Optimal Power and Data Rate Assignment when Power-Limited Data and Media Terminals share a 3G CDMA cell: The 3 Terminal Scenario

Virgilio Rodriguez, David J. Goodman
ECE Department, Polytechnic University
5 Metrotech Center
Brooklyn, NY 11201
e-mail: vr@ieee.org

***Abstract*— We study a cell shared by data and media transmitting terminals, within a model relevant to the uplink of a VSG-CDMA cell, a technique part of 3G standards. Each media terminal has a fixed data rate and an inflexible SIR requirement. But the data terminals are delay-tolerant, and their power and data rates can be assigned at will within specified limits, for the benefit of network efficiency. We seek power and data rate allocations that maximize the weighted sum of the throughput of each data terminal, while satisfying the bit rate and SIR requirements of the media terminals. In this paper, 3 terminals are considered: one media transmitting terminal, and two data terminals, one more "important" than the other. Our analysis is based on classical optimization theory, and should accommodate many physical layer configurations.**

I. Introduction

Modern wireless networks will accommodate simultaneous transceivers operating at very different bit rates. Some of the transceivers may be transferring data, while others transfer media content, such as voice, images, or video. Previous work of ours ([1], [2], [3]) study throughput maximization in a model relevant to a single-cell VSG-CDMA system in which each *data* terminal can operate within a range of bit rates, assumed continuous for tractability. This paper discusses how to extend our previous work to consider three additional items: (i) transmission power limits, (ii) non-negligible out-of-cell interference, and (iii) the presence of media-transmitting terminals with fixed bit rates and inflexible signal-to-interference ratio (SIR) requirements.

Power limitations are important for obvious reasons. However, when out-of-cell interference is negligible (system is "interference limited"), the noise term in the SIR expression may be neglected, and the power allocation reduces to finding a vector of carrier-to-interference ratios expressing received power ratios. For example, with only two terminals, power allocation reduces to finding the optimal ratio between the received powers of the two terminals. Theoretically, the optimal power levels are arbitrary, as long as they maintain the optimal ratio. However, when the noise term includes strong out-of-cell interference, the power limitations of the terminals need to be explicitly considered. Additionally, there may be media-transmitting terminals operating at fixed bit rates and SIR. To the data terminals, these media terminals appear as additional sources of "noise".

In the present work, a power-limited media terminal interacts with two data terminals. The cell seeks to maximize a *weighted* sum of the throughputs of the data terminals. We refer to the terminal whose throughput is weighted more heavily as "important". The weights admit various interpretations, including levels of importance or priority among the *data* terminals, "utility" per bit each terminal derives, or monetary prices paid by the terminals. The data terminals are delay-tolerant, and operate at power and data rates that can be assigned at will within specified limits. However, the media terminal operates at a fixed data rate, and has an inflexible SIR requirement.

The throughput maximization problem with media terminals present does not appear to have been treated in the literature, especially within an analytical model. However, in addition to references [1], [2], [3], other references treat throughput maximization among data terminals. Reference [4] represents a related strand of work involving these authors, in which the data rates are fixed and identical, and the throughput-maximizing number of terminals is found, along with the power allocation. Reference [5] has much in common with our previous works, except that this reference does not consider weights and applies a simplifying linearization. Reference [6] also seeks data rates and power allocation, and consider a "sigmoidal-like" frame-success function, but focuses on the downlink, does not consider weights, and provides a sub-optimal algorithmic solution based on pricing. Reference [7] maximizes a fairly general "capacity function", does not consider weights, and assume that the terminal's data rates are different but fixed. Other related works seek decentralized solutions.

Another significant difference between our models and the literature is our characterization of the frame-success function (FSF), which gives the probability that a data packet is received successfully in terms of the terminal's received signal-to-interference ratio (SIR). This function depends on many physical attributes of the system, such as the modulation technique, the forward error detection scheme, the nature of the channel, and properties of the receiver. We do *not* impose any particular algebraic functional form ("equation") on the FSF. Rather, we assume that *all that is known* about this function is that its graph is a smooth S-shaped curve, as displayed in fig. 1, and base our analysis on properties derived from this shape. Hence, our analysis should apply to many physical layer configurations of practical interest. Reference [8] discusses further this modeling approach.

