Wireless Netw (2009) 15:709–726
DOI 10.1007/s11276-007-0070-5

A novel approach to smart multi-cell radio resource management based on load gradient calculations

Jordi Pérez-Romero · Oriol Sallent · Ramon Agustí

Published online: 20 September 2007

Abstract This paper presents a novel methodology for capturing the coupling between the different cells in both the uplink and downlink directions in a Wideband Code Division Multiple Access (WCDMA) scenario. It is based on the definition and computation of the gradient of the uplink cell load factor and the downlink transmitted power, which are the two main parameters that reflect the actual cell load in the two link directions. The paper shows that the gradient is able to capture the relevant information about the spatial distribution of traffic, which has an impact on cell performance. The proposed methodology is also used as the basis for defining and evaluating new Radio Resource Management (RRM) strategies that operate at a multi-cell level.

Keywords WCDMA · Radio resource management · Cellular systems · Spatial traffic distribution

1 Introduction

The mobile communications industry is currently shifting its focus from second-generation (2G) to third-generation (3G) technology. While current 2G wireless networks continue to evolve by bringing new facilities and services onto the market aided by packet switching capabilities, Wideband Code Division Multiple Access (WCDMA) radio technology has been made a reality through the launching of 3G networks [1, 2]. To date, a great deal of effort has been put into understanding the peculiarities of a WCDMA in cellular network deployment and optimisation. However, extensive research is still needed to devise a comprehensive framework for managing 3G and beyond WCDMA-based radio access technologies operating in different environments and supporting a range of different services.

System deployment must be preceded by careful network planning, in which the expected traffic load must be taken into account. In the case of 3G, this traffic load needs also to be defined in terms of traffic class distribution (e.g., expected conversational traffic, interactive traffic, etc.) due to the various Quality of Service (QoS) requirements. If QoS guarantees are to be provided, a reduction in the randomness associated with wireless cellular scenarios is required in order to make the communication channel as predictive as possible. Amongst other factors, this randomness is due to propagation, mobility and traffic generation processes. The only way to achieve a tight control over the air interface is by means of Radio Resource Management (RRM) strategies, which include admission and congestion control, packet scheduling, power control and handover [3, 4]. The whole set of RRM functions is devoted to guaranteeing a certain level of QoS that maintains the planned coverage area and offers high capacity, while using the radio resources efficiently. Interactions between the various RRM strategies play a key role in WCDMA network performance, which relies heavily on the amount of interference on the air interface. In this framework, load control is a key element behind the different RRM functions in a WCDMA scenario, which can be summarised as follows:

J. Pérez-Romero (✉) · O. Sallent · R. Agustí
Department of Signal Theory and Communications, Universitat Politècnica de Catalunya (UPC), c/ Jordi Girona, 1-3, Campus Nord, Barcelona, Spain
e-mail: jorperez@tsc.upc.edu

O. Sallent
e-mail: sallent@tsc.upc.edu

R. Agustí
e-mail: ramon@tsc.upc.edu

1. Admission control is intended to ensure the average cell load at acceptable levels.
2. Congestion control is intended to react in overload situations resulting from inefficient control over the air interface.
3. Packet scheduling is intended to order users' transmissions suitably, thus preserving certain load levels.
4. Power control operates on a per-link basis in order to limit the interference of each specific connection or, equivalently, to limit the load contribution from each link, while keeping the promised quality.
5. Handover is intended to ensure the best possible connectivity for each link while optimising the interference (or load) added to the air interface.

The first contribution of this paper is to introduce an innovative mathematical framework for capturing the air interface coupling between the different cells in a WCDMA scenario. This framework is presented in a compact formulation for both the uplink and downlink directions. To the authors' best knowledge, this is the first time this approach has been taken. The multiple issues that affect the behaviour of radio interfaces and the greater degree of coupling between them due to the nature of the WCDMA, in which users transmit at the same time using the same carrier, will be demonstrated much more clearly in comparison with classical formulations.

One of the main characteristics of traffic in cellular networks is the non-homogeneous spatial traffic distribution. As 3G multimedia-intensive traffic demand profiles are expected to be different from those in 2G, spatial traffic distributions are even more important than in 2G networks. Although to some extent network planning is able to take this into consideration, the high dynamics associated with traffic clearly requires additional mechanisms that are able to cope with the potential problems on the network performance for real traffic profile distributions, which are significantly different from those expected in the network planning phase. Traffic variations in time and space, which result in different load levels in different cells and at different times, can, for example, be smoothed with the aid of the handover algorithm, which attempts to balance the load between cells. Multi-cell admission control and/or congestion control mechanisms are other examples.

In this framework, mechanisms supporting smart load control actions could be highly useful and could be applied at different levels. A smart load control for a reference cell would be based on the ability to detect those neighbouring cells that most affect reference cell performance. In this context, the second main contribution of this paper is the exploitation of the analytical model developed. This is done by obtaining the gradient of the reference cell uplink load factor (or alternatively the transmitted power in the downlink) with respect to any neighbouring cell. This gradient makes it possible to identify the most critical cells and users that influence a reference cell, as a result of which smart RRM algorithms may be defined. The third contribution of the paper is the use of the proposed load gradient methodology in the definition of specific congestion and admission control algorithms and a multi-carrier allocation method, which expands on the preliminary work of the authors presented in [5]. However, it should be pointed out that the aim of this paper is not to provide an extensive analysis or an optimisation of the RRM algorithms presented. Instead, by means of various examples we aim to illustrate the potential range of applicability of the proposed gradient-based framework.

The rest of this paper is organised as follows. Section 2 presents a summary of related work in order to establish the framework of the paper. In Sect. 3, a novel analytical model is developed to obtain the uplink cell load factor and downlink transmitted power level in terms of the neighbouring cells' corresponding load factors and transmitted power levels. Section 4 introduces the gradients in the uplink and downlink directions as the basis for a range of different RRM algorithms. This section also discusses some of the considerations regarding the practical implementation of the proposed methodology. Section 5 describes the different scenarios in which the proposed framework was evaluated by means of simulations. Section 6 presents some of the results of the gradient evaluation in order to show how it is able to capture the different spatial traffic distributions. Section 7 presents some examples of the applications of the proposed approach. Most notably, specific congestion control, admission control and multi-carrier assignment algorithms that exploit the developed framework are presented. Finally, Sect. 8 summarises the main conclusions reached in the present study.

2 Related work and contributions of this paper

Since the early 1990s, several works [6, 7] have dealt with the estimation of the capacity of cellular CDMA networks, particularly with regard to the development of WCDMA technologies for 3G systems such as the UMTS. One key consideration consists in limiting the interference generated by each connection on the air interface, thus keeping the total load below the specific limits with regard to the maximum power available. Strategies for estimating the load in WCDMA systems are presented in [8, 9], in which the concept of the load factor is introduced as an indication of the total interference measured in one cell. Taking this

into account, RRM strategies for WCDMA networks are based on dynamically controlling the load in the different cells in order to ensure that the services can be provided satisfactorily while the desired coverage is guaranteed and the capacity is maximised.

From the point of view of the different RRM strategies in the literature, in [10] an admission control algorithm based on the required SIR for each connection is proposed, while in [11, 12] admission control strategies based on measuring the total amount of received power at the node B are presented. In [13–15], the load factor is considered as an indicator for admission control in the uplink direction. Among the various strategies presented, some give an overview of use in WCDMA systems, while others are more aligned to specific systems, such as the UMTS [16, 17]. In [18], mathematical models for various call admission schemes are built and an effective linear programming technique for finding a better admission control scheme is presented. In [19], a classification of call admission control strategies depending on whether they are based on modelling or on measurements is presented. Insofar as the downlink direction is concerned, admission control strategies are normally based on transmission power measurements [20, 21]. In [22], a joint uplink and downlink strategy based on received and transmitted powers, respectively, is presented. However, in most of the cases mentioned admission control is executed at the cell level, whereby the load/power increase of the new user in its serving cell is considered. Furthermore, these cases either ignore the impact of the mobile on the rest of the cells or use a simple model for estimating the intercell interaction [23].

In the case of congestion control strategies, which are designed to deal with situations in which the system has reached an overload status, the QoS guarantees are at risk due to the evolution of system dynamics (mobility aspects, increase in interference, etc.). As a result of this, they try to reduce the load in the scenario by overriding current connections [24]. One way of doing this is by reducing the maximum bit rate of the currently admitted users depending on their service requirements. In the literature, there are few works addressing the congestion control problem for WCDMA networks [5, 25–28]. Moreover, to the authors' best knowledge all the approaches are based on actions that only affect the cell in which congestion is triggered, but do not exploit the possible interactions between cells should traffic loads and service mixes differ.

Taking the above into account, this paper differs from previous works in that an accurate model to capture the interactions and couplings between cells on a WCDMA cellular network is presented, which is based on a new concept of load gradients in both the uplink and downlink directions. This modelling has applications in the development of RRM strategies at a multi-cell level, as will be illustrated through various examples in Sect. 7. Specifically, the contributions of this paper can be summarised as follows:

1. It provides closed-form expressions for the relationships between the load factors and transmitted powers of the different cells in a WCDMA network, which to the authors' knowledge is not available in the open literature.
2. It provides a new framework, based on the gradient concept, for computing the interactions between cells, which has proved to be valid in detecting different types of non-homogeneous user distributions.
3. It illustrates through various examples the applicability of the new framework in the development of RRM algorithms.

3 Multiple cell coupling in a WCDMA scenario

3.1 The uplink case

In a WCDMA cell, all the users share a common bandwidth and each new connection increases the interference level of other connections, which affects their quality. Let us consider a scenario with $(K + 1)$ cells, $j = 0,1,\ldots, K$, and let us focus on the central cell, which has been chosen as the reference cell and numbered as the 0-th cell. The average bit energy to noise power spectral density (Eb/No) for each user transmitting in this cell is given by:

$$\left(\frac{E_b}{N_o}\right)_{i_0} = \frac{P_{0,i_0} \times \frac{W}{R_{i_0}}}{P_N + \chi_0 + \left[P_{R,0} - P_{0,i_0}\right]} \quad i_0 = 1\ldots n_0 \qquad (1)$$

where P_{0,i_0} is the uplink received power from the i-th user in the 0-th cell, R_{i_0} is the transmission bit rate of the i-th user in the 0-th cell, W is chip rate, P_N is the thermal noise power, χ_0 is the intercell interference observed in the 0-th cell and $P_{R,0}$ is the total received own-cell power at this cell. In addition, n_0 is the total number of users transmitting in the 0-th cell. In practice, the closed loop power control in WCDMA will try to maintain the value of $(E_b/N_o)_{i_0}$ equal to a target value for most of the time. This ensures that specific QoS requirements are met in terms of bit error rate or block error rate. However, the formulation presented is general in order to account for the fact that the actual $(E_b/N_o)_{i_0}$ may fall below the target. It should also be pointed out that the formulation presented is general in order to include multiple services (i.e., users with different requirements in terms of bit rate and Eb/No).

