecommunications
University of Quebec, Montreal, Canada
{maaref, aissa, affes}@inrs-telecom.uquebec.ca

Abstract—**We consider packet scheduling and rate assignment on the uplink of a packet data wireless CDMA network in the presence of imperfect Interference Cancellation (IC) and limited user transmission rates, and subject to in-cell and out-of-cell resource limitations. The objective is to propose and implement a system level position-based flow control algorithm that accounts for a limited IC capability provided by power control for multi-user detection. The proposed algorithm assigns packets to be transmitted to separate queues, one for each spatial zone within which packets generate roughly the same in-cell interference and impose equal interference on a neighboring base station. Given the cell partitioning into zones, the algorithm dynamically adapts to the resource constraints and efficiently uses IC to provide for fairness in serving the various queues without giving up the objective of maximizing data throughput. Throughput and fairness are two conflicting objectives to be optimized. We show that the joint use of IC and location-based scheduling is able to achieve complete fairness with negligible loss in throughput even under stringent resource limitations.**

I. Introduction

Code Division Multiple Access (CDMA) systems are interference limited. Managing the interference generated by packet transmissions is expected to improve transmission performance both in terms of throughput and fairness. The goal of this study is to provide an uplink flow control algorithm for packet data transmission where the control accounts for IC to achieve a better throughput-fairness tradeoff curve. The transmit layer algorithm exploits useful information that is made available by the physical layer and adapts easily to the resource availability.

The uplink flow control problem is tightly related to power control and can be formulated as the selection of packets to be transmitted from mobiles that have previously made a transmission request to the serving base station. The selection is made such that time variation in the available resources is exploited while ensuring fairness among the active mobiles irrespective of their location within the cell. Fairness is a concern since mobiles that are near the edge of a cell need more transmission power per packet than those that are closer to the base station. These mobiles are, therefore, the ones that generate more interference to a neighboring base station, which may result in excessive outage there if that cell is heavily loaded.

In a previous study, we proposed a flow control algorithm that adapts to the existing resource availability and results in significantly higher network utilization [1], [2]. The study was for the downlink only. We explored the advantages of dividing the cells into regions defined by equal resource requirements and showed that the algorithm responds to short-time resource variations to achieve high throughput with a low likelihood of overload. A similar approach has been applied to the uplink [3]. However, resources available were considered time-invariant with no selectable user transmission rates and no implementation of IC. In this paper, we adapt our previously proposed formulation to the uplink, considering time-variant resources and accounting for the available user transmission rates. We also propose a location-based control for packet flow at the base stations of power-controlled CDMA networks in the presence of imperfect IC. The new allocation algorithm is designed to take advantage of interference reduction capabilities to provide any desired tradeoff between throughput and fairness. The transmit strategy proposed is shown capable of achieving a better tradeoff between throughput and fairness compared to the case with no IC.

Despite the notable increase in capacity offered by various multi-user detection techniques, industry has been slow to offer IC in practical systems. One sub-optimal but of reduced complexity technique is Interference Subspace Rejection (ISR) [4]. ISR is an IC technique that is able to operate at complexity levels as low as those offered by Successive or Parallel IC detectors (SIC, PIC), while providing higher interference suppression efficiency. ISR can be performed either Successively (ISR-S), or in Parallel (ISR-P). Herein, we consider ISR-S. Incorporation of hybrid modes is left for future work. ISR-S successively nulls the interference originating from previously decoded users in the composite signal received at the base station. Subsequently decoded users will thereby experience reduced interference.

The remainder of this paper is organized as follows. In section II, we state the system model and describe the problem. Section III characterizes the resource consumptions and interference limitations used by the scheduling algorithm that is described in section IV. Finally, we give some application results in section V. Concluding remarks are provided in the last section.

II. System Model and Problem Description

We consider a hexagonal cell geometry with a single layer of surrounding cells as illustrated in Figure 1. A cell is divided into three 120° sectors with transmissions from a pair of regions each consisting of n_z zones. Each region of the pair generates interference to the base station opposing it. Consider BS_0 as the target base station; two neighboring base stations are identified, BS_1 and BS_2, each affected by transmissions from one of the two regions of BS_0. A simple example of such a configuration is shown in Figure 1 with $n_z = 2$. In this case, a region will be referred to subsequently as a pair of inner and outer zones.

Work supported by the Mission Universitaire de Tunisie en Amérique du Nord and the Canadian NSERC Research Grants Program.

0-7803-7802-4/03/$17.00 © 2003 IEEE

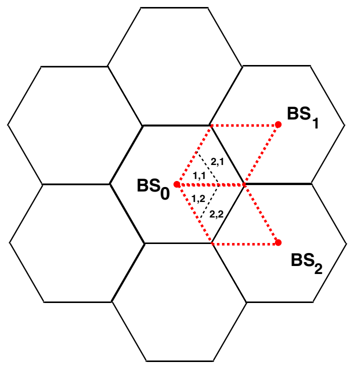

Fig. 1. The hexagonal cell geometry with first layer of surrounding cells and sector partioning into zones.

We consider a discrete set of allowable rates that are multiples of a basic rate R_b. Data users are assumed to require the same quality of service. The different rates are accommodated by varying the spreading gain so that all the transmitted signals occupy the same total bandwidth W.