Below, a relatively simple optimization model relevant to uplink data and media transmission in one VSG-CDMA cell is built. The first-order necessary optimizing conditions (FONOC) are presented, and two possible solutions to FONOC are discussed: one in which only the important data terminal operates at the highest available data rate, and another solution in which both data terminals operate at this rate.

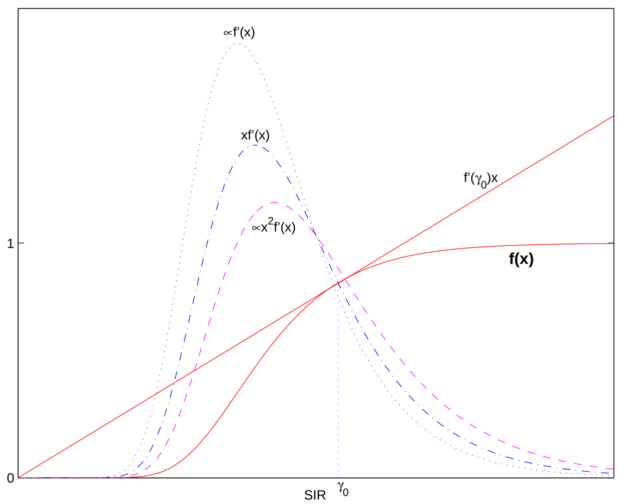

Fig. 1. A particular $f(x)$, $xf'(x)$, and *scaled* versions of $f'(x)$, and $x^2 f'(x)$. γ_0 satisfies $xf'(x) = f(x)$

II. General Formulation

A. System Model

We seek to solve:

$$\text{Maximize } T_1(G_1, \alpha_1) + \beta T_2(G_2, \alpha_2) \tag{1}$$

subject to

$$\begin{aligned} \frac{\alpha_1}{1+\alpha_1} + \frac{\alpha_2}{1+\alpha_2} &\leq 1 - \epsilon_3 &(2)\\ G_i &\geq G_0 \;,\; i \in \{1,2\} &(3)\\ G_3 &= \bar{G}_3 &(4)\\ \alpha_3 &= \frac{\bar{\gamma}_3}{\bar{G}_3} &(5)\end{aligned}$$

In this simple model,

1) The throughput of *data* terminal i is defined as $R_C T_i(G_i, \alpha_i)$, with

$$T_i(G_i, \alpha_i) := \frac{f(G_i \alpha_i)}{G_i} \tag{6}$$

2) $G_i = R_C / R_i$, is the spreading gain of terminal i; i.e., the ratio of the channel's chip rate , R_C to the terminal's data transmission rate R_i (bits per second). $G_0 \geq 1$ is the lowest available spreading gain (determined by the highest available data rate).

3) α_i is the carrier-to-interference ratio (CIR) of the signal from terminal i received at the base station. α_i is defined as,

$$\alpha_i := \frac{P_i h_i}{\sum_{\substack{j=1\\ j\neq i}}^{N} P_j h_j + \sigma^2} = \frac{Q_i}{\sum_{\substack{j=1\\ j\neq i}}^{N} Q_j + \sigma^2} \tag{7}$$

with P_i the transmission power of terminal i, h_i its path gain, $h_i P_i := Q_i$ its received power, and σ^2 a representative of the average noise power and, possibly, out-of-cell interference.

4) Each terminal has an upper bound on its transmission power, $\bar{P}_i$. For convenience, we set $h_i \bar{P}_i = \bar{Q}_i$.

5) Constraint (2) ensures that a set of achievable received powers exists that produce the given α_i's. [9], [10], [11]. This is discussed below.