The total received power $P_{R,0}$ can also be expressed as:

$$P_{R,0} = \sum_{i=1}^{n_0} P_{0,i_0} \tag{2}$$

In the uplink direction, the relative amount of load in a WCDMA cell is usually given by the cell load factor [8], which measures the spectral efficiency of a WCDMA cell by linking up the total received power from transmissions to the total received power including thermal noise. Specifically, for the 0-th cell the uplink cell load factor η_0 is given by:

$$\begin{aligned}\eta_0 &= \frac{P_{R,0}+\chi_0}{P_{R,0}+\chi_0+P_N} = 1-\frac{P_N}{P_{TOT,0}} \\ &= \left(1+\frac{\chi_0}{P_{R,0}}\right)\sum_{i=1}^{n_0}\frac{1}{\frac{W}{\left(\frac{E_b}{No}\right)_{i_0}R_{i_0}}+1}\end{aligned} \tag{3}$$

The first part of this equality reflects the fact that the load factor can be obtained from the noise power and the total received power $P_{TOT,0} = P_{R,0} + \chi_0 + P_N$. Likewise, the latter part of this equality corresponds to the expression of the load factor according to the number of transmissions and can be easily obtained from (1) and (2).

By introducing the uplink cell load factor definition in (1), the following is derived:

$$P_{0,i_0} = \frac{\left(P_N+\chi_0+P_{R,0}\right)}{\left[\frac{W}{\left(\frac{E_b}{No}\right)_{i_0}R_{i_0}}+1\right]} = \frac{P_N\frac{1}{1-\eta_0}}{\left[\frac{W}{\left(\frac{E_b}{No}\right)_{i_0}R_{i_0}}+1\right]} \tag{4}$$

From the transmitter side, the power transmitted by the i-th user in 0-th cell may be expressed as:

$$P_{T,i_0} = L_{i_0,0} \times P_{0,i_0} \tag{5}$$

where $L_{i_0,0}$ is the path loss from the i-th user of the 0-th cell to the 0-th cell site including shadowing and antenna gains. At this point, (2) can be expressed as:

$$P_{R,0} = \sum_{i_0=1}^{n_0}\frac{P_N\frac{1}{1-\eta_0}}{\frac{W}{\left(\frac{E_b}{No}\right)_{i_0}R_{i_0}}+1} \tag{6}$$

The intercell interference observed in the reference cell can be expressed as:

$$\chi_0 = \sum_{j=1}^{K}\sum_{i_j=1}^{n_j}\frac{P_{T,i_j}}{L_{i_j,0}} \tag{7}$$

where P_{T,i_j} is the power transmitted by the i-th user in the j-th cell, n_j is the total number of users transmitting in the j-th cell and $L_{i_j,0}$ is the path loss from the i-th user in the j-th cell to the reference 0-th cell. By considering (4) and (5) as well as extending the cell load factor definition in (3) to the rest of the neighbouring cells to obtain η_j, (7) can be rearranged as:

$$\chi_0 = \sum_{j=1}^{K}\sum_{i_j=1}^{n_j}\frac{L_{i_j,j}\frac{P_N\frac{1}{1-\eta_j}}{\frac{W}{\left(\frac{E_b}{No}\right)_{i_j}R_{i_j}}+1}}{L_{i_j,0}} = \sum_{j=1}^{K}P_N\frac{1}{1-\eta_j}\sum_{i_j=1}^{n_j}\frac{L_{i_j,j}}{L_{i_j,0}}\frac{1}{\frac{W}{\left(\frac{E_b}{No}\right)_{i_j}R_{i_j}}+1} \tag{8}$$

At this stage, it is useful to introduce the following definitions:

$$S_{0,0}^{UL} = \sum_{i_0=1}^{n_0}\frac{1}{\frac{W}{\left(\frac{E_b}{No}\right)_{i_0}R_{i_0}}+1} \tag{9}$$

$$S_{j,0}^{UL} = \sum_{i_j=1}^{n_j}\frac{L_{i_j,j}}{L_{i_j,0}}\frac{1}{\frac{W}{\left(\frac{E_b}{No}\right)_{i_j}R_{i_j}}+1} \tag{10}$$

Notice that (9) can be interpreted as the cell load factor in the reference cell should the 0-th cell be isolated. Additionally, the term $S_{j,0}^{UL}$ in (10) can be interpreted as a weighted cell load factor contribution from the j-th cell to the reference cell, which depends on how the users in the j-th cell are distributed with respect to the 0-th cell. Thus, (3) can be finally expressed as:

$$\eta_0 = \frac{S_{0,0}^{UL}+\frac{\chi_0}{P_N}}{1+\frac{\chi_0}{P_N}} = \frac{S_{0,0}^{UL}+\sum_{j=1}^{K}\frac{S_{j,0}^{UL}}{1-\eta_j}}{1+\sum_{j=1}^{K}\frac{S_{j,0}^{UL}}{1-\eta_j}} \tag{11}$$

In the above form, the coupling in a WCDMA cellular system is explicitly reflected. In this case, the resulting cell load in a reference cell depends on all the users in the scenario or, more specifically, on all the other cells' respective load factors.

3.2 The downlink case

In the downlink direction, the Eb/No of the i-th user in the 0-th cell can be expressed as:

$$\left(\frac{E_b}{N_o}\right)_{i_0} = \frac{\frac{P_{T,i_0}}{L_{i_0,0}}\times\frac{W}{R_{i_0}}}{P_N+\chi_{i_0}+\rho\times\left[\frac{P_{T,0}-P_{T,i_0}}{L_{i_0,0}}\right]} \qquad i_0 = 1\ldots n_0 \tag{12}$$

$$P_{T,0} = P_{p,0} + \sum_{i_0=1}^{n_0}P_{T,i_0} \tag{13}$$

where P_{T,i_0} is the power transmitted to the i-th user, χ_{i_0} represents the intercell interference observed by the i-th user, $L_{i_0,0}$ is the path loss from the i-th user of the 0-th cell to the 0-th cell site, R_{i_0} is the i-th user transmission rate, ρ is the orthogonality factor (since orthogonal codes are used in the downlink to identify the different transmissions in a given cell, but a certain degree of orthogonality is lost due to multipath), P_N is the downlink background noise and W the bandwidth. $P_{T,0}$ is the base station transmitted power and $P_{p,0}$ the power devoted to common control channels. It is worth mentioning that as in the case of the uplink direction, the downlink closed loop power control will try to maintain the measured $(E_b/N_o)_{i_0}$ equal to a given target value.

Notice that there are certain differences in the downlink direction in comparison to the uplink direction. Specifically, the intercell interference is user-specific since it depends on user location and the power transmitted by the base station is shared by all users.

By considering (12) and (13) it can be obtained that the total transmitted power required to satisfy all users' demands should be:

$$P_{T,0} = \frac{P_{p,0} + \sum_{i_0=1}^{n_0} \frac{(P_N+\chi_{i_0})}{\frac{W/R_{i_0}}{\left(\frac{E_b}{N_o}\right)_{i_0}}+\rho} L_{i_0,0}}{1 - \sum_{i_0=1}^{n_0} \frac{\rho}{\frac{W/R_{i_0}}{\left(\frac{E_b}{N_o}\right)_{i_0}}+\rho}} \tag{14}$$

where the intercell interference seen by the i-th user of the 0-th cell is given by:

$$\chi_{i_0} = \sum_{j=1}^{K} \frac{P_{T,j}}{L_{i_0,j}} \tag{15}$$

Substituting (15) in (14), the following will be given:

$$P_{T,0} = \frac{P_{p,0} + P_N \sum_{i_0=1}^{n_0} \frac{L_{i_0,0}}{\frac{W/R_{i_0}}{\left(\frac{E_b}{N_o}\right)_{i_0}}+\rho} + \sum_{i_0=1}^{n_0} \sum_{j=1}^{K} \frac{L_{i_0,0}}{L_{i_0,j}} \frac{P_{Tj}}{\frac{W/R_{i_0}}{\left(\frac{E_b}{N_o}\right)_{i_0}}+\rho}}{1 - \sum_{i_0=1}^{n_0} \frac{\rho}{\frac{W/R_{i_0}}{\left(\frac{E_b}{N_o}\right)_{i_0}}+\rho}} \tag{16}$$

At this stage, it is useful to introduce the following definitions:

$$S_{0,0}^{DL} = \sum_{i_0=1}^{n_0} \frac{1}{\frac{W/R_{i_0}}{\left(\frac{E_b}{N_o}\right)_{i_0}} + \rho} \tag{17}$$

$$S_{0,0}^{DL*} = \sum_{i_0=1}^{n_0} \frac{L_{i_0,0}}{\frac{W/R_{i_0}}{\left(\frac{E_b}{N_o}\right)_{i_0}} + \rho} \tag{18}$$

$$S_{0,j}^{DL} = \sum_{i_0=1}^{n_0} \frac{L_{i_0,0}}{L_{i_0,j}} \frac{1}{\frac{W/R_{i_0}}{\left(\frac{E_b}{N_o}\right)_{i_0}} + \rho} \tag{19}$$

Notice that (17) and (18) depend on the own-cell transmissions without and with propagation conditions, respectively, similarly to the contributions of the uplink cell load factor in (9). In turn, (19) denotes the influence of the j-th cell over the reference 0-th cell, which is similar to Expression (10) for the uplink. Thus, (16) can be finally expressed as:

$$P_{T,0} = \frac{P_{p,0} + P_N S_{0,0}^{DL*} + \sum_{j=1}^{K} P_{T,j} S_{0,j}^{DL}}{1 - \rho S_{0,0}^{DL}} \tag{20}$$

In the above form, the coupling existing in a WCDMA cellular system is explicitly reflected. In this case, the resulting transmitted power of a given cell depends on all the other cells' respective transmitted power levels.

Notice that the coupling between the 0-th cell and the j-th cell in the downlink is given by the term $S_{0,j}^{DL}$ in (19), which depends on how the users are distributed in the 0-th cell in relation to the j-th cell and in relation to the 0-th cell. However, the coupling in the uplink direction given by the term $S_{j,0}^{UL}$ in (10) depends on how the users in the j-th cell are distributed with respect to the 0-th cell and with respect to their own j-th cell.

4 Load gradient evaluation

Based on the previous formulation this section presents the computation of the load gradient for both the uplink and downlink directions. This computation is given as a measurement of the variation of the load in the reference cell, which is numbered as the 0-th cell, with respect to the load variation in the neighbouring cells, which are numbered $j = 1,\ldots,K$. Finally, certain considerations regarding the practical implementation of the algorithm are provided.