The objective of packet flow control is to determine the best transmission assignment per time slot to mobiles requesting packet transmissions, given the time-varying resource availability, mobility and time-varying transmission demands. For this purpose, mobiles are assigned to zones based on their current power requirements and are periodically reassigned to sectors and zones as a result of mobility. The flow control algorithm determines a transmit matrix denoting the allocated rates at each time slot, such that the current in-cell and out-of-cell interference limitations are not exceeded. However, when resource limitations are stringent, transmissions from mobiles in the outer zones cannot be allowed since they generate most of the interference to the neighboring base stations and hence delay for these users builds up. As a result, not only does unfairness increase, but also persists when the available resources vary slowly, and improves only when the interference is properly managed. We rely on successive IC to reduce the cost in interference associated to mobiles in outer zones. In the decoding process, these mobiles will be considered after the ones in the inner zones, thus reducing their transmission powers and thereby the interference they generate. This is expected to achieve more fairness while still striving to maximize throughput.

III. Resource Characterization

A. Resource Constraints

For a given 120°sector, three resource constraints are identified: one in-cell, and two out-of-cell. The in-cell resource utilization corresponds to the total power received at the base station and is represented as a linear function of the number of packets transmitted from the inner and outer zones. The out-of-cell resource utilizations correspond to the out-of-cell interference generated in the facing neighboring base stations by the packets transmitted to the target base station. We assume that resource availabilities can be predicted adequately based on the resource utilization measurements for the current time slot and communicated between base stations at each time slot.

For target BS_0, let IC^l be the in-cell power limit during time slot l, and $OC^l_{j,\{j=1,2\}}$ the out-of-cell interference margins respectively allowed by $\text{BS}_{j,\{j=1,2\}}$ for transmissions originating from mobiles in zones $(i,j)_{\{i=1,\ldots,n_z\}}$. We normalize these in-cell and out-of-cell interference margins by the interference generated at the target BS_0 by an arriving packet with the minimum SIR required [6]. This power corresponds to the equal-power solution provided by the power control module. Thus, the limits for the considered sector respectively translate into corresponding tolerable numbers of packets per time slot, say NI^l and $NOC^l = [NOC^l_1, NOC^l_2]$. These limits actually stand for the average maximum number of packets that can be transmitted from mobiles in the target sector after support of the ongoing stream services and without giving rise to excessive outage in the neighboring facing sectors.

B. Power Control with Interference Cancellation

Users are decoded with the same Signal-to-Interference Ratio (SIR). We assume perfect power control so that signals originating from mobiles with a rate of m times the basic rate R_b are received at the base with m times the power level that corresponds to transmission at R_b. For one 120° sector with 2 n_z zones and transmissions at R_b, the SIRs of the N users can be written as follows:

$$\Gamma_k = G\,\frac{S_k}{\sum_{i=k+1}^{N} S_i + \sum_{i=1}^{k-1}\theta S_i + N_0} \quad k=1,\ldots,N \qquad (1)$$

where, G is the spreading factor relative to the basic rate, $\{S_k\}_{\{k=1,\ldots,N\}}$ are the receive powers relative to the N users, N_0 is the background noise (includes the other-cell interference) and $0 \le \theta \le 1$ stands for the estimation error in rejecting the interference relative to the signal of user k from the composite signal received at the target BS_0. θ is assumed known, equal for all i, and independent of k. This can be seen as the worst case scenario when choosing θ equal to the maximum of all the estimation errors. The interference rejection efficiency is then defined by $\eta = 1-\theta$. Setting $\Gamma_k = \Gamma$ for $k = 1,\ldots,N$, the set of optimal powers $\{S_k\}_{k=1,\ldots,N}$ satisfies a recursive solution [5] given by:

$$S_k = S_{k-1} - \frac{\eta S_{k-1}^2}{V_{k-1} + N_0} \quad k=2,\ldots,N \qquad (2)$$

$$V_k = \sum_{i=1}^{N} S_i - \sum_{i=1}^{k-1}\eta S_i \quad k=1,\ldots,N-1 \qquad (3)$$

These equations are solved iteratively to an arbitrary accuracy, starting with initializing S_1 to S_T/N, where $S_T = \sum_{i=1}^{N} S_i$, and increasing it with some step size $\delta \ll S_1$. Convergence of this algorithm is assured as long as $\theta \le 1$, the only constraint being to properly choose the step size δ which determines the number of iterations needed to converge to the set of optimal powers. In the remainder, we will use $\{S_k(\eta)\}_{\{k=1,\ldots,N\}}$ to denote this set of powers corresponding to a given IC efficiency η.

C. Differential Resource Requirements

Transmissions at the basic rate R_b from mobiles in a zone (i,j) are considered to arrive at the target BS_0 with the same average power level and generate, on average, the same amount of out-of-cell interference to the facing BS_j. For the purpose of flow control, we differentiate the resource requirements among mobiles and zones on a per-packet basis. Hence, $\{S_k(\eta)\}_{\{k=1,\dots,N\}}$ will be used to denote the set of optimum powers per packet.