6) The product $G_i \alpha_i$, denoted as γ_i, is terminal i's signal to interference (SIR) ratio. For media terminals, a specific SIR value must be provided. For data terminal, the SIR is to be determined optimally, along with the data rates, to maximize the network's weighted throughput. Notice that

$$\alpha_i/(1+\alpha_i) \equiv 1/(1+\alpha_i^{-1}) \equiv 1/(1+G_i/\gamma_i) \tag{8}$$

7) Each terminal has an upper bound on its transmission power, $\bar{P}_i$. For convenience, we set $h_i \bar{P}_i = \bar{Q}_i$. For media terminals, we define $\hat{h}_i = (1 + \bar{G}_i/\bar{\gamma}_i) h_i$ as the terminal's "effective" path gain, because the analysis shows that the terminal with the lowest $\hat{h}_i \bar{P}_i$ has the greatest difficulty in reaching the power level leading to its desired SIR. The greater limitation to network performance is imposed by the terminal in the worst situation. Because of the inflexible SIR requirement of media terminals, it is less favorable for the cell that the terminal in the worst situation be a media terminal, as opposed to a data terminal, and we assume so, below.

8) $\beta_i \geq 1$ is a weight, which admits various practical interpretations. With no loss of generality, we set $1 = \beta_1 \leq \beta_2 = \beta$.

9) There is a frame-success function (FSF), f_S, which gives the probability of the correct reception of a data packet in terms of the received SIR. We assume that *all that is known* about this function is that it is an "S-curve", as discussed in [8] (see fig. 1), and that it has a continuous second derivative. For technical reasons, we work with $f(x) := f_S(x) - f_S(0)$, ($f_S(0)$ is very small but not zero). To provide numerical examples, we use the FSF corresponding, under suitable assumptions, to non-coherent FSK modulation, with no FEC, and packet

size 80, which is,

$$f_s(x) = \left[1 - \frac{1}{2}\exp\left(-\frac{x}{2}\right)\right]^{80} \tag{9}$$

In the development below, an asterisk used as a superscript on a variable denotes a specific value of the variable which satisfies certain optimality condition. We refer to a data terminal operating at maximal data rate as "favored" or "favorite".

B. Power Limitations

For a given set of α_i, the resulting received power levels are such that

$$Q_i = \frac{\sigma^2}{1 - s_0}\frac{\alpha_i}{1+\alpha_i} \tag{10}$$

with

$$s_0 := \sum_{i=1}^{N}\frac{\alpha_i}{1+\alpha_i} \equiv \sum_{i=1}^{N}\frac{1}{1+G_i/\gamma_i} \tag{11}$$

[9], [10], [11].

It is sometimes convenient to observe that s_0 can be written as

$$s_0 := \sum_{i=1}^{N}\frac{\alpha_i}{1+\alpha_i} \equiv N - \sum_{i=1}^{N}\frac{1}{1+\alpha_i} \tag{12}$$

Whenever $s_0 < 1$, eq. (10) gives positive values. But some of these values may be too high, given the power limitations of the terminals. To prevent this from happening, we proceed as in [10], and obtain the feasibility condition given by inequality (2) as follows. We want, for each i :

$$Q_i = \frac{\sigma^2}{1 - s_0}\frac{\alpha_i}{\alpha_i + 1} \leq h_i\bar{P}_i$$

Simple algebra indicates that, in order for this inequality to hold, s_0 must satisfy:

$$s_0 \leq 1 - \frac{\sigma^2}{h_i\bar{P}_i}\frac{\alpha_i}{\alpha_i+1} \equiv 1 - \frac{\sigma^2}{\left(1+\alpha_i^{-1}\right)h_i\bar{P}_i} \text{ for all } i\text{, OR}$$