4.1 The uplink case

The gradient of the reference cell load factor in the uplink is defined here as:

$$\vec{\nabla}\eta_0 = \left(\frac{\partial\eta_0}{\partial\eta_1}, \frac{\partial\eta_0}{\partial\eta_2}, \ldots, \frac{\partial\eta_0}{\partial\eta_K}\right) \tag{21}$$

where the derivative with respect to the k-th cell is obtained from (11) as:

$$\frac{\partial \eta_0}{\partial \eta_k} = \frac{1 - S_{0,0}^{UL}}{\left(1 + \sum_{j=1}^{K} \frac{S_{j,0}^{UL}}{1-\eta_j}\right)^2} \frac{\partial}{\partial \eta_k} \left(\sum_{j=1}^{K} \frac{S_{j,0}^{UL}}{1-\eta_j} \right) = \frac{1 - S_{0,0}^{UL}}{\left(1 + \sum_{j=0 \atop j\neq 0}^{K} \frac{S_{j,0}^{UL}}{1-\eta_j}\right)^2} \sum_{\substack{j=0 \\ j\neq 0}}^{K} \frac{S_{j,0}^{UL}}{\left(1-\eta_j\right)^2} \frac{\partial \eta_j}{\partial \eta_k} \quad (22)$$

Since $\partial\eta_0/\partial\eta_k$ in turn depends on the derivatives of the load factor in the rest of the cells, it is useful to generalise (22) for any two cells m, k in which $m \neq k$, thus obtaining:

$$\frac{\partial \eta_m}{\partial \eta_k} = C_m \sum_{\substack{j=0 \\ j\neq m}}^{K} \frac{S_{j,m}^{UL}}{\left(1-\eta_j\right)^2} \frac{\partial \eta_j}{\partial \eta_k} \quad k = 0, 1, \ldots, K, \quad m = 0, 1, \ldots, K, \ m \neq k \quad (23)$$

where the following has been defined:

$$C_m = \frac{1 - S_{m,m}^{UL}}{\left(1 + \sum_{j=0 \atop j\neq m}^{K} \frac{S_{j,m}^{UL}}{1-\eta_j}\right)^2} \quad (24)$$

However, since $\partial\eta_j/\partial\eta_k = 1$ for $j = k$, the following is yielded based on (23):

$$\sum_{\substack{j=0 \\ j\neq m,k}}^{K} \frac{C_m S_{j,m}^{UL}}{\left(1-\eta_j\right)^2} \frac{\partial \eta_j}{\partial \eta_k} - \frac{\partial \eta_m}{\partial \eta_k} = -\frac{C_m S_{k,m}^{UL}}{\left(1-\eta_k\right)^2} \quad m = 0, 1, \ldots, K, \ m \neq k \quad (25)$$

Expression (25) shows that a set of K linear equations are defined for each k cell. In these equations, the K unknown quantities are the derivatives $\partial\eta_j/\partial\eta_k$ for $j = 0,1,\ldots,K$, in which $j \neq k$. This linear system can be rewritten using matrix notation as:

$$\mathbf{A}_k^{UL} \cdot \mathbf{d}_k^{UL} = \mathbf{b}_k^{UL} \quad (26)$$

where matrix $\mathbf{A}_k^{UL}$ is given by:

$$\mathbf{A}_k^{UL} = \begin{bmatrix} -1 & \frac{C_0 S_{1,0}^{UL}}{(1-\eta_1)^2} & \cdots & \frac{C_0 S_{k-1,0}^{UL}}{(1-\eta_{k-1})^2} & \frac{C_0 S_{k+1,0}^{UL}}{(1-\eta_{k+1})^2} & \cdots & \frac{C_0 S_{K,0}^{UL}}{(1-\eta_K)^2} \\ \frac{C_1 S_{0,1}^{UL}}{(1-\eta_0)^2} & -1 & & & & & \\ \vdots & & -1 & & & & \\ \frac{C_{k-1} S_{0,k-1}^{UL}}{(1-\eta_0)^2} & & & & & & \vdots \\ \frac{C_{k+1} S_{0,k+1}^{UL}}{(1-\eta_0)^2} & & & & \ddots & & \\ \vdots & & & & & & \\ \frac{C_K S_{0,K}^{UL}}{(1-\eta_0)^2} & & & \cdots & & & -1 \end{bmatrix} \quad (27)$$

Vectors $\mathbf{b}_k^{UL}$ and $\mathbf{d}_k^{UL}$ are defined as:

$$\mathbf{b}_k^{UL} = \left[\frac{-C_0 S_{k,0}^{UL}}{(1-\eta_k)^2}, \frac{-C_1 S_{k,1}^{UL}}{(1-\eta_k)^2}, \ldots, \frac{-C_{k-1} S_{k,k-1}^{UL}}{(1-\eta_k)^2}, \frac{-C_{k+1} S_{k,k+1}^{UL}}{(1-\eta_k)^2}, \ldots \frac{-C_K S_{k,K}^{UL}}{(1-\eta_k)^2} \right]^T \quad (28)$$

$$\mathbf{d}_k^{UL} = \left[\frac{\partial \eta_0}{\partial \eta_k}, \frac{\partial \eta_1}{\partial \eta_k}, \ldots, \frac{\partial \eta_{k-1}}{\partial \eta_k}, \frac{\partial \eta_{k+1}}{\partial \eta_k}, \ldots, \frac{\partial \eta_K}{\partial \eta_k} \right]^T \quad (29)$$

The solution of the linear equation system gives vector $\mathbf{d}_k^{UL}$ as:

$$\mathbf{d}_k^{UL} = (\mathbf{A}_k^{UL})^{-1} \cdot \mathbf{b}_k^{UL} \quad (30)$$

Notice that by considering all the values of k = 0,1,…,K, a total of $(K + 1)$ linear equation systems such as the one defined in (26) can be formulated, thus obtaining all the derivative vectors for all the cells $\mathbf{d}_k^{UL}$, $k = 0, 1, \ldots, K$, which finally provide the gradient in (21).

4.2 The downlink case

In the downlink direction, the coupling between the different cells and the reference 0-th cell is captured by means of the derivatives of the power transmitted by the reference cell with respect to the power transmitted by each of the other cells. Thus, the gradient of the downlink transmitted power is defined as:

$$\vec{\nabla} P_{T,0} = \left(\frac{\partial P_{T,0}}{\partial P_{T,1}}, \frac{\partial P_{T,0}}{\partial P_{T,2}}, \ldots, \frac{\partial P_{T,0}}{\partial P_{T,K}} \right) \quad (31)$$

Specifically, the derivative with respect to the k-th cell is obtained from (20):

$$\frac{\partial P_{T,0}}{\partial P_{T,k}} = \frac{1}{1 - \rho S_{0,0}^{DL}} \frac{\partial}{\partial P_{T,k}} \left(\sum_{j=1}^{K} S_{0,j}^{DL} P_{T,j} \right) = \frac{1}{1 - \rho S_{0,0}^{DL}} \sum_{\substack{j=0 \\ j\neq 0}}^{K} S_{0,j}^{DL} \frac{\partial P_{T,j}}{\partial P_{T,k}} \quad (32)$$

This expression can be generalised for any two cells m, k, in which $m \neq k$, thus obtaining:

$$\frac{\partial P_{T,m}}{\partial P_{T,k}} = X_m \sum_{\substack{j=0 \\ j\neq m}}^{K} S_{m,j}^{DL} \frac{\partial P_{T,j}}{\partial P_{T,k}} \quad k = 0, 1, \ldots, K, \ m = 0, 1, \ldots, K, \ m \neq k \quad (33)$$

where the following has been defined:

$$X_m = \frac{1}{1 - \rho S_{m,m}^{DL}} \tag{34}$$

Given that $\partial P_{T,j}/\partial P_{T,k} = 1$ for $j = k$, the following is yielded:

$$\sum_{\substack{j=0 \\ j \neq m,k}}^{K} X_m S_{m,j}^{DL} \frac{\partial P_{T,j}}{\partial P_{T,k}} - \frac{\partial P_{T,m}}{\partial P_{T,k}} = -X_m S_{m,k}^{DL}$$
$$m = 0, 1, \ldots, K, \; m \neq k \tag{35}$$

As in the uplink case, (35) defines a set of K linear equations where the K unknown quantities are the derivatives $\partial P_{T,j}/\partial P_{T,k}$ $j = 0, 1, \ldots, K$, $j \neq k$. Again, using matrix notation:

$$\mathbf{A}_k^{DL} \cdot \mathbf{d}_k^{DL} = \mathbf{b}_k^{DL} \tag{36}$$

where matrix $\mathbf{A}_k^{DL}$ is given by:

$$\mathbf{A}_k^{DL} = \begin{bmatrix} -1 & X_0 S_{0,1}^{DL} & \cdots & X_0 S_{0,k-1}^{DL} & X_0 S_{0,k+1}^{DL} & \cdots & X_0 S_{0,K}^{DL} \\ X_1 S_{1,0}^{DL} & -1 & & & & & \\ \vdots & & -1 & & & & \\ X_{k-1} S_{k-1,0}^{DL} & & & & & & \vdots \\ X_{k+1} S_{k+1,0}^{DL} & & & & \ddots & & \\ \vdots & & & & & & \\ X_K S_{K,0}^{DL} & & \cdots & & & & -1 \end{bmatrix} \tag{37}$$

In turn, vectors $\mathbf{b}_k^{DL}$ and $\mathbf{d}_k^{DL}$ are defined as:

$$\mathbf{b}_k^{DL} = \Big[-X_0 S_{0,k}^{DL}, -X_1 S_{1,k}^{DL}, \ldots, -X_{k-1} S_{k-1,k}^{DL}, -X_{k+1} S_{k+1,k}^{DL}, \ldots, -X_K S_{K,k}^{DL}\Big]^T \tag{38}$$

$$\mathbf{d}_k^{DL} = \left[\frac{\partial P_{T,0}}{\partial P_{T,k}}, \frac{\partial P_{T,1}}{\partial P_{T,k}}, \cdots, \frac{\partial P_{T,k-1}}{\partial P_{T,k}}, \frac{\partial P_{T,k+1}}{\partial P_{T,k}}, \cdots, \frac{\partial P_{T,K}}{\partial P_{T,k}}\right]^T \tag{39}$$

Thus, the solution is given by:

$$\mathbf{d}_k^{DL} = (\mathbf{A}_k^{DL})^{-1} \cdot \mathbf{b}_k^{DL} \tag{40}$$

Again, by considering all the values of k = 0,1,…,K, a total of (K + 1) linear equation systems such as the one defined in (36) should be solved to obtain the downlink derivative vectors $\mathbf{d}_k^{DL}$, $k = 0, \ldots, K$ for all the cells and the corresponding downlink transmitted power gradient.

4.3 Practical implementation considerations

The measurements that are required to compute the load gradient from a practical point of view are detailed below. Specifically, in the uplink direction, these measurements are the following:

- Path loss values between each user and the neighbouring cells. Notice that this measurement is available in current WCDMA systems such as the UMTS [29] through measurement reports. Specifically, mobile stations take this measurement from the pilot channel transmitted by each monitored cell. Measurement reports can be compiled regularly, typically in the order of every 0.5 s, which should be more than enough to cope with mobility and traffic variations.
- Uplink load factor measurements. This is one of the parameters that is used by WCDMA systems to take RRM decisions (admission control and congestion control, for example) [8], and can be obtained from the total power received at the base station and the noise power level, as shown in (3).
- Transmission bit rate and Eb/No of each user, which are also known to the system.