Let $\overline{\alpha}_{i,j}$ denote the normalized average power of an arriving packet at BS_0 from a given zone (i,j):

$$\overline{\alpha}_{i,j}(\eta) = \overline{S}_{i,j}(\eta) \;/\; S \;, \tag{4}$$

where, $\overline{S}_{i,j}(\eta) = \frac{\sum_{k \in \aleph_i} S_k(\eta)}{N_i}$, $\aleph_i$ is the set of indices of mobiles in zones $(i,j)_{\{j=1,2\}}$, and N_i is the number of these mobiles[1].

Now, let a packet be transmitted from zone (i,j) and calculate the average amount of interference generated by this transmission to the facing BS_j. Considering the path loss between the mobile and the target BS_0 proportional to $10^{(\xi/10)} d^{-4}$ (d is the distance from a mobile in zone (i,j) to target BS_0 and ξ is a Gaussian random variable with zero mean and standard deviation $\sigma = 8$dB), the interference contributed by this packet transmission to BS_j is given by $\overline{S}_{i,j}(\eta)\,(d/d_j)^4\, 10^{(\xi/10)}$ where d_j is the distance from the mobile in zone (i,j) to BS_j as shown in Figure 2. For simplicity, we consider 4 zones in each sector as shown in Figure 1. Following a transmission of a packet from zone (i,j) in BS_0, the normalized interference generated to the facing BS_j is given by $\beta_{i,j}(\eta) = \overline{\alpha}_{i,j}(\eta)\,(d/d_j)^4\, 10^{(\xi/10)}$.

If (x,y) are the mobile's coordinates (Figure 2), and given that $d = (x^2+y^2)^{1/2}$ and $d_j = ((R\sqrt{3}-x)^2+y^2)^{1/2}$, we calculate an average value of $\beta_{i,j}$ as

$$\overline{\beta}_{i,j}(\eta) = \overline{\alpha}_{i,j}(\eta)\,\overline{\gamma}_{i,j} \;. \tag{5}$$

Given that $\gamma(x,y) = (x^2+y^2)^2 \;/\; ((R\sqrt{3}-x)^2+y^2)^2$, and denoting $A_{i,j}$ as the area of zone (i,j), the zone average coefficient $\overline{\gamma}_{i,j}$ is found by numerically evaluating the integral $\frac{1}{A_{i,j}} \iint_{\text{zone }(i,j)} \gamma(x,y)\,dx\,dy$. In the presence of significant shadowing, the power-based zone assignment may result in complex zone boundaries, thus for convenience we have omitted shadowing considerations from our calculations. The separation between the inner and outer zones is determined so as to minimize the mean-squared error between the actual zone coefficient values at any point in the zone and the averaged value relative to each zone. For a pair of inner and outer zones opposing BS_j, it is easy to show that the coefficient γ_s at the separation line between inner and outer zones is constant and lies on a circle with radius $r_s = \frac{R\sqrt{3}\sqrt{\gamma_s}}{1-\sqrt{\gamma_s}}$ and centered at $(-r_s\,{\gamma_s}^{1/4}, 0)$, where R is the cell radius. The intersection between this circle and the 60° region constitutes the separation between the zones (Figure 2). Given a separation line defined by the contour of constant γ_s, the mean squared error is expressed in (6), where the region area A_j is given by $A_j = \sum_i A_{i,j}$.

$$\mathcal{E}(\gamma_s) = \frac{1}{A_j} \sum_i \iint (\gamma(x,y) - \overline{\gamma}_{i,j})^2\,dx\,dy \;. \tag{6}$$

Let γ_s^* be the value of γ_s that corresponds to the line of separation that minimizes $\mathcal{E}$ (Figure 3). γ_s^* is found to be 0.44 (-3.57dB). This yields $\overline{\gamma}_{1,1} = \overline{\gamma}_{1,2} = 0.08$, and $\overline{\gamma}_{2,1} = \overline{\gamma}_{2,2} = 0.62$. These values will be used in the simulations.

[1]Note that the actual transmit powers are assigned by power control. The equal-resource assumption is made for the purposes of flow control only.

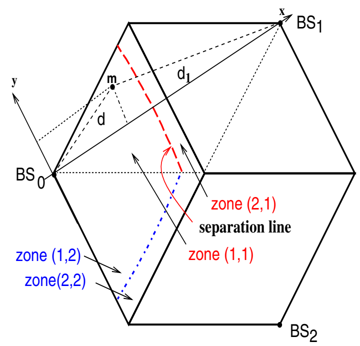

Fig. 2. Separation line between inner and outer zones.

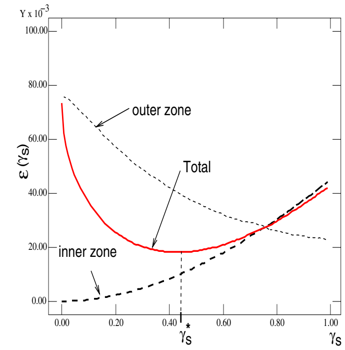

Fig. 3. Mean square error minimization.

IV. Packet Scheduler Design

Packet scheduling is formulated as a constrained integer optimization problem following the method proposed in [2] for the downlink. The formulation uses an objective function composed of a weighted sum of throughput, fairness, and a function which quantifies the proximity to the available remaining resources.