$$s_0 \leq 1 - \frac{\sigma^2}{\min_i\{(1+1/\alpha_i)h_i\bar{P}_i\}}$$

Because of the inflexible SIR requirement of media terminals, it is less favorable for the cell that the terminal in the worst situation be a media terminal, as opposed to a data terminal. Therefore, let us assume that that terminal 3 is in the "worst situation" in the sense that $\left(1+\bar{G}_3/\bar{\gamma}_3\right)h_3\bar{P}_3 < \left(1+\alpha_i^{-1}\right)h_i\bar{P}_i$ for $i \in \{1,2\}$. For example, the power limits of the data terminals may be such that $h_i\bar{P}_i \geq \left(1+\bar{G}_3/\bar{\gamma}_3\right)h_3\bar{P}_3$ for $i \in \{1,2\}$. This guarantees that regardless of the optimal choice of α_i, a data terminal will not minimize $(1+1/\alpha_i)h_i\bar{P}_i$. Thus, we obtain:

$$s_0 \leq 1 - \frac{\sigma^2}{\left(1+\bar{G}_3/\bar{\gamma}_3\right)h_3\bar{P}_3} \tag{13}$$

We obtain constraint (2) from (13), by moving $\alpha_3/(1+\alpha_3)$ to the right-hand side, and defining:

$$\epsilon_3 = \left(1 + \frac{\sigma^2}{h_3\bar{P}_3}\right)\frac{1}{1+\bar{G}_3/\bar{\gamma}_3} \tag{14}$$

Notice that each of the symbols on the right-hand side of equation (14) represents a known quantity. Thus, ϵ_3 is known.

III. Obtaining the Optimal Values

A. Optimization Model

We seek to solve:

$$\max_{G_i,\alpha_i} \frac{f(G_1\alpha_1)}{G_1} + \beta\frac{f(G_2\alpha_2)}{G_2} \tag{15}$$

subject to

$$\begin{aligned}
\frac{\alpha_1}{1+\alpha_1} + \frac{\alpha_2}{1+\alpha_2} &\leq 1 - \epsilon_3 && (16)\\
G_i &\geq G_0 \ i \in \{1,2\} && (17)\\
G_3 &= \bar{G}_3 && (18)\\
\alpha_3 &= \bar{\gamma}_3/\bar{G}_3 && (19)
\end{aligned}$$

Some reflection indicates that constraint (16) should be satisfied with equality. Otherwise, we could increase the weighted throughput by raising either α_i, while still satisfying constraint (16). However, it is not clear a priori whether either or both of constraints (17) should be satisfied with equality.

B. First-Order Necessary Optimizing Conditions (FONOC)

The Lagrangian corresponding to this problem can be written as

$$\begin{aligned}
T_1(G_1\alpha_1) + \beta T(G_2\alpha_2) \quad + \\
\lambda\left(\sum_{i=1}^{2}\frac{\alpha_i}{1+\alpha_i} - 1 + \epsilon_3\right) + \sum_{i=1}^{2}\mu_i(G_0 - G_i)
\end{aligned} \tag{20}$$

The FONOC can be expressed in vector form, with $\gamma_i = G_i\alpha_i$, as:

$$\begin{bmatrix} \partial T_1(G_1,\alpha_1)/\partial G_1 - \mu_1 \\ \beta\partial T_2(G_2,\alpha_2)/\partial G_2 - \mu_2 \\ f'(\gamma_1) + \lambda(1+\alpha_1)^{-2} \\ \beta f'(\gamma_2) + \lambda(1+\alpha_2)^{-2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{21}$$

with

$$\begin{aligned}
\frac{\alpha_1}{1+\alpha_1} + \frac{\alpha_2}{1+\alpha_2} &= 1 - \epsilon_3 && (22)\\
\mu_1(G_0 - G_1) &= 0 && (23)\\
\mu_2(G_0 - G_2) &= 0 && (24)
\end{aligned}$$

Notice that

$$\frac{\partial T_i(G_i,\alpha_i)}{\partial G_i} = \frac{\gamma_i f'(\gamma_i) - f(\gamma_i)}{G_i^2} \tag{25}$$