In the downlink direction, the measurements that are required to compute the gradient according to (32) are the same as those in the uplink direction; namely, the path loss between each mobile and the corresponding base stations, and the each user's bit rate and Eb/No. In addition, the orthogonality factor ρ should be considered as a parameter of the algorithm, depending on the environment in which it operates.

Notice also that in a practical situation, the various load factor measurements, the transmitted powers and the corresponding derivatives should be averaged in order to smooth short-term effects (such as Rayleigh, interpath interference) and traffic fluctuations.

As far as computational complexity is concerned, the operations required are feasible in real time so that they can be applied in practical WCDMA systems. This is mainly possible because the RRM algorithms, such as those used for admission or congestion control that it is possible to base on this formulation, will operate on a long-term basis (i.e., in the order of several seconds). However, based on Expressions (22) and (32), an alternative method of approximation for computing the gradient in both the uplink and downlink directions is presented in the Appendix, which makes it possible to further reduce computational complexity.

5 Scenario description and simulation methodology

The framework presented together with the applicability examples, which will be described in Sect. 7, were evaluated by means of simulations in several representative scenarios. Each scenario consisted of an area of 9×9 km^2, in which 23 omni-directional cells with a cell radius of 1 km were arranged in a hexagonal layout. Results were taken for the 0-th cell, which was the central cell and was called BS0. The six cells surrounding BS0 were numbered from BS1 to BS6. In all the scenarios, the path loss was computed following the macrocell model described in [30], which is a common reference in WCDMA evaluations, given by $L_p(\text{dB}) = 128.1 + 37.6 \log(d(\text{km}))$, where d is the distance to the base station. An additional log-normal shadowing with standard deviation of 6 dB was used. Other parameters common to all the scenarios considered here were the chip rate of W = 3.84 Mchips/s and the background noise of –103 dBm for the uplink and –99 dBm for the downlink [30]. The maximum transmitted power in the uplink direction was 21 dBm, while the minimum transmitted power was –44 dBm. In the downlink direction, the maximum power available at the base station was 43 dBm, and the power devoted to the common pilot channel was 30 dBm. Each user was connected to the cell that received the highest pilot power. The antenna gain was 0 dB, and the downlink orthogonality factor was ρ = 0.4.

The specific characteristics of each scenario are described below.

- Scenario 1 (homogeneous scenario). This scenario examined a uniform spatial traffic distribution. A single service was considered with a bit rate of 64 kb/s in both the uplink and downlink directions, which is representative of a typical access bearer for a videophone service [31]. The Eb/No target was 2.9 dB for the uplink and 3.2 dB for the downlink. These values were obtained from a link layer simulator that takes into account the detailed link level characterisation of the physical layer in terms of channel coding, spreading and modulation, transmit diversity and channel impulse response [32].
- Scenario 2 (non-homogeneous scenario at the spatial level). This scenario examined the same service as in Scenario 1 but a non-homogeneous spatial traffic distribution was considered. It was assumed that the neighbouring cell, BS5, had three times as much traffic as the central reference cell, BS0, and that BS4 had twice as much traffic as BS0. The rest of cells had the same amount of traffic as BS0. The numbering of the various neighbouring base stations surrounding the central cell, BS0, is shown in Fig. 1(a).
- Scenario 3 (non-homogeneous scenario at both the spatial and service levels). In this scenario, two types of users were considered: Real Time (RT) and Non Real Time (NRT) users. NRT users are delay-tolerant users whose throughput can be degraded if necessary in order to maintain the throughput of RT users. The RT users had the same characteristics as in Scenario 1 and 2. The bit rate of NRT users was 64 kb/s for the uplink and 384 kb/s for the downlink, which is representative of a typical access bearer for a data service [31]. The Eb/No target of NRT users was 4.2 dB for the uplink and 4.6 dB for the downlink. There were 10 RT users in the scenario. On average, 15% of the users were distributed in the central cell, while the rest of the RT users were homogeneously distributed in the other cells. Furthermore, a variable number was considered for NRT users, whilst BS4 and BS5 (see the layout in Fig. 1(a)) had twice as many NRT users as the rest of cells. No NRT users were distributed in the BS0 reference cell.

Following a similar simulation approach as described in [30], the simulations were based on a number of snapshots. In each snapshot, a number of users were scattered across the whole area according to its specific spatial distribution. It was assumed that all these users were simultaneously active in both the uplink and downlink directions. Power control was then simulated by means of an iterative method to compute the power transmitted per user in the uplink and downlink directions. From this computation, the total uplink load factor in each cell and the total transmitted power per cell was obtained. Besides these general comments about the approach taken in the simulations, the specific characteristics of each simulation study will be detailed in Sects. 6 and 7.

6 Load gradient analysis for different spatial traffic distributions

In order to gain an insight into the potential applicability of the framework presented, a set of illustrative results are shown in this section. Various spatial distributions are considered based on Scenarios 1 and 2 in order to find the relationship between the spatial traffic distribution and the uplink/downlink gradients. Results are taken from the BS0 central cell. In each snapshot, 200 users were randomly scattered across the whole scenario.

6.1 Homogeneous spatial traffic distribution

In this case, which corresponds to Scenario 1 as described in Sect. 5, a uniform traffic distribution was considered in all the cells and within each cell. Figure 1(a) plots, the location

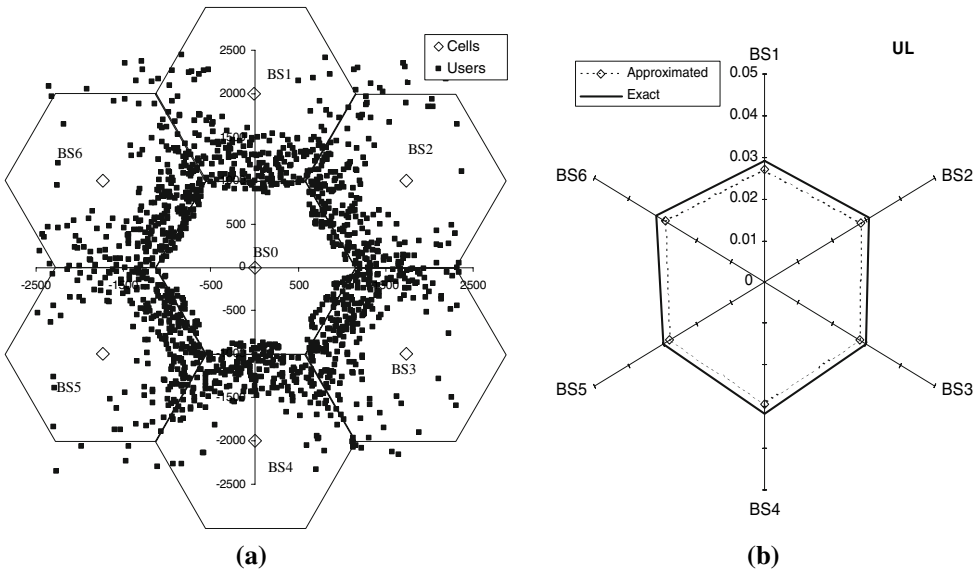

Fig. 1 Spatial representation of the highest contribution to $S_{j,0}^{UL}$ for $j = 1,\ldots, K$ (**a**) and radial plot of $\partial\eta_0/\partial\eta_j$ for the first ring of interfering cells in the uplink (**b**) in a scenario with a homogeneous traffic distribution

of the mobile with the highest contribution to the term $S_{j,0}^{\mathrm{UL}}$ $(j = 1,\ldots,K)$ in each snapshot. According to (11), the users plotted in Fig. 1(a) are therefore those who have the greatest effect on the reference cell load factor, or those that cause the highest intercell interference in the BS0 cell. Notice that in the uplink direction, these users tend to be those who are closer to the reference cell site. Figure 1(b) is a radial plot showing the derivatives $\partial\eta_0/\partial\eta_j$ for the six neighbouring cells ($j = 1,\ldots, 6$). It can be observed that since the traffic is homogeneously distributed, the derivatives are similar in the 6 cells. Figure 1(b) also shows a comparison between the exact computation of the derivatives that appear in Sect. 4 and the approximate computation, which is presented in the Appendix. This comparison shows that the approximated method is a good estimation of the derivative.

Similarly, Fig. 2(a) shows the users in the reference cell having the highest contribution to the term $S_{0,j}^{DL}$ in the downlink direction (i.e., those experiencing the highest downlink intercell interference). The corresponding downlink derivatives for the six neighbouring cells are presented in Fig. 2(b). Similar comments as those made on the uplink apply in this homogeneous case, and all the derivatives are approximately the same.

6.2 Non-homogeneous spatial traffic distribution

At this stage, two different types of non homogeneous spatial traffic distributions were considered and are referred to as inter-cell and intra-cell non-homogeneity.

6.2.1 Inter-cell non-homogeneity

This type of spatial non-homogeneity, which corresponds to Scenario 2 described in Sect. 5, reflects the situation in which the number of users differs from cell to cell but users are uniformly distributed inside a given cell. In this case, Fig. 3(a) shows the users that contributed the most to the inter-cell interference in the reference cell BS0 in the uplink direction. Likewise, the uplink gradient components for the six neighbouring cells are shown in Fig. 3(b). It can be seen that BS5 (and also BS4 to a lesser extent) are the cells to which BS0 is most sensitive, because they have the highest derivative. Consequently, the gradient in the uplink direction is able to identify non-homogeneous spatial distributions between different cells.

With respect to the downlink direction, since the gradient in (32) depends on the term $S_{0,j}^{DL}$, that is, on how the users of the 0-th cell are distributed with respect to the neighbouring cells, the derivatives are approximately the same for all the neighbouring cells because users in the 0-th cell are homogeneously distributed inside this cell. Therefore, the corresponding plots, which are not shown for the sake of brevity, would be similar to those in Fig. 2.