For one 120° sector with 2 n_z zones, the resource constraints for a time slot l can be expressed as

$$\begin{array}{l} \sum_{i=1}^{n_z} \left(\overline{\alpha}_{i,1}(\eta) \sum_{u \in \mathcal{U}_{i,1}^l} n_{i,1}^u + \overline{\alpha}_{i,2}(\eta) \sum_{u \in \mathcal{U}_{i,2}^l} n_{i,2}^u\right) \leq NI^l \\ \sum_{i=1}^{n_z} \overline{\beta}_{i,1}(\eta) \sum_{u \in \mathcal{U}_{i,1}^l} n_{i,1}^u \leq NOC_1^l \\ \sum_{i=1}^{n_z} \overline{\beta}_{i,2}(\eta) \sum_{u \in \mathcal{U}_{i,2}^l} n_{i,2}^u \leq NOC_2^l \\ 0 \;\leq\; n_{i,j}^u \;\leq\; M;\, u = 1,\dots,U_{i,j}\,;\, i = 1,\cdots,n_z\,;\, j = 1,2 \end{array} \tag{7}$$

where, $U_{i,j}$ is the number of users in zone (i,j) and $\mathcal{U}_{i,j}^l$ is the set of indices denoting the users in zone (i,j) which are allowed access to the available resources in the current time slot, $n_{i,j}^u$ is the number of packets transmitted by user u in zone (i,j), and M is the user's maximum allowable transmission rate expressed in terms of number of packets per time slot.

We define a resource proximity function that measures the resource availability associated with an assignment matrix $\mathbf{n}(l) = \{\mathbf{n}_{1,1},\cdots,\mathbf{n}_{n_z,1},\mathbf{n}_{1,2},\cdots,\mathbf{n}_{n_z,2}\}$ in time slot l where each column vector, of dimension $U = max_{i,j} U_{i,j}$, indicates the number of packets transmitted by each user. An element vector $\mathbf{n}_{i,j}$ is given by $\mathbf{n}_{i,j} = [n_{i,j}^1, n_{i,j}^2, \cdots, n_{i,j}^{U_{i,j}}, 0, \cdots, 0]^T$ with $n_{i,j}^u = 0$ if $u \notin \mathcal{U}_{i,j}^l$. The resource proximity is defined as the proximity to the nearest resource limit, measured in terms of the additional packets that may be transmitted from the most tightly constrained zone and is expressed as

$$\mathbf{P_n}(l) \;=\; \min_{i,j}\; \hat{n}_{i,j}(\mathbf{n}) \tag{8}$$

where, $\hat{n}_{i,j}$ is the maximum number of packets that could be transmitted from zone (i,j) given the available remaining resources at time slot l, and expressed as

$$\hat{n}_{i,j}(\mathbf{n}) = \min\left\{\left\lfloor RNI^l(\mathbf{n})/\overline{\alpha}_{i,j}(\eta)\right\rfloor, \left\lfloor RNOC_j^l(\mathbf{n})/\overline{\beta}_{i,j}(\eta)\right\rfloor\right\} \tag{9}$$

where $\lfloor\;.\;\rfloor$ denotes the integer part of a number, RNI^l and $RNOC_{\{j=1,2\}}^l$ being respectively the in-cell and out-of-cell available remaining resources corresponding to an assignment matrix $\mathbf{n}$.

Packet scheduling consists of finding, at each time slot l, the transmit matrix $\mathbf{n}^f(l)$ which jointly maximizes throughput and fairness, while ensuring that the resource constraints (Eq. 7) are satisfied and rates are properly allocated to each user. To provide fair allocation of resources among users for equitable levels of service while maintaining an acceptable throughput, we define the optimization criterion as the maximization of the functional $\mathcal{OF}_{\mathbf{n}}(l)$ corresponding to an assignment $\mathbf{n}$ at time slot l

$$\mathcal{OF}_{\mathbf{n}}(l) = \mathbf{T}_{\mathbf{n}}(l) + \mathbf{P}_{\mathbf{n}}(l) + \lambda\, \mathbf{F}_{\mathbf{n}}(l) \tag{10}$$

where, $\mathbf{T}_{\mathbf{n}}(l)$ is the throughput expressed in total number of packets transmitted in time slot l, $\mathbf{P}_{\mathbf{n}}(l)$ is the resource proximity resulting from assignment $\mathbf{n}(l)$ (Eq. 8), $\mathbf{F}_{\mathbf{n}}(l)$ is the fairness obtained with assignment $\mathbf{n}(l)$, and the coefficient λ is chosen to tune the trade-off between throughput and fairness. The fairness metric is defined in terms of the variance of delays on the remaining head-of-queue packets [2].

The optimization problem is solved in two parts. The first function of the algorithm consists in searching for the optimal assignment matrix $\mathbf{n}^*(l)$ given no limitation on the availability of rates. This assignment matrix is the one that maximizes the objective function (Eq. 10) under the identified constraints (Eq. 7). The second function of the algorithm is to find the final assignment matrix $\mathbf{n}^f(l)$, given $\mathbf{n}^*(l)$ and a set of allowable transmission rates $\mathcal{R}$.