Also, from eq. (12), condition (22) can be equivalently stated as

$$\frac{1}{1+\alpha_1} + \frac{1}{1+\alpha_2} = 1 + \epsilon_3 \tag{26}$$

C. Solving FONOC

1) A single-favorite boundary solution: From our previous experience with similar problems [1], [2], [3], we will first explore a solution to FONOC in which the important data terminal operates at maximal data rate ($G_2 = G_0$), with the data rate of the ordinary terminal somewhere within its allowable range (i.e., $\mu_1 = 0$, which allows any $G_1 \geq G_0$ per "complementary slackness" condition (23)). (With only 2 data terminals, the phrase "single favorite" is redundant, since there can be at most one favorite. But the phrase is used because it has a similar usage in the many-terminal scenario)

Working with the top row of the matrix equation (21), we obtain $\gamma_1 f'(\gamma_1) = f(\gamma_1)$, an equation of the general form:

$$x f'(x) = f(x) \tag{27}$$

With f an S-curve, there is a unique positive value γ_0 which satisfies equation (27), which can be seen in figure (1) at the tangency point between the graph of f and a straight line from the origin. Therefore,

$$G_1^* \alpha_1^* = \gamma_0 \tag{28}$$

Combining eq. (28) with the bottom half of the matrix equation (21), we obtain

$$\left(\frac{1+\alpha_2}{1+\alpha_1}\right)^2 \frac{f'(\gamma_2)}{f'(\gamma_0)} = \frac{1}{\beta} \tag{29}$$

Equation (26) can be written as

$$\frac{1+\alpha_2}{1+\alpha_1} = (1+\epsilon_3)\alpha_2 + \epsilon_3 \tag{30}$$

Combining eqs. (29) and (30), we obtain, with x in place of γ_2,:

$$\left((1+\epsilon_3)\frac{x}{G_0} + \epsilon_3\right)^2 \frac{f'(x)}{f'(\gamma_0)} = \frac{1}{\beta} \tag{31}$$

In eq. (31), all quantities, except for x, are presumed known. Thus, this is a single-variable equation. Notice that $G_0 \geq 2$; and values of $\epsilon_3$greater than or equal to 1 are useless, because if $\epsilon_3 \geq 1$ condition (21) cannot possibly be satisfied; thus, $(1+\epsilon_3)(x/G_0) + \epsilon_3 \leq x + 1$. This fact is useful in arguing that $((1+\epsilon_3)(x/G_0) + \epsilon_3)^2 f'(x)/f'(\gamma_0)$ has the same "bell-shaped" graph of the function $x^2 f'(x)$ (fig. 1). This implies that, if G_0 is "too large", the "top" of this bell may fall below $1/\beta$, unless β is also "very large". Thus, eq. (31) may have no solution. On the other hand, when G_0 is sufficiently small and/or β and/or ϵ_3 is sufficiently large, two values of x, on either side of the peak, will satisfy eq. (31). We choose the larger value, δ_0, as the FONOC-solving SIR for the important terminal. Now, $\alpha_2 = \delta_0/G_0$; with this value, we obtain α_1 directly from eq. (26), and by plugging this α_1 value into eq. (28), we obtain G_1. Thus, a complete "single-favorite" solution to FONOC is found. However, if the resulting α_1^* is negative, or if $G_1^* < G_0$, this solution is useless, and we must consider a "dual-favorite" solution, with both data terminals operating at the highest available data rate.

2) Dual-favorite Boundary Solution: In the preceding section, we considered the SFBS, in which only the important terminal operates at the lowest available spreading gain (highest data rate). We observed that the SFBS may fail depending on the values of the parameters G_0, β. In this section we seek a "greedy" solution to FONOC, in which both terminals operate at the highest available data rate.