6.2.2 Intra-cell non-homogeneity

This type of non-homogeneity reflects the situation in which there is the same amount of traffic in each cell but the user distribution is non-uniform inside a given cell. As an example, let us consider a variant of Scenario 1 in which all the cells have the same load but the users in BS0 are distributed non-homogeneously inside the cell, 80% of whom are concentrated in the lower right-hand area (i.e., in the region closest to BS3). In the uplink direction, the results, which are not shown for the sake of brevity, are very similar to those obtained for the homogeneous case shown in Fig. 1. This demonstrates that intra-cell spatial traffic non-homogeneity does not significantly affect the

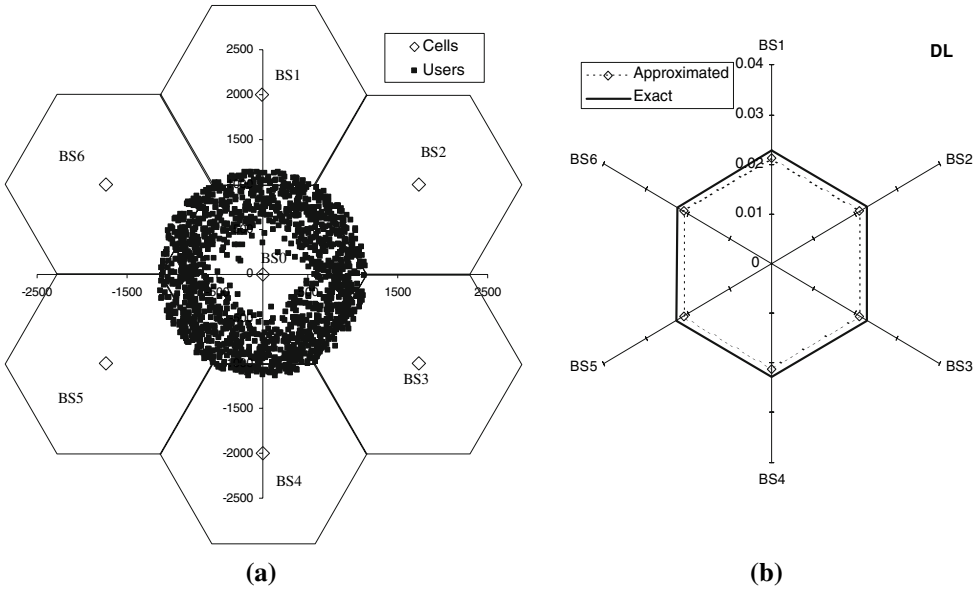

Fig. 2 Spatial representation of the user having the highest downlink intercell interference (**a**) and radial plot of $\partial P_{T0}/\partial P_{Tk}$ for the first ring of interfering cells in the downlink (**b**) with a homogeneous traffic distribution

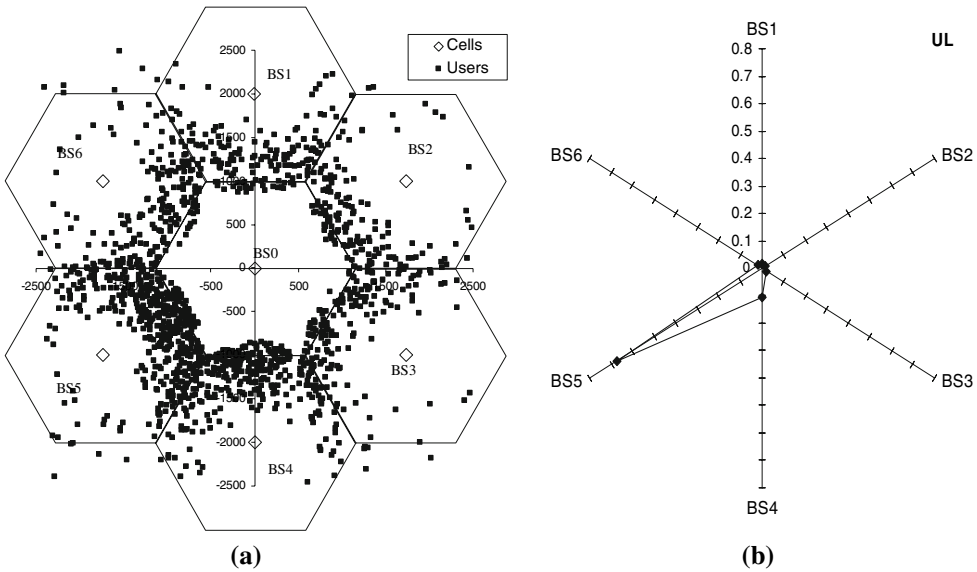

Fig. 3 Spatial representation of the highest contribution to $S_{j,0}^{UL}$ for $j = 1,\ldots, K$ (**a**) and radial plot of $\partial\eta_0/\partial\eta_j$ for the first ring of interfering cells in the uplink (**b**) in a scenario with intercell non-homogeneity

sensitivity in the uplink direction of the reference cell. It furthermore shows that the most relevant users continue to be those in the neighbouring cells that are closer to the reference cell site. It should be mentioned that the user distribution inside the BS0 reference cell affects the intercell distribution in the neighbouring cells. It therefore modifies the power transmitted by the users in these neighbouring cells, thus causing an indirect variation in the load factor in the reference cell. However, simulation results in this situation have revealed that this indirect variation is negligible.

Nevertheless, different conclusions can be drawn from the downlink analysis. As is shown in Fig. 4(a), the most relevant users from the perspective of the reference cell are mainly located in the lower right-hand area of this cell, in which user density is the highest. This results in a higher coupling between cell BS3 and the reference cell, because most of the users in the reference cell are affected by interference from the power transmitted by BS3. As a result of this, BS3 has the largest component on the power gradient, followed by BS2 and BS4, as shown in Fig. 4(b). Consequently, the gradient of the downlink transmitted power is able to identify non-homogeneous spatial distributions inside the reference cell.

7 Applicability examples: RRM algorithms

The gradient framework for WCDMA uplink and downlink characterisation together with the explicit formulation of

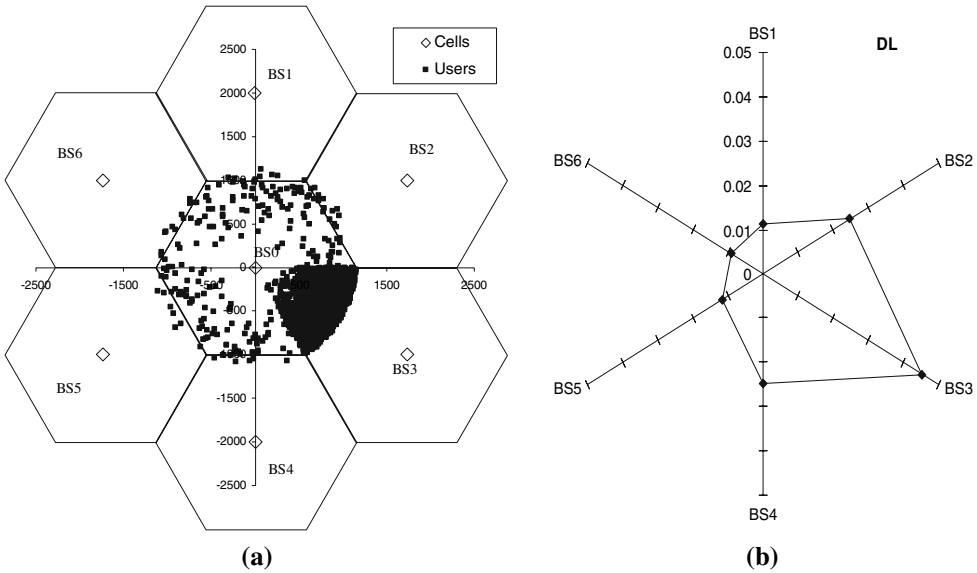

Fig. 4 Spatial representation of the user having the highest downlink intercell interference (**a**) and radial plot of $\partial P_{T0}/\partial P_{Tk}$ for the first ring of interfering cells (**b**) with an intracell non-homogeneous traffic distribution

the coupling between the different cells in the environment is claimed to have a wide range of applications, mostly in the context of RRM strategies. Some examples of these applications are presented below in addition to the results of their performance. Here, it should be noted that the aim of this paper is not to provide an extensive analysis of the different algorithms, but to show the potential applicability of the proposed framework by means of a few examples.

7.1 Congestion control algorithms

Congestion occurs when it is not possible to provide the admitted users with the normal services agreed on for a given percentage of time because of an overload. Usually the network is planned to operate below a certain maximum load factor (alternatively in the downlink it is planned to operate below a certain fraction of the maximum power available at the cell site). Thus, the congestion control procedure will include first the congestion detection phase (e.g., when the load factor increases over a certain threshold for a certain amount of time). The congestion resolution phase will then be triggered when congestion has been detected and it consists in the actions that must be undertaken in order to maintain the network stability, such as user/services prioritisation, the reduction of the maximum transmission rate for delay-tolerant services, the blocking of new connections or the dropping of on-going connections. Finally, when the congestion is overcome (that is, when the load factor has dropped below the maximum level for a certain period of time), the congestion recovery algorithm is triggered in order to restore the transmission capabilities the various mobiles had before the congestion resolution was executed. The algorithm and results presented below, which take advantage of the gradient framework presented in Sects. 3 and 4, concern the congestion resolution phase. Congestion detection and congestion recovery algorithms, such as the ones presented in [24], would be readily applicable to this context.

7.1.1 Uplink direction

Let us consider that an overload situation occurs at the reference cell, so that the uplink load factor η_0 at a given time is above a desired target value, η_T. The objective of the congestion control algorithm is to reduce the load by inhibiting delay-tolerant transmissions (i.e., non real time users) smartly, so that the minimum number of users is affected by the congestion resolution actions, thus minimizing the throughput reduction and the signalling load. In order to avoid load reductions beyond the necessary level, the procedure is envisaged on a step-by-step basis and, consequently, the number of algorithm iterations is an additional useful measurement. Thus, certain load reduction decisions are taken at the time the congestion is triggered and, after a period that should be long enough for the signalling and execution of layer-3 messages at the mobile terminals plus the time necessary for the power control to converge to the new air interface conditions, the load would be measured again and new actions would be taken at that time if necessary. The algorithm stops after a number of periods or iterations, when the measured cell load factor is below η_T. In the above framework, the proposed algorithm carries out the following steps per iteration, whenever the congestion control is triggered for the reference cell:

Step 0. Compute the gradient of the uplink load factor given by (21)

Step 1. Select the cell $k = 0,1,\ldots,K$ with maximum value of the product $(\eta_{k,NRT} \cdot \frac{\partial \eta_0}{\partial \eta_k})$ where $\eta_{k,NRT}$ is the estimated amount of load factor devoted to non-real time (NRT) traffic in cell k, given by [8]:

$$\eta_{k,NRT} = (1+f) \sum_{i_k=1}^{n_{k,NRT}} \frac{1}{\frac{W}{\left(\frac{E_b}{N_o}\right)_{i_k} R_{i_k}} + 1} \tag{41}$$

where $n_{k,NRT}$ is the number of NRT users in the k-th cell, f is a parameter that reflects the ratio between intercell and intracell power and $(E_b/N_o)_{i_k}$ is the Eb/No target of the i-th user in the k-th cell. Notice that this step ensures that the cell that has the greatest influence on the reference cell is selected, thus giving rise to more efficient load control actions.

The f parameter takes values that are typically around 0.6 [7], but depends on the specific user distribution and on the propagation conditions. In practice, the real value may exhibit high variations over time, so it may be possible to optimise this parameter for each specific situation, as is discussed in [13]. However, such an optimisation is beyond the scope of this paper and the value $f = 0.6$, which is considered sufficiently representative of the simulations carried out, is retained here for purposes of illustration.

Step 2. The target reduction to be achieved in the selected cell k is given by

$$\Delta\eta_k = \frac{\Delta\eta_0}{\left(\frac{\partial \eta_0}{\partial \eta_k}\right)} \tag{42}$$

where $\Delta\eta_0 = \eta_0 - \eta_T$ is the desired reduction in cell 0.