For a given fairness multiplier λ and an initial assignment matrix, the algorithm iteratively updates the assignment matrix $\mathbf{n}^{(m)}(l)$, increasing index m until the stopping criterion is met. Index m counts the iterations until the optimal assignment matrix $\mathbf{n}^*(l)$ is reached. If the starting point is the zero matrix, the index counts the packets in the assignment matrix.

As long as the resource constraints are satisfied and the objective function increases, the algorithm iterates on m according to the following steps:

1) Define up to 2 n_z possible assignments that include one additional packet to be transmitted from non-empty queues.
2) Inhibit the assignments that violate the constraints.
3) If there are no feasible assignments, stop. Else, continue.
4) Determine the $\mathcal{OF}$ value associated with each assignment and select the assignment that results in the highest value.
5) Update the functional $\mathcal{OF}^{(m)}_{\mathbf{n}}(l)$ and the delay set corresponding to the heads of active queues.
6) Set $m = m + 1$ and repeat from 1).

The final assignment matrix for the current time slot $\mathbf{n}^f(l)$ is equal to $\mathbf{n}^*(l)$ in case the transmission rates found can be accomodated by variable spreading. Consider a more realistic situation where a number of allowable rates is available. Denote by $\mathcal{R}$ this set of rates expressed in terms of numbers of packets per time slot. Given the optimal assignment matrix $\mathbf{n}^*(l)$ and the set of allowable rates $\mathcal{R}$, user rates should be allocated so that to result in the least spare resources. For this purpose, we proceed separately along each column vector of $\mathbf{n}^*(l)$ to find the final assignment matrix $\mathbf{n}^f(l)$ with elements in $\mathcal{R}$.

For a given zone (i,j), the algorithm first finds a set of lower and upper assignment vectors (referred to by l for low, and h for high in superscripts), $\underline{\mathbf{n}}^l_{i,j}$ and $\underline{\mathbf{n}}^h_{i,j}$ that satisfy the following criteria

$$\begin{aligned} \underline{\mathbf{n}}^l_{i,j} &= arg_{\{\underline{\mathbf{n}} \in \mathcal{R}^U\ :\underline{\mathbf{n}}\ \leq\ \underline{\mathbf{n}}^*_{i,j}\}}\ min\ \| \underline{\mathbf{n}} - \underline{\mathbf{n}}^*_{i,j} \|^2 \\ \underline{\mathbf{n}}^h_{i,j} &= arg_{\{\underline{\mathbf{n}} \in \mathcal{R}^U\ :\underline{\mathbf{n}}\ \geq\ \underline{\mathbf{n}}^*_{i,j}\}}\ min\ \| \underline{\mathbf{n}} - \underline{\mathbf{n}}^*_{i,j} \|^2 \end{aligned} \tag{11}$$

where $\|\ .\ \|$ designates the Eucledian norm of a vector.

Having $\underline{\mathbf{n}}^l_{i,j}$ and $\underline{\mathbf{n}}^h_{i,j}$, we calculate the number of packets $N_{i,j}$ that is to be reallocated while meeting the constraints on the available user transmission rates. Let $N_{i,j}$ be this number of packets which is given by

$$N_{i,j} = \sum_{u \in \mathcal{U}^l_{i,j}} (n^{*u}_{i,j} - n^{lu}_{i,j}). \tag{12}$$

The final allocation vector $\underline{\mathbf{n}}^f_{i,j}$ satisfies

$$\underline{\mathbf{n}}^f_{i,j} = arg_{\{\underline{\mathbf{n}} \in \mathcal{R}^U\}}\ min\ \{\Delta_{i,j} = N_{i,j} - \sum_{u \in \mathcal{U}^l_{i,j}} (n^u - n^{lu}_{i,j})\}\ . \tag{13}$$

If there is more than one suitable vector satisfying Eq. 13, we select the closest one (in the sense of the Eucledian norm) to the optimal assignment vector $\underline{\mathbf{n}}^*_{i,j}$.

Therefore, for a given zone (i,j), the discretization strategy may be described as follows [2] :

1) Find the vectors $\underline{\mathbf{n}}^l_{i,j}$ and $\underline{\mathbf{n}}^h_{i,j}$ that satisfy Eq. 11.
2) Initialize $\underline{\mathbf{n}}^f_{i,j}$ to $\underline{\mathbf{n}}^h_{i,j}$ and $\underline{\mathbf{v}}_{i,j}$ to $(\underline{\mathbf{n}}^f_{i,j} - \underline{\mathbf{n}}^l_{i,j})$.
3) Evaluate $\underline{\mathbf{t}} = \underline{\mathbf{v}}_{i,j} - ([\underline{\mathbf{v}}_{i,j} - (\underline{\mathbf{n}}^*_{i,j} - \underline{\mathbf{n}}^l_{i,j})]/\underline{\mathbf{v}}_{i,j})^2$.
4) Determine user $u = arg_{\{u \in \mathcal{U}^l_{i,j}\}}\ min\ t^u$.
5) For this user, set $n^{fu}_{i,j} = n^{lu}_{i,j}$.
6) Evaluate the number of reallocated packets $N = \sum_{u \in \mathcal{U}^l_{i,j}} (n^{fu}_{i,j} - n^{lu}_{i,j})$.
7) If $N < N_{i,j}$, stop (the number of packets not transmitted is $\Delta_{i,j} = N_{i,j} - N$); else, repeat from step 2).