Working with the last two rows of equation (21) we establish, with $x = \gamma_2$ and $y = \gamma_1$, that:

$$f'(y)\left(1 + \frac{y}{G_0}\right)^2 = \beta f'(x)\left(1 + \frac{x}{G_0}\right)^2 \tag{32}$$

Eq. (26) can be re-written as

$$\frac{1}{1 + x/G_0} + \frac{1}{1 + y/G_0} = 1 + \epsilon_3 \tag{33}$$

Eqs. (32 and 33) form a system of two non-linear equations in two unknowns. This system can be solved. Its solution is characterized through fig. 2.

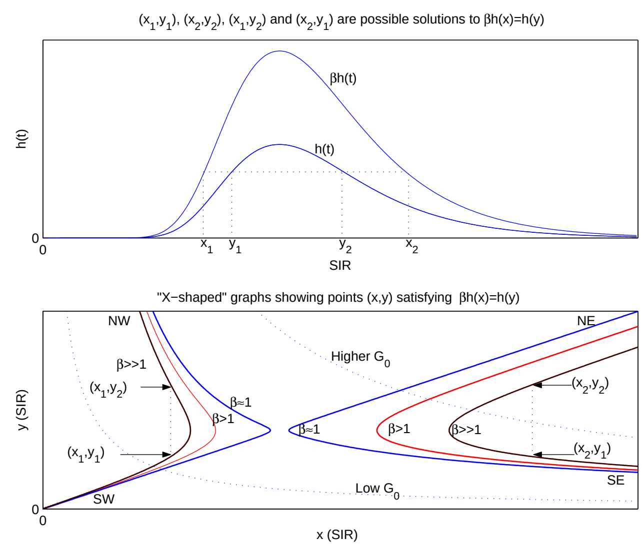

Fig. 2. With x and y respectively the SIR of the important and ordinary terminal, FONOC requires that $\beta h(x) = h(y)$, with $h(t) = f'(t)\,(1 + t/G_0)^2$. Any of the pairs (x_1, y_1), (x_2, y_2), (x_1, y_2), or (x_2, y_1) (top) satisfies this equation, but may not be feasible. We plot all such points, which reveals an "X-shaped" graph (NE, NW, SW and SE are directional labels). On the same axes, we plot the hyperbolic curves (dotted) which represent the constraint equation (33). The intersection of the hyperbola with the NE leg of the X yields the maximizer. The term ϵ_3 (representing the resources used up by the media terminal) has the effect of "pulling down" the hyperbola, whereas G_0 tends to raise it. When G_0 is sufficiently low or ϵ_3 is sufficiently high, the hyperbola may *only* intersect the SW leg of the X-curve, which leads to a minimum.

IV. Discussion

The optimal power levels and data rates for data terminals that share one base station with media terminals, which have fixed bit rates and inflexible SIR requirements, have been investigated. This scenario is relevant to 3G CDMA. The objective is to maximize the *weighted* sum of the *data* terminal's throughput, while honoring QoS commitments made to the media terminal. Two weights, which admit various interpretations, including levels of importance, "utilities", or monetary prices, are considered. The properties of the physical layer are embodied in the frame success function (FSF), which gives, in terms of received signal-to-interference ratio (SIR), the probability that a data packet is correctly received. No specific functional form ("equation") is imposed on the FSF. It is assumed that *all that is known* about the FSF is that its graph is "S-shaped", and the analysis follows from properties derived from this shape (some additional technical assumptions are needed by certain results). Therefore, many physical layer configurations of practical interest are accommodated. Each physical layer has a preferred SIR, γ_0, easily identified in the graph of the FSF.

Our primary aim in the present paper was to start extending the analysis in [1], [2], [3], in which neither the presence of media-transmitting terminals nor out-of-cell interference are considered, to the richer and more interesting scenario discussed herein. Our main conclusion is that much of our previous analysis can be applied to the present scenario. The effects of the media terminal, the out-of-cell interference (noise), and the power limitations of the terminals, combine into a single term, ϵ_3, that reduces the right-hand-side of the constraint on the carrier-to-interference ratios. This term represents the "resources" consumed by the media terminal.