Step 3. Order the NRT users of the k-th cell in a table in decreasing order of their contribution to the term $S_{k,0}^{UL}$, given by the factor $I_{i_k,0}^{UL}$:

$$I_{i_k,0}^{UL} = \frac{L_{i_k,k}}{L_{i_k,0}} \frac{1}{\frac{W}{\left(\frac{E_b}{N_0}\right)_{i_k} R_{i_k}} + 1} \tag{43}$$

Step 4. Inhibit the transmissions of the users in the table until the desired reduction of $\Delta\eta_k$ is reached, or until all the users have been inhibited. Notice that, after one user is inhibited, the selected cell may no longer be the one with the steepest gradient. However, this is not considered in the algorithm in order to achieve greater speed. In order to account for this fact, the algorithm should inhibit only one user per iteration. It should then measure the loads again and compute the gradients, which slows down the capability of the algorithm.

Step 5. Measure η_0 and if it is still higher than η_T return to Step 0 and perform another iteration.

7.1.2 Downlink direction

Similarly, the objective of the algorithm in the downlink direction is to limit the fraction of transmitted power with respect to the maximum available power (i.e., P_{T0}/P_{max0}) to a given bound ϕ_T. Thus, the steps to be taken by the algorithm when congestion is triggered in the reference cell would be as follows:

Step 0. Compute the gradient of the downlink transmitted power given by (31)

Step 1. Select the cell $k = 0,1,\ldots,K$ with maximum value of the product: $(P_{k,NRT} \cdot \frac{\partial P_{T0}}{\partial P_{Tk}})$ where $P_{k,NRT}$ is the amount of transmitted power devoted to NRT traffic in the k-th cell.

Step 2. The power reduction to be achieved in the selected cell k is given by

$$\Delta P_{Tk} = \frac{\Delta P_{T0}}{\left(\frac{\partial P_{T0}}{\partial P_{Tk}}\right)} \tag{44}$$

where ΔP_{T0} is the desired reduction in cell 0, given by $\Delta P_{T0} = P_{T0} - \phi_T \cdot P_{max0}$

Step 3. Order the NRT users in cell k in a table in decreasing order of transmitted power.

Step 4. Inhibit the transmissions of the users in the table until reaching the desired reduction of ΔP_{Tk} or until all the users have been inhibited.

Step 5. Measure $\phi_0 = P_{T0}/P_{max0}$ and if it is still higher than ϕ_T return to Step 0 and perform another iteration.

In order to show the suitability of the envisaged algorithms, two reference algorithms were considered for purposes of comparison in both the uplink and downlink directions. The first is denoted as a "random algorithm" and operates using the same steps as the previously described algorithm. However, in Step 1 the base station is randomly selected from the six neighbouring cells and in Step 3 the users are not ordered. The second is denoted as an "interactive-based algorithm" and operates by selecting in Step 1 the base station with the largest number of interactive users.

The performance evaluation was carried out for Scenario 3 described in Sect. 5. A total of 15,000 snapshots were simulated. For each snapshot, a mixture of RT and NRT users were scattered across the whole area according to the specific spatial traffic distribution given in the scenario description. The objective of the load control algorithms was to keep both the uplink cell load factor

and the DL power fraction below $\eta_T = 0.8$ and $\phi_T = 0.8$, respectively, in the BS0 reference cell. Consequently, the congestion control algorithm was executed for each snapshot in which either the uplink cell load factor or the downlink power fraction in the reference cell was above the limit of 0.8. It is worth mentioning that although the scenario considered a total of 23 cells, only the six cells surrounding the reference cell were considered for the algorithm (i.e., $K = 6$ in Step 1 of the algorithm description).

Figure 5 shows the corresponding average values of the uplink cell load factor and downlink power fraction in the reference cell before applying the congestion algorithm as a function of the number of NRT users in the scenario under examination. Figures 6 and 7 show the uplink and the downlink performances in terms of the required NRT throughput reduction in the neighbouring cells in order to solve the congestion situations in the reference cell. The results are presented for the three algorithms studied. Furthermore, in the derivative-based case, the results are presented for the exact gradient computation method explained in Sect. 4 as well as for the method of approximation, as explained in the Appendix. The first feature to note is that the use of the method of approximation does not reveal significant differences compared to the exact method, thus validating this method for radio engineering purposes. The reason is that the gradient is in fact mainly used in the algorithm to select the appropriate cell to execute the congestion actions, so it depends more on cell with the highest value in the gradient than on the exact value of the derivatives. It can be observed that the algorithm based on derivatives achieves the load reduction in the reference cell with the lowest NRT throughput reduction, thus executing a more efficient congestion control. The reason is that the derivative-based algorithm is able to select the most appropriate cell to execute the load reduction actions in each situation. Notice also that in some situations in the uplink, the congestion control resolution with the derivative-based algorithm is achieved with NRT throughput reductions below 0, which means that the throughput after congestion is higher than the throughput before congestion. This is because some users who were in outage before triggering congestion—and thus their transmissions were erroneous—are able to transmit correctly thanks to the load reduction achieved by the algorithm during congestion resolution. In contrast, this effect is not observed with the interactive-based and the random algorithm, because the throughput reduction due to the number of inhibited transmissions is higher than the throughput increase due to the reduction in outage.

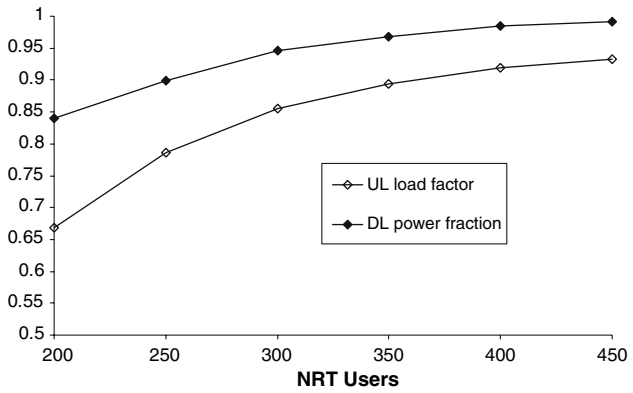

Fig. 5 Average uplink load factor and downlink power fraction in the reference cell

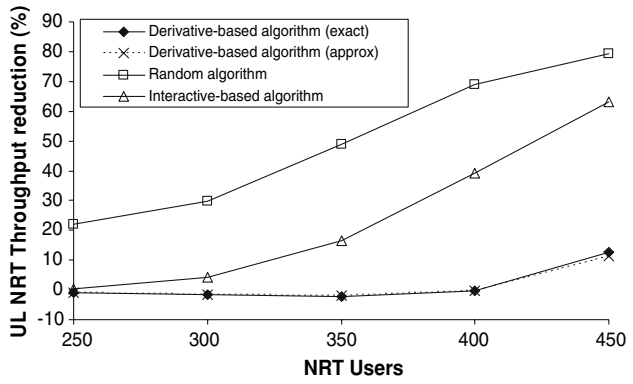

Fig. 6 UL NRT Throughput reduction in the neighbouring cells

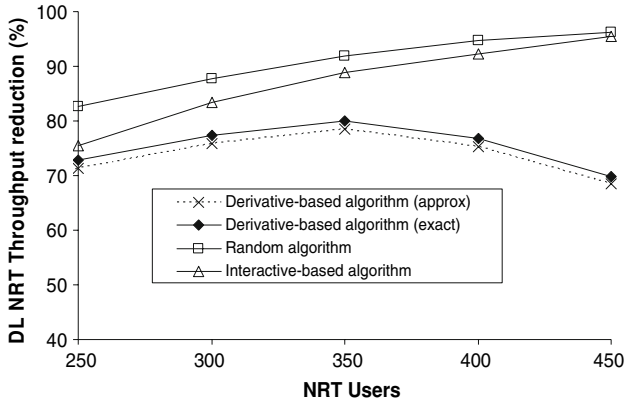

Fig. 7 DL NRT Throughput reduction in the neighbouring cells

The duration of the algorithm is also much shorter in the case of the derivative-based algorithm. This is reflected in Table 1, which presents the average number of iterations required to solve the congestion situations in the different situations. The lower the number of iterations, the more effective the algorithm is, which indicates that it is able to identify those cells and users with the highest influence on the air interface load. Notice also that when focus is placed on the downlink direction, the derivative-based algorithm still performs better. However, in this case it requires a higher NRT throughput reduction and a greater number of

Table 1 Average number of iterations of the algorithms considered

NRT users	UPLINK			DOWNLINK		
	Derivative	Random	Interactive	Derivative	Random	Interactive
250	1.82	4.64	3.85	23.19	24.56	23.94
300	3.39	6.68	7.45	32.67	35.61	35.87
400	10.41	17.24	17.11	50.98	69.00	69.02

iterations. This is due to the higher load in the downlink direction, as depicted in Fig. 5.

7.2 Admission control algorithms

The admission control procedure is used to decide whether to accept or reject a new connection depending on the interference it adds to the existing connections. Admission control principles make use of the uplink load factor (or the power transmitted in the downlink) and the estimate of the load increase that a new bearer request would cause in the radio network [14–20]. Usually, the algorithm is applied at a single cell level. However, the impact of the user's potential acceptance should be assessed for both the serving cell and the neighbouring cells. This would therefore guarantee that the set-up of the new connection is affordable to system overall. In order to illustrate the improvements that can be achieved by means of the gradient-based solution with respect to another approach, the algorithm presented in [23] is taken as a reference below. It is denoted as the "multi-cell reference algorithm", which proves to be an efficient solution for multi-cell admission control.

If we examine the uplink direction, in which n_o users have already been admitted to the reference cell ($j = 0$), let us assume that the $(n_o + 1)$-th user requests admission to this cell, denoted below as the serving cell. The user requirements are bit rate R_{n_o+1} and Eb/No target $(E_b/N_o)_{n_0+1}$. The multi-cell reference algorithm estimates the increase in received power due to the new admission request. It accepts the new user provided that the following condition holds in the serving cell ($j = 0$) and in the neighbouring cells $j = 1,\ldots, K$:

$$\frac{P_{TOT,j} + \Delta P_{est,j}}{P_N} \leq NR_{\max} = \frac{1}{1 - \eta_{\max}} \tag{45}$$

where $P_{TOT,j}$ is the total received power in the j-th cell before admission, P_N is the noise power, $NR_{\max}$ is the admission control threshold expressed in terms of noise rise (i.e., the ratio between total received power with respect to noise power), which is equivalent to maximum load factor $\eta_{\max} = 1 - 1/NR_{\max}$, and $\Delta P_{est,j}$ is the power increase estimation due to the new user in the j-th cell. For the serving cell, the power increase estimation is given by:

$$\Delta P_{est,0} = \frac{P_{TOT,0}\Delta\eta_{est,0}}{1 - \eta_0 - \Delta\eta_{est,0}} \tag{46}$$

where the load increase estimation of the new user in the serving cell is given by:

$$\Delta\eta_{est,0} = \frac{1}{\frac{W}{\left(\frac{E_b}{N_o}\right)_{n_0+1} R_{n_0+1}} + 1}\left(\omega + (1-\omega)\frac{P_N}{P_N + \chi_0}\right) \tag{47}$$

where χ_0 is the total inter-cell interference power at cell $j = 0$ and w is a weight factor that takes the value 0.5 in [23].