V. Simulations and Discussions

Consider a four zone subdivision in each sector and let $N = 20$ be the number of active users in the target sector with a distribution of $[7\ 3\ 7\ 3]$, respectively in zones $[z_{1,1}\ z_{1,2}\ z_{2,1}\ z_{2,2}]$. The average load per user is assumed to be M packets per time slot. For purposes of comparison, we use a time-to-completion measure defined as the number of time slots required to transmit all packets that arrive up to a given time slot. Herein, results are presented for a packet arrival interval of 100 time slots.

Given the traffic load offered, different operating conditions can be examined to evaluate the performance of the transmission strategy. Such conditions can be simulated by setting the average available resources to result in a system limited in terms of in-cell, out-of-cell or both resources. The algorithm has been studied for a wide range of operating conditions but results provided herein correspond to a worst-case scenario. The latter chosen so that the traffic load exceeds the available resources both in in-cell and out-of-cell. The example depicted is represented by the following parameters: $\overline{NI}$ = 150 packets per time slot, $\overline{NOC}$ = [7,7], M =

[2]Operations on vectors are element-wise, only the non-zero elements on these vectors are considered.

10 packets per user per time slot and the set of available rates is $\mathcal{R} = \{0, 1, 5, 10\}$.

Two values of the parameter η for error estimation are considered, namely $\eta = 0$ for no IC, and $\eta = 0.8$ for an imperfect IC that corresponds to an amount of residual interference $\theta = 0.2$. The value $\eta = 0.8$ can be enabled in practice [7] using the IC technique implemented in this work. We present the results in terms of throughput and queuing delays for different values of the fairness multiplier λ. Four values of λ are considered: $\lambda = 0$ for throughput maximization only, λ approaching ∞ for extreme importance assigned to fairness, and two intermediate values.

Results are organized in two parts. First, we show the effect of fairness on throughput with and without IC, given a set $\{\overline{NI}, \overline{NOC}\}$ of average available resources. In the second part, we show the enhancement in performance achieved when exploiting the increase in the available resources resulting from the implementation of IC.

Consider first the case where no IC is performed. We can see in Figure 4 that during the packet arrival interval, the mean throughput decreases as λ increases. This decrease is traded off for an increase in fairness. This improvement is achieved by striving to equalize the delays of the head-of-queue packets. Delay equalization is indeed improved as λ increases, as can be seen in Figure 5 showing the maximum delay at the head of each queue for $\lambda = \infty$, compared to the results corresponding to $\lambda = 0$ and shown in Figure 6. Using different operating conditions, the performance has been evaluated for different values of λ to allow operation with two intermediate values that we define by $\lambda = 20$ for modest fairness and $\lambda = 50$ for high fairness. Fairness results for these values are not provided in terms of delay, and would be discussed in terms of time-to-completion.

Our scheduling algorithm is capable of ensuring reasonable fairness for both intermediate values of λ without a significant decrease in throughput from the maximum achievable corresponding to $\lambda = 0$. However, when the available resources are very stringent as in the example depicted here, the maximum fairness that can be achieved not only cannot perfectly equalize the delays but results in a loss in throughput of 60% compared to throughput maximization only. This decrease is of 6% for $\lambda = 20$ and 10% for $\lambda = 50$. Consider that a loss of 10% is tolerated and compare the time-to-completion corresponding to the inner and outer zones. The results provided in Table I show how acceptable equalization cannot be achieved under the stringent out-of-cell limits.

TABLE I
COMPARISON OF THE ALGORITHM TIME-TO-COMPLETION FOR $\eta = 0$ AND $\eta = 0.8$, $\overline{NI} = 150$, $\overline{NOC} = [7, 7]$.

	$\lambda = 0$	$\lambda = 20$	$\lambda = 50$	$\lambda = \infty$
$\eta = 0$	$(107, 438)^3$	$(108, 437)$	$(114, 435)$	$(270, 430)$
$\eta = 0.8$	$(115, 333)$	$(117, 330)$	$(123, 330)$	$(247, 329)$

Considering the same values of λ, throughput results with IC used with an efficiency $\eta = 0.8$ are represented in Figure 7. Take $\lambda = 0$, since the system is in-cell resource limited, the use of IC in favor of the users in the outer zones decreases the total throughput compared to no IC. However, as we can see in Figure 8 fairness is considerably increased. As can be seen in Table I, the time-to-completion corresponding to the outer users is reduced from 438 to 333 when the one corresponding to the inner users increases by 8 time slots only. A result that comes at the cost of a reduction in throughput by only 6%. This percentage also corresponds to using $\lambda = 20$ with no IC, but yielding a patently unfair service.

TABLE II
COMPARISON OF THE ALGORITHM TIME-TO-COMPLETION FOR $\eta = 0.8$, $\overline{NI} = 150$, $\overline{NOC} = [14, 14]$.