This analysis focuses on two allocations satisfying the first-order necessary optimality conditions (FONOC): (i) an "unbalanced" assignment in which the important terminal operates at the highest available data rate, with the other terminal achieving the SIR, γ_0; and (ii) a "greedy" assignment in which both terminals operate at the highest available data rate. In our previous work, we have analyzed the second-order conditions for these allocations, for the case in which there are no media terminals, or noise. We proved that the unbalanced allocation is a maximizer whenever it exists, while the greedy allocation can lead to either a maximum or a minimum depending upon the system parameters. We expect those conclusions to continue to hold in the present scenario.

The important terminal should always operate at maximal data rate. From our previous work we know that, without the media terminal, only when G_0 is "large" relative to β should both data terminals operate at maximal data rate (G_0 is the smallest available spreading gain and β is the weight of the important terminal). With the media terminal consuming resources, the G_0 value at which both terminals should operate at maximal rate for a given β *increases*. All this makes intuitive sense, because when G_0 is "large", the highest available data rate is relatively small, and keeping only one terminal operating at maximal data rate is not appealing, unless that terminal has "a lot of weight". However, when a media terminal is taking up resources, and/or when the highest available data rate is very high (G_0 is low), an allocation in which *only one* terminal operates at the highest rate is more appealing. The "greedy" allocation is particularly treacherous, because it can lead to either a maximum or a minimum, depending upon how large G_0 is, and the amount of resources consumed by the the media terminal (ϵ_3).

It is significant that the greedy and the unbalanced allocations are complementary in this sense: the factors that tend to turn the greedy allocation into a minimizer (a low G_0 and/or a high ϵ_3), tend to make feasible the unbalanced allocation, which is a maximizer.

References

[1] V. Rodriguez and D. Goodman, "Power and data rate assignment for maximal weighted throughput in 3G CDMA," *IEEE WCNC*, vol. 1, pp. 525–31, 2003.

[2] V. Rodriguez and D. Goodman, "Power and data rate assignment for maximal weighted throughput: A dual-class 3G CDMA scenario," *IEEE ICC*, vol. 1, pp. 397–401, 2003.

[3] V. Rodriguez and D. Goodman, "Power and data rate assignment for maximal weighted throughput in 3G CDMA: A global solution with two classes of users," *IEEE WCNC*, vol. Forthcoming, 2004.

[4] D. Goodman, Z. Marantz, P. Orenstein, and V. Rodriguez, "Maximizing the throughput of cdma data communications," *IEEE VTC-Fall*, 2003.

[5] S. Ulukus and L. Greenstein, "Throughput maximization in CDMA uplinks using adaptive spreading and power control," *IEEE ISSSTA*, vol. 2, pp. 565 –9, 2000.

[6] J. Lee, R. Mazumdar, and N. Shroff, "Joint power and data rate allocation for the downlink in multi-class CDMA wireless networks," *Proc. of the 40th Allerton Conference on Communications, Control and Computing*, 2002.

[7] C. W. Sung and W. S. Wong, "Power control and rate management for wireless multimedia CDMA systems," *IEEE Trans. Commun.*, vol. 49, pp. 1215–26, Jul. 2002.

[8] V. Rodriguez, "Robust modeling and analysis for wireless data resource management," *IEEE WCNC*, vol. 2, pp. 717–722, 2003.

[9] V. Rodriguez, "From power levels to power ratios and back: A change of domain and its modeling implications," WICAT 02-011, Polytechnic University, Brooklyn, New York, 2002. Available at http://wicat.poly.edu/reports.

[10] A. Sampath, P. Kumar, and J. Holtzman, "Power control and resource management for a multimedia CDMA wireless system," *IEEE PIMRC*, vol. 1, pp. 21–5, 1995.

[11] S. Ataman and A. Wautier, "Perfect power control in a multiservice CDMA cellular system," *IEEE PIMRC*, vol. 3, pp. 1222 –6, 2002.