In turn, the power increase estimation in the neighbouring cells, $j = 1,\ldots,K$, is given by:

$$\Delta P_{est,j} = \Delta P_{est,j}\frac{L_{n_0+1,0}}{L_{n_0+1,j}} \tag{48}$$

where $L_{n_0+1,j}$ is the total path loss between the new user and the j-th cell.

Based on this framework, the following alternatives are compared below:

- Multi-cell reference algorithm. This corresponds to the application of Condition (45) with a power increase estimation given by (47) for the serving cell and by (48) for the neighbouring cells.
- Single-cell algorithm. This corresponds the application of the Condition (45) to the serving cell only.
- Multi-cell gradient-based algorithm. This corresponds to a modification in the reference algorithm by making use of the proposed gradient-based framework. Specifically, the power increase estimation for the neighbouring cells $j = 1,\ldots,K$, is given by:

$$\Delta P_{est,j} = \frac{P_{TOT,j}\Delta\eta_{est,j}}{1 - \eta_j - \Delta\eta_{est,j}} = \frac{P_{TOT,j}\Delta\eta_{est,0}\frac{\partial\eta_j}{\partial\eta_0}}{1 - \eta_j - \Delta\eta_{est,0}\frac{\partial\eta_j}{\partial\eta_0}} \tag{49}$$

To illustrate the performance of the proposed strategy, the three alternatives were evaluated and compared in Scenario 1, as described in Sect. 5, in which a total of 15,000 snapshots were simulated. The admission of a new

user to the reference cell was simulated in each snapshot. The maximum load factor was $\eta_{\max} = 0.8$. The method of approximation for computing the derivatives was used in the simulations. Only the uplink direction was considered.

Figure 8 shows the admission probability (i.e., the probability that the new user will be accepted, provided that a number of users have already been accepted) as a function of the total number of users in the scenario. It can be observed that the admission probability is higher for the single-cell algorithm than for the multi-cell algorithms, because only one condition is checked, while the multi-cell algorithms check the conditions for all the cells and therefore there are more possibilities a call being rejected. Likewise, the admission probability using the multi-cell reference algorithm and the gradient-based algorithm is very similar. However, in order to check that admission is being properly carried out, the so-called false admission probability is also plotted in Fig. 9. It is defined as the probability of measuring a load factor above the desired threshold $\eta_{\max}$ in either the reference cell or the neighbouring cells after accepting a new user, which means that

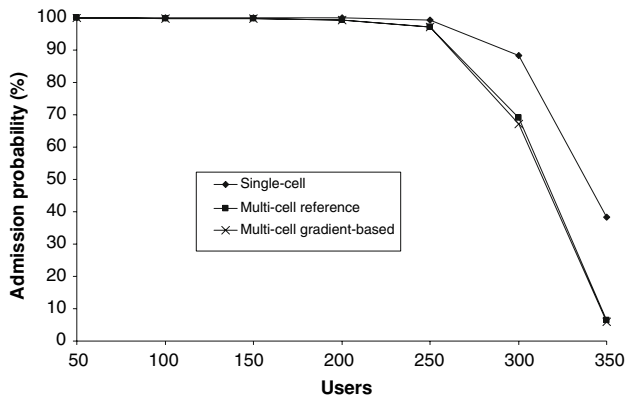

Fig. 8 Admission probability for the considered admission control algorithms

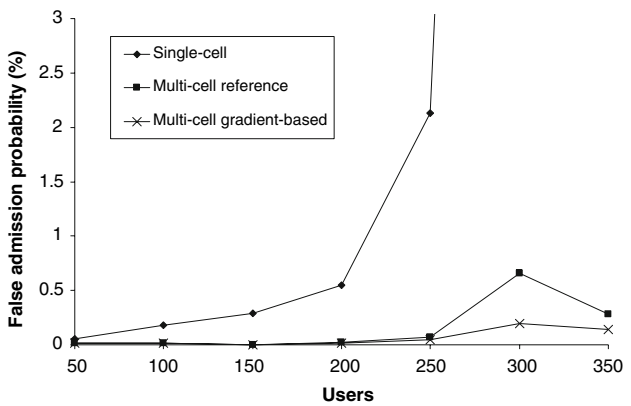

Fig. 9 False admission probability for the considered admission control algorithms

the new user should in fact have been rejected. It can be observed that the single-cell algorithm has a very high rate of false admissions, while the multi-cell gradient-based algorithm offers the lowest rate of false admissions, which demonstrates that the proposed framework is able to achieve a better load control during the admission phase. Notice that the behaviour of the false admission probability in Fig. 9 shows an initial increase in the number of users followed by a decrease when the number of users is very high, which means that in this latter case there are very few admissions and consequently very few false admissions.

Another possible performance indicator would be the false rejection probability (i.e., the probability of rejecting a user in spite of the fact that the load factor in the reference and neighbouring cells would have remained below the threshold $\eta_{\max}$ if the user had been accepted). Notice that in general this effect is not as important as a false admission when comparing schemes exhibiting in practice near the same admission probability (like the multi-cell reference algorithm and the multi-cell gradient-based algorithm in Fig. 8) because it does not result in a load increase above the threshold. The results are not shown here for the sake of brevity, but they indicate that the false rejection probability takes low values and there are not significant differences between the multi-cell reference algorithm and the multi-cell gradient-based algorithm.

7.3 Multiple carrier planning

Although the WCDMA allows for complete frequency reuse, it is usual in many countries that multiple frequency carriers are available for a given operator (i.e., typically two or three). In this case, a frequency planning is needed [33]. In this exercise, given that in practice non-homogeneous traffic distributions through cells will be present, the frequency planning algorithm is able to use the gradient framework, which will eventually lead to a semi-dynamic planning algorithm. Thus, the basic idea behind this approach is that if $\partial\eta_0/\partial\eta_k > \beta$, where β is a certain threshold, both the 0-th and the k-th cell would have different frequencies in order to decouple their mutual interference. Furthermore, given that traffic profiles will also vary over time, the gradient evaluation could be regularly carried out in order to find the optimum network configuration for each situation.

In order to illustrate this approach with some results, Scenario 2 described in Sect. 5 was considered, in which a total of 15,000 snapshots were simulated. In this scenario, BS5 had more traffic than the rest of cells and therefore presented the highest value for $\partial\eta_0/\partial\eta_k$ (see Fig. 10, which presents the load factor derivative of BS0 with respect to BS5 and to another cell, BS2, for example). BS5 was therefore considered to be most suited for changing the

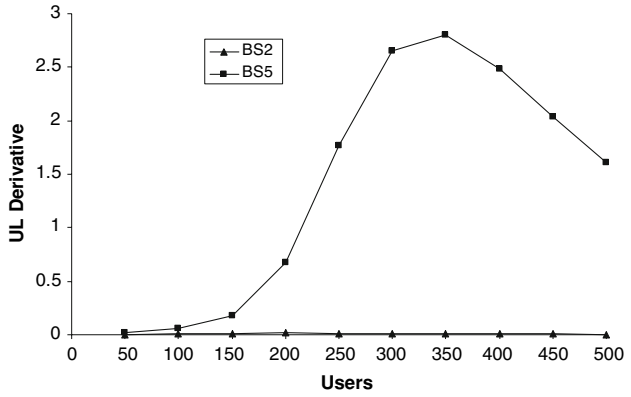

Fig. 10 Uplink load factor derivatives with respect to BS5 and BS2

carrier frequency according to the derivative criterion. Figures 11 and 12 show the outage probability (i.e., the probability that the Eb/No will fall below the target) and the average load factor in the reference cell, respectively. The results in these figures are shown for the case in which all the cells use the same carrier (i.e., complete frequency reuse), for the case in which BS5 changes to a different carrier—according to the derivative-based criterion—and

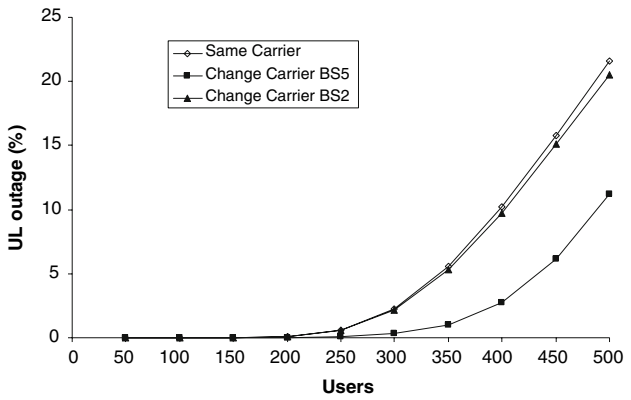

Fig. 11 UL outage in the reference cell for different carrier allocations

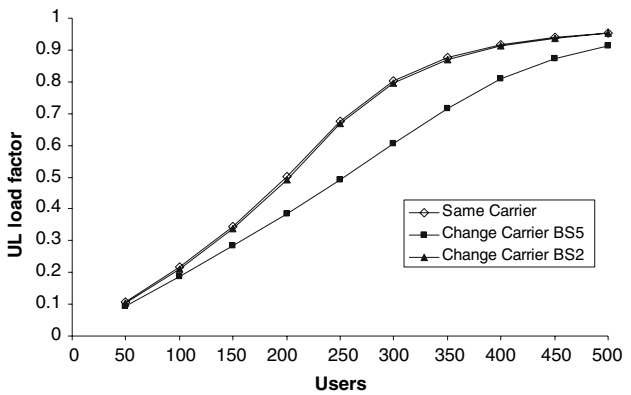

Fig. 12 Uplink load factor in the reference cell as a function of the number of users in the scenario

for the case in which BS2 is the cell with a different carrier. It can be seen that the frequency change in BS2 does not provide a significant reduction in interference. In contrast, when BS5 is the cell that changes to a different frequency, both the outage and the load factor show a considerable reduction in the reference cell.

8 Conclusions

This paper has presented an innovative framework for capturing the coupling between different cells in a WCDMA multi-cell scenario. In the uplink direction, the relationship between the uplink cell load factors in the different cells was explicitly obtained. Based on the expressions that were derived, the gradient of the load factor (i.e., the derivatives of the load factor of a reference cell with respect to the load factor of the neighbouring cells) was computed. A similar analysis was carried out in the downlink direction in terms of base station transmitted power. It has been shown that the gradient in the uplink direction is able to identify the influence of non-homogeneous spatial distributions at the inter-cell level (i.e., neighbouring cells that have varying amounts of traffic) over the reference cell. Likewise, the gradient in the downlink direction is able to identify the influence between the different cells depending on the spatial distribution at the intra-cell level (i.e., how the users are distributed in the reference cell).