$\lambda = 0$	$\lambda = 20$	$\lambda = 50$	$\lambda = \infty$
$(132, 178)^3$	$(133, 177)$	$(133, 174)$	$(153, 160)$

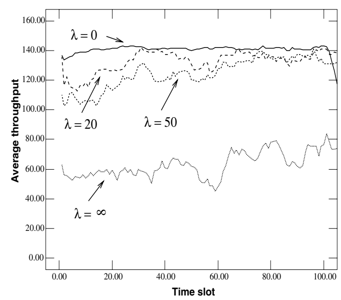

Fig. 4. Average throughput as a function of the fairness multiplier λ for $\eta = 0$.

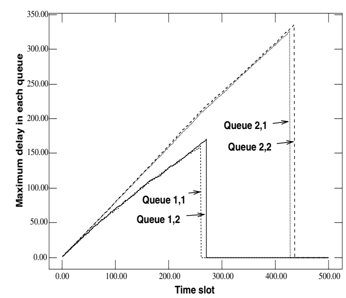

Fig. 5. Maximum delay at the head of each queue resulting from fairness maximization ($\lambda = \infty$).

For $\lambda = \infty$, comparing the results of Figure 9 to those shown in Figure 5 for no IC, we can see how the algorithm is capable of approaching complete fairness. Complete equalization of the delays cannot be achieved due to the fact that the out-of-cell resource limits are very tight. In this case, the gap between the time-to-completion of the inner and outer zones goes from 160 without IC to 82 when IC is implemented, when at the same time, the average throughput in the arrival interval increases by 20%. Taking the extreme cases of $\lambda = 0$ and $\lambda = \infty$, the loss in throughput is 60% without IC, while it is only 48% when IC is implemented.

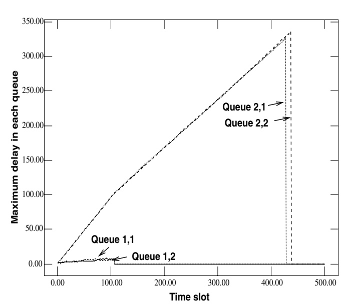

Fig. 6. Maximum delay at the head of each queue resulting from throughput maximization ($\lambda = 0$).

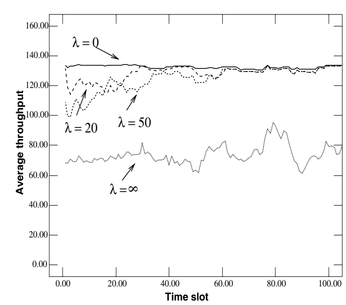

Fig. 7. Average throughput as a function of the fairness multiplier λ.

We study now the advantages of our scheduling algorithm in the presence of IC as a function of the availability of resources. For the given average available resources ($\overline{NI} = 150$ and $\overline{NOC} = [7, 7]$), the target sector is subject to stringent out-of-cell limits. With no IC implemented, remote users experience unacceptable delays and complete fairness cannot be achieved even

[3]Pairs correspond to the algorithm time-to-completion for the inner and outer zones.

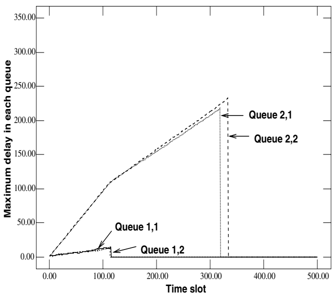

Fig. 8. Maximum delay at the head of each queue resulting from throughput maximization ($\lambda = 0$) for $\eta = 0.8$.

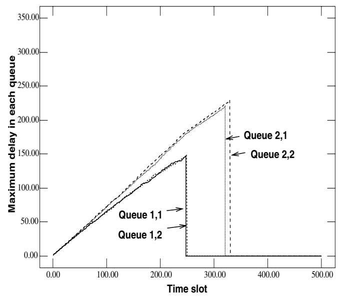

Fig. 9. Maximum delay at the head of each queue resulting from fairness maximization ($\lambda = \infty$) for $\eta = 0.8$.

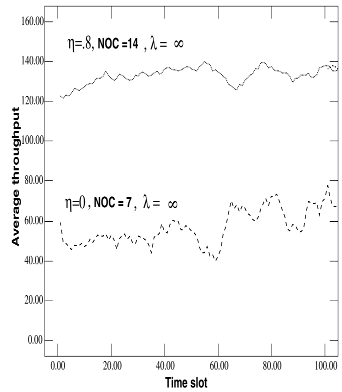

Fig. 10. Comparison of the average throughput resulting from fairness maximization for $\eta = 0$ and $\overline{NOC} = [7, 7]$ versus $\eta = 0.8$ and $\overline{NOC} = [14, 14]$.