Some examples of the applicability of the proposed methodology in the field of RRM have also been proposed and analysed. A particular example is a congestion control algorithm that selects the appropriate cell for executing congestion control actions depending on the gradient, as a result of which it is possible to obtain a faster and more efficient reaction capability in congestion situations. Similarly, with respect to admission control it has been shown that by making use of the derivatives, the effect of a new user on the rest of cells can be better captured in order to reduce the false admission probability. Finally, the possibility of making use of the gradient when the appropriate carriers are selected for allocation to the different cells in WCDMA scenarios with multiple carriers has also been analysed.

Acknowledgements This work was carried out in the framework of project IST-AROMA (http://www.aroma-ist.upc.edu), which was partly funded by the European Community and by the Spanish Research Council (CICYT) under grant TEC2006-26873-E and COSMOS grant TEC2004-00518.

9 Appendix: Method of approximation for gradient computation

The method described in Sect. 4 for the gradient computation of the uplink cell load factor and the downlink base

station transmitted power gradient involves the solution of $(K + 1)$ linear equation systems. Although the computational complexity would be feasible in real time, from a radio network engineering point of view a simpler formulation that maintains a sufficient degree of accuracy may be preferred. In this framework, this Appendix provides a method of approximation for gradient computation that can be used as an alternative to the method described in Sect. 4. Specifically, the method of approximation for Expression (22) is given by:

$$\frac{\partial \eta_0}{\partial \eta_k} \approx \frac{S_{k,0}^{UL}\left(1 - S_{0,0}^{UL}\right)}{(1-\eta_k)^2\left(1 + \sum_{j=1}^{K} \frac{S_{j,0}^{UL}}{1-\eta_j}\right)^2} \tag{50}$$

where it has been assumed that the term in the summation in (22) that most contributes to $\partial\eta_0/\partial\eta_k$ is $\partial\eta_k/\partial\eta_k = 1$.

For the downlink direction, a similar argument can be made, therefore, the approximation of (32) is given by:

$$\frac{\partial P_{T0}}{\partial P_{Tk}} \approx \frac{S_{0,k}^{DL}}{1 - \rho S_{0,0}^{DL}} \tag{51}$$

In order to assess the accuracy of this approximation, various simulations were carried out in different scenarios using different traffic distributions. Various examples are shown in Figs. 1(b) and 2(b), which correspond to the uplink and downlink directions under the conditions discussed in Sect. 6.1. Notice that in both links the approximation underestimates the exact derivative due to the terms that were neglected when expressions (50) and (51) were obtained. In general, for other load conditions and spatial distributions the approximation holds quite well and errors below 10% were observed. Furthermore, as is shown in Sect. 7.1, the use of the exact or the approximate gradient has a very low impact on the performance that is observed with the gradient-based algorithms.

References

1. Bannister, J., Mather, P., & Coope, S. (2004). *Convergence technologies for 3G networks*. John Wiley and Sons.
2. Tanner, R., & Woodard, J. (2004). *WCDMA requirements and practical design*. John Wiley & Sons.
3. 3GPP TR 25.922 v5.3.0, Radio resource management strategies.
4. Sallent, O., Pérez-Romero, J., Agustí, R., & Casadevall, F. (2003). Provisioning multimedia wireless networks for better QoS: RRM strategies for 3G W-CDMA. *IEEE Communications Magazine, 41*(2), 100–106.
5. Pérez-Romero, J., Sallent, O., & Agustí, R. (2004). A novel approach for multi-cell load control in WCDMA. *IEE 3G Conference*, London.
6. Gilhousen, K. S., Jacobs, I. M., Padovani, R., Viterbi, A. J., Weaber, L. A., & Wheatley, C. E. III. (1991). On the capacity of a cellular CDMA system. *IEEE Transactions on Vehicular Technology, 40*(2), 303–312.
7. Viterbi, A. J., Viterbi, A. M., & Zehavi, E. (1994). Other-cell interference in cellular power-controlled CDMA. *IEEE Transactions on Communications, 42*(2/3/4), 1501–1504.
8. Holma, H., & Toskala, A. (2002). *WCDMA for UMTS* (2nd ed.). John Wiley & Sons.
9. Lundin, E. G., Gunnarsson, F., & Gustafsson, F. (2003). Uplink load estimation in WCDMA. *IEEE Wireless Communications and Networking*. WCNC 2003, 1669–1674.
10. Lui, Z., & El Zarki, M. (1994). SIR-based call admission control for DS-CDMA cellular systems. *IEEE Journal on Selected Areas in Communications, 12*(4), 638–644.
11. Badia, L., Zorzi, M., & Gazzini, A. (2002). On the impact of user mobility on call admission control in WCDMA systems. *56th IEEE VTC Fall Conference*, Vancouver, pp. 121–126.
12. Redana, S., & Capone, A. (2002). Received power-based call admission control techniques for UMTS uplink. *56th IEEE VTC Fall Conference*, Vancouver, pp. 2206–2210.
13. Sallent, O., Pérez-Romero, J., & Agusti, R. (2003). Optimizing statistical uplink admission control for W-CDMA. *57th IEEE VTC fall conference*, Orlando, USA.
14. Holma, H., & Laakso, J. (1999). Uplink admission control and soft capacity with MUD in CDMA. *IEEE Vehicular Technology Conference in Fall 1999*, Amsterdam, pp. 431–435.
15. Gunnarsson, F., Geijer Lundin, E., Bark, G., & Wiberg, N. (2002). Uplink admission control in WCDMA based on relative load estimates. *IEEE International Conference on Communications, ICC-2002*, pp. 3091–3095.
16. Capone, A., & Redana, S. (2001). Call admission control techniques for UMTS. *54th IEEE VTC Fall Conference*, Atlantic City, pp. 959–929.
17. Dimitriou, N., Sfikas, G., & Tafazolli, R. (2000). Call admission policies for UMTS. *51st IEEE VTC Spring Conference*, Tokyo, pp. 1420–1424.
18. Ho, C. J., Copeland, J. A., Lea, C. T., & Stuber, G. L. (2001). On call admission control in DS/CDMA cellular networks. *IEEE Transactions on Vehicular Technology, 50*(6), 1328–1343.
19. Phan-Van, V., & Glisic, S. (2001). Radio resource management in CDMA cellular segments of multimedia wireless IP networks. *The 4th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Aalborg, Denmark.
20. Knutsson, J., Butovitsch, P., Persson, M., & Yates, R. D. (1998). Downlink admission control strategies for CDMA systems in a Manhattan environment. *IEEE Vehicular Technology Conference VTC*, pp.1453–1457.
21. Kazmi, M., Godlewski, P., & Cordier, C. (2000). Admission control strategy and scheduling algorithms for downlink packet transmission in WCDMA. *52nd IEEE Vehicular Technology Conference Fall*, Boston, pp. 674–680.
22. Aïssa, S., Kuri, J., & Mermelstein, P. (2004). Call admission on the uplink and downlink of a CDMA system based on total received and transmitted powers. *IEEE Transactions on Wireless Communications, 3*(6), 2407–2416.
23. Outes, J., Nielsen, L., Pedersen, K., & Mogensen, P. (2001). Multi-cell admission control for UMTS. *VTC Spring Conference*, Vol. 2, pp. 987–991.
24. Pérez-Romero, J., Sallent, O., Agustí, R., & Sánchez, J. (2002). Managing radio network congestion in UTRA FDD. *IEE Electronics Letters, 38*, 1384–1386.
25. Liu, T. K., & Silvester, J. A. (1998). Joint admission/congestion control for wireless CDMA systems supporting integrated services. *IEEE Journal on Selected Areas in Communications, 16*(6), 845–857.

26. Passas, N., & Merakos, L. (1996). A graceful degradation method for congestion control in wireless personal communication networks. *Proceedings of the IEEE Vehicular Technology Conference VTC*, pp. 126–130.
27. De Bernardi, R., Imbeni, D., Vignali, L., & Karlsson, M. (2000). Load control strategies for mixed services in WCDMA. *51st IEEE Vehicular Technology Conference (VTC) Spring*, Tokio, pp. 825–829.
28. Rave, W., Kohler, T., Voigt, J., & Fettweis, G. (2001). Evaluation of load control strategies in an UTRA/FDD network. *IEEE 53rd Vehicular Technology Conference Spring*, Rhodes, Greece, pp. 2710–2714.
29. 3GPP TS 25.331 Radio Resource Control (RRC); Protocol Specification.
30. 3GPP TR 25.942 v5.3.0, Radio Frequency (RF) System Scenarios.
31. 3GPP TS 34.108 Common Test Environments for User Equipment (UE); conformance testing.
32. Olmos, J. J., & Ruiz, S. (2002). Transport block error rates for UTRA-FDD downlink with transmission diversity and turbo coding. *13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications PIMRC-2002*, Vol. 1, pp. 31–35.
33. Rautiainen, T. (2002). Breaking the hierarchical cell structure in WCDMA networks. In *Proceedings of the 55th IEEE VTC Spring*, Vol. 1, pp.110–114.

Author Biographies

Jordi Pérez-Romero received the Telecommunications Engineering degree from the ETSETB of the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in December 1997 and the Ph.D. at the Department of Signal Theory and Communications (TSC) of the UPC in April 2001. He joined the Radio Communications Group in January 1998, with a grant of the Spanish Educational Ministry, and is currently associate professor. His research interests are in the field of mobile communication systems, especially packet radio techniques, spread-spectrum systems, radio resource and QoS management, and heterogeneous wireless networks. He has been involved in different European Projects (WINEGLASS, ARROWS, EVEREST, E2R, NEWCOM, and AROMA) as well as in projects for private companies. He has published several papers in international journals and conferences and has co-authored one book on mobile communications.

Oriol Sallent is Associate Professor at the Universitat Politècnica de Catalunya. His research interests are in the field of mobile communication systems, especially radio resource and spectrum management for heterogeneous networks. He has published more than 70 papers in international journals and conferences. He has participated in several research projects of the 5th and 6th Framework Programme of the European Commission and served as a consultant for a number of private companies.

Ramon Agustí received the Engineer of Telecommunications degree from the Universidad Politécnica de Madrid, Spain, in 1973, and the Ph.D. degree from the Universitat Politècnica de Catalunya (UPC), Spain, 1978. In 1973 he joined the Escola Tècnica Superior d'Enginyers de Telecomunicació de Barcelona, Spain, where he became Full Professor in 1987. After graduation he was working in the field of digital communications with particular emphasis on transmission and development aspects in fixed digital radio, both radio relay and mobile communications. For the last 15 years he has been mainly concerned with the performance analysis, development of planning tools and equipment for mobile communication systems and he has published about two hundred papers in that areas. He participated in the European program COST 231 and in the COST 259 as Spanish representative delegate. He has also participated in the RACE and ACTS European research programs and currently in the IST as well as in many private and public funded projects. He received the Catalonia Engineer of the year prize in 1998 and the Narcis Monturiol Medal issued by the Government of Catalonia in 2002 for his research contributions to the mobile communications field. He is part of the editorial board of several Scientific International Journals and since 1995 is conducting a post graduate annual course on mobile communications. He co-authored two books on Mobile communications.