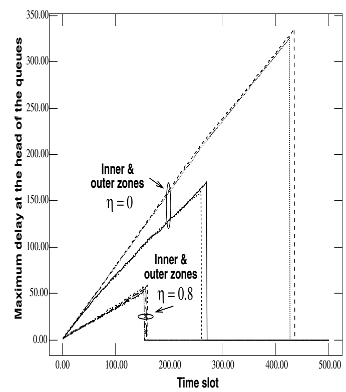

Fig. 11. Comparison of the maximum delay at the head of each queue resulting from fairness maximization for $\eta = 0$ and $\overline{NOC} = [7, 7]$ versus $\eta = 0.8$ and $\overline{NOC} = [14, 14]$.

with $\lambda = \infty$ (Table I). As can be seen in Table I, the time-to-completion for the outer zones, using $\lambda = \infty$, is only 8 time slots lower than that of $\lambda = 0$. The resources being stringent, transmissions from mobiles in the outer zones cannot be allowed. IC on the other hand, when applied to all the sectors in the network allows lowering the transmit power of the outer zone users, thus translating into more available resources to handle the out-of-cell interference. This increase in capacity results in less stringent out-of-cell limits. The new limits can be approximated as: $\overline{NOC}(\eta) = \frac{(1+f)}{(1+f-\eta)} \overline{NOC}(\eta = 0)$, where f is the other-cell to in-cell interference ratio for which a typical value $f = 0.55$ is chosen, assuming a path loss exponent of 4, shadowing standard deviation of $\sigma = 8$dB, and equally loaded cells [8]. This yields an average out-of-cell limit of $\overline{NOC}(\eta = 0.8) = [14, 14]$ allowed for the target sector when the IC efficiency is $\eta = 0.8$.

We show in Figure 10 the throughput values resulting from fairness maximization for both sets of out-of-cell resources. Denote for simplicity the resource limits $\overline{NOC} = [7, 7]$ by Case$_A$ and $\overline{NOC} = [14, 14]$ by Case$_B$. As can be seen in the figure, throughput is considerably increased for Case$_B$ compared to Case$_A$. For $\lambda = \infty$, we observe how the algorithm exploits the availability of resources to increase throughput and considerably decrease the completion time as shown in Table II. If fairness is of importance, a value of $\lambda = \infty$ is used. In this case, while a loss in throughput of 60% for Case$_A$ would reduce the completion time for the outer zones from 438 to 430, the use of IC allows maximum achievable fairness with a loss of only 3% compared to throughput maximization only. As can be seen in Figure 11, while the use of $\lambda = \infty$ and no IC yields high delay values both for the inner and outer zones, the implementation of IC allows more resources to be available allowing higher performance. It is important to mention that the results provided here under heavy load are chosen to emphasize the flexibility of our algorithm in achieving any desirable trade-off between throughput and fairness, and its capability of providing complete fairness that is difficult to achieve when the resources vary slowly and under stringent out-of-cell resource limits.

VI. Conclusion and Future Work

In this paper, we considered the joint use of packet scheduling and Interference Cancellation (IC) to maximize data throughput and fairness on the uplink of wireless CDMA networks. The objective was to devise a low-complexity flow control algorithm that takes advantage of IC and dynamically adapts to the resource constraints to provide adequate compromise between throughput and fairness under limited IC capability. Our scheme assigns packets to be transmitted to separate queues, one for each spatial zone defining a group of users, and characterized by equal average resource requirements. We showed that using flow control with IC can indeed provide for fairness among users without a loss in throughput even under stringent resource limitations. The algorithm is designed to provide adequate compromise between throughput and fairness even under limited IC capability. While we focused on a uniform distribution of mobiles and provided results for a single set of average available resources, our formulation is general enough to account for these situations. Further work includes the benefits of non-homogeneous organization of zones, effects of mobility and operation under hybrid modes of ISR.

References

[1] S. Aïssa and P. Mermelstein, "Downlink Flow Control in WCDMA Networks," *In Proc. IEEE International Conf. on Telecommunications, ICT'2000 Acapulco, Mexico* , pp 41-45, May 2000.

[2] S. Aïssa and P. Mermelstein, "Downlink Flow Control for Wireless CDMA Packet Data Networks," *IEEE Trans. on Vehicular Technology*, vol. 51, no. 5, pp. 1193–1205, September 2002.

[3] E. Haddad, P. Mermelstein, and C. Despins, "Transmit Assignment Algorithm for Uplink Flow Control in WCDMA Networks," in *Proc. IEEE Conf. Vehicular Technology*, pp. 416–420, Fall 1999.

[4] S. Affes, H. Hansen, and P. Mermelstein, "Interference Subspace Rejection: A Framework for Multiuser Detection in Wideband CDMA", *IEEE Jour. on Selec. Areas in Comm.*, vol. 20, no. 2, pp. 287–302, February 2002.

[5] J. Andrews and T. Meng, "Amplitude and Phase Estimation Considerations for Asynchronous CDMA with Successive Interference Cancellation", 52^{nd} *IEEE Veh Tech. Conf.*, vol. 3, pp. 1211–15, September 2000.

[6] J. Kuri, P. Mermelstein, "Call Admission on the Uplink of a CDMA System based on Total Received Power," in *IEEE International Conf. on Communications*, pp. 1431–1436, 1999.

[7] S. Affes, K. Cheikhrouhou, and P. Mermelstein, "High-Speed Data Access with Enhanced Interference Suppression over Wideband CDMA Networks", Invited Paper , *Proc. of IEEE Intl. Conf. on Systems, Man and Cybernetics SMC'2002, Hammamet, Tunisia*, to appear, October 6-9, 2002.

[8] A. J. Viterbi, A. Viterbi, and E. Zehavi, "Other-Cell Interference in Cellular Power Controlled CDMA ," *IEEE Trans. on Vehi. Techno.*,vol. 40, no. 2, 1